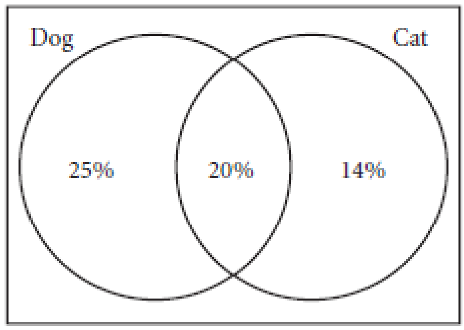
Chapter 9

Discussion Question Solutions

D1. Use a Venn diagram to summarize the given information



a. From the diagram, households that contain either a dog, a cat, or both comprise 59% of the sample, so 41% own neither a cat or a dog.

b. No, $P(own\ a\ dog) = 0.45$, but

$$P(own\ a\ dog \mid own\ a\ cat) = \frac{P(own\ a\ dog \text{ and } own\ a\ cat)}{P(own\ a\ cat)} = \frac{0.2}{0.34} \approx 0.588$$

So, households that have a cat are more likely to have a dog than households are overall. Alternatively, students can look at these probabilities:

$P(own\ a\ dog \text{ and } own\ a\ cat) = 0.20$, but

$P(own\ a\ dog) \cdot P(own\ a\ cat) = (0.45)(0.34) = 0.153$.

Because the Multiplication Rule for Independent Events doesn't hold, the two events aren't independent.

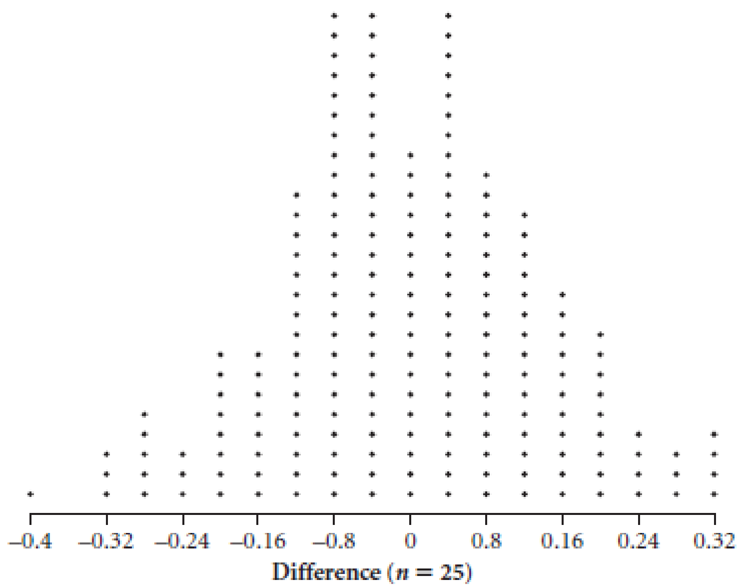
c. No. The two percentages, 45% and 34%, did not come from independent samples.

D2. The margin of error is given by $z^*\sqrt{\frac{p(1-p)}{n}}$. For a 95% confidence interval z^* is 1.96 or just less than 2. The maximum standard error occurs when $p = 0.5$ so the margin of error cannot be greater than $2\sqrt{\frac{0.5(1-0.5)}{n}} = 2\frac{\sqrt{0.25}}{\sqrt{n}} = \frac{2\cdot 0.5}{\sqrt{n}} = \frac{1}{\sqrt{n}}$.

Similarly for the difference of two population proportions, the 95% margin of error is $1.96\sqrt{\frac{p_1(1-p_1)}{n} + \frac{p_2(1-p_2)}{n}}$. The margin of error cannot be larger than

$$2\sqrt{\frac{0.5(1-0.5)}{n} + \frac{0.5(1-0.5)}{n}} = 2\sqrt{\frac{0.25}{n} + \frac{0.25}{n}} = 2\sqrt{\frac{0.5}{n}} = 2\frac{\sqrt{1/2}}{\sqrt{n}} = \frac{\sqrt{2}}{\sqrt{n}}.$$

D3. a. A sample dot plot is shown here.



b. The mean of the distribution in the dot plot in part a is –0.00240. Students' answers also should be close to 0. The standard error of the distribution in the dot plot is 0.13941. Students' answers should be close to 0.135.

c. The theoretical value of the mean is $\mu_{\hat{p}_1-\hat{p}_2} = p_1 - p_2 = 0.35 - 0.35 = 0, \hat{p}_1$ which is close to $-0.00240,$ the estimate in part b. The theoretical value of the standard error is

$$\sigma_{\hat{p}_1-\hat{p}_2} = \sqrt{\frac{p_1(1-p_1)}{n_1} + \frac{p_2(1-p_2)}{n_2}} = \sqrt{\frac{0.35(1-0.35)}{25} + \frac{0.35(1-0.35)}{25}} \approx 0.135$$

which is quite close to 0.139, the estimate in the simulation.

D4. a. The null hypothesis was rejected that there is no difference between the proportion of all men and the proportion of all women who are lefthanded. If this hypothesis is true, a Type I error was made. The chance of making this type of error is equal to the significance level, which in most cases is small.

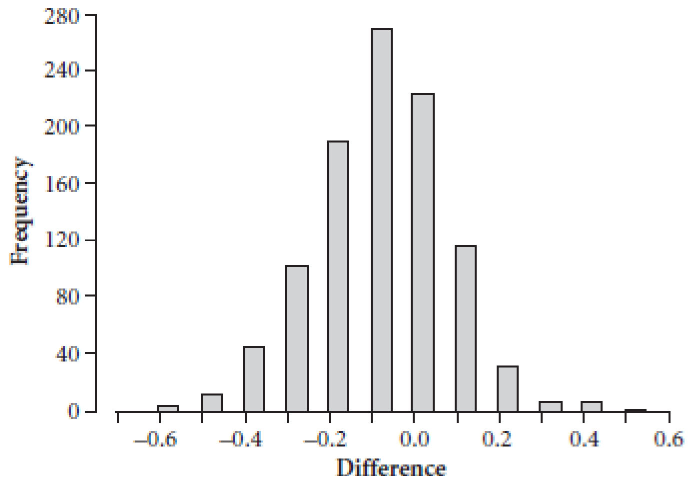
b. These are the counts of left-handers and righthanders in each population, and they clearly are all at least 5.

D5. In Example 9.4, the hypothesis was not rejected. If the hypothesis is actually false then a Type II error was made. The hypothesis was that the percentage of overweight adults in 2004 is not significantly different from the percentage in 2009. Committing a Type II error could cause a growing problem to be over looked or taken too lightly.

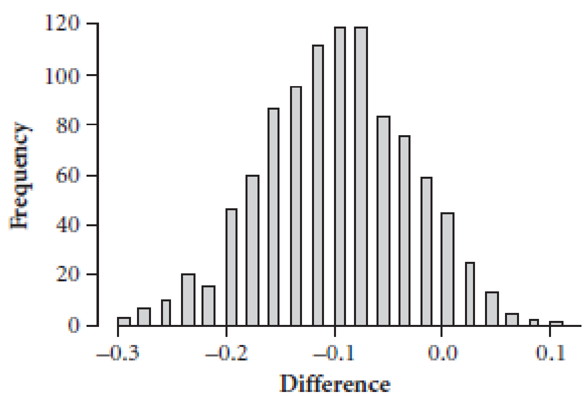
D6. **a.** The difference in the proportions in the two samples is small enough that it could reasonably have come from two populations with equal proportions of successes. This possibility suggests that either the two population proportions are equal or the sample sizes aren't large enough to distinguish between the two populations.

b. The difference in the proportions in the two samples is large enough that it isn't reasonable
to assume that the samples came from two populations with equal proportions of successes.

D7. For the sample of size 10, n_1p_1 and n_2p_2 are equal to 1 and 2, respectively. And the distribution is somewhat skewed, as shown by this histogram. In addition, the gaps between possible sample proportions are large and the tails are quite short, which makes calculating probabilities by using a continuous normal distribution not very accurate.



On the other hand, for the samples of size 50, each of n_1p_1, $n_1(1-p_1)$, n_2p_2, and $n_2(1-p_2)$ is at least 5. The distribution of $\hat{p}_1-\hat{p}_2$ is more "filled in." This makes the distribution of the difference closer to a continuous graph, and generally looks approximately normal. The table of the standard normal distribution works reasonably well to estimate probabilities.



D8. In a two-population sample survey, a sample is randomly and independently selected from each population being studied. Conclusions can then be made about the populations from which the samples were taken. In a two-treatment experiment, the treatments are randomly assigned to the population of volunteers or experimental units, which are not a random sample from a larger population. Here, conclusions can be made about the effects of the treatments on this group of experimental units only.

D9. In both cases you compare two proportions, say, p_1 and p_2. The null hypothesis is usually that $p_1 = p_2$. The difference between two-population sample surveys and two-treatment experiments is what the proportions represent. In a two-population sample survey, each proportion is the proportion you would get if you could ask everyone in that particular population the survey question. In a two-treatment experiment, each proportion is the proportion of success you would observe if *all* the experimental units (in all treatment groups) could be given each treatment.

D10. No, the proportions who improve could be 0.07 and 0.02. No, the proportions could be 0.99 and 0.94, making Treatment B good in almost all situations.

D11. "Double-blind" means that neither the physician (who will decide whether the patient has developed AIDS) nor the patient know whether the patient is receiving AZT or AZT + ACV. "Randomized" means that the patients are assigned randomly to the two treatments. "Clinical trial" means a comparative experiment to evaluate a medical treatment that is based on actual patients in realistic situations.

D12. In Example 9.5, the hypothesis was not rejected. If the hypothesis is actually false then a Type II error was made. In this case, the hypothesis was that two types of respiratory protection do not differ in their ability to protect someone from the flu virus. If a Type II error is committed, the hypothesis is false and one type of protection *is* better than the other. Thus, users of the other type are more at risk.

D13. The null hypothesis was rejected that there is no difference between the responses to the two treatments. If this hypothesis is true, a Type I error was made. This study might affect the treatment of patients. If a Type I error was made, that means AZT + ACV really was no more effective than AZT alone, but decisions will now be made as if AZT + ACV really was more effective. This could be costly at best, and dangerous at worst. For this reason, medical studies will usually be repeated in a variety of contexts to be sure the conclusions are valid.

D14. When the treatments are randomly assigned to subjects, the only consistent difference between subjects is the treatment, provided the researchers are careful to treat subjects alike, except for the treatments. This allows you to draw conclusions about cause and effect. With observational studies, the conditions of interest come already built into the subjects being studied, so the groups of subjects frequently come with other conditions in common. This means it is impossible to tell whether a difference in the response is due to the condition of interest or confounding variables. However, you can still use these procedures to answer the question, "Could the result I see in the observed data have reasonably happened by chance?" If the answer is no, then there is evidence of an effect that should be investigated further.

D15. Ideally, they should be set up as randomized experiments. Randomized experiments are needed to draw any valid conclusion about the effects of the treatments. Randomized experiments are difficult to implement in an educational setting, but probably more problematic is that parents do not want researchers "experimenting" on their children (even though the teaching methods currently in use have not been established as effective).

Practice Problem Solutions

P1. a. As before, the samples can be considered random samples, and the samples were selected independently of each other. Each of

$$n_1\hat{p}_1 = 100\cdot 0.56 = 56 \qquad n_1(1-\hat{p}_1) = 100\cdot 0.44 = 44$$
$$n_2\hat{p}_2 = 100\cdot 0.63 = 63 \qquad n_2(1-\hat{p}_2) = 100\cdot 0.37 = 37$$

are at least 5, where n_1 and n_2 are the numbers of households sampled in 1994 and this year, respectively, and $\hat{p}_1$ and $\hat{p}_2$ are the proportions of households in 1994 and this year, respectively, that had a pet. The number of U.S. households in each year is larger than 10 times 100 or 1000.

b. The confidence interval is

$$(\hat{p}_1-\hat{p}_2)\pm z^*\sqrt{\frac{\hat{p}_1(1-\hat{p}_1)}{n_1}+\frac{\hat{p}_2(1-\hat{p}_2)}{n_2}} = (0.56-0.63)\pm 1.96\sqrt{\frac{0.56(1-0.56)}{100}+\frac{0.63(1-0.63)}{100}}$$
$$\approx\ -0.07\ \pm\ 0.136$$

or about –0.206 to 0.066. You are 95% confident that the difference in the two rates of pet ownership is between –0.206 to 0.066. This means that it is plausible that *20.6% less* households owned pets in 1994 than own pets now, and it is also plausible that 6.6% more households owned a pet in 1994 than own a pet now.

c. Yes. A difference of 0 does lie within the confidence interval. This means that if the difference in the proportion of pet owners now and in 1994 is actually 0, getting a difference of –0.07 in the samples is reasonably likely. Thus, it is plausible that there is no difference between the proportion of all households that owned a pet in 1994 and the proportion of all households that own a pet now. There is insufficient evidence to support a claim that there was a change in the percentage of households that own a pet between 1994 and now.

P2. a. The Gallup poll uses what can be considered a simple random sample. The populations are binomial (answering "yes" or "no"), and the samples would be independent of each other. Each of $n_1\hat{p}_1 = 325 \bullet 0.5 = 162.5$, $n_1(1-\hat{p}_1) = 325 \bullet 0.5 =$

162.5, $n_2 \hat{p}_2 = 224 \cdot 0.28 = 62.72$, $n_2(1 - \hat{p}_2) = 224 \cdot 0.72 = 161.28$ are at least five. There are more than 325 • 10 = 3,250 13 to 15-year olds and more than 224 • 10 = 2,240 16 to 17-year olds in the U.S. The conditions for a confidence interval for the difference of two proportions are met.

b.

$$(\hat{p}_1 - \hat{p}_2) \pm z^* \sqrt{\frac{\hat{p}_1(1-\hat{p}_1)}{n_1} + \frac{\hat{p}_2(1-\hat{p}_2)}{n_2}} = (0.50 - 0.28) \pm 1.96\sqrt{\frac{0.50(1-0.50)}{325} + \frac{0.28(1-0.28)}{224}}$$
$$\approx 0.22 \pm 0.08$$

or about 0.14 to 0.30.

c. You are 95% confident that the difference between the proportion of all 13 to 15-year olds who respond, "yes" and the proportion of all 16 to 17-year olds who would respond, "yes" to the question that it was appropriate for parents to install a special device on the car to allow parents to monitor teenagers' driving speeds is between 14% and 30%.

d. 0 is not in the confidence interval, which implies that it is *not* plausible that there is no difference between these proportions. This means that if the difference in the proportion of 13 to 15-year olds who respond, "yes" and the proportion of all 16 to 17-year olds who would respond, "yes" is actually 0, getting a difference of 0.14 in the samples is not at all likely. Thus, you are convinced that there is a difference in opinion between these two ages groups on this question.

P3. As before, the samples can be considered random samples, and the samples were selected independently of each other. Here,

$$n_1 = 1000,\ \ \hat{p}_1 = 0.34,\ n_2 = 800,\ \ \hat{p}_2 = 0.18$$

and certainly each of the products $n_1\hat{p}_1,\ n_1(1-\hat{p}_1),\ n_2\hat{p}_2,$ and $n_2(1-\hat{p}_2)$ is at least 5.

The 95% confidence interval is

$$(\hat{p}_1 - \hat{p}_2) \pm z^* \sqrt{\frac{\hat{p}_1(1-\hat{p}_1)}{n_1} + \frac{\hat{p}_2(1-\hat{p}_2)}{n_2}} = (0.34 - 0.18) \pm 1.96\sqrt{\frac{0.34(1-0.34)}{1000} + \frac{0.18(1-0.82)}{800}}$$

or about (0.121, 0.199).

This means that you are 95% confident that the difference between the proficiency percentages for 4-year and 2-year colleges is between 12% and 20%.

P4. As before, the samples can be considered random samples, and the samples were selected independently of each other. Here,

$$n_1 = 273,\ \ \hat{p}_1 = \tfrac{48}{273} = 0.176,\ n_2 = 442,\ \ \hat{p}_2 = \tfrac{28}{442} = 0.063$$

and certainly each of the products $n_1\hat{p}_1,\ n_1(1-\hat{p}_1),\ n_2\hat{p}_2,$ and $n_2(1-\hat{p}_2)$ is at least 5.

The 95% confidence interval is

$$(\hat{p}_1 - \hat{p}_2) \pm z^* \sqrt{\frac{\hat{p}_1(1-\hat{p}_1)}{n_1} + \frac{\hat{p}_2(1-\hat{p}_2)}{n_2}} = (0.176 - 0.063) \pm 1.96\sqrt{\frac{0.176(1-0.176)}{273} + \frac{0.063(1-0.063)}{442}}$$

or about (0.062, 0.164).

This means that you are 95% confident that the difference between the percentages playing video games online for DC versus DV types is between 6.2% and 16.4%.

P5. a. As before, the samples can be considered random samples, and the samples were selected independently of each other. Here,

$$n_1 = 663,\ \hat{p}_1 = 0.35,\ n_2 = 1591,\ \hat{p}_2 = 0.29$$

and certainly each of the products $n_1\hat{p}_1,\ n_1(1-\hat{p}_1),\ n_2\hat{p}_2$, and $n_2(1-\hat{p}_2)$ is at least 5.

b. The 95% confidence interval is

$$(\hat{p}_1 - \hat{p}_2) \pm z^* \sqrt{\frac{\hat{p}_1(1-\hat{p}_1)}{n_1} + \frac{\hat{p}_2(1-\hat{p}_2)}{n_2}} = (0.35 - 0.29) \pm 1.96\sqrt{\frac{0.35(0.65)}{663} + \frac{0.29(0.71)}{1591}}$$

or about (0.017, 0.103).

c. You are 95% confident that the difference between the proportion of those who Twitter who live in urban areas the proportion of Internet users who live in urban areas is between 0.017 and 0.103.

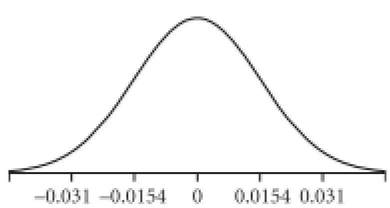
d. No; yes, because only positive differences are in the CI

e. If you were to repeat the process of taking two random samples and constructing a confidence interval for the difference over and over, in the long run, you expect that 95% of them contain the true difference in the proportion of people who live in urban areas.

P6. a. The expected value would be $p_1 - p_2 = 0.12 - 0.12 = 0$.

b. The *SE* is $\sqrt{\frac{p_1(1-p_1)}{n_1} + \frac{p_2(1-p_2)}{n_2}} = \sqrt{\frac{0.12 \cdot 0.88}{1000} + \frac{0.12 \cdot 0.88}{800}} \approx 0.0154$

c. Since n_1p_1, $n_1(1-p_1)$, n_2p_2, and $n_2(1-p_2)$ are all at least 5 you can approximate the sampling distribution of the difference with a normal model with mean 0 and *SD* 0.514.



d. You can use your calculator. **normalcdf(0.05,1E99,0,0.0154)** will give approximately 0.00058. Alternatively, you can use the *z*-score.

$$z = \frac{(\hat{p}_1 - \hat{p}_2) - (p_1 - p_2)}{\sqrt{\frac{p_1(1-p_1)}{n_1} + \frac{p_2(1-p_2)}{n_2}}} = \frac{0.05 - 0}{\sqrt{\frac{0.12 \cdot 0.88}{1000} + \frac{0.12 \cdot 0.88}{800}}} \approx 3.244.$$

According to Table A, the probability of a z-score greater than 3.24 is approximately 0.0006.

P7. a. False. The values of $\hat{p}_1$ and $\hat{p}_2$ vary from sample to sample.

b. True. We are told that the proportion of successes in the two populations are equal.

c. True.

d. True. We have $\mu_{\hat{p}_1 - \hat{p}_2} = p_1 - p_2 = 0$

e. True. As you can see from the dot plots in the display, the sample differences $\hat{p}_1 - \hat{p}_2$ have less variability in the samples of size 100 than in the samples of size 30. They cluster more closely to 0, so we can see there is more of a chance of having the sample difference $\hat{p}_1 - \hat{p}_2$ nearer 0 with a larger sample size.

P8. This situation calls for a one-sided significance test for the difference of two proportions because we are asked whether the data support the conclusion that there was a decrease in voter support for the candidate.

Check conditions. You are told that you have two random samples from a large population (potential voters in some city). It's reasonable to assume that the samples are independent. For the first survey n_1= 600 and $\hat{p}_1 = \frac{321}{600} = 0.535$. For the second survey, n_2 = 750 and $\hat{p}_1 = \frac{382}{750} \approx 0.509$. Each of $n_1 \hat{p}_1$ = 321, $n_1(1 - \hat{p}_1)$ = 279, $n_2 \hat{p}_2$ = 382, and $n_2(1 - \hat{p}_2)$ = 368 is at least 5. The number of potential voters at both times is much larger than 10 times the sample size for both samples.

State your hypotheses.

H_0: The proportion, p_1, of potential voters who favored the candidate in the first survey is equal to the proportion, p_2, of potential voters who favored the candidate one week before the election, or $p_1 = p_2$.

H_a: The proportion, p_1, of potential voters who favored the candidate in the first survey is greater than the proportion, p_2, of potential voters who favored the candidate one week before the election, or $p_1 > p_2$.

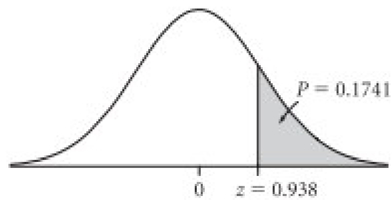
Compute the test statistic and draw a sketch. The test statistic is

$$z = \frac{(\hat{p}_1 - \hat{p}_2) - (p_1 - p_2)}{\sqrt{\hat{p}(1-\hat{p})\left(\frac{1}{n_1} + \frac{1}{n_2}\right)}} \approx \frac{(0.535 - 0.5093) - 0}{\sqrt{0.521(1-0.521)\left(\frac{1}{600} + \frac{1}{750}\right)}} \approx 0.939$$

where

$$\hat{p} = \frac{\textit{total number of successes in both samples}}{n_1 + n_2} = \frac{321 + 382}{600 + 750} \approx 0.521$$

Using the table, the P-value for this one-sided test is 0.1736. From the TI-84+, the test statistic is $z = 0.938$ and the P-value is 0.1741. In this case, the **2-PropZTest** gives us the most accurate answer because there is less rounding.



Write a conclusion in context. If there is no difference between the proportion of potential voters who favored the candidate at three weeks and the proportion of potential voters who favored the candidate at one week, then there is a 0.1741 chance of getting a difference of 0.0257 or larger with samples of these sizes. This difference is not statistically significant—it can reasonably be attributed to chance variation. We do not reject the null hypothesis and can not conclude that there has been a drop in support for the new candidate.

P9. a. H_o: $p_1 - p_2 = 0$, Ha: $p_1 - p_2 \neq 0$, where p_1 is the proportion of conforming pellets from Method A and p_2 is the proportion of conforming pellets from Method B.

b. Here,

$$n_1 = n_2 = 100,\ \ \hat{p}_1 = 0.38 \text{ (Method A)},\ \ \hat{p}_2 = 0.29\,.$$

and certainly each of the products $n_1\hat{p}_1$, $n_1(1-\hat{p}_1)$, $n_2\hat{p}_2$, and $n_2(1-\hat{p}_2)$ is at least 5. Also, the pooled estimate is

$$\hat{p} = \frac{\text{total number of successes from both treatments}}{n_1 + n_2} = \frac{67}{200} \approx 0.335\,.$$

So, the test statistic is

$$z = \frac{(\hat{p}_1 - \hat{p}_2) - (p_1 - p_2)}{\sqrt{\hat{p}(1-\hat{p})\left(\frac{1}{n_1} + \frac{1}{n_2}\right)}} = \frac{(0.38 - 0.29) - 0}{\sqrt{0.335(1-0.335)\left(\frac{1}{100} + \frac{1}{100}\right)}} \approx 1.348\,.$$

c. Since this is a two-tailed test, the p-value is $2P(Z > 1.348) = 2(0.0889) = 0.1778$.

d. There is insufficient evidence to conclude that the proportion of conforming pellets from Method A differs from the proportion of conforming pellets from Method B.

P10. a. Construct 95% confidence intervals for each of the 5 behaviors. If the confidence interval contains 0, then that behavior does not yield a statistically significant difference.

Registered to vote:

$$(\hat{p}_1 - \hat{p}_2) \pm z^* \cdot \sqrt{\frac{\hat{p}_1(1-\hat{p}_1)}{n_1} + \frac{\hat{p}_2(1-\hat{p}_2)}{n_2}} = (0.89 - 0.78) \pm 1.96\sqrt{\frac{(0.89)(0.11)}{3011} + \frac{(0.78)(0.22)}{1055}}$$

$$= 0.11 \pm 0.027, \text{ or } \ 0.083 \text{ to } 0.137$$

Active in community:

$$(\hat{p}_1 - \hat{p}_2) \pm z^* \cdot \sqrt{\frac{\hat{p}_1(1-\hat{p}_1)}{n_1} + \frac{\hat{p}_2(1-\hat{p}_2)}{n_2}} = (0.68 - 0.41) \pm 1.96\sqrt{\frac{(0.68)(0.32)}{3011} + \frac{(0.41)(0.59)}{1055}}$$
$$= 0.27 \pm 0.034, \text{ or } 0.236 \text{ to } 0.304$$

Suffer from personal addiction:

$$(\hat{p}_1 - \hat{p}_2) \pm z^* \cdot \sqrt{\frac{\hat{p}_1(1-\hat{p}_1)}{n_1} + \frac{\hat{p}_2(1-\hat{p}_2)}{n_2}} = (0.12 - 0.13) \pm 1.96\sqrt{\frac{(0.12)(0.88)}{3011} + \frac{(0.13)(0.87)}{1055}}$$
$$= -0.01 \pm 0.023, \text{ or } -0.033 \text{ to } 0.013$$

Overweight:

$$(\hat{p}_1 - \hat{p}_2) \pm z^* \cdot \sqrt{\frac{\hat{p}_1(1-\hat{p}_1)}{n_1} + \frac{\hat{p}_2(1-\hat{p}_2)}{n_2}} = (0.41 - 0.26) \pm 1.96\sqrt{\frac{(0.41)(0.59)}{3011} + \frac{(0.26)(0.74)}{1055}}$$
$$= 0.15 \pm 0.032, \text{ or } 0.118 \text{ to } 0.182$$

Stressed out:

$$(\hat{p}_1 - \hat{p}_2) \pm z^* \cdot \sqrt{\frac{\hat{p}_1(1-\hat{p}_1)}{n_1} + \frac{\hat{p}_2(1-\hat{p}_2)}{n_2}} = (0.26 - 0.37) \pm 1.96\sqrt{\frac{(0.26)(0.74)}{3011} + \frac{(0.37)(0.63)}{1055}}$$
$$= -0.11 \pm 0.033, \text{ or } -0.143 \text{ to } -0.077$$

So, all are significant except “suffer from personal addiction.”

b. The strongest evidence is given by “active in community” since it is the furthest away from 0.

c. Activist Christians may have been more willing to participate than non-Christian activists, especially those from higher economic strata.

P11. a. These are not independent because men and women are paired.
b. These are independent because the sampling is done with replacement.

c. These are nearly independent because the men and women are not paired, and can be considered independent for this large population.

P12. Proceed as follows:

Check conditions. First, the conditions for an experiment are met, which allows the computing of a confidence interval for the difference of two proportions: Treatments were randomly assigned to subjects. Each of $n_1\hat{p}_1$ = 169, $n_1(1 - \hat{p}_1)$ = 10,868, $n_2\hat{p}_2$ = 138, and $n_2(1 - \hat{p}_2)$ = 10,896 is at least 5.

Do computations. The 95% confidence interval is:

$$(\hat{p}_1 - \hat{p}_2) \pm z^* \cdot \sqrt{\frac{\hat{p}_1(1-\hat{p}_1)}{n_1} + \frac{\hat{p}_2(1-\hat{p}_2)}{n_2}} = (0.0153 - 0.0125) \pm 1.96\sqrt{\frac{(0.0153)(0.9847)}{11{,}037} + \frac{(0.0125)(0.9875)}{11{,}034}}$$
$$= 0.0028 \pm 0.0031, \text{ or } -0.0003 \text{ to } 0.0059$$

Write a conclusion in context. Suppose all of the subjects could have been given the aspirin treatment and all of the subjects could have been given the placebo treatment.

Then you are 95% confident that the difference in the proportion who would get ulcers is in the interval (-0.0003, 0.0059). Because 0 is in this interval, it is plausible that there is no difference in the proportions who would get ulcers. The term "95% confident" means that this method of constructing confidence intervals results in $p_1 - p_2$ falling in an average of 95 out of every 100 confidence intervals you construct.

P13. The hypotheses are:
H_0: $p_1 - p_2 = 0$, H_a: $p_1 - p_2 \neq 0$, where p_1 is the proportion of successes if all subjects could have been asked for a quarter and p_2 is the proportion of successes if all subjects could have been asked for 17 cents.

Here,

$$n_1 = 72,\ \ \hat{p}_1 = 0.306,\ n_2 = 72,\ \ \hat{p}_2 = 0.431$$

and certainly each of the products $n_1\hat{p}_1,\ n_1(1-\hat{p}_1),\ n_2\hat{p}_2$, and $n_2(1-\hat{p}_2)$ is at least 5. The randomness conditions are also met.

Also, the pooled estimate is

$$\hat{p} = \frac{\text{total number of successes from both treatments}}{n_1 + n_2} = \frac{22.032 + 31.032}{72 + 72} \approx 0.369 .$$

So, the test statistic is

$$z = \frac{(\hat{p}_1 - \hat{p}_2) - (p_1 - p_2)}{\sqrt{\hat{p}(1-\hat{p})\left(\frac{1}{n_1} + \frac{1}{n_2}\right)}} = \frac{(0.431 - 0.306) - 0}{\sqrt{0.369(1-0.369)\left(\frac{1}{72} + \frac{1}{72}\right)}} \approx 1.555$$

Since this is a two-sided test, the p-value is $2P(Z > 1.555) = 0.1212$. Hence, there is insufficient evidence, at the 5% level, to say that asking for 17 cents will increase the percentage of success over asking for 25 cents.

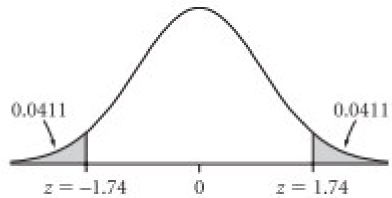
P14. a. Because we are simply looking for a difference we will use a two-sided significance test for a difference in proportions.

Check conditions. The problem does not state whether treatments were randomly assigned. The other condition is met, however. Each of $n_1\hat{p}_1 = 411$, $n_1(1 - \hat{p}_1) = 4009$, $n_2\hat{p}_2 = 463$, and $n_2(1 - \hat{p}_2) = 3989$ is at least 5.

State your hypotheses.
H_0: If all patients could have been given Lipitor, the proportion p_1 of them that had heart attacks would be the same as the proportion p_2 that would have had heart attacks had they all been given Zocor.
H_a: $p_1 \neq p_2$

Calculate the test statistic and draw a sketch.

$$z=\frac{(\hat{p}_1-\hat{p}_2)-(p_1-p_2)}{\sqrt{\hat{p}(1-\hat{p})\left(\frac{1}{n_1}+\frac{1}{n_2}\right)}}=\frac{(0.093-0.104)-0}{\sqrt{0.0985(1-0.0985)\left(\frac{1}{4420}+\frac{1}{4452}\right)}}\approx-1.738$$



Here the pooled estimate is

$$\hat{p}=\frac{\text{total number of successes from both treatments}}{n_1+n_2}=\frac{874}{8872}\approx 0.0985$$

The z-score of –1.738 corresponds to a P-value of 2 • 0.0411 = 0.0822.
A 2-PropZTest on the TI-84+ gives a z-score of –1.7402 and a P-value of 0.0818.

State your conclusion in context. Because the P-value is greater than 0.05 you would not reject the null hypothesis. There is not sufficient evidence to conclude that, if all experimental units would have been treated with Lipitor, the proportion who had heart attacks would have been different than if all patients had been treated with Zocor.

b. The conditions and test statistic will be the same as in part a. You need to restate your hypotheses, calculate the new P-value, and state your conclusions.

State your hypotheses.
H_0: If all patients could have been given Lipitor, the proportion p_1 of them that had heart attacks would be the same as the proportion p_2 that would have had heart attacks had they all been given Zocor.
H_a: $p_1 < p_2$ (If Lipitor is more effective, you would expect the proportion of patients having heart attacks to be lower.)

P-value. The test statistic is still –1.738. The P-value is now half what it was for a two-sided test. A 2-PropZTest for alternative hypothesis $p_1 < p_2$ now shows a P-value of 0.0409. The sketch in part (a) would be shaded only in the left tail.

State your conclusion in context. The P-value of 0.0409 is less than 0.05. You would reject the null hypothesis that if all patients could have been given Lipitor, the proportion of them that had heart attacks would be the same as the proportion that would have had heart attacks had they all been given Zocor. If there would have been no difference in the proportion of patients who had heart attacks if they had all taken Lipitor and the proportion of patients who had heart attacks if they had all taken Zocor, then there is a 0.0409 chance of getting a difference of –1.74 or smaller in the proportions from random assignment of these treatments to the subjects. This difference can not be reasonably attributed to chance variation. There is evidence that Lipitor is more effective than Zoloc.

A one-sided test makes it easier to reject the null hypothesis if the difference is in the direction your alternative hypothesis states. Mathematically, this happens because the entire 5% rejection region is on that side, meaning a less extreme z-score will allow rejection. Philosophically, the fact that you suspect one direction may be due to evidence in favor of that alternative hypothesis. Less additional evidence is needed to verify this.

P15. Because the null hypothesis was not rejected in Part A, a Type II error could have been made. In part B the null hypothesis was rejected, so a Type I error could have occurred. A Type I error would mean the patient receives a different drug even though there is no actual difference in their effectiveness. A Type II error means a patient is not given a new drug that would actually have a better chance of success. Both could be serious errors as both mean the patient is not receiving the most effective drug.

P16. In all cases below, the hypotheses (in symbols) are:
H_0: $p_1 - p_2 = 0$, H_a: $p_1 - p_2 \neq 0$.

TV in bedroom: Here,

$$n_1 = 92,\ \hat{p}_1 = 0.435,\ n_2 = 100,\ \hat{p}_2 = 0.43$$

and certainly each of the products $n_1\hat{p}_1,\ n_1(1-\hat{p}_1),\ n_2\hat{p}_2,$ and $n_2(1-\hat{p}_2)$ is at least 5. The randomness conditions are also met.

Also, the pooled estimate is

$$\hat{p} = \frac{\text{total number of successes from both treatments}}{n_1 + n_2} = \frac{40.02 + 43}{100 + 92} \approx 0.432.$$

So, the test statistic is

$$z = \frac{(\hat{p}_1 - \hat{p}_2) - (p_1 - p_2)}{\sqrt{\hat{p}(1-\hat{p})\left(\frac{1}{n_1} + \frac{1}{n_2}\right)}} = \frac{(0.435 - 0.43) - 0}{\sqrt{0.432(1-0.432)\left(\frac{1}{92} + \frac{1}{100}\right)}} \approx 0.070$$

Since this is a two-sided test, the p-value is $2P(Z > 0.070) = 0.9442$. Hence, there is insufficient evidence, at the 5% level, to say these proportions are different.

College grads: Here,

$$n_1 = 92,\ \hat{p}_1 = 0.45,\ n_2 = 100,\ \hat{p}_2 = 0.21$$

and certainly each of the products $n_1\hat{p}_1,\ n_1(1-\hat{p}_1),\ n_2\hat{p}_2,$ and $n_2(1-\hat{p}_2)$ is at least 5. The randomness conditions are also met.

Also, the pooled estimate is

$$\hat{p} = \frac{\text{total number of successes from both treatments}}{n_1 + n_2} = \frac{21 + 41.4}{100 + 92} \approx 0.325.$$

So, the test statistic is

$$z = \frac{(\hat{p}_1 - \hat{p}_2) - (p_1 - p_2)}{\sqrt{\hat{p}(1-\hat{p})\left(\frac{1}{n_1} + \frac{1}{n_2}\right)}} = \frac{(0.45 - 0.21) - 0}{\sqrt{0.325(1-0.325)\left(\frac{1}{92} + \frac{1}{100}\right)}} \approx 3.547$$

Since this is a two-sided test, the p-value is $2P(Z > 3.547) = 0.0004$. Hence, there is sufficient evidence, at the 5% level, to the difference in these proportions is not attributed to chance.

Female participant: Here,

$$n_1 = 92,\ \ \hat{p}_1 = 0.45,\ n_2 = 100,\ \ \hat{p}_2 = 0.485$$

and certainly each of the products $n_1\hat{p}_1,\ n_1(1-\hat{p}_1),\ n_2\hat{p}_2$, and $n_2(1-\hat{p}_2)$ is at least 5. The randomness conditions are also met.

Also, the pooled estimate is

$$\hat{p} = \frac{\text{total number of successes from both treatments}}{n_1 + n_2} = \frac{41.4 + 48.5}{100 + 92} \approx 0.468.$$

So, the test statistic is

$$z = \frac{(\hat{p}_1 - \hat{p}_2) - (p_1 - p_2)}{\sqrt{\hat{p}(1-\hat{p})\left(\frac{1}{n_1} + \frac{1}{n_2}\right)}} = \frac{(0.45 - 0.485) - 0}{\sqrt{0.468(1-0.468)\left(\frac{1}{92} + \frac{1}{100}\right)}} \approx -0.4855$$

Since this is a two-sided test, the p-value is $2P(Z < -0.4855) = 0.6273$. Hence, there is insufficient evidence, at the 5% level, to say these proportions are different.

Thus, we see that only the "college grads" issue shows a difference that could not be easily relegated to chance alone.

b. Randomization to the larger units (schools) rather than to the smaller units (students) generally is not a good idea because it reduces the number of randomly assigned units and that reduces the effective sample size. In the extreme, if all the students in one school of 500 students acted alike, the result would be one new piece of information for the school rather than 500 pieces of information that could have been obtained if the 500 students had been randomly selected from a large population of students.

P17. a. This is an observational study. There was no random sampling done, and no random assignment of treatments.

b. ***Check conditions.*** We already know there was no random assignment of treatments. Each of $n_1 \cdot \hat{p}_1 = 103$, $n_1(1 - \hat{p}_1) = 805$, $n_2\hat{p}_2 = 53$, and $n_2(1 - \hat{p}_2) = 614$ is at least 5, where $\hat{p}_1$ is the proportion of the n_1 people observed who had been abused as children who later went on to commit violent crime, and $\hat{p}_2$ is the proportion of the n_2 people observed who had not been abused as children who later went on to commit violent crime. This second condition is met.

Calculate the interval.

$$(\hat{p}_1-\hat{p}_2)\pm z^*\sqrt{\frac{\hat{p}_1(1-\hat{p}_1)}{n_1}+\frac{\hat{p}_2(1-\hat{p}_2)}{n_2}}=(0.113-0.079)\pm 1.645\sqrt{\frac{0.113\cdot 0.887}{908}+\frac{0.079\cdot 0.921}{667}}$$

$$\approx 0.034 \pm 0.024$$

or about from about 0.01 to 0.058.

c. We can conclude that the difference in proportions of the people in this study that were abused as children who later committed crimes and the people in this study who were not abused as children who later committed crimes cannot be reasonably attributed to chance. There may be, and probably are, many other factors that contributed to this difference, so we cannot conclude from this study alone that abuse of children causes them to be more likely to commit violent crime later in life.

P18. The riders on greenways show the strongest association between helmet use and the law. Note that this still does not imply causation.

Exercise Solutions

E1. B and C

E2. Here,

$$n_1 = 332,\ \ \hat{p}_1 = 0.69,\ n_2 = 798,\ \ \hat{p}_2 = 0.46$$

and certainly each of the products $n_1\hat{p}_1,\ n_1(1-\hat{p}_1),\ n_2\hat{p}_2$, and $n_2(1-\hat{p}_2)$ is at least 5.

The 95% confidence interval is given to be (0.169, 0.290).

Choice A is a reasonable interpretation of the confidence interval given that you expect 95% of the time for the difference in the proportions to lie in this interval. Choice C is true because both factors used to compute the standard error would be larger, so that the product $z^* \cdot SE$ would be larger, thereby widening the interval.

E3. a. You do not know that this is a random sample of all purebred dog owners or all mutt owners in the San Diego area. However, you could consider guessing to be a random event and you want to compare the probability of guessing correctly with purebred dog owners and mutt owners. Each of

$$n_1\hat{p}_1 = 16 \qquad n_1(1-\hat{p}_1) = 9$$

$$n_2\hat{p}_2 = 7 \qquad n_2(1-\hat{p}_2) = 13$$

are at least 5, where n_1 and n_2 are the numbers of guesses made with pure-bred dog owners and with mutt owners, respectively, and $\hat{p}_1$ and $\hat{p}_2$ are the proportions of correct guesses made with purebred dog owners and with mutt owners, respectively. There are probably more than 25 • 10 = 250 purebred dog owners and more than 20 • 10 = 200 mutt owners in the San Diego area.

b.

$$(\hat{p}_1-\hat{p}_2)\pm z^*\sqrt{\frac{\hat{p}_1(1-\hat{p}_1)}{n_1}+\frac{\hat{p}_2(1-\hat{p}_2)}{n_2}}=(0.64-0.35)\pm 1.96\sqrt{\frac{0.64(1-0.64)}{25}+\frac{0.35(1-0.35)}{20}},$$
$$\approx 0.29\pm 0.281$$
or about 0.009 to 0.571.

c. You are 95% confident that if the judges had been given a choice of two dogs for each owner, the difference in the proportion of correct guesses for all purebred owners and the proportion of correct guesses for all mutt owners in the San Diego area would be between 0.009 and 0.571.
d. No. This implies that the difference in the proportions in the study were probably not due to chance. You have sufficient evidence that there is a higher probability of judges guessing correctly with purebred owners than with mutt owners.

e. The researchers need to enlarge their sample sizes and to select dog-owners and judges randomly from their location of interest. As it is, we can't know whether these results mean anything in terms of guessing correctly or whether there is something distinctive about either San Diego dogs or the judges from San Diego that led to these results. Or perhaps there was something distinctive about the dogs and owners that were chosen that influenced the judges' guesses.

f. 95% of all possible samples would yield a difference in proportions that is between 0.009 and 0.571.

E4. a. You are told you can assume that the samples were independently and randomly selected. The samples were taken from binomial populations (they either feel comfortable or they don't). Letting n_1 and n_2 represent the sample size for 2009 and 2001, respectively, and $\hat{p}_1$ and $\hat{p}_2$ represent the sample proportions of people that were uncomfortable with the lack of face-to-face contact in those respective years, Here,
$$n_1=2000,\ \ \hat{p}_1=0.304,\ n_2=2000,\ \ \hat{p}_2=0.249$$
and certainly each of the products $n_1\hat{p}_1,\ n_1(1-\hat{p}_1),\ n_2\hat{p}_2,$ and $n_2(1-\hat{p}_2)$ is at least 5. The number of Americans in each year was more than ten times the given sample sizes.

b.
$$(\hat{p}_1-\hat{p}_2)\pm z^*\sqrt{\frac{\hat{p}_1(1-\hat{p}_1)}{n_1}+\frac{\hat{p}_2(1-\hat{p}_2)}{n_2}}=(0.304-0.249)\pm 2.576\sqrt{\frac{0.304(1-0.304)}{2000}+\frac{0.249(1-0.249)}{2000}}$$
$$\approx 0.055\pm 0.0364$$
or about 0.0186 to 0.0914.

c. You are 99% confident that the difference between the proportion of all Americans that were uncomfortable with the lack of face-to-face contact in 2009 and the proportion of all Americans who were uncomfortable with the lack of face-to-face contact in 2001 is in the interval 0.0186 and 0.0914 or between 1.86% and 9.14%.

d. No, 0 is not in the interval, which implies that the statement, "the proportion of all

Americans who were uncomfortable with the lack of face-to-face contact did not change from 2001 to 2009" is not plausible. Because this confidence interval does not include zero, you are confident that the percentage of those who were uncomfortable with the lack of face-to-face contact has increased in the U.S. **e.** 99% of all possible samples would yield a difference in proportions that is between 0.0186 and 0.571.
E5. a. Here, $n_1 = 4775,\ \hat{p}_1 = 0.52,\ n_2 = 2685,\ \hat{p}_2 = 0.61$ and certainly each of the products $n_1\hat{p}_1,\ n_1(1-\hat{p}_1),\ n_2\hat{p}_2$, and $n_2(1-\hat{p}_2)$ is at least 5. The 90% confidence interval is $(\hat{p}_1 - \hat{p}_2) \pm z^* \sqrt{\frac{\hat{p}_1(1-\hat{p}_1)}{n_1} + \frac{\hat{p}_2(1-\hat{p}_2)}{n_2}} = (0.61-0.52) \pm 1.645\sqrt{\frac{0.61(1-0.61)}{2685} + \frac{0.52(1-0.52)}{4775}}$ or about 0.070 to 0.110. This means that you are 90% confident that the true difference is in the interval (0.070, 0.110). **b.** More of those who have a negative view of hazing may have responded, biasing the reported percentages toward the high side. If both male and female samples are similarly biased, the difference in sample percentages may be a valid estimate.
E6. a. Drinking Game: Here, $n_1 = 640,\ \hat{p}_1 = 0.47,\ n_2 = 1295,\ \hat{p}_2 = 0.53$ and certainly each of the products $n_1\hat{p}_1,\ n_1(1-\hat{p}_1),\ n_2\hat{p}_2$, and $n_2(1-\hat{p}_2)$ is at least 5. The 95% confidence interval is $(\hat{p}_1 - \hat{p}_2) \pm z^* \sqrt{\frac{\hat{p}_1(1-\hat{p}_1)}{n_1} + \frac{\hat{p}_2(1-\hat{p}_2)}{n_2}} = (0.53-0.47) \pm 1.96\sqrt{\frac{0.47(1-0.47)}{640} + \frac{0.55(1-0.53)}{1295}}$ or about 0.013 to 0.107. This means that you are 95% confident that the true difference is in the interval (0.013, 0.107). Sing or Chant: Here, $n_1 = 640,\ \hat{p}_1 = 0.27,\ n_2 = 1295,\ \hat{p}_2 = 0.32$ and certainly each of the products $n_1\hat{p}_1,\ n_1(1-\hat{p}_1),\ n_2\hat{p}_2$, and $n_2(1-\hat{p}_2)$ is at least 5. The 95% confidence interval is $(\hat{p}_1 - \hat{p}_2) \pm z^* \sqrt{\frac{\hat{p}_1(1-\hat{p}_1)}{n_1} + \frac{\hat{p}_2(1-\hat{p}_2)}{n_2}} = (0.32-0.27) \pm 1.96\sqrt{\frac{0.32(1-0.32)}{1295} + \frac{0.27(1-0.27)}{640}}$ or about 0.007 to 0.093. This means that you are 95% confident that the true difference is in the interval (0.007, 0.093).

b.

Drinking Game:

Here,

$$n_1 = 544,\ \ \hat{p}_1 = 0.26,\ n_2 = 818,\ \ \hat{p}_2 = 0.23$$

and certainly each of the products $n_1\hat{p}_1,\ n_1(1-\hat{p}_1),\ n_2\hat{p}_2$, and $n_2(1-\hat{p}_2)$ is at least 5.

The 90% confidence interval is

$$(\hat{p}_1-\hat{p}_2)\pm z^*\sqrt{\frac{\hat{p}_1(1-\hat{p}_1)}{n_1}+\frac{\hat{p}_2(1-\hat{p}_2)}{n_2}} = (0.26-0.23)\pm 1.645\sqrt{\frac{0.26(1-0.26)}{544}+\frac{0.23(1-0.23)}{818}}$$

or about -0.009 to 0.069. This means that you are 90% confident that the true difference is in the interval (-0.009, 0.069).

Sing or Chant:

Here,

$$n_1 = 544,\ \ \hat{p}_1 = 0.18,\ n_2 = 818,\ \ \hat{p}_2 = 0.25$$

and certainly each of the products $n_1\hat{p}_1,\ n_1(1-\hat{p}_1),\ n_2\hat{p}_2$, and $n_2(1-\hat{p}_2)$ is at least 5.

The 90% confidence interval is

$$(\hat{p}_1-\hat{p}_2)\pm z^*\sqrt{\frac{\hat{p}_1(1-\hat{p}_1)}{n_1}+\frac{\hat{p}_2(1-\hat{p}_2)}{n_2}} = (0.25-0.18)\pm 1.645\sqrt{\frac{0.25(1-0.25)}{818}+\frac{0.18(1-0.18)}{544}}$$

or about 0.007 to 0.093. This means that you are 90% confident that the true difference is in the interval (0.033, 0.107).

E7. a. The conditions are met for constructing a confidence interval for the difference of two proportions:

- You were told that you may assume that the samples are equivalent to simple random samples.
- The number of men and number of women in the United States are more than 425 • 10 = 4250.
- Each of

$$n_1\hat{p}_1 = 425(0.23) = 98$$
$$n_1(1 - \hat{p}_1) = 425(1 - 0.23) = 327$$
$$n_2\hat{p}_2 = 425(0.34) = 145$$
$$n_2(1 - \hat{p}_2) = 425(1 - 0.34) = 281$$

is at least 5, where n_1 and n_2 are the sample sizes for men and women respectively, and $\hat{p}_1$ and $\hat{p}_2$ are the proportions of men and women, respectively, in the sample who said they would prefer to be addressed by their last name.

b.

$$(\hat{p}_1-\hat{p}_2)\pm z^*\cdot\sqrt{\frac{\hat{p}_1(1-\hat{p}_1)}{n_1}+\frac{\hat{p}_2(1-\hat{p}_2)}{n_2}} = (0.23-0.34)\pm 2.576\sqrt{\frac{(0.23)(1-0.23)}{425}+\frac{(0.34)(1-0.34)}{425}}$$
$$= -0.11 \pm 0.079$$

or about (-0.189, -0.031).

c. You are 99% confident that the difference in the percentage of all men and the percentage of all women who prefer to be addressed by their last name is in the interval –0.189 to –0.031. (Alternatively, you are 99% confident that the difference in the percentage of all women and the percentage of all men who prefer to be addressed by their last name is in the interval 0.031 to 0.189.)

d. 0 is not in the confidence interval. This means that the statement, "There is no difference in the proportions of all men who would prefer to have their last name used and the proportion of all women who would prefer to have their last name used" is not plausible. If the difference in the proportion of men who prefer being addressed by their last name and the proportion of women who prefer being addressed by their last name is actually 0, getting a difference of 11% in the samples is not at all likely. Thus, you are convinced that there is a difference between the percentage of women and percentage of men who prefer to be addressed by their last name.

E8. a. The samples are independently and randomly selected from two binomial populations (they either believe 16 is the correct age or they do not). Each of

$$n_1\hat{p}_1 = 1{,}000 \cdot 0.46 = 460 \qquad n_1(1-\hat{p}_1) = 1{,}000 \cdot 0.54 = 540$$

$$n_2\hat{p}_2 = 1{,}000 \cdot 0.35 = 350 \qquad n_2(1-\hat{p}_2) = 1{,}000 \cdot 0.65 = 650$$

are at least five, where n_1 and n_2 represent the sample size for 1995 and 2004, respectively, and $\hat{p}_1$ and $\hat{p}_2$ represent the proportions of people in the sample who believed 16 is the correct age for being permitted to have a driver's license in those respective years.

b. At the 95% confidence level,

$$(\hat{p}_1 - \hat{p}_2) \pm z^* \sqrt{\frac{\hat{p}_1(1-\hat{p}_1)}{n_1} + \frac{\hat{p}_2(1-\hat{p}_2)}{n_2}} = (0.46 - 0.35) \pm 1.96\sqrt{\frac{0.46(1-0.46)}{1{,}000} + \frac{0.35(1-0.35)}{1{,}000}}$$

$$\approx 0.11 \pm 0.043$$

or between 0.067 and 0.153.

c. You are 95% confident that the difference between the proportion of all adults in the U.S. favoring 16 as the correct age to begin driving in 1995 and the proportion of all adults in the U.S. favoring 16 as the correct age to begin driving in 2004 is in the interval 0.067 and 0.153.

d. No, 0 is not in the interval, which implies that the statement, "The proportion of adults in the United States who believe 16 is the correct age to be permitted to have a driver's license has not changed between 1995 and 2004" is not plausible. You have evidence that this proportion has dropped. If the difference in the proportion of adults in 2004 and the proportion of all adults in 1995 favoring 16 as the driving age is actually 0, getting a difference of 11% in the samples is not reasonably likely. Thus, you are convinced that there is a difference in opinion between these years on this question.

E9. ***Check conditions.*** The conditions are met for constructing a confidence interval for the difference of two proportions:

- You were told that you may assume that the samples are equivalent to simple random samples.
- There are more than 76,000 male students and more than 76,000 female students in the United States.
- Each of

$$n_1\hat{p}_1 = 7{,}600 \bullet 0.59 = 4484$$
$$n_1(1 - \hat{p}_1) \approx 7{,}600 \bullet 0.41 = 3116$$
$$n_2\hat{p}_2 = 7{,}600 \bullet 0.48 = 3648$$
$$n_2(1 - \hat{p}_2) = 7{,}600 \bullet 0.52 = 3952$$

is at least 5, where n_1 and n_2 are the numbers of male and female high school seniors sampled, respectively, and $\hat{p}_1$ and $\hat{p}_2$ are the proportions of male and female high school seniors sampled, respectively, who have played on sports teams run by their school during the 12 months preceding the survey.

Do computations. The 95% confidence interval for the difference in the proportions of male and female seniors who have played on sports teams run by their school during the 12 months preceding the survey is

$$(\hat{p}_1 - \hat{p}_2) \pm z^* \sqrt{\frac{\hat{p}_1(1-\hat{p}_1)}{n_1} + \frac{\hat{p}_2(1-\hat{p}_2)}{n_2}} = (0.59 - 0.48) \pm 1.96\sqrt{\frac{0.59(1-0.59)}{7,600} + \frac{0.48(1-0.48)}{7,600}}$$
$$\approx 0.11 \pm 0.016$$

or between 0.094 and 0.126.

Write a conclusion in context. Because 0 isn't included in this confidence interval, it is acceptable to say that senior boys are "significantly more likely" than senior girls to have played on sports teams run by their school in the previous 12 months.

E10. ***Check Conditions.*** The conditions are met for constructing a confidence interval for the difference of two proportions:

- You were told that you may assume that the samples are equivalent to simple random samples.
- There were more than 77,000 female seniors in the United States in 1991 and more than 76,000 female seniors in the United States recently. (Here you are assuming that the study in 1991 had the same number of males and female seniors)
- Each of

$$n_1\hat{p}_1 = 7600 \bullet 0.48 = 3648$$

$$n_1(1 - \hat{p}_1) = 7600 \bullet 0.52 = 3952$$
$$n_2\hat{p}_2 = 7700 \bullet 0.47 = 3619$$
$$n_2(1 - \hat{p}_2) = 7700 \bullet 0.53 = 4081$$

is at least 5, where n_1 and n_2 are the numbers of female high school seniors sampled recently and in 1991, respectively, and $\hat{p}_1$ and $\hat{p}_2$ are the proportions of female high school seniors sampled in those years who have played on sports teams run by their school during the 12 months preceding the survey.

Do Computations. Using a 95% confidence level, the confidence interval is

$$(\hat{p}_1 - \hat{p}_2) \pm z^* \sqrt{\frac{\hat{p}_1(1-\hat{p}_1)}{n_1} + \frac{\hat{p}_2(1-\hat{p}_2)}{n_2}} = (0.48 - 0.47) \pm 1.96\sqrt{\frac{0.48(1-0.48)}{7,600} + \frac{0.47(1-0.47)}{7,700}}$$
$$\approx 0.01 \pm 0.016$$

or between –0.006 and 0.026.

State conclusion in context. Because this confidence interval contains 0, it is plausible that there was no change between 1991 and recently in the proportion of female high school students who had played on a sports team run by their school during the 12 months preceding the survey. So, no, this increase does not represent a significant change in the level of participation of female seniors.

E11. In general, as sample sizes get larger, the length of the confidence interval gets smaller. (If the sample size for only one sample gets larger, that part of the formula for the standard error goes to zero. This alone won't make the standard error itself go to zero unless the other sample size also gets larger.)

E12. I. A 95% confidence interval for the difference of two proportions $p_1 - p_2$ consists of those differences for which the observed difference of the two sample proportions $\hat{p}_1 - \hat{p}_2$ is a reasonably likely outcome. (That is, the confidence interval contains any differences in population proportions that could have produced the observed difference in sample proportions within the middle 95% of all possible outcomes.)

II. If you construct one hundred 95% confidence intervals, you expect that the difference of the population proportions $p_1 - p_2$ will be in 95 of them.

E13. You can use z in this way only because the sampling distribution of the estimate $\hat{p}_1 - \hat{p}_2$ is approximately normal. How do you know that the distribution of $\hat{p}_1 - \hat{p}_2$ is approximately normal? A theorem in mathematical statistics given in text says that the sampling distribution of the difference of two normally distributed random variables is normal. So the sampling distribution of $\hat{p}_1 - \hat{p}_2$ will be approximately normal if the separate sampling distributions of $\hat{p}_1$ and $\hat{p}_2$ are normal. They are approximately normal if each of $n_1\hat{p}_1, n_1(1-\hat{p}_1), n_2\hat{p}_2,$ and $n_2(1-\hat{p}_2)$ is at least 10. However, this condition is stronger than necessary in the case of a difference—the sampling distribution of the difference will be approximately normal as long as each one of these is at least 5.

E14. a. It is plausible that the two samples came from populations with the same proportion of successes because the observed value of $\hat{p}_1 - \hat{p}_2$ is a reasonably likely result if $p_1 - p_2 = 0$.

b. It suggests that the two samples didn't come from populations with the same proportion of successes unless a rare event occurred.

E15. The method is not correct because the respondents weren't selected independently from two different populations. These people were all from the same population and are differentiated only by their answer to the question. The appropriate method to use is a confidence interval for a proportion from a single population.

$$\hat{p} \pm 1.96\sqrt{\frac{\hat{p}(1-\hat{p})}{n}} = 0.75 \pm 1.96\sqrt{\frac{0.75(1-0.75)}{1008}} \approx 0.75 \pm 0.027\,.$$

You are 95% confident that the proportion of online respondents who would favor the legal drinking age as 21 is in the interval 0.723 to 0.777.

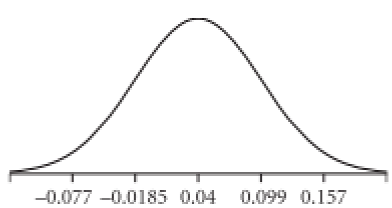
E16. This method can not be used because the respondents weren't selected independently from two different populations. These people were all from the same population and are differentiated only by their answer to the question. You could use a confidence interval for a single proportion separately to estimate the proportion of all adults that would give a particular response for those who chose age 16 or for those who chose age 18 but not for both at once.

E17. a. $\mu_{\hat{p}_1-\hat{p}_2} = 0.24 - 0.20 = 0.04$.

b. Don't use the pooled variance here because the two populations do not have a common variance.

$$\sigma_{\hat{p}_1-\hat{p}_2} = \sqrt{\frac{p_1(1-p_1)}{n_1} + \frac{p_2(1-p_2)}{n_2}} = \sqrt{\frac{0.24 \cdot 0.76}{100} + \frac{0.20 \cdot 0.80}{100}} \approx 0.0585$$

c.



d. You can use your calculator. **normalcdf(0.05,1E99,.04,.0585)** will give approximately 0.432. Alternatively, you can use the z-score.

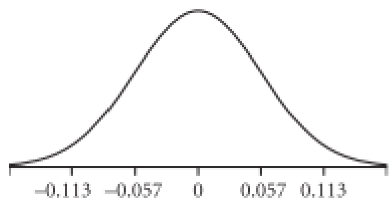
$$z = \frac{(\hat{p}_1 - \hat{p}_2) - (p_1 - p_2)}{\sqrt{\frac{p_1(1-p_1)}{n_1} + \frac{p_2(1-p_2)}{n_2}}} = \frac{0.05 - 0.04}{\sqrt{\frac{0.24 \cdot 0.76}{100} + \frac{0.20 \cdot 0.80}{100}}} \approx 0.171.$$

According to Table A, the probability of a z-score greater than 0.17 is approximately 0.4325.

E18. a. $\mu_{\hat{p}_1 - \hat{p}_2} = 0.2 - 0.2 = 0.$

b. $\sigma_{\hat{p}_1 - \hat{p}_2} = \sqrt{\frac{p_1(1-p_1)}{n_1} + \frac{p_2(1-p_2)}{n_2}} = \sqrt{\frac{0.2 \cdot 0.8}{100} + \frac{0.2 \cdot 0.8}{100}} \approx 0.0566$

c.



d. You can use your calculator. **normalcdf(0.05,1E99,0,.0566)** will give approximately 0.189. Alternatively, you can use the z-score.

$$z = \frac{(\hat{p}_1 - \hat{p}_2) - (p_1 - p_2)}{\sqrt{\frac{p_1(1-p_1)}{n_1} + \frac{p_2(1-p_2)}{n_2}}} = \frac{0.05 - 0}{\sqrt{\frac{0.2 \cdot 0.8}{100} + \frac{0.2 \cdot 0.8}{100}}} \approx 0.884.$$

According to Table A, the probability of a z-score greater than 0.88 is approximately 0.1894.

E19. ***Check conditions.*** Although the situation probably is actually more complicated, you can assume that you have two independent random samples. All of $n_1\hat{p}_1 = 177(0.30) = 53.1$, $n_1(1 - \hat{p}_1) = 177(1 - 0.30) = 123.9$, $n_2\hat{p}_2 = 616(0.24) = 147.84$, and $n_2(1 - \hat{p}_2) = 616(1 - 0.24) = 468.16$ are at least 5. The number of people in each age group is much larger than 10 times the sample size.

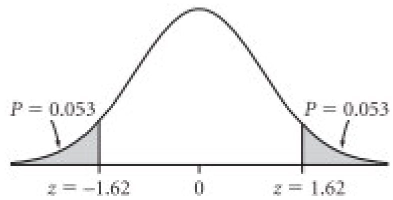
State your hypotheses.

H_0: The proportion, p_1, of all people aged 18 to 29 who sleep eight hours or more on a weekday is equal to the proportion, p_2, of all people aged 30 to 49 who sleep eight hours or more on a weekday.

H_a: $p_1 \neq p_2$

Compute the test statistic and draw a sketch. The test statistic is

$$z=\frac{(\hat{p}_1-\hat{p}_2)-(p_1-p_2)}{\sqrt{\hat{p}(1-\hat{p})\left(\frac{1}{n_1}+\frac{1}{n_2}\right)}}=\frac{(0.30-0.24)-0}{\sqrt{0.253\cdot 0.747\left(\frac{1}{177}+\frac{1}{616}\right)}}\approx 1.62$$



Here the pooled estimate, $\hat{p}$, is

$$\hat{p}=\frac{\text{total number of successes in both samples}}{n_1+n_2}=\frac{53+148}{177+616}\approx 0.253.$$

The P-value from the calculator for a two-sided test is
2 • **normalcdf(–1E99,–1.62)** ≈ 0.105.

Using the table with $z = -1.62$ gives a P-value of 2(0.0526) = 0.1052. Using the **2-PropZTest** command on your calculator with $x_1 = 53$ and $x_2 = 148$ gives a test statistic $z = 1.5951$ and a P-value of 0.1107.

Write a conclusion in context. No significance level was given, so you can assume 0.05. The P-value is larger than 0.05 so this difference is not statistically significant and you do not reject the null hypothesis. If the proportion of all Americans aged 18 to 29 who sleep eight hours or more on a workday is equal to the proportion of all Americans aged 30 to 49 who do so, the probability of getting a difference in sample proportions of 6% or larger from samples of these sizes is 0.11. Because this P-value is larger than 0.05, you can reasonably attribute the difference to chance variation. You have no evidence that the proportions would be different if you were to ask everyone in each of these two age groups whether they sleep more than eight hours on a workday.

E20. a. Here,

$$n_1=680\ (men),\ \ \hat{p}_1=0.33,\ n_2=698\ (women),\ \ \hat{p}_2=0.36$$

and certainly each of the products $n_1\hat{p}_1$, $n_1(1-\hat{p}_1)$, $n_2\hat{p}_2$, and $n_2(1-\hat{p}_2)$ is at least 5. The randomness conditions are also met.

The hypotheses, in symbols, are:

$$H_0: p_2 - p_1 = 0,\ H_a: p_2 - p_1 > 0.$$

Note that the pooled estimate is

$$\hat{p}=\frac{\text{total number of successes from both treatments}}{n_1+n_2}=\frac{224.4+251.28}{680+698}\approx 0.345.$$

So, the test statistic is

$$z = \frac{(\hat{p}_1 - \hat{p}_2) - (p_1 - p_2)}{\sqrt{\hat{p}(1-\hat{p})\left(\frac{1}{n_1} + \frac{1}{n_2}\right)}} = \frac{(0.36 - 0.33) - 0}{\sqrt{0.345(1-0.345)\left(\frac{1}{680} + \frac{1}{698}\right)}} \approx 1.171$$

Since this is a one-sided test, the p-value is $P(Z > 1.171) = 0.1210$. Hence, there is insufficient evidence, at the 10% level, that women are cyberbullied more than men.

b. There is no reason to necessarily expect bias in the samples based simply on the fact that it is a convenience sample, so it might still be valid.

E21. The question asks "Was NASA being looked upon *more favorably* by the American public in 2007 than in 1999?" This suggests you should do a one-sided significance test for the difference of two proportions.

This was a Gallup poll so we can assume these samples are equivalent to simple random samples. Each of $n_1\hat{p}_1$ = 1000 • 0.46 = 460, $n_1(1 - \hat{p}_1)$ = 1000(1 – 0.46) =640, $n_2\hat{p}_2$ = 1010 • 0.56 = 565.6, and $n_2(1 - \hat{p}_2)$ = 1010(1 – 0.56) =444.4 is at least 5. There are more than 10 • 1010 = 10,100 adult Americans. The conditions for inference are met.

The hypotheses are:

H_0: The proportion, p_1, of all adult Americans who gave NASA a favorable rating in 1999 is equal to the proportion, p_2, of all adult Americans who gave NASA a favorable rating in 2007.
H_a: $p_1 < p_2$

Note that the pooled estimate is

$$\hat{p} = \frac{\text{total number of successes from both treatments}}{n_1 + n_2} = \frac{460 + 565.6}{1000 + 1010} \approx 0.510.$$

So, the test statistic is

$$z = \frac{(\hat{p}_1 - \hat{p}_2) - (p_1 - p_2)}{\sqrt{\hat{p}(1-\hat{p})\left(\frac{1}{n_1} + \frac{1}{n_2}\right)}} = \frac{(0.56 - 0.46) - 0}{\sqrt{0.510(0.490)\left(\frac{1}{1{,}000} + \frac{1}{1{,}010}\right)}} \approx 4.484$$

The p-value is < 0.0001.

Hence, with a p-value as low as 0.0001, which is well below 0.05, you would reject the null hypothesis. If the proportion of all adult Americans who gave NASA a favorable rating in 2007 is equal to the proportion of all adult Americans who gave NASA a favorable rating in 1999, then there is at most a 1 out of 10,100 chance of getting a difference in sample proportions of 10% or larger. There is strong evidence that NASA was being looked upon more favorably in 2007 than it was in 1999.

E22. a. Here,

$$n_1 = 400 \ (men), \ \ \hat{p}_1 = 0.37, \ n_2 = 600 \ (women), \ \ \hat{p}_2 = 0.27$$

and certainly each of the products $n_1\hat{p}_1, \ n_1(1-\hat{p}_1), \ n_2\hat{p}_2$, and $n_2(1-\hat{p}_2)$ is at least 5. The randomness conditions are also met.

The hypotheses, in symbols, are:

$$H_0: p_1 - p_2 = 0, \ H_a: p_1 - p_2 > 0.$$

Note that the pooled estimate is

$$\hat{p} = \frac{\text{total number of successes from both treatments}}{n_1 + n_2} = \frac{148+162}{400+600} \approx 0.310.$$

So, the test statistic is

$$z = \frac{(\hat{p}_1 - \hat{p}_2) - (p_1 - p_2)}{\sqrt{\hat{p}(1-\hat{p})\left(\frac{1}{n_1} + \frac{1}{n_2}\right)}} = \frac{(0.37 - 0.27) - 0}{\sqrt{0.310(1-0.310)\left(\frac{1}{400} + \frac{1}{600}\right)}} \approx 3.350$$

Since this is a one-sided test, the p-value is $P(Z > 3.350) = 0.0004$. Hence, there is strong evidence that the proportion of men favoring the lower drinking age is greater than the proportion of women that does.

b. Here,

$$n_1 = 100 \ (men < 40), \ \ \hat{p}_1 = 0.52, \ n_2 = 300 \ (men \geq 40), \ \ \hat{p}_2 = 0.32$$

and certainly each of the products $n_1\hat{p}_1, \ n_1(1-\hat{p}_1), \ n_2\hat{p}_2$, and $n_2(1-\hat{p}_2)$ is at least 5. The randomness conditions are also met.

The hypotheses, in symbols, are:

$$H_0: p_1 - p_2 = 0, \ H_a: p_1 - p_2 > 0.$$

Note that the pooled estimate is

$$\hat{p} = \frac{\text{total number of successes from both treatments}}{n_1 + n_2} = \frac{52+96}{100+300} \approx 0.37.$$

So, the test statistic is

$$z = \frac{(\hat{p}_1 - \hat{p}_2) - (p_1 - p_2)}{\sqrt{\hat{p}(1-\hat{p})\left(\frac{1}{n_1} + \frac{1}{n_2}\right)}} = \frac{(0.52 - 0.32) - 0}{\sqrt{0.37(1-0.37)\left(\frac{1}{100} + \frac{1}{300}\right)}} \approx 3.587$$

Since this is a one-sided test, the p-value is $P(Z > 3.587) = 0.0001$. Hence, there is strong evidence that the proportion of men < 40 favoring the lower drinking age is greater than the proportion of men ≥ 40 that does.

c. No, because the two populations from which the samples are drawn are not independent.

E23. a. Here,

$$n_1 = 663,\ \hat{p}_1 = 0.35,\ n_2 = 1591,\ \hat{p}_2 = 0.29$$

and certainly each of the products $n_1\hat{p}_1$, $n_1(1-\hat{p}_1)$, $n_2\hat{p}_2$, and $n_2(1-\hat{p}_2)$ is at least 5. The randomness conditions are also met.

The hypotheses, in symbols, are:

$$H_0: p_1 - p_2 = 0,\ H_a: p_1 - p_2 > 0.$$

Note that the pooled estimate is

$$\hat{p} = \frac{\text{total number of successes from both treatments}}{n_1 + n_2} = \frac{232.05 + 461.39}{663 + 1591} \approx 0.308.$$

So, the test statistic is

$$z = \frac{(\hat{p}_1 - \hat{p}_2) - (p_1 - p_2)}{\sqrt{\hat{p}(1-\hat{p})\left(\frac{1}{n_1} + \frac{1}{n_2}\right)}} = \frac{(0.35 - 0.29) - 0}{\sqrt{0.308(1-0.308)\left(\frac{1}{663} + \frac{1}{1591}\right)}} \approx 2.811$$

Since this is a one-sided test, the *p*-value is $P(Z > 2.811) = 0.0024$. Hence, there is sufficient evidence to say that a larger proportion of those who twitter live in urban areas.

b. Here,

$$n_1 = 663,\ \hat{p}_1 = 0.76,\ n_2 = 1591,\ \hat{p}_2 = 0.60$$

and certainly each of the products $n_1\hat{p}_1$, $n_1(1-\hat{p}_1)$, $n_2\hat{p}_2$, and $n_2(1-\hat{p}_2)$ is at least 5. The randomness conditions are also met.

The hypotheses, in symbols, are:

$$H_0: p_1 - p_2 = 0,\ H_a: p_1 - p_2 > 0.$$

Note that the pooled estimate is

$$\hat{p} = \frac{\text{total number of successes from both treatments}}{n_1 + n_2} = \frac{503.88 + 954.6}{663 + 1591} \approx 0.647.$$

So, the test statistic is

$$z = \frac{(\hat{p}_1 - \hat{p}_2) - (p_1 - p_2)}{\sqrt{\hat{p}(1-\hat{p})\left(\frac{1}{n_1} + \frac{1}{n_2}\right)}} = \frac{(0.76 - 0.60) - 0}{\sqrt{0.647(1-0.647)\left(\frac{1}{663} + \frac{1}{1591}\right)}} \approx 7.241$$

Since this is a one-sided test, the *p*-value is $P(Z > 7.241) < 0.0001$. Hence, there is strong evidence to say that those who twitter read newspapers online at a higher percentage than those who do not twitter.

c. No; you need the number of people sampled in each age group.

E24. a. Here,

$$n_1 = 992\ (\textit{conservative}),\ \hat{p}_1 = 0.21,\ n_2 = 511\ (\textit{liberal}),\ \hat{p}_2 = 0.30$$

and certainly each of the products $n_1\hat{p}_1$, $n_1(1-\hat{p}_1)$, $n_2\hat{p}_2$, and $n_2(1-\hat{p}_2)$ is at least 5. The randomness conditions are also met.

Note that the pooled estimate is

$$\hat{p} = \frac{\text{total number of successes from both treatments}}{n_1 + n_2} = \frac{208.32 + 153.3}{992 + 511} \approx 0.241.$$

a. You are testing (A).

b. The hypotheses, in symbols, are: H_o: $p_1 - p_2 = 0$, Ha: $p_1 - p_2 \neq 0$.
So, the test statistic is

$$z = \frac{(\hat{p}_1 - \hat{p}_2) - (p_1 - p_2)}{\sqrt{\hat{p}(1-\hat{p})\left(\frac{1}{n_1} + \frac{1}{n_2}\right)}} = \frac{(0.30 - 0.21) - 0}{\sqrt{0.241(1-0.241)\left(\frac{1}{992} + \frac{1}{511}\right)}} \approx 3.865$$

Since this is a two-sided test, the p-value is $2P(Z > 3.865) < 0.0002$. Hence, there is strong evidence to say that the proportion of Catholics is significantly different for conservatives versus liberals.

c. Here,

$$n_1 = 992 \ (\textit{conservative}), \ \ \hat{p}_1 = 0.12, \ n_2 = 511 \ (\textit{liberal}), \ \ \hat{p}_2 = 0.06$$

and certainly each of the products $n_1\hat{p}_1$, $n_1(1-\hat{p}_1)$, $n_2\hat{p}_2$, and $n_2(1-\hat{p}_2)$ is at least 5. The randomness conditions are also met.

The hypotheses, in symbols, are: H_0: $p_1 - p_2 = 0$, H_a: $p_1 - p_2 \neq 0$.

Note that the pooled estimate is

$$\hat{p} = \frac{\text{total number of successes from both treatments}}{n_1 + n_2} = \frac{119.04 + 30.66}{992 + 511} \approx 0.100.$$

So, the test statistic is

$$z = \frac{(\hat{p}_1 - \hat{p}_2) - (p_1 - p_2)}{\sqrt{\hat{p}(1-\hat{p})\left(\frac{1}{n_1} + \frac{1}{n_2}\right)}} = \frac{(0.12 - 0.06) - 0}{\sqrt{0.100(1-0.100)\left(\frac{1}{992} + \frac{1}{511}\right)}} \approx 3.673$$

Since this is a two-sided test, the p-value is $2P(Z > 3.673) < 0.0002$. Hence, there is strong evidence to say that the two groups differ in terms of their rates of participation in mission trips are concerned.

d. No, because the two populations from which the samples are drawn are not independent.

E25. Observe that $z = 7.51$ and the p-value is near zero. So, there is strong evidence that

the male and female populations differ with respect to the percentages that have experienced hazing.
E26. a. Observe that $z = 2.46$ with p-value = 0.014. So, at the 5% level, there is strong evidence that these proportions are different. **b.** Observe that $z = 6.70$ with p-value near zero. So, certainly at the 5% level, there is strong evidence that these proportions are different.

E27. a. We wish to test:
H_0: $p_1 - p_2 = 0$, H_a: $p_1 - p_2 \neq 0$, where p_1 is the proportion of domestic fruit showing no residue and p_2 is the proportion of imported fruit showing no residue.

Here,

$$n_1 = 344,\ \ \hat{p}_1 = 0.442,\ n_2 = 1136,\ \ \hat{p}_2 = 0.704$$

and certainly each of the products $n_1\hat{p}_1,\ n_1(1-\hat{p}_1),\ n_2\hat{p}_2$, and $n_2(1-\hat{p}_2)$ is at least 5. The randomness conditions are also met.

Note that the pooled estimate is

$$\hat{p} = \frac{\text{total number of successes from both treatments}}{n_1 + n_2} = \frac{152.048 + 799.744}{344 + 1136} \approx 0.643\,.$$

So, the test statistic is

$$z = \frac{(\hat{p}_1 - \hat{p}_2) - (p_1 - p_2)}{\sqrt{\hat{p}(1-\hat{p})\left(\dfrac{1}{n_1} + \dfrac{1}{n_2}\right)}} = \frac{(0.704 - 0.442) - 0}{\sqrt{0.643(1-0.643)\left(\dfrac{1}{344} + \dfrac{1}{1136}\right)}} \approx 8.89$$

The p-value is near zero. Hence, there is strong evidence of a difference between the domestic and imported fruits with regard to the proportions showing no residue.

b. We wish to test:
H_0: $p_1 - p_2 = 0$, H_a: $p_1 - p_2 \neq 0$, where p_1 is the proportion of domestic vegetables showing no residue and p_2 is that proportion of imported vegetables showing no residue.

Here,

$$n_1 = 672,\ \ \hat{p}_1 = 0.738,\ n_2 = 2447,\ \ \hat{p}_2 = 0.604$$

and certainly each of the products $n_1\hat{p}_1,\ n_1(1-\hat{p}_1),\ n_2\hat{p}_2$, and $n_2(1-\hat{p}_2)$ is at least 5. The randomness conditions are also met.

Note that the pooled estimate is

$$\hat{p} = \frac{\text{total number of successes from both treatments}}{n_1 + n_2} = \frac{495.936 + 1477.988}{672 + 2447} \approx 0.633\,.$$

So, the test statistic is

$$z = \frac{(\hat{p}_1 - \hat{p}_2) - (p_1 - p_2)}{\sqrt{\hat{p}(1-\hat{p})\left(\frac{1}{n_1} + \frac{1}{n_2}\right)}} = \frac{(0.738 - 0.604) - 0}{\sqrt{0.633(1-0.633)\left(\frac{1}{672} + \frac{1}{2447}\right)}} \approx 6.384$$

The p-value is near zero. Hence, there is strong evidence of a difference between the domestic and imported vegetables with regard to the proportions showing no residue (but in a different direction for the result for fruit).

c. No, the differences may be valid estimates even though the proportions are biased toward the higher values.

E28. **a.** We wish to test:
H_0: $p_1 - p_2 = 0$, H_a: $p_1 - p_2 \neq 0$, where p_1 is the proportion of domestic vegetables showing residue in violation of a standard and p_2 is the proportion of imported vegetables showing residue in violation of a standard.

Here,

$$n_1 = 672,\ \ \hat{p}_1 = 0.024,\ n_2 = 2447,\ \ \hat{p}_2 = 0.054$$

and certainly each of the products $n_1\hat{p}_1$, $n_1(1-\hat{p}_1)$, $n_2\hat{p}_2$, and $n_2(1-\hat{p}_2)$ is at least 5. The randomness conditions are also met.

Note that the pooled estimate is

$$\hat{p} = \frac{\text{total number of successes from both treatments}}{n_1 + n_2} = \frac{16.128 + 132.138}{672 + 2447} \approx 0.048.$$

So, the test statistic is

$$z = \frac{(\hat{p}_1 - \hat{p}_2) - (p_1 - p_2)}{\sqrt{\hat{p}(1-\hat{p})\left(\frac{1}{n_1} + \frac{1}{n_2}\right)}} = \frac{(0.054 - 0.024) - 0}{\sqrt{0.048(1-0.048)\left(\frac{1}{672} + \frac{1}{2447}\right)}} \approx 3.222$$

The p-value is $2P(Z > 3.222) = 0.0012$. Hence, there is strong evidence of a difference between the domestic and imported vegetables with regard to the proportions showing residue in violation of a standard.

The population for which the inference is relevant is the domestic and imported vegetable supplies.

b. We wish to test:
H_0: $p_1 - p_2 = 0$, H_a: $p_1 - p_2 \neq 0$, where p_1 is the proportion of domestic fruit showing residue in violation of a standard and p_2 is the proportion of imported fruit showing residue in violation of a standard.

Here,

$$n_1 = 344,\ \ \hat{p}_1 = 0.009,\ n_2 = 1136,\ \ \hat{p}_2 = 0.036.$$

Note that the pooled estimate is

$$\hat{p} = \frac{\text{total number of successes from both treatments}}{n_1 + n_2} = \frac{3.096 + 40.896}{344 + 1136} \approx 0.030.$$

So, the test statistic is

$$z = \frac{(\hat{p}_1 - \hat{p}_2) - (p_1 - p_2)}{\sqrt{\hat{p}(1-\hat{p})\left(\frac{1}{n_1} + \frac{1}{n_2}\right)}} = \frac{(0.036 - 0.009) - 0}{\sqrt{0.030(1-0.030)\left(\frac{1}{344} + \frac{1}{1136}\right)}} \approx 2.57$$

The p-value is $2P(Z > 2.57) = 0.102$. Hence, there is strong evidence of a difference between the domestic and imported fruit with regard to the proportions showing residue in violation of a standard, but not quite at the 10% level.

The population for which the inference is relevant is the domestic and imported fruit supplies.

c. Note that $n_1 p_1 = 344(0.009) = 3.096$ is not greater than 5. This is a violation of one of the conditions required for inference. This might have led to an invalid conclusion.

E29. a. The question "Is this a significant increase?" not, "Is this significantly different?" You would do a one-sided test.

The polls were conducted by Gallup, who uses what can be considered a simple random sample. Each of

$$n_1\hat{p}_1 = 1{,}000 \bullet 0.48 = 480,\ n_1(1 - \hat{p}_1) = 1{,}000(1 - 0.48) = 520,$$
$$n_2\hat{p}_2 = 1{,}000 \bullet 0.43 = 430, \text{ and } n_2(1 - \hat{p}_2) = 1{,}000(1 - 0.43) = 570$$

is at least 5, where n_1 and n_2 are the numbers of adults polled in 2009 and 2008, respectively, and $\hat{p}_1$ and $\hat{p}_2$ are the proportions of polled adults in 2009 and 2008, respectively, that logged onto the Internet for an hour or more daily. There were more than 10 • 1,000 = 10,000 adults both years in the United States. Hence, the conditions for inference are met.

We wish to test the hypotheses:

H_0: The proportion, p_1, of adults in 2009 who logged onto the Internet for at least an hour daily is equal to the proportion, p_2, of adults in 2008 who logged onto the Internet for at least an hour daily.
H_a: $p_1 < p_2$

Note that the pooled estimate is

$$\hat{p} = \frac{\text{total number of successes from both treatments}}{n_1 + n_2} = \frac{480 + 430}{1000 + 1000} \approx 0.455.$$

So, the test statistic is

$$z = \frac{(\hat{p}_1 - \hat{p}_2) - (p_1 - p_2)}{\sqrt{\hat{p}(1-\hat{p})\left(\frac{1}{n_1} + \frac{1}{n_2}\right)}} = \frac{(0.48 - 0.43) - 0}{\sqrt{0.455(1-0.545)\left(\frac{1}{1000} + \frac{1}{1000}\right)}} \approx 2.245$$

The p-value is $P(Z > 2.245) = 0.0123$. As such, you would reject the null hypothesis that the proportion of adults in 2009 who logged onto the Internet for at least an hour daily is equal to the proportion of adults in 2008 who did so. You have sufficient evidence that this proportion has increased over the year.

b. The one-sided test of H_o: $p = 0.5$, Ha: $p < 0.5$ has $z \approx -1.265$ and a P-value of about 0.1030. You cannot conclude that less than a majority use the Internet more than an hour per day in 2009.

E30. a. We wish to test:
H_0: $p_1 - p_2 = 0$, H_a: $p_1 - p_2 \neq 0$, where p_1 is the proportion in the 30-49 age group who report using the Internet more than 1 hour per day and p_2 is the proportion in the 18-29 age group who report using the Internet more than 1 hour per day.

Here,

$$n_1 = 100,\ \hat{p}_1 = 0.62,\ n_2 = 300,\ \hat{p}_2 = 0.54$$

and certainly each of the products $n_1\hat{p}_1$, $n_1(1-\hat{p}_1)$, $n_2\hat{p}_2$, and $n_2(1-\hat{p}_2)$ is at least 5. The randomness conditions are also met.

Note that the pooled estimate is

$$\hat{p} = \frac{\text{total number of successes from both treatments}}{n_1 + n_2} = \frac{62 + 162}{100 + 300} \approx 0.56.$$

So, the test statistic is

$$z = \frac{(\hat{p}_1 - \hat{p}_2) - (p_1 - p_2)}{\sqrt{\hat{p}(1-\hat{p})\left(\frac{1}{n_1} + \frac{1}{n_2}\right)}} = \frac{(0.6 - 0.54) - 0}{\sqrt{0.56(1-0.56)\left(\frac{1}{100} + \frac{1}{300}\right)}} \approx 1.047$$

The p-value is $2P(Z > 1.047) = 2(0.1476) = 0.2952$. Hence, there is not strong evidence of a difference between these two groups.

b. No, because the samples are not independent.

E31. B

E32. We made use of three facts that we learned earlier:

I. The mean of the distribution of the difference of two random variables is the difference of their individual means. We used this fact in stating that the mean of the sampling distribution of the difference $\hat{p}_1 - \hat{p}_2$ is equal to the difference of the means of the sampling distributions of $\hat{p}_1$ and $\hat{p}_2$, or

$$\mu_{\hat{p}_1-\hat{p}_2} = \mu_{\hat{p}_1} - \mu_{\hat{p}_2} = p_1 - p_2$$

II. The variance of the distribution of the difference of two independent random variables is the sum of their individual variances. We used this fact in stating that the value of the variance of the sampling distribution of the difference $\hat{p}_1 - \hat{p}_2$ when $p_1 = p_2 = p$ is

$$\frac{p_1(1-p_1)}{n_1} + \frac{p_2(1-p_2)}{n_2} = p(1-p)\left(\frac{1}{n_1}+\frac{1}{n_2}\right)$$

III. Under some not-very-restrictive assumptions, the distribution of the difference of two independent random variables is approximately normally distributed. Specifically, all of the values $n_1\hat{p}_1, n_1(1-\hat{p}_1),\ n_2\hat{p}_2,$ and $n_2(1-\hat{p}_2)$ must be at least 5.

E33. a. The question asks you to determine if you have statistically significant evidence that more males are left-handed than females, so use a one-sided test.

Check conditions. You saw that the conditions were met in the example in the text.

State your hypotheses.
H_0: $p_1 - p_2 = 0.02$, where p_1 is the proportion of all males who are left-handed and p_2 is the proportion of all females who are left-handed.
H_a: $p_1 - p_2 > 0.02$.

Calculate the test statistic and* P*-value.

$$z = \frac{(\hat{p}_1-\hat{p}_2)-(p_1-p_2)}{\sqrt{\frac{\hat{p}_1(1-\hat{p}_1)}{n_1}+\frac{\hat{p}_2(1-\hat{p}_2)}{n_2}}} = \frac{(0.106-0.079)-0.02}{\sqrt{\frac{0.106\cdot 0.894}{1{,}067}+\frac{0.079\cdot 0.921}{1{,}170}}} \approx 0.570$$

According to Table A, the z-score for 0.570 corresponds to a 1-sided P-value of 0.2843.

State your conclusion in context. Since the P-value of 0.2843 is greater than 0.05 you would not reject the null hypothesis that the proportion of all males who are left-handed is 2% more than the proportion of all females who are left-handed. If the proportion of all males who are left-handed is equal to 2% more than the proportion of all females who are left-handed, then, the chance of seeing a difference in sample proportions greater than the observed 2.7% is 0.2843. Because this probability is so large, the observed difference can be attributed to chance alone and there is insufficient evidence to support the claim that the proportion of males who are left-handed is at least 2% greater than the proportion of females who are left-handed.

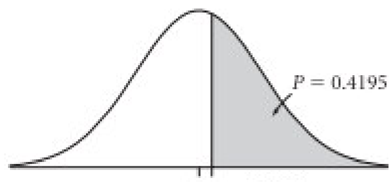
E34. ***Check conditions.*** The conditions are met for doing a test of significance of the difference of two proportions. You have a random sample from each of two large populations. Each of $n_1\hat{p}_1 = 5, n_1(1-\hat{p}_1) = 95,\ n_2\hat{p}_2 = 5,$ and $n_2(1-\hat{p}_2) = 59$ is at least 5.

State your hypotheses.
H_0: $p_1 - p_2 = 0.02$, where p_1 is the proportion of all mornings where Bus B is late and p_2 is the proportion of all mornings Bus A is late.
H_a: $p_1 - p_2 .> 0.02$

Compute the test statistic and draw a sketch. The test statistic is

$$z = \frac{(\hat{p}_1 - \hat{p}_2) - (p_1 - p_2)}{\sqrt{\frac{\hat{p}_1(1-\hat{p}_1)}{n_1} + \frac{\hat{p}_2(1-\hat{p}_2)}{n_2}}} = \frac{\left(\frac{5}{64} - \frac{5}{100}\right) - 0.02}{\sqrt{\frac{\frac{5}{64}\left(1-\frac{5}{64}\right)}{64} + \frac{\frac{5}{100}\left(1-\frac{5}{100}\right)}{100}}} = 0.2031$$

The one-sided P-value is 0.4195.



Write a conclusion in context. Suppose the proportion of all mornings that Bus B is late is equal to 2% more than the proportion of all mornings that Bus A is late. Under this assumption, the chance of seeing a difference in sample proportions greater than the observed 2.8% is 0.4195. Because this probability is so large, the observed difference can be attributed to chance alone. You take Bus B.

E35. a. We wish to test:
H_0: $p_1 - p_2 = 0$, H_a: $p_1 - p_2 \neq 0$, where p_1 is the proportion of subjects getting colds if all subjects could have been given vitamin C and p_2 is the proportion of subjects getting colds if all subjects could have been given the placebo.

Here,

$$n_1 = 139 \text{ (vitamin C)}, \ \hat{p}_1 = 0.122, \ n_2 = 140, \ \hat{p}_2 = 0.221$$

and certainly each of the products $n_1\hat{p}_1$, $n_1(1-\hat{p}_1)$, $n_2\hat{p}_2$, and $n_2(1-\hat{p}_2)$ is at least 5. The randomness conditions are also met.

Note that the pooled estimate is

$$\hat{p} = \frac{\text{total number of successes from both treatments}}{n_1 + n_2} = \frac{17+31}{139+140} \approx 0.172.$$

So, the test statistic is

$$z = \frac{(\hat{p}_1 - \hat{p}_2) - (p_1 - p_2)}{\sqrt{\hat{p}(1-\hat{p})\left(\frac{1}{n_1} + \frac{1}{n_2}\right)}} = \frac{(0.221 - 0.122) - 0}{\sqrt{0.172(1-0.172)\left(\frac{1}{139} + \frac{1}{140}\right)}} \approx 2.191$$

The p-value is $2P(Z > 2.191) =$ 0.0285. So, there is sufficient evidence, at the 5% level, to conclude that the proportions getting colds differs for the two treatments, and vitamin

C seems to have a positive effect.

b. There is insufficient evidence of a difference at the 1% level because the p-value is larger than 0.01.

E36. **a.** We wish to test:
H_0: $p_1 - p_2 = 0$, H_a: $p_1 - p_2 \neq 0$, where p_1 is the proportion of the placebo group who got a cold and p_2 is the proportion of the vitamin C group who got a cold.

Here,

$$n_1 = 411 \text{ (placebo)},\ \ \hat{p}_1 = 0.815,\ n_2 = 407 \text{ (vitamin C)},\ \ \hat{p}_2 = 0.742$$

and certainly each of the products $n_1\hat{p}_1,\ n_1(1-\hat{p}_1),\ n_2\hat{p}_2,$ and $n_2(1-\hat{p}_2)$ is at least 5. The randomness conditions are also met.

Note that the pooled estimate is

$$\hat{p} = \frac{\text{total number of successes from both treatments}}{n_1 + n_2} = \frac{335+302}{411+407} \approx 0.779\,.$$

So, the test statistic is

$$z = \frac{(\hat{p}_1 - \hat{p}_2) - (p_1 - p_2)}{\sqrt{\hat{p}(1-\hat{p})\left(\frac{1}{n_1} + \frac{1}{n_2}\right)}} = \frac{(0.815 - 0.742) - 0}{\sqrt{0.779(1-0.77)\left(\frac{1}{411} + \frac{1}{407}\right)}} \approx 2.516$$

The p-value is $2P(Z > 2.516) = 0.0118$. So, there is sufficient evidence, at the 5% level, to conclude that the proportions getting colds differs for the two treatments.

b. There is not quite sufficient evidence at the 1% level because the p-value is larger than 0.01.

E37. We wish to test:
H_0: $p_1 - p_2 = 0$, H_a: $p_1 - p_2 \neq 0$, where p_1 is the proportion of subjects getting polio if all subjects could have been given the Salk vaccine and p_2 is the proportion of subjects getting polio if all subjects could have been given the placebo.

Here,

$$n_1 = 200{,}745 \text{ (Salk vaccine)},\ \ \hat{p}_1 = 0.0004,\ n_2 = 201{,}229 \text{ (placebo)},\ \ \hat{p}_2 = 0.0008$$

and certainly each of the products $n_1\hat{p}_1,\ n_1(1-\hat{p}_1),\ n_2\hat{p}_2,$ and $n_2(1-\hat{p}_2)$ is at least 5. The randomness conditions are also met.

Note that the pooled estimate is

$$\hat{p} = \frac{\text{total number of successes from both treatments}}{n_1 + n_2} = \frac{82+162}{200{,}745 + 201{,}229} \approx 0.0006\,.$$

So, the test statistic is

$$z = \frac{(\hat{p}_1 - \hat{p}_2) - (p_1 - p_2)}{\sqrt{\hat{p}(1-\hat{p})\left(\frac{1}{n_1} + \frac{1}{n_2}\right)}} = \frac{(0.0008 - 0.0004) - 0}{\sqrt{0.0006(1-0.0006)\left(\frac{1}{200{,}745} + \frac{1}{201{,}229}\right)}} \approx 5.178$$

The p-value is $2P(Z > 5.178)$ is near 0; there is strong evidence to conclude that the proportion getting polio is smaller among those getting the vaccine. (The difference in proportions may seem small, but the vaccine cut the incidence of polio about in half.)

E38. We wish to test:
H_0: $p_1 - p_2 = 0$, H_a: $p_1 - p_2 \neq 0$, where p_1 is the proportion of larvae that died using Method A and p_2 is the proportion of larvae that died using Method B.

Here,

$$n_1 = 20,\ \ \hat{p}_1 = 0.40,\ n_2 = 20,\ \ \hat{p}_2 = 0.60$$

and certainly each of the products $n_1\hat{p}_1,\ n_1(1-\hat{p}_1),\ n_2\hat{p}_2,$ and $n_2(1-\hat{p}_2)$ is at least 5. The randomness conditions are also met.

Note that the pooled estimate is

$$\hat{p} = \frac{\text{total number of successes from both treatments}}{n_1 + n_2} = \frac{8+12}{20+20} \approx 0.50\,.$$

So, the test statistic is

$$z = \frac{(\hat{p}_1 - \hat{p}_2) - (p_1 - p_2)}{\sqrt{\hat{p}(1-\hat{p})\left(\frac{1}{n_1} + \frac{1}{n_2}\right)}} = \frac{(0.60 - 0.40) - 0}{\sqrt{0.5(1-0.5)\left(\frac{1}{20} + \frac{1}{20}\right)}} \approx 1.265$$

The p-value is $2P(Z > 1.265) = 0.2058$. So, there is not sufficient evidence that the two methods are significantly different in how effective they are in killing larvae.

E39. Because it was hypothesized before the experiment began that aspirin was beneficial, you should conduct a one-sided significance test for the difference of two proportions. (If you do a two-sided test, the computations and conclusion will be the same in this case.)

Check conditions. The subjects weren't selected randomly from a larger population, but the treatments were randomly assigned to the subjects, so you can use this significance test for the difference of two proportions. For the group taking aspirin, $n_1 = 11{,}037$ and $\hat{p}_1 = \frac{139}{11{,}037} \approx 0.0126$. For the group taking a placebo, $n_2 = 11{,}034$ and $\hat{p}_2 = \frac{239}{11{,}034} \approx 0.0217$.
Therefore, each of $n_1\,\hat{p}_1 = 139$, $n_1(1 - \hat{p}_1) = 10{,}898$, $n_2\hat{p}_2 = 239$, and $n_2(1 - \hat{p}_2) = 10{,}795$ is at least 5.

State your hypotheses.
H_0: If all of the men could have been given aspirin, the proportion, p_1, who had a heart attack would have been equal to the proportion, p_2, of the men who had a heart attack if

all could have been given the placebo.
H_a: $p_1 < p_2$

Write a conclusion in context. Using the output, we conclude that if there is no difference in the proportion of men who would have had a heart attack if they had all taken aspirin and the proportion who would have had a heart attack if they had all taken the placebo, then there is almost no chance of getting a difference of 0.0091 or smaller in the two proportions from a random assignment of these treatments to the subjects. This difference can not reasonably be attributed to chance variation. You reject the null hypothesis.

Note that although the difference in proportions is very small, only 0.0091, this difference is statistically significant because of the large sample sizes. Further, men who take low-dose aspirin cut their chance of a heart attack almost in half.

E40. Because it was hypothesized before the experiment began that aspirin was beneficial, we will conduct a one-sided test of the significance of the difference of two proportions.

Check conditions. The subjects weren't selected randomly from a larger population, but the treatments were randomly assigned to the subjects, so we can use this significance test for the difference of two proportions. (But note that the randomization wasn't done within the group of men who had heart attacks—that would have been impossible.) For the group taking aspirin, $n_1 = 139$ and $\hat{p}_1 \approx 0.072$. For the group taking a placebo, $n_2 = 239$ and $\hat{p}_2 \approx 0.109$. Therefore, each of $n_1\hat{p}_1 = 10$, $n_1(1 - \hat{p}_1) = 129$, $n_2\hat{p}_2 = 26$, and $n_2(1 - \hat{p}_2) = 213$ is at least 5.

State your hypotheses.
H_0: If all of the men who had a heart attack could have been given the same treatment, the proportion, p_1, who would have died of the heart attack after taking aspirin would be equal to the proportion, p_2, who would have died after taking the placebo.
H_a: $p_1 < p_2$

Write a conclusion in context. Using the output, we conclude that if there would have been no difference in the proportion of men who would have died from their heart attack if they had all taken aspirin and the proportion who would have died if they had all taken the placebo, then there is a 0.1197 chance of getting a difference of –0.037 or smaller in the proportions from random assignment of these treatments to the subjects. This difference can reasonably be attributed to chance variation. You do not reject the null hypothesis.

E41. a. This is an observational study.

b. *Check Conditions.* There was no randomization, so this is an observational study. Each of $n_1\hat{p}_1 \approx 912$, $n_1(1 - \hat{p}_1) \approx 273$, $n_2\hat{p}_2 = 106$, and $n_2(1 - \hat{p}_2) = 52$ is at least 5.

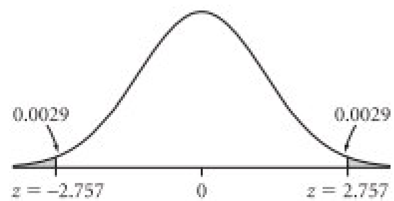
State hypotheses.

H$_0$: The difference between the proportion p_1 of dementia-free people who exercise three or more times a week, and the proportion p_2 of those with signs of dementia who exercise three or more times a week can be reasonably attributed to chance variation.

H$_a$: The difference cannot be reasonably attributed to chance variation.

Calculate the test statistic and draw a sketch.

$$z = \frac{(\hat{p}_1 - \hat{p}_2) - (p_1 - p_2)}{\sqrt{\hat{p}(1-\hat{p})\left(\frac{1}{n_1} + \frac{1}{n_2}\right)}} = \frac{(0.77 - 0.67) - 0}{\sqrt{0.758 \cdot 0.242\left(\frac{1}{1185} + \frac{1}{158}\right)}} \approx 2.757$$



Here, the pooled estimate $\hat{p}$ is

$$\frac{\text{the total number of exercisers in both groups}}{n_1 + n_2} = \frac{912 + 106}{1185 + 158} \approx 0.758$$

A z-score of 2.757 corresponds to a P-value of 2(0.0029) = 0.0058.

State conclusion in context. Because this P-value is so low, much less than 0.05, you would reject the null hypothesis that the difference in proportions could be reasonably attributed to chance. There is evidence of an association between exercise and a delay of dementia for this group of persons in this study.

This study cannot demonstrate any causal relationship due to the lack of randomization, but the issue may warrant more study. This is an example of a newspaper reporting a result that is not indicated by the study, and demonstrates the importance of clearly stating what your study shows and what it does not show.

E42. a. This is an observational study.

b. An observational study cannot provide clear evidence of causation. All you can do is see if the association that appears among your subjects might be reasonably attributed to chance variation or if there might be some other explanation.

Check conditions. There is no randomization. Each of $n_1\hat{p}_1$ = 950, $n_1(1 - \hat{p}_1)$ = 117, $n_2\hat{p}_2$ = 348, and $n_2(1 - \hat{p}_2)$ = 54 is at least 5.

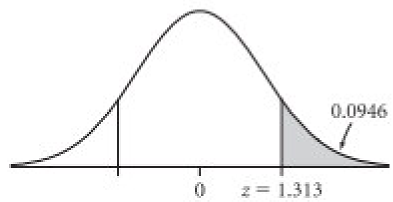
State your hypotheses.
H$_0$: The difference in proportions of deaths between non-smokers and pipe smokers can

be reasonably attributed to chance variation.

H_a: The difference cannot be reasonably attributed to chance variation alone.

Calculate the test statistic and draw a sketch.

$$z = \frac{(\hat{p}_1 - \hat{p}_2) - (p_1 - p_2)}{\sqrt{\hat{p}(1-\hat{p})\left(\frac{1}{n_1} + \frac{1}{n_2}\right)}} = \frac{(0.8903 - 0.8657) - 0}{\sqrt{0.884 \cdot 0.116\left(\frac{1}{1067} + \frac{1}{402}\right)}} \approx 1.313$$



Here the pooled estimate

$$\hat{p} = \frac{\text{the total number of experimental units still alive}}{n_1 + n_2} = \frac{950 + 348}{1067 + 402} \approx 0.884$$

The *P*-value for a *z*-score of 1.313 is 2(0.0946) = 0.1892. Using the TI-84+, $z = 1.3147$ and the *P*-value is 0.1886.

State your conclusion in context. Because the *P*-value is more than 0.05 you would not reject the null hypothesis. There is insufficient evidence that the difference between the proportion of non-smokers who died and the proportion of pipe smokers who died is due to anything other than chance.

E43. a. H_0: $p_1 - p_2 = 0$, H_a: $p_1 - p_2 > 0$, where p_1 is the proportion of subjects eating goldfish if all subjects could have seen the host eating goldfish and p_2 is the proportion of subjects eating goldfish if all subjects could have seen the host eating animal crackers.

b. Here,

$$n_1 = 29,\ \ \hat{p}_1 = 0.724,\ n_2 = 26,\ \ \hat{p}_2 = 0.462$$

and certainly each of the products $n_1\hat{p}_1$, $n_1(1-\hat{p}_1)$, $n_2\hat{p}_2$, and $n_2(1-\hat{p}_2)$ is at least 5. The randomness conditions are also met.

Note that the pooled estimate is

$$\hat{p} = \frac{\text{total number of successes from both treatments}}{n_1 + n_2} = \frac{21 + 12}{29 + 26} \approx 0.60\ .$$

So, the test statistic is

$$z = \frac{(\hat{p}_1 - \hat{p}_2) - (p_1 - p_2)}{\sqrt{\hat{p}(1-\hat{p})\left(\frac{1}{n_1} + \frac{1}{n_2}\right)}} = \frac{(0.724 - 0.462) - 0}{\sqrt{0.6(1-0.6)\left(\frac{1}{29} + \frac{1}{26}\right)}} \approx 1.980$$

The p-value is $P(Z > 1.980) = 0.0238$. So, there is sufficient evidence, at the 5% level, to conclude that students watching the host eat goldfish have a higher proportion of goldfish eaters than if they watch the host eat animal crackers.

E44. a. Children were randomly assigned to a treatment, and each of $n_1\hat{p}_1 = 19$, $n_1(1-\hat{p}_1) = 42$, $n_2\hat{p}_2 = 32$, and $n_2(1-\hat{p}_2) = 30$ is at least 5, where $\hat{p}_1$ is the proportion of the children in the treatment group of size n_1 who went to preschool that were arrested, and $\hat{p}_2$ is the proportion of the children in the treatment group of size n_2 who did not go to preschool that were arrested. The conditions for constructing a confidence interval are met.

b. The confidence interval is

$$(\hat{p}_1 - \hat{p}_2) \pm z^* \sqrt{\frac{\hat{p}_1(1-\hat{p}_1)}{n_1} + \frac{\hat{p}_2(1-\hat{p}_2)}{n_2}} = \left(\tfrac{19}{61} - \tfrac{32}{62}\right) \pm 1.645\sqrt{\frac{\frac{19}{61}\cdot\frac{42}{61}}{61} + \frac{\frac{32}{62}\cdot\frac{30}{62}}{62}} \approx -0.205 \pm 0.143$$

or about –0.348 to –0.062.

c. Suppose all of the children went to preschool and all of the children did not go to preschool. Then you are 90% confident that the difference in the proportion who would get arrested is in the interval –0.348 to –0.062. Because 0 is not in this interval, it is not plausible that there is no difference in the proportions who would get arrested. The term "90% confident" means that this method of constructing confidence intervals results in $p_1 - p_2$ falling in an average of 90 out of every 100 confidence intervals you construct.

d. No. A great deal happens to people between the ages of 3 and 19, and it would be a big stretch to say that not going to preschool *caused* more children to get arrested later on in their lives. This study does raise some interesting questions worthy of further research, but this study alone is not enough to establish cause.

E45. a. It seems reasonable that larger tumors would be more likely to spread than smaller tumors.

b. No. There is not a random sample of patients with tumors of either size. Instead there is a group of patients enrolled in a particular program. It is true that the other conditions have been met. Each of

$$n_1\hat{p}_1 = 234,\ n_1(1-\hat{p}_1) = 24$$
$$n_2\hat{p}_2 = 98,\ n_2(1-\hat{p}_2) = 20$$

is at least 5. The number of cancer patients with tumors of each given size is more than ten times the respective sample size given in the problem.

Note: You could do a test to see whether the observed difference in proportions can be reasonably attributed to chance. The conditions *are* met for such a test.

c. $z = \frac{(\hat{p}_1 - \hat{p}_2) - (p_1 - p_2)}{\sqrt{\hat{p} \cdot (1-\hat{p})\left(\frac{1}{n_1} + \frac{1}{n_2}\right)}} = \frac{(234/258 - 98/118) - 0}{\sqrt{\frac{332}{376} \cdot \frac{44}{376}\left(\frac{1}{258} + \frac{1}{118}\right)}} \approx 2.14$.

Table A gives a one-sided P-value of 0.0162. 0.02 would be a correct conservatively rounded approximation. Here, $\hat{p} = \frac{234+98}{258+118} = \frac{332}{376}$.

d. If tumors measuring 15 mm or less and tumors measuring 16-25 mm are equally likely to metastasize, then there is about a 2% probability of seeing a difference in proportions of metastases at least as large as that seen in this study.

E46. a. Answers will vary, but researchers can't look at the data before deciding this and so students shouldn't give the sample proportions in their reasons for their choice One of the study's authors did not know which way it might go; "In one sense it seems ironic that something like taking a natural substance is being used by people getting plastic surgery. But when you look at it carefully, that population is looking for self-improvement. They are using both herbs and plastic surgery to rejuvenate themselves."

b. This is likely a sample survey experiment since it is based simply on a yes/no question asked of participants.

c. It does not appear to that these are random samples, so that condition is not met. However, the other conditions are met. Each of

$$n_1\hat{p}_1 = 0.55 \cdot 100 = 55,\ n_1(1-\hat{p}_1) = 45$$
$$n_2\hat{p}_2 = 0.24 \cdot 100 = 24,\ n_2(1-\hat{p}_2) = 76$$

is at least five, where n_1 and n_2 are the numbers of plastic surgery patients and non-patients, respectively, in the study and $\hat{p}_1$ and $\hat{p}_2$ are the proportions of each group in this study that take herbal supplements. There are more than 10 • 100 = 1000 plastic surgery patients and non-plastic surgery patients in the Los Angeles area.

Note: You could do a test to see whether the observed difference in proportions can be reasonably attributed to chance. The conditions *are* met for such a test.

The test statistic is

$$z = \frac{(\hat{p}_1 - \hat{p}_2) - (p_1 - p_2)}{\sqrt{\hat{p}(1-\hat{p})\left(\frac{1}{n_1} + \frac{1}{n_2}\right)}} = \frac{(0.55 - 0.24) - 0}{\sqrt{0.395 \cdot 0.605\left(\frac{1}{100} + \frac{1}{100}\right)}} \approx 4.48.$$

Here, the pooled estimate is

$$\hat{p} = \frac{55+24}{100+100} = \frac{79}{200} \approx 0.395.$$

Table A does not extend to z-scores greater than 3.8 but a calculator will give the values. For a one-sided test with alternative hypothesis $p_1 > p_2$ the P-value is about 0.0000037. For a two-sided test the P-value is about 0.0000073. (For a one-sided test with alternative hypothesis $p_1 < p_2$ the P-value is about $1 - 0.0000037 \approx 0.9999963$.)

d. If plastic surgery patients and non-patients are equally likely to use herbal supplements, a difference in the observed proportions as extreme as or more extreme than the result given in this study would occur less than 0.0001 of the time. This provides strong evidence that the difference can not reasonably be attributed to chance variation.

E47. ***Is a significance test legal in this case?*** Purists would say that we should not use a test of significance in this situation. They have two reasons. The first is that the numbers given are not a random sample from any population—in fact, they are the population of Reggie Jackson's "at bats." (He is retired, so there will be no further at bats.) We know all of his at bats, and we can see that, in fact, he did have a higher batting average in the World Series than in regular season play.

The second reason is that this is a classic example of "data snooping." There are hundreds of baseball players. Even if some underlying batting average is the same in regular season play as in the World Series for all players, by definition some players are certain to be rare events and do better in the World Series than in regular season play. Reggie is simply the player that stands out as the rarest of the predictable rare events.
Note that the question asks whether Reggie's better average in the World Series can reasonably be attributed to chance. This is the first question we should ask before assigning him the nickname "Mr. October." If it turns out that we can't reasonably attribute this to chance, then we have to look for some other explanation. That explanation might in fact be that we did some data snooping and ended up with a Type I error. On the other hand, the explanation might be that he came through in the World Series. At any rate, the data must pass the test that the results can't reasonably be attributed to chance before we take any further steps in comparing the performance of Reggie Jackson in the World Series to regular season play.

Check conditions. The two samples aren't random; they are the entire populations. Thus, a significance test will tell us only whether such a difference can reasonably be attributed to chance. Each of $n_1\hat{p}_1 \approx 2584$, $n_1(1-\hat{p}_1) \approx 7280$, $n_2\hat{p}_2 \approx 35$, and $n_2(1-\hat{p}_2) \approx 63$ is at least 5.

State your hypotheses.

H_0: The difference between the proportion of hits in regular season play and the proportion of hits in the World Series can reasonably be attributed to chance variation.
H_a: The difference between the proportion of hits in regular season play and the proportion of hits in the World Series is too large to be attributed to chance variation.

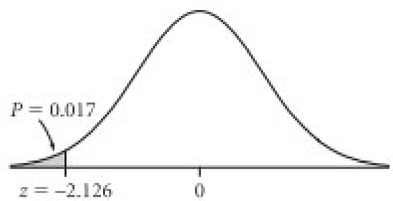
Compute the test statistic and draw a sketch.
The test statistic is

$$z = \frac{(\hat{p}_1 - \hat{p}_2) - (p_1 - p_2)}{\sqrt{\hat{p}(1-\hat{p})\left(\frac{1}{n_1} + \frac{1}{n_2}\right)}} = \frac{(0.262 - 0.357) - 0}{\sqrt{0.263(1-0.263)\left(\frac{1}{9864} + \frac{1}{98}\right)}} \approx -2.126,$$

or, using the TI-84+'s **2-PropZTest**, z =2.1299. Here,

$$\hat{p} = \frac{\textit{total number of successes in both samples}}{n_1 + n_2} = \frac{2584 + 35}{9864 + 98} \approx 0.263$$

The one-sided P-value is about 0.017.



Write a conclusion in context. A difference as large as Reggie's between regular season and World Series play would happen by chance to fewer than 17 players in 1000. Therefore, Reggie's record is indeed unusual. (It is interesting that Reggie hit only 0.227 in 163 bats in crucial League Championship Series play.)

E48. a. ***Check conditions.*** The two samples aren't random; they are the entire populations. Thus, a significance test will tell us only whether the difference in the percentages of casualties (43.9% vs. 27.4%) can reasonably be attributed to chance. Each of $n_1\hat{p}_1$ = 1054, $n_1(1 - \hat{p}_1)$ = 1346, $n_2\hat{p}_2$ = 411, and $n_2(1 - \hat{p}_2)$ = 1089 is at least 5.

State your hypotheses.

H_0: The difference in the proportions of British and American troops wounded can reasonably be attributed to chance variation.

H_a: The difference is too large to be attributed to chance alone.

Compute the test statistic. The test statistic is

$$z = \frac{(\hat{p}_1 - \hat{p}_2) - (p_1 - p_2)}{\sqrt{\hat{p}(1-\hat{p})\left(\frac{1}{n_1} + \frac{1}{n_2}\right)}} = \frac{(0.439 - 0.274) - 0}{\sqrt{0.376(1-0.376)\left(\frac{1}{2400} + \frac{1}{1500}\right)}} \approx 10.35$$

or 10.36 without rounding.

Here,

$$\hat{p} = \frac{\textit{total number of wounded on both sides}}{n_1 + n_2} = \frac{1054 + 411}{2400 + 1500} \approx 0.376$$

The P-value is close to 0.

Write a conclusion in context. The difference in the proportion of casualties can not reasonably be attributed to chance alone. There is almost no chance of getting a difference this large unless British soldiers were more likely to be wounded.

b. ***Check conditions.*** The two samples aren't random; they are the entire populations. Thus, a significance test will tell us only whether the difference in the percentages of deaths (9.417% vs. 9.333%) can reasonably be attributed to chance. Each of $n_1\hat{p}_1 = 226$, $n_1(1 - \hat{p}_1) = 2174$, $n_2\hat{p}_2 = 140$, and $n_2(1 - \hat{p}_2) = 1360$ is at least 5.

State your hypotheses.

H_0: The difference in the proportion of British and American troops killed can reasonably be attributed to chance variation.

H_a: The difference is too large to be attributed to chance alone.

Compute the test statistic and draw a sketch. The test statistic is

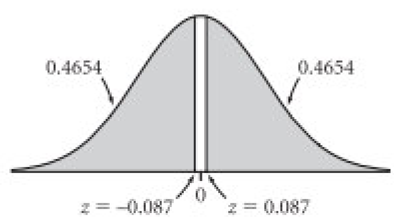
$$z = \frac{(\hat{p}_1 - \hat{p}_2) - (p_1 - p_2)}{\sqrt{\hat{p}(1-\hat{p})\left(\frac{1}{n_1} + \frac{1}{n_2}\right)}} = \frac{(0.09417 - 0.09333) - 0}{\sqrt{0.0938(1-0.0938)\left(\frac{1}{2400} + \frac{1}{1500}\right)}} \approx 0.0875$$

Here,

$$\hat{p} = \frac{\textit{total number of successes in both samples}}{n_1 + n_2} = \frac{226 + 140}{2400 + 1500} \approx 0.0938$$

The TI-84+ gives a test statistic $z = 0.087$ and a two-sided P-value of 0.9308.

Note: Here if $\hat{p}_1 = 0.094$ and $\hat{p}_2 = 0.093$ are used in the formula, the test statistic would be $z = 0.1042$ and the P-value would be 0.9170.



Write a conclusion in context. The difference of only 0.001 in the proportions of deaths can reasonably be attributed to chance alone. There is no reason to conclude that British troops were more or less likely to be killed than American troops. This condition is true even though they were wounded at a much greater rate. Either the British wounds were less severe or they had better medical care available.

E49. ***Check conditions.*** The two samples may be considered random samples. They were taken independently from the population of U.S. adults in 2008 and in 1974 The number of adults in each year is larger than ten times 1702. Finally, each of $n_1\hat{p}_1 = 1702(0.48) = 817$, $n_1(1 - \hat{p}_1) = 1702(1 - 0.48) = 885$, $n_2\hat{p}_2 = 1002(0.46) = 461$, and $n_2(1 - \hat{p}_2) = 1002(1 - 0.46) = 541$ is at least 5.

Do computations. The 90% confidence interval for the difference of the two population proportions p_1 and p_2 is

$$(\hat{p}_1 - \hat{p}_2) \pm z^* \cdot \sqrt{\frac{\hat{p}_1(1-\hat{p}_1)}{n_1} + \frac{\hat{p}_2(1-\hat{p}_2)}{n_2}} = (0.48 - 0.46) \pm 1.65\sqrt{\frac{(0.48)(1-0.48)}{1702} + \frac{(0.46)(1-0.46)}{1002}}$$

$$= 0.02 \pm 0.033$$

Alternatively, you can write this confidence interval as (-0.013, 0.053).

Write a conclusion in context. You are 90% confident that the difference between the proportion of all adults who would assign a grade of A or B in 2008 and in 1974 is between -0.013 and 0.053. Because 0 is in this confidence interval, the increase is insignificant.

E50. a. This is an observational study.

b. ***Check conditions.*** The two samples aren't random, they are the entire populations. Thus, a significance test will tell us only whether the difference in the proportions of snowboard injuries and ski injuries that were fractures (27.9% vs. 15.3%) can reasonably be attributed to chance. Each of

$$n_1\hat{p}_1 = 148,\ n_1(1 - \hat{p}_1) = 383,\ n_2\hat{p}_2 = 146, \text{ and } n_2(1 - \hat{p}_2) = 806$$

is at least 5.

State your hypotheses.

H_0: The difference in the proportion of snowboard injuries and ski injuries that were fractures can reasonably be attributed to chance variation.

H_a: The difference is too large to be attributed to chance alone.

Compute the test statistic and draw a sketch. The test statistic is

$$z = \frac{(\hat{p}_1 - \hat{p}_2) - (p_1 - p_2)}{\sqrt{\hat{p}(1-\hat{p})\left(\frac{1}{n_1} + \frac{1}{n_2}\right)}} = \frac{(0.279 - 0.153) - 0}{\sqrt{0.198(1-0.198)\left(\frac{1}{531} + \frac{1}{952}\right)}} \approx 5.838$$

or 5.805 without rounding.

Here,

$$\hat{p} = \frac{\textit{total number of successes in both samples}}{n_1 + n_2} = \frac{148 + 146}{531 + 952} \approx 0.198$$

The *P*-value is close to 0.

Write a conclusion in context. The difference in the proportion of injuries that were fractures can not reasonably be attributed to chance variation. (One possible explanation is that snowboarders take more risks and so any injury is more likely to be serious. Another explanation may be that because skiers tend to be older, the same fall that would injure them slightly might not injure a younger person at all. Thus, the skiers have more injuries, but they tend to be less serious.)

c. In part b, you may have noticed that the number of fractures was about the same for skiers as for snowboarders. But now you are told that there are about twice as many skiers. Thus, snowboarders were twice as likely to have a fracture as a skier. However, you can't carry out a test to determine whether this is statistically significant without knowing about how many snowboarders and how many skiers there were.

E51. a. ***Check conditions.*** The treatments were assigned randomly to the subjects, so you can use this significance test for the difference of two proportions. For the group taking the medication, n_1 = 25 and $\hat{p}_1 = 13/25 = 0.52$. For the group taking a placebo, n_2 = 26 and $\hat{p}_2 = 10/26 = 0.385$. Each of

$$n_1\hat{p}_1 = 13,\ n_1(1-\hat{p}_1) = 12,\ n_2\hat{p}_2 = 10, \text{ and } n_2(1-\hat{p}_2) = 16$$

is at least 5.

State your hypotheses.
H_0: The proportion, p_1, of people who would have responded if everyone had been given medication is equal to the proportion, p_2, of people who would have responded if everyone had been given the placebo.
H_a: $p_1 \neq p_2$

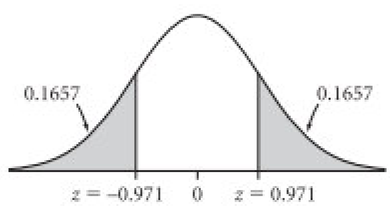
Compute the test statistic and draw a sketch. The test statistic is

$$z = \frac{(\hat{p}_1 - \hat{p}_2) - (p_1 - p_2)}{\sqrt{\hat{p}(1-\hat{p})\left(\frac{1}{n_1} + \frac{1}{n_2}\right)}} = \frac{(0.52 - 0.385) - 0}{\sqrt{0.451(1-0.451)\left(\frac{1}{25} + \frac{1}{26}\right)}} \approx 0.9686$$

or, using **2-PropZTest**, $z = 0.9713$.
Here,

$$\hat{p} = \frac{\text{total number who responded}}{n_1 + n_2} = \frac{13+10}{25+26} \approx 0.451$$

The two-sided *P*-value is 0.3314.



Write a conclusion in context. If there had been no difference between the proportion of people who would have responded if they had all taken the medication and the proportion who would have responded if they had all taken the placebo, then there is a 0.331 chance of getting a difference of 0.135 or larger in the two proportions from a random assignment of subjects to treatment groups. This difference can reasonably be attributed to chance variation. You do not reject the null hypothesis.

Note: Although the difference in the proportion who responded isn't statistically significant, the main point of the study was that "Brain physiology in placebo responders was altered in a different manner than in the medication responders."

b. Neither the subject nor the examining physician knew which treatment they were getting. This can be done by making sure the antidepressant and placebo look alike, and having the random assignment made by a third party.

E52. Here,

$$n_1 = 480,\ \ \hat{p}_1 = 0.48,\ n_2 = 520,\ \ \hat{p}_2 = 0.34$$

and certainly each of the products $n_1\hat{p}_1,\ n_1(1-\hat{p}_1),\ n_2\hat{p}_2$, and $n_2(1-\hat{p}_2)$ is at least 5. The randomness conditions are also met.

The 90% confidence interval for the difference of the two population proportions p_1 and p_2 is

$$(\hat{p}_1 - \hat{p}_2) \pm z^* \cdot \sqrt{\frac{\hat{p}_1(1-\hat{p}_1)}{n_1} + \frac{\hat{p}_2(1-\hat{p}_2)}{n_2}} = (0.48 - 0.34) \pm 1.645\sqrt{\frac{(0.48)(1-0.48)}{480} + \frac{(0.34)(1-0.34)}{520}}$$

$$= 0.14 \pm 0.051$$

Alternatively, you can write this confidence interval as (0.089, 0.191). Since 0 is not in the confidence interval, the data support the claim that a higher proportion of men favor legalization.

E53. a. We wish to test:

H_0: $p_1 - p_2 = 0$,

H_a: $p_1 - p_2 < 0$, where p_1 is the population proportion favoring stricter gun control laws in 2009 and p_2 is that proportion in 1990.

Here,

$$n_1 = 1023,\ \ \hat{p}_1 = 0.39\ (2009),\ n_2 = 1023,\ \ \hat{p}_2 = 0.78\ (1990)$$

and certainly each of the products $n_1\hat{p}_1,\ n_1(1-\hat{p}_1),\ n_2\hat{p}_2$, and $n_2(1-\hat{p}_2)$ is at least 5. The randomness conditions are also met.

Note that the pooled estimate is

$$\hat{p} = \frac{\text{total number of successes from both treatments}}{n_1 + n_2} = \frac{398.97 + 797.94}{1023 + 1023} \approx 0.585.$$

So, the test statistic is

$$z = \frac{(\hat{p}_1 - \hat{p}_2) - (p_1 - p_2)}{\sqrt{\hat{p}(1-\hat{p})\left(\frac{1}{n_1} + \frac{1}{n_2}\right)}} = \frac{(0.39 - 0.78) - 0}{\sqrt{0.585(1-0.585)\left(\frac{1}{1023} + \frac{1}{1023}\right)}} \approx -17.90$$

The p-value $P(Z < -17.90)$ is near zero. So, there is strong evidence to conclude that the proportion favoring stricter gun control laws has decreased between 1990 and 2009.

b. Here,

$$n_1 = 1023,\ \ \hat{p}_1 = 0.39\ (2009),\ n_2 = 100,\ \ \hat{p}_2 = 0.78\ (1990)$$

and certainly each of the products $n_1\hat{p}_1,\ n_1(1-\hat{p}_1),\ n_2\hat{p}_2$, and $n_2(1-\hat{p}_2)$ is at least 5. The randomness conditions are also met.

Note that the pooled estimate is

$$\hat{p} = \frac{\text{total number of successes from both treatments}}{n_1 + n_2} = \frac{78 + 398.97}{100 + 1023} \approx 0.425 .$$

So, the test statistic is

$$z = \frac{(\hat{p}_1 - \hat{p}_2) - (p_1 - p_2)}{\sqrt{\hat{p}(1-\hat{p})\left(\frac{1}{n_1} + \frac{1}{n_2}\right)}} = \frac{(0.39 - 0.78) - 0}{\sqrt{0.425(1-0.425)\left(\frac{1}{100} + \frac{1}{1023}\right)}} \approx -7.53$$

The p-value $P(Z < -7.53)$ is near zero. So, there is still strong evidence of a decrease.

c. We construct two 95% confidence intervals, one for the samples from (a) and one for the samples from (b):

Samples from (a):

$$(\hat{p}_1 - \hat{p}_2) \pm z^* \cdot \sqrt{\frac{\hat{p}_1(1-\hat{p}_1)}{n_1} + \frac{\hat{p}_2(1-\hat{p}_2)}{n_2}} = (0.78 - 0.39) \pm 1.96\sqrt{\frac{(0.39)(1-0.39)}{1023} + \frac{(0.78)(1-0.78)}{1023}}$$

or about (.351, 429).

Samples from (b):

$$(\hat{p}_1 - \hat{p}_2) \pm z^* \cdot \sqrt{\frac{\hat{p}_1(1-\hat{p}_1)}{n_1} + \frac{\hat{p}_2(1-\hat{p}_2)}{n_2}} = (0.78 - 0.39) \pm 1.96\sqrt{\frac{(0.39)(1-0.39)}{1023} + \frac{(0.78)(1-0.78)}{100}}$$

or about (0.304, 0.476).

The interval formed using the smaller sample size is about twice as wide as the other.

E54. The idea is to construct the confidence intervals and determine if they contain zero. Use the formula:

$$(\hat{p}_1 - \hat{p}_2) \pm 1.96 \cdot \sqrt{\hat{\sigma}^2_{p_1} + \hat{\sigma}^2_{p_2}}$$

a. "For profit" vs. "public hospitals" with comprehensive EHR:

$$(1.0 - 2.7) \pm 1.96\sqrt{0.4^2 + 0.7^2} = -1.7 \pm 1.96(0.806),$$

or about (-3.28, -0.120). Since this interval doesn't contain 0, there is sufficient evidence to conclude that the proportions are different.

"Private" vs. "public hospitals" with comprehensive EHR:

$$(1.5 - 2.7) \pm 1.96\sqrt{0.3^2 + 0.7^2} = -1.2 \pm 1.96(0.762),$$

or about (-2.69, 0.293). Since this interval contains 0, there is not sufficient evidence to conclude that the proportions are different.

b. "For profit" vs. "public hospitals" with basic EHR:

$$(5.0 - 6.9) \pm 1.96\sqrt{1.1^2 + 1.1^2} = -1.9 \pm 3.049 ,$$

or about (-4.95, 1.15). Since this interval contains 0, there is not sufficient evidence to conclude that the proportions are different.

"Private" vs. "public hospitals" with basic EHR:

$$(8.0 - 6.9) \pm 1.96\sqrt{0.7^2 + 1.1^2} = 1.1 \pm 2.556 ,$$

or about (-1.456, 3.656). Since this interval contains 0, there is not sufficient evidence to conclude that the proportions are different.

E55. First, we wish to test:
H_0: $p_1 - p_2 = 0$, H_a: $p_1 - p_2 \neq 0$, where p_1 is the proportion of deaths if all subjects could have been given the intensive treatment and p_2 is the proportion of deaths if all subjects could have been given the conventional treatment.
Here,
$n_1 = 3054$ (intensive treatment), $\hat{p}_1 = 0.271$, $n_2 = 3050$ (conventional treatment), $\hat{p}_2 = 0.246$ and certainly each of the products $n_1\hat{p}_1$, $n_1(1-\hat{p}_1)$, $n_2\hat{p}_2$, and $n_2(1-\hat{p}_2)$ is at least 5.
The randomness conditions are also met.

Note that the pooled estimate is

$$\hat{p} = \frac{\text{total number of successes from both treatments}}{n_1 + n_2} = \frac{829+751}{3054+3050} \approx 0.259 .$$

So, the test statistic is

$$z = \frac{(\hat{p}_1 - \hat{p}_2) - (p_1 - p_2)}{\sqrt{\hat{p}(1-\hat{p})\left(\frac{1}{n_1} + \frac{1}{n_2}\right)}} = \frac{(0.271 - 0.246) - 0}{\sqrt{0.259(1-0.259)\left(\frac{1}{3054} + \frac{1}{3050}\right)}} \approx 2.229$$

The p-value is $2P(Z > 2.229) = 0.0258$. So, there is sufficient evidence to conclude that the death proportions differ for the two treatments.

Next, assume that
$n_1 = 300$ (intensive treatment), $\hat{p}_1 = 0.271$, $n_2 = 300$ (conventional treatment), $\hat{p}_2 = 0.246$ and certainly each of the products $n_1\hat{p}_1$, $n_1(1-\hat{p}_1)$, $n_2\hat{p}_2$, and $n_2(1-\hat{p}_2)$ is at least 5.
The randomness conditions are also met.

Note that the pooled estimate is

$$\hat{p} = \frac{\text{total number of successes from both treatments}}{n_1 + n_2} = \frac{81.3+73.8}{300+300} \approx 0.259 .$$

So, the test statistic is

$$z = \frac{(\hat{p}_1 - \hat{p}_2) - (p_1 - p_2)}{\sqrt{\hat{p}(1-\hat{p})\left(\frac{1}{n_1} + \frac{1}{n_2}\right)}} = \frac{(0.271 - 0.246) - 0}{\sqrt{0.259(1-0.259)\left(\frac{1}{300} + \frac{1}{300}\right)}} \approx 0.699$$

The p-value is $2\,P(Z > 0.699) = 0.485$. So, there is not insufficient evidence to conclude that the death rates for the treatments differ.

E56. Yes. Suppose that, in fact, the proportions of males and females who pass the bar exam are exactly equal in each state. Because we are doing 50 hypothesis tests at 0.05, we expect 0.05(50) = 2.5 Type I errors. We will conclude that there is some inequity in an average of 2.5 states even if there is no inequity in any state. In addition, students may say that the pass rates may be unequal for perfectly equitable reasons, such as females study harder.

So what do statisticians do in such a situation? There are three options:
1. If you have n hypotheses to check, reduce α to $\frac{\alpha}{n}$. Because n is 50 in this case, you would use $\alpha = \frac{0.05}{50} = 0.001$ This makes the overall significance level less than or equal to 0.05. (This is called the Bonferroni method.)

2. In those states with statistically significant results, go out and get another sample to verify the results from the first one.

3. If it is impossible to get new data, randomly divide the sample from each state into two parts. Use the first sample in the first round of tests. For those states with a statistically significant result, verify this result in the second half of the sample.

E57. No, because the two sample proportions making up the difference are dependent. If one is very large, the other has to be small.

E58. a. More than 0.50 because using p = 0.5 would indicate an equal chance of touching the mark.
b. If you treat the "with mirror" and "without mirror" as independent samples, you can use the test for differences of proportions from this chapter.

c. Here,

$$n_1 = 12,\ \ \hat{p}_1 = 1.00,\ n_2 = 12,\ \ \hat{p}_2 = 0.583$$

and certainly each of the products $n_1\hat{p}_1,\ n_1(1-\hat{p}_1),\ n_2\hat{p}_2$, and $n_2(1-\hat{p}_2)$ is at least 5. The randomness conditions are also met.

Note that the pooled estimate is

$$\hat{p} = \frac{\text{total number of successes from both treatments}}{n_1 + n_2} = \frac{12+7}{12+12} \approx 0.792\,.$$

So, the test statistic is

$$z = \frac{(\hat{p}_1 - \hat{p}_2) - (p_1 - p_2)}{\sqrt{\hat{p}(1-\hat{p})\left(\frac{1}{n_1} + \frac{1}{n_2}\right)}} = \frac{(1.00 - 0.583) - 0}{\sqrt{0.792(1-0.792)\left(\frac{1}{12} + \frac{1}{12}\right)}} \approx 2.517$$

The p-value $2P(Z > 2.517)$ is about 0.0118. So, there is strong evidence to conclude that the proportions are different.

Concept Review Solutions

C1. B. Here, $$n_1 = 1000,\ \hat{p}_1 = 0.26\ (in\ 1998),\ n_2 = 1000,\ \hat{p}_2 = 0.30\ (in\ 2008)$$ and certainly each of the products $n_1\hat{p}_1$, $n_1(1-\hat{p}_1)$, $n_2\hat{p}_2$, and $n_2(1-\hat{p}_2)$ is at least 5. The randomness conditions are also met. The 90% confidence interval is $$(\hat{p}_1 - \hat{p}_2) \pm z^* \cdot \sqrt{\frac{\hat{p}_1(1-\hat{p}_1)}{n_1} + \frac{\hat{p}_2(1-\hat{p}_2)}{n_2}} = (0.30 - 0.26) \pm 1.645\sqrt{\frac{(0.26)(1-0.26)}{1000} + \frac{(0.30)(1-0.30)}{1000}} = 0.04 \pm 0.033$$ or about (0.007, 0.073). Since the interval does not contain 0, there is evidence of a change in proportion from 1998 to 2008.
C2. E. By definition of Type II error.
C3. C
C4. A. Because 0 is in the confidence interval, it is plausible that there is no difference between the two teachers. Choice E is correct because although the difference in pass rates was 20%, that is not statistically significant because of the width of the confidence interval, which is a result of sample sizes of only 25 students.
C5. C. The probability that the null hypothesis will be rejected when it is true is equal to α.
C6. B. All conditions are satisfied for a two-sample test for the difference of two proportions. The *z*-score is about -1.92 with a one-sided *p*-value of 0.027. Therefore, reject the null hypothesis that there is no difference between the proportion of yellow M&M's and the proportion of yellow Skittles. The difference in the two proportions cannot reasonably be attributed to chance variation alone.
C7. a. We wish to test: H_o: $p_1 - p_2 = 0$, Ha: $p_1 - p_2 \neq 0$, where p_1 is the proportion of people who received the real acupuncture treatment who reported a reduction in migraines and p_2 is the proportion of people who received the sham acupuncture treatment who reported a reduction in migraines. **b.** Here, $\hat{p}_1 = \frac{12}{19}$, $\hat{p}_2 = \frac{8}{17}$, so that $\hat{p}_1 - \hat{p}_2 = \frac{12}{19} - \frac{8}{17} \approx 0.161$. **c.** The number of successes would be $$\hat{p}_1 \cdot n_1 + \hat{p}_2 \cdot n_2 = \hat{p}_1(36)\ \left(\text{assuming } \hat{p}_1 - \hat{p}_2 = 0\right).$$ **d.** Here, $\hat{p}_1 = \frac{9}{19}$, $\hat{p}_2 = \frac{11}{17}$, so that $\hat{p}_1 - \hat{p}_2 = \frac{9}{19} - \frac{11}{17} \approx -0.173$. **e.** Just identify the values in (b) and (d) along the *x*-axis. **f.** No. The distribution is centered around 0 and one would be prompted to reject the null hypothesis if only the vast majority of the differences were 0.

C8. a. 7 times out of 200, or 3.5% of the time. Given that the null hypothesis is that there is no difference between the proportions, we would reject the null hypothesis and claim that there is a difference in the proportions.

b. This conclusion is the same as the one in the example, albeit with a different p-value.