Chapter 8

Discussion Question Solutions

D1. You don't need a confidence interval for $\hat{p}$ because you already know exactly what that is from our sample and you know that it probably would have been different if you had taken a different sample. What you want is an interval that has a good chance of capturing the true but unknown proportion of successes *p* in the population from which the sample was taken.

D2. The quantity $n\hat{p} = 31,357.69$ or 31,358 is the number of first-year students in the sample that spend more than 20 hours per week preparing for class, and the quantity $n(1-\hat{p}) = 153,099.31$ or 153,099 is the number of first-year students in the sample that spend 20 hours or less preparing for class. In general, $n\hat{p}$ is the number of "successes" in the sample, and $n(1-\hat{p})$ is the number of "failures."

D3. The large sample sizes help establish the necessary conditions for the test of proportions. A low response rate from a survey is fairly common and should not surprising. The reason why it is low is the issue that could impact the test result. For instance, if many non-respondents were embarrassed to admit they spent little time preparing for class, then the observed percentages may not reflect the population.

D4. The width of the confidence interval is

$$2E = 2z^* \sqrt{\frac{\hat{p}(1-\hat{p})}{n}}$$

Assume *n* and z^* stay constant. A look at various values of $\hat{p}(1-\hat{p})$ at and around $\hat{p} = 0.5$

$\hat{p}$	0.2	0.3	0.4	0.5	0.6	0.7	0.8
$\hat{p}(1-\hat{p})$	0.16	0.21	0.24	0.25	0.24	0.21	0.16

shows $\hat{p}(1-\hat{p})$ decreases, and hence the width of the confidence interval decreases, away from $\hat{p} = 0.5$. In fact, it can be shown $\hat{p}(1-\hat{p})$ is maximum when $\hat{p} = 0.5$. Thus, in general, the interval width will decrease away from 0.5.

D5. The term *error attributable to sampling* means the same thing as *sampling error* or *variation due to sampling*. It means that when we take random samples from a given population, the values
of $\hat{p}$ do not turn out to be the same each time and usually aren't equal to *p*. However, these values do tend to cluster around *p*.

D6. In addition to the variation in sampling that results from taking a random sample from a given population, general categories of sources of error include

- not getting a random sample in the first place (such as voluntary response surveys where people phone in to talk shows or write in to newspapers)
- errors in coding or recording the responses (a poll taker misunderstands what a person is saying)
- getting an invalid response from the people surveyed because they misunderstood the question or because they did not tell the truth about a controversial issue

The formula for the margin of error takes into account only the variation that results from looking at a random sample and not at the entire population. It is generally impossible to predict the bias that may result from the three sources of error above.

D7. The width of the confidence interval is

$$2E = 2z^* \sqrt{\frac{\hat{p}(1-\hat{p})}{n}}$$

If $\hat{p}$ and n are constant, then the interval width is directly related to z^* and z^* determines the capture rate. If the interval width increases, the capture rate increases, and vice versa. For example, a value of $z^* = 1.96$ corresponds to a capture rate of 95%. To increase the interval width by a factor of 1.5, multiply 1.96 by 1.5 to get $z^* = 2.94$. This value of z^* corresponds to a capture rate of 99.7%.

D8. From Chapters 6 and 7, students should understand that the spread of the sampling distribution of the proportion of successes decreases as the sample size n increases. Examining the formula for the margin of error

$$E = z^* \sqrt{\frac{\hat{p}(1-\hat{p})}{n}}$$

shows that if z^* and $\hat{p}$ remain the same then an increase in n will increase the denominator making the value of E smaller.

D9. For $E = 10\%$, use

$$n = 1.96^2 \left(\frac{0.5(1-0.5)}{0.1^2} \right) = 96.04$$

So you need a sample size of 97, which costs 5(97), or \$485. (It is customary always to round up when computing sample size.)
For $E = 1\%$, use

$$n = 1.96^2 \left(\frac{0.5(1-0.5)}{0.01^2} \right) = 9604$$

which costs 5(9604), or \$48,020.
For $E = 0.1\%$, use

$$n = 1.96^2\left(\frac{0.5(1-0.5)}{0.001^2}\right) = 960,400$$

which costs 5(960,400), or $4,802,000.

To cut the margin of error by $\frac{1}{10}$ requires multiplying the sample size, and the cost, by 100.

D10. a. The test is to see if evidence indicates whether people can correctly identify the gourmet coffee or do they just guess. Assuming they guess, then a person has a one out of three chance of being correct so $p_0 = \frac{1}{3}$. This is a one-sided test as the question is whether tasters do better than just guessing. The hypotheses are then

$H_0: p = \frac{1}{3},\ H_a: p > \frac{1}{3}$ where p is the probability a person correctly identifies the gourmet coffee. For the given sample, $\hat{p} = \frac{52}{100} = 0.52.$

b. If there is no discrimination then the probability p that a woman is hired will be 0.4, since 40% of the applicants are women. Thus $p_0 = 0.4.$ This is a two-sided test because we are looking for evidence of gender discrimination. If p is significantly less than 0.4, then there is discrimination towards women, while if p is significantly greater than 0.4, then the discrimination is toward men. The hypotheses are

$H_0: p = 0.4,\ H_a: p \neq 0.4$ where p is the probability a woman is hired. For the given sample, $\hat{p} = 15\% = 0.15.$

D11. A z-score of 0 tells you that your sample proportion is equal to the hypothesized standard. In other words, $\hat{p} = p_0$.

D12. a. The test statistic z will increase. An increase in the sample size n makes the denominator of the test statistic smaller, which makes the test statistic larger. This means that you are more likely to have a statistically significant result.

b. The test statistic z will increase because its numerator will increase. This means that you are more likely to have a statistically significant result.

D13. Yes, they could be wrong. The inescapable fact about statistical inference is that, no matter what decision rule you use, you can never be certain whether you are right or wrong. Additional testing with other groups would help support or disprove the conclusion.

D14. **a.** Here

$$z = \frac{\hat{p} - p_0}{\sqrt{\frac{p_0(1-p_0)}{n}}} = \frac{0.52 - 1/3}{\sqrt{\frac{1/3(1-1/3)}{100}}} = 3.96.$$

Using software, this corresponds to *P*-value of about 0.00004. This very small *P*-value is much less than the 0.05 level of significance so one would conclude the evidence supports the claim that people do have some ability to distinguish gourmet coffee from ordinary and instant brands.

b. In this situation we are not given a sample size but since we are given percentages we may take n to be 100. Then

$$z = \frac{\hat{p} - p_0}{\sqrt{\frac{p_0(1-p_0)}{n}}} = \frac{0.15 - 0.40}{\sqrt{\frac{0.40(1-0.40)}{100}}} = -5.10.$$

Again, using software, we find an extremely small *P*-value of 3.4×10^{-7}. So the evidence supports the claim that there is discrimination in the hiring practices of the local police department and the discrimination is toward women.

D15. The null hypothesis gives the parameter value upon which the probability is based. Without the model, you cannot calculate a probability. For example, suppose you spin a coin and record whether it falls heads up or tails up. You cannot determine the probability of getting a result as extreme as or more extreme than 17 heads out of 40 when spinning a coin because you don't know the probability of getting heads on a single spin of a coin. But you can ask, "What would the probability be if spinning a coin were fair?" Because that probability is reasonably large, it's plausible that spinning a coin is fair.

D16. No. Either the null hypothesis is true or it is false. The only probabilities possible for this are 1 or 0. The *P*-value is the conditional probability *P*(*a sample statistic as extreme as or more extreme than the result you got from your sample* | *the null hypothesis is true*).

D17. A *P*-value of 1 would indicate that seeing a value as extreme or more extreme than the value of the test statistic is guaranteed. There is no way you could not see a test statistic at least that extreme. In the case of a z-test, a *P*-value of 1 is theoretically not possible although values very close to 1 could occur. In later chapters, students will see tests where theoretically a *P*-value could be 1.

D18. a. Across the United States, the difference in percentages amounts to quite a few more boy babies than girl babies. There are approximately 4,000,000 births in the United States each year. So we would expect about 2,048,000 boys and 1,952,000 girls, or 96,000 more boys. People planning government or commercial services that depend on gender, such as manufacturers of clothing or other items for children, would find this an important difference.

b. The probability that all four children would all be boys is $(0.512)^4 \approx 0.069$ while the probability they would all be girls is $(0.488)^4 \approx 0.057$. Both probabilities are small and considerably close so there is little difference in the probability of four boys and four girls.

D19. When α is larger, z^* is smaller (in absolute value). This relationship makes it more likely to get a *z*-score as extreme as or more extreme than z^*. Thus it is easier to reject the null hypothesis if α is larger.

D20. a. A larger level of significance, α, makes it <u>easier</u> to reject the null hypothesis.

b. A smaller level of significance, α, makes it <u>harder</u> to reject the null hypothesis.

c. A one-sided test (on the correct side) makes it <u>easier</u> to reject the null hypotheses than a two-sided test.

D21. *p* is the true population proportion. It does not vary with the sample and is unknown.

p_0 is the hypothesized value of the population proportion. It does not vary with the sample and it is known in the sense that one knows what value you desire to test in test null hypothesis.

$\hat{p}$ is the sample proportion, that is, the proportion of observations in a random sample that possess the characteristic being studied. It varies with the sample and is "known" or computed from each sample.

D22. a. about 0.025; about 0.05

b. about 0.05; about 0.10

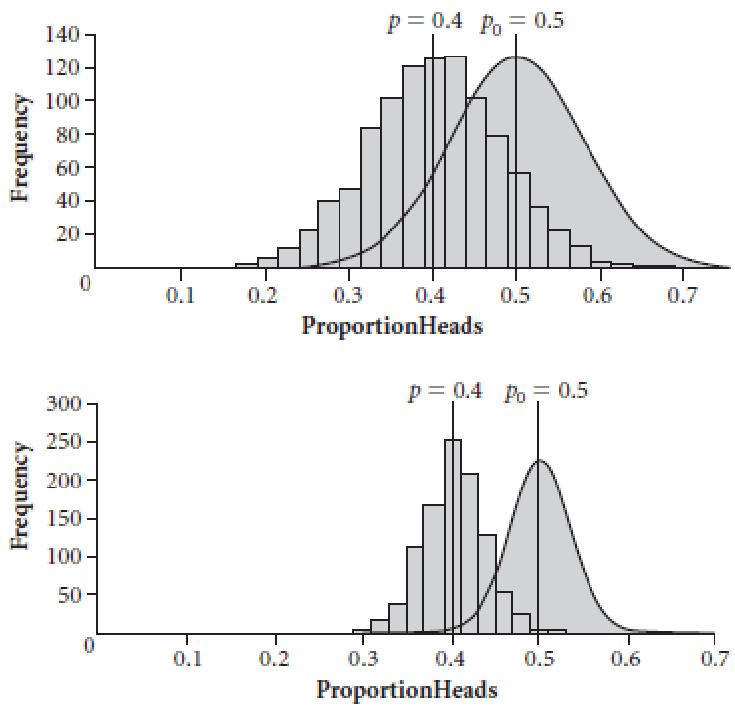
c. For example, for a fixed null hypothesis, the probability that z is larger than 1.96 depends on whether that null hypothesis is true. If p is indeed equal to p_0, that probability is 0.025. But if p is much larger than p_0, then the difference between $\hat{p}$ and p_0 will tend to be quite large, making z quite large, and so the probability that z will be greater than 1.96 could be almost certain.

D23. To reject a null hypothesis, you need a small *P*-value. It is easier to reject a false null hypothesis with a one-sided test (if you have the right direction). For a given test statistic z, the *p*-value will be half that of a two-sided test.

D24. The probability of rejecting a true null hypothesis is equal to the level of significance and so is equal to 0.02 here. This is a Type I error because the null hypothesis is actually true and it has been rejected.

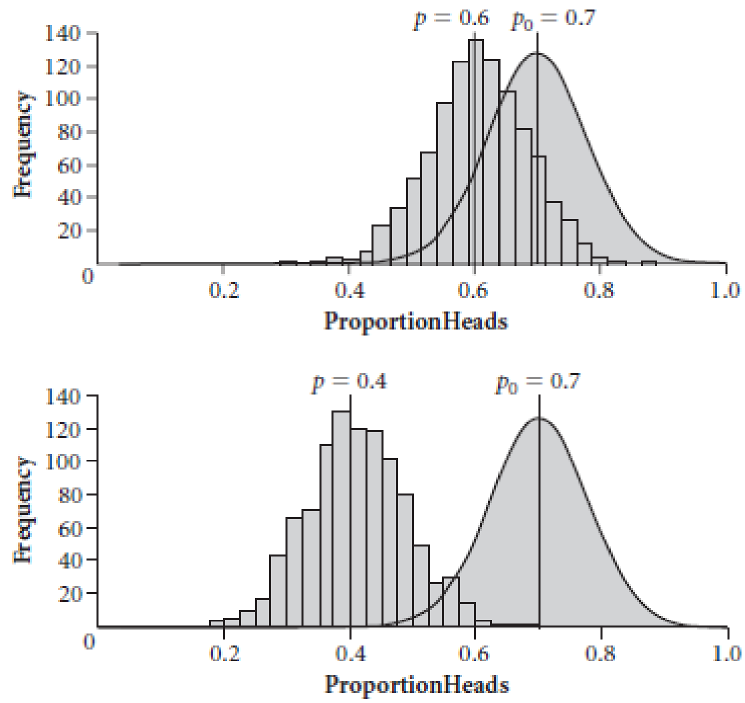
D25. Because there are too many situations in the world where the null hypothesis is actually false. If you use a large critical value, you will fail to reject the null hypothesis in many of these cases and so make many Type II errors.

D26. Power is the probability of rejecting a null hypothesis. When you increase the sample size, the sampling distribution of the sample proportion tends to cluster more closely around the actual population proportion *p*. Suppose you have a penny and hypothesize that the probability it lands heads when spun is 0.5. However, the actual probability is 0.4. The plots below illustrate this. The histogram in the top plot shows repeated trials of spinning the penny 40 times and then recording the proportion of spins that were heads. The histogram in the bottom plot shows repeated trials of spinning the penny 200 times and then recording the proportion of spins that were heads.



The normal curves show what the sampling distribution would look like assuming that the null hypothesis is true. Notice that with a sample size of 200 a much larger portion of the simulated sampling distribution (which is shown by the histogram) is in the rejection region of the normal curve than with a sample size of 40. This means you are more likely to reject the null hypothesis (therefore, increasing the power) with the larger sample size. See, also, the answer to P36.

D27. The power of the test would increase. The plots below show simulated sampling distributions for $p = 0.6$ and for $p = 0.4$. In both cases, the null hypothesis was $p = 0.7$. The normal curves show what the sampling distributions would look like assuming that the null hypothesis is true of $p = 0.7$. In the bottom plot, where *p* is farther from p_0, a much larger portion of the simulated sampling distribution (which is shown by the histogram) is in the rejection region for $p_0 = 0.7$. This means you would be more likely to reject the null hypothesis when p is farther from p_0, so the test has greater power.



D28. If you set $\alpha = 1$, then the critical values would be ± 0. Because any test statistic is as extreme as or more extreme than 0, you would then always reject the null hypothesis and the power would be 1. Similarly, if $\alpha = 0$, the critical values would not exist because no matter how far out in the tails you go in the normal distribution, the tail area has some nonzero area. So, in this case, you would never be able to reject the null hypothesis and the power would be 0. From a practical standpoint, nobody would bother to do a test under these conditions.

D29. Compute:
$(0.5)^5 = 0.03125 > 0.01$, $(0.5)^6 = 0.015625 > 0.01$, $(0.5)^7 = 0.0078125 < 0.01$. After 7 Sundays, you will have statistically significant evidence at the 0.01 level of significance to confront your uncle.

Practice Problem Solutions

P1. Note that $\hat{p} = \frac{14}{40} = 0.35$. So, the middle 95% of the sampling distribution of $\hat{p}$ is

$$0.35 \pm 1.96\sqrt{\frac{(0.35)(0.65)}{40}}$$

or about 0.202 to 0.498. Hence, we easily classify the given values as plausible or not plausible as follows:

a. Yes **b.** Yes **c.** No.

P2. Note that $\hat{p} = 0.08$. So, the middle 95% of the sampling distribution of $\hat{p}$ is $$0.08 \pm 1.96\sqrt{\tfrac{(0.08)(0.92)}{1166}}$$ or about 0.064 to 0.096. Hence, we easily classify the given values as plausible or not plausible as follows: **a.** No **b.** No **c.** Definitely not.
P3. a. Using n = 40 and $\hat{p} = \frac{25}{40} = 0.625$, we see that the 95% confidence interval is $$0.625 \pm 1.96\sqrt{\tfrac{(0.625)(0.375)}{40}}$$ or about 0.47 to 0.78. **b.** p is the proportion of males in the population from which the sample was selected who would give shocks up to the danger level.
P4. a. The population is the collection of all college-bound students at the time of the Survey. The parameter being estimated is the proportion of the population that reported reading about college rankings. And, $\hat{p} = 0.20$. **b.** Yes. Note that $$np = 500(0.20) = 100, \quad n(1-p) = 500(0.80) = 400.$$ The sample is taken from a binomial population (since there are only two responses allowed) and the population size certainly exceeds 5,000. **c.** Using n = 500 and $\hat{p} = 0.20$, we see that the 90% confidence interval is $$0.20 \pm 1.645\sqrt{\tfrac{(0.20)(0.80)}{500}}$$ or about 0.17 to 0.23. This means that among the college-bound students in the population from which this sample was selected, you are 90% confident that the percentage who would report they have read about college rankings is between 17% and 23%.
P5. You would expect 0.95(365) = 346.75 of the confidence intervals to contain the proportion.
P6. You would expect 0.95(350) = 332.50 of the confidence intervals to contain the proportion.
P7. a. Using n = 1144 and $\hat{p} = 0.71$, observe that $np \geq 10, \quad n(1-p) \geq 10$. The sample is taken from a binomial population (since there are only two responses allowed) and the population size certainly exceeds 11,400. So, the conditions are met. But, there might be a non-response bias. **b.** The 95% confidence interval is

$$0.71 \pm 1.96\sqrt{\tfrac{(0.71)(0.29)}{1144}}$$

or about 0.68 to 0.74.

c. Among the population of physicians from which this sample was selected, you are 95% confident that the percentage who would say that they had an obligation to refer the patent is between 68% and 74%.

d. The method used to produce the confidence interval captures the population proportion 95% of the time in repeated usage.

P8. Observe that the confidence interval is given by 0.51 ± 0.031, or about 0.479 to 0.541.

a. Yes
b. Note that

$$E = z^* \sqrt{\tfrac{(0.51)(0.49)}{1011}} = 0.016z^* .$$

So, if $z^* = 1.96$ (for 95% confidence), then the margin of error is 0.031. So, it is reported correctly.

P9. Observe that the confidence interval is given by 0.46 ± 0.03, or about 0.43 to 0.49.

a. No
b. Note that

$$E = z^* \sqrt{\tfrac{(0.46)(0.54)}{1400}} = 0.013z^* .$$

So, if $z^* = 1.96$ (for 95% confidence), then the margin of error is $0.026 \approx 0.03$. So, it is reported correctly.

P10. **a.** $z^* = 1.285$ for 80% confidence level.
b. The formula for margin of error is $E = z^* \sqrt{\tfrac{p(1-p)}{n}}$. Since the value of z^* increases as the confidence level increases, the margin of error will be larger for 95% confidence level than for 80% confidence level.

P11. **a.** Decrease because $z^* \sqrt{\tfrac{p(1-p)}{400}} > z^* \sqrt{\tfrac{p(1-p)}{800}}$.
b. Decrease because $1.96\sqrt{\tfrac{p(1-p)}{n}} > 1.645\sqrt{\tfrac{p(1-p)}{n}}$.
c. Decrease (for same reason as given in (b)).

P12. Using $\hat{p} = 0.36$, we see that the desired value of n is given by

$$n = (1.96)^2 \left[\frac{(0.36)(0.64)}{(0.03)^2} \right] \approx 984 .$$

P13. Using $\hat{p} = 0.50$ (since a value is not specified), we see that the desired value of n is given by

$$n = (1.96)^2 \left[\frac{(0.5)(0.5)}{(0.03)^2} \right] \approx 1067 .$$

P14. Because you have no estimate for p, use $p = 0.5$.

a. For $E = 2\%$ with 95% confidence, use

$$n = z^2 \left(\frac{p(1-p)}{E^2} \right) = 1.96^2 \left(\frac{0.5(1-0.5)}{0.02^2} \right) = 2401 .$$

b. For $E = 1\%$ with 99% confidence, use

$$n = z^2 \left(\frac{p(1-p)}{E^2} \right) = 2.576^2 \left(\frac{0.5(1-0.5)}{0.01^2} \right) = 16{,}589.44 \approx 16{,}590 .$$

c. For $E = 0.5\%$ with 90% confidence, use

$$n = z^2 \left(\frac{p(1-p)}{E^2} \right) = 1.645^2 \left(\frac{0.5(1-0.5)}{0.005^2} \right) = 27{,}060.25 \text{ or } 27{,}061 .$$

P15. a. The standard $p_0 = 0.5$.

b. Since it is asking if there is a change (period), the test should be two-sided. Hence, the null and alternative hypotheses are:

H_o: $p = 0.50$ H_a: $p \neq 0.50$, where p is the proportion of all of this year's juniors who would say they will want extra tickets

c. $\hat{p} = \frac{16}{40} = 0.40$

d. There is not strong evidence of change.

P16. a. The student is guessing, and so the probability the student gets a question right is 0.5. Or, using symbols, $p = 0.5$, where p is the probability a student gets a question right.

b. The test should be one-sided. The null and alternative hypotheses are:
H_o: $p = 0.50$ H_a: $p > 0.50$, where p is the probability that the student gets a question correct

c. The sample proportion is $\hat{p} = \frac{30}{40} = 0.75$.

d. The result is statistically significant. The vertical line leading from a sample proportion of $\hat{p} = 0.75$ does not intersect the middle 95% of the possible values of $\hat{p}$ when $p = 0.5$. Thus 0.75 is not a reasonably likely result if $p = 0.5$. We reject the hypothesis that the student was guessing.

e. This does not prove beyond any doubt that the student was not guessing. The student may have been guessing and been extremely lucky to get a score this high.

P17. For P15: $\dfrac{0.40-0.50}{\sqrt{\frac{(0.50)(0.50)}{40}}}=-1.265$ For P16: $\dfrac{0.75-0.50}{\sqrt{\frac{(0.50)(0.50)}{40}}}=3.162$

P18. a. The standard $p_0 = 0.5$.

b. The test should be one-sided. Hence, the null and alternative hypotheses are:

H_o: $p = 0.50$ H_a: $p > 0.50$, where p is the probability that the judge selects the correct dog

c. $\hat{p} = \frac{23}{45} = 0.511$

d. $z = \dfrac{0.511-0.50}{\sqrt{\frac{(0.50)(0.50)}{45}}} = 0.1476$

P19. If the 2-year-olds are guessing then $p_0 = 0.5$ and $\hat{p} = \frac{28}{50} = 0.56$. The test statistic is

$$z = \frac{0.56-0.5}{\sqrt{\frac{0.5\cdot 0.5}{50}}} \approx 0.849$$

The P-value is 0.198. If the 2-year-olds could not tell a difference between the rakes, the probability of getting the rake with the food 28 times or more is almost 20%. If it is true that 2-year-olds were selecting a rake at random, which would give them a 50% chance of selecting the rake that would get them food, the probability is 0.196 that the 2-year-olds would rake in food as often as they did. This does not provide sufficient evidence that the 2-year-olds were able to deliberately choose the rake that got them the food.

P20. a. It is given that the standard is $p_0 = \frac{2}{3} \approx 0.67$.

b. The test should be two-sided. Hence, the null and alternative hypotheses are:

H_o: $p = 0.67$ H_a: $p \neq 0.67$, where p is the proportion of all U.S. undergraduates who receive financial aid.

c. $\hat{p} = \frac{302}{400} = 0.755$

d. $z = \dfrac{0.755-0.67}{\sqrt{\frac{(0.67)(0.33)}{400}}} = 3.615$. The p-value is $2P(Z > 3.615) \approx 0.0003$.

e. There is strong evidence against H_o significant at 0.05.

P21. a. First, we test: H_o: $p = 0.50$ versus H_a: $p < 0.50$, where p is the probability of inattentional blindness.

Observe that the test statistic is

$$z = \frac{0.46 - 0.50}{\sqrt{\frac{(0.50)(0.50)}{192}}} \approx -1.109 .$$

The p-value is P(Z<-1.109) = 0.1561. As such, there is insufficient evidence to reject the null hypothesis that the inattentional blindness rate is (at least) 50%.

b. The plausible proportions are given by the confidence interval centered around the value of p in H_0. That is,

$$0.5 \pm 1.96\sqrt{\frac{(0.50)(0.50)}{192}} = 0.5 \pm 0.0707 ,$$

or about 0.4293 to 0.5707.

P22. a. Note that $np = 266(.02) = 5.3$, which is less than 10, and the sampling distribution is skewed toward the larger values. So, the conditions are not met for this test. As such, a normal approximation might not be the best approximation to Display 8.10.

b. No, because the skewness is not drastic.

P23. a. The alternative hypothesis is:
H_a: $p < 0.05$, where p is the defective rate for the new supplier.

b. Using $n = 300$, $\hat{p} = 0.03$, and $p_0 = 0.05$, we see that the test statistic is

$$z = \frac{0.03 - 0.05}{\sqrt{\frac{(0.05)(0.95)}{300}}} \approx -1.589 .$$

The p-value is P(Z< -1.589) =0.056.
c. No, there is not sufficient evidence at the 1% level since the p-value is greater than 0.01.

P24. a. The test should be one-sided since you want to make certain the proportion is at least a certain value.

b. The null and alternative hypotheses are:
H_o: $p = 0.87$ H_a: $p > 0.87$, where p is the proportion of Down syndrome detected by the new test

c. Using $n = 200$, $\hat{p} = \frac{182}{200}$, and $p_0 = 0.87$, we see that the test statistic is

$$z = \frac{\frac{182}{200} - 0.87}{\sqrt{\frac{(0.87)(0.13)}{200}}} \approx 1.68 .$$

The p-value is P(Z>1.68) =0.0465.

d. Yes, there is evidence to support the alternative that the detection rate exceeds 87% since the p-value is less than 0.05.

P25. a. The test should be one-sided.

b. The null and alternative hypotheses are:
H_o: $p = 0.55$ H_a: $p < 0.55$, where *p* is the proportion of all likely voters who support the bond issue

c. Using $n = 700$, $\hat{p} = 0.53$, and $p_0 = 0.55$, we see that the test statistic is

$$z = \frac{0.53 - 0.55}{\sqrt{\frac{(0.55)(0.45)}{700}}} \approx -1.064 .$$

The p-value is P(Z<-1.064) =0.1437.

d. No, there isn't statistically significant evidence to support the alternative that the support has dropped below 55% since the p-value exceeds 0.01.

P26. a. The student is guessing, and so the probability that he or she gets any one question correct is 0.2.
Or, $p = 0.2$, where *p* is the probability that the student gets any one question correct.

b. $p = 0.104$, where *p* is the proportion of veterans in your county.

c. Of the people who wash their car once a week, $\frac{1}{7}$ wash them on Saturday.
Or, $p = \frac{1}{7}$, where *p* is the proportion of people who wash their cars once a week who wash them on Saturday.

P27. B

P28. The null and alternative hypotheses are:
H_o: $p = (18/38)$ H_a: $p \neq (18/32)$, where *p* is the probability of a red

Using $n = 500$, $\hat{p} = \frac{194}{500}$, and $p_0 = \frac{18}{38}$, we see that the test statistic is

$$z = \frac{\frac{194}{500} - \frac{18}{38}}{\sqrt{\frac{\left(\frac{18}{38}\right)\left(\frac{20}{38}\right)}{500}}} \approx -3.837 .$$

The p-value is 2P(Z<-3.837) nearly 0.0001.

Hence, we reject the null hypothesis that the probability of red is (18/38) for this wheel.

P29. The null and alternative hypotheses are:
H_o: $p = 0.07$, H_a: $p > 0.07$, where *p* is the probability that a bear will select a mini-van to break into.

Using $n = 412$, $\hat{p} = \frac{120}{412}$, and $p_0 = 0.07$, we see that the test statistic is

$$z = \frac{\frac{120}{412} - 0.07}{\sqrt{\frac{(0.07)(0.93)}{412}}} \approx 17.60 .$$

The p-value is P(Z >17.60), which is nearly 0.

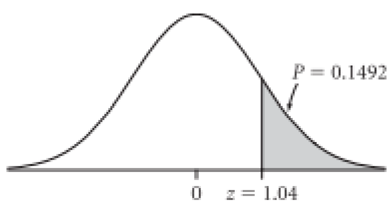
Hence, we reject the null hypothesis that the probability that a bear will select a mini-van to break into is 0.07 and conclude that bears break into mini-vans more often than would be predicted by knowing that 7% of all vehicles parked overnight are mini-vans.

P30. ***Check conditions.*** The conditions are met for doing a test of significance for a proportion because you have a random sample, both $np_0 = 169(0.51) = 86.19$ and $n(1 - p_0) = 169(1 - 0.51) = 82.81$ are at least 10, and the number of teens in your community (probably) is at least 10(169) = 1690. If this latter condition doesn't hold, the test will be conservative.

State your hypotheses. The null hypothesis is that 51% (or even fewer) of teens in your community know this fact. The alternate hypothesis is that the percentage is greater than 51%.

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

$$z = \frac{\hat{p} - p_0}{\sqrt{\frac{p_0(1-p_0)}{n}}} = \frac{0.55 - 0.51}{\sqrt{\frac{0.51(1-0.51)}{169}}} \approx 1.04$$



The *P*-value from Table A is about 0.1492. Thus, 0.55 is much the sort of sample proportion you might get when taking a random sample of size 169 from a population with $p = 0.51$. Doing the test on the calculator, $z = 1.0479$ and the *P*-value is about 0.1473.

Write a conclusion in context. You should not reject the null hypothesis. If it is true that the percentage of teens in your community who know epilepsy is not contagious is only 51%, then it is reasonably likely to get 55% who know this in a random sample of 169 teens.

P31. a. Because Hila is using a significance level of 0.05, she should reject the null hypothesis if her test statistic is larger than 1.96 or smaller than –1.96. The value of her test statistic is

$$z = \frac{\hat{p} - p_0}{\sqrt{\frac{p_0(1-p_0)}{n}}} = \frac{0.25 - (1/6)}{\sqrt{\frac{0.167(1-0.167)}{100}}} \approx 2.234$$

Hila rejects the null hypothesis.

b. She has made a Type I error. The null hypothesis is true, and she rejected it.

P32. a. The two-sided p-value is about 0.088. Hence, he should not reject the null hypothesis that the lottery is fair with regard to the number 1.

b. Jack did not make an error, at the 5% level of significance.

P33. a. The choices are random; since there are five such choices, we have $p = 0.2$, where p is the probability of a correct choice. The researchers failed to reject the null hypothesis.

b. The researchers could have made a Type II error because the null hypothesis was not rejected.

c. The researchers would have failed to discover that the subjects could choose the dog food more often than chance would predict.

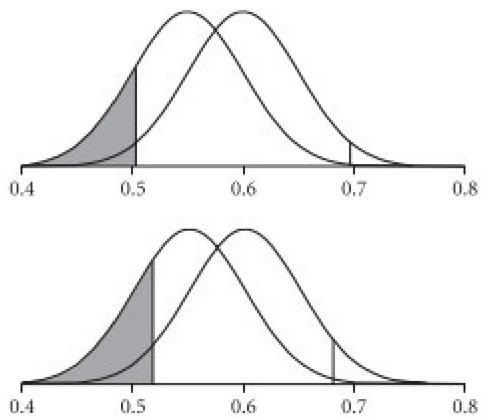
P34. a. A Type-I error would be made when concluding that the proportion of children living near freeways in Los Angeles have a higher rate of asthma than the general population of children when they don't.

A serious consequence of this would be scaring parents into moving away from freeways.

b. A Type-II error would be made when concluding that it's plausible that the rates are equal when children living near freeways really have a higher rate.

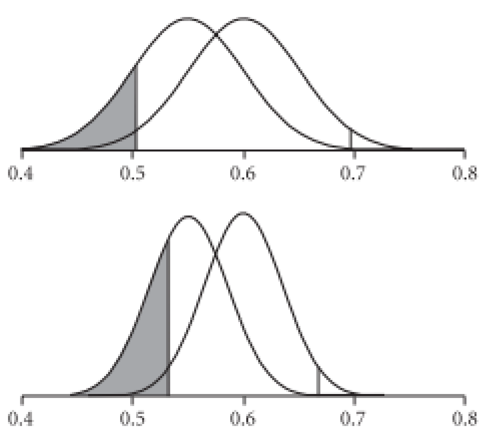
A serious consequence of this would be failing to warn parents about a serious health problem for their children.

P35. The test with the higher significance level ($\alpha = 0.10$) has greater power because there is a greater chance of rejecting the null hypothesis. In the graphs below, the hypothesized value p_0 is 0.6 and the unknown population proportion p is 0.55. The solid shaded region shows the reasonably likely values of the normal approximation for the binomial distribution if the null hypotheses were true. The top graph has the middle 95% shaded ($\alpha = 0.05$) and the bottom graph has the middle 90% shaded ($\alpha = 0.10$). The normal distribution to the left in each graph is the normal approximation for the actual sampling distribution of the sample proportion for $p = 0.55$. The region indicated by the striped shading shows the portion of that normal distribution that lies in the rejection region for the hypothesized value p_0. This region is larger in the bottom graph indicating a greater likelihood of rejecting the null hypothesis and thus, greater power.

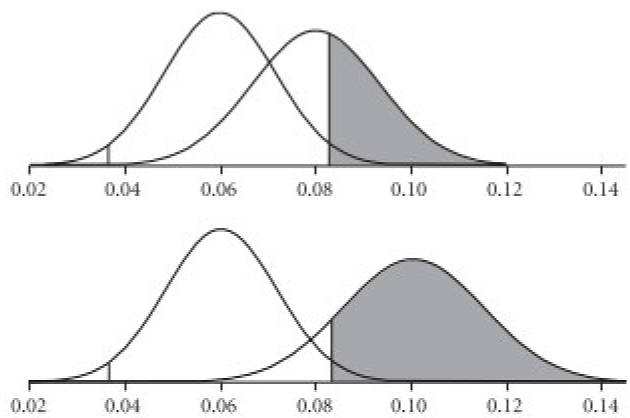


P36. The test with the larger sample size (200) has greater power. In the graphs below, the hypothesized value p_0 is 0.6 and the unknown population proportion p is 0.55. The solid shaded region shows the reasonably likely values of the normal approximation for the binomial distribution if the null hypotheses were true. Both graphs have the middle 95% shaded (α = 0.05). The normal distribution to the left in each graph is the normal approximation for the actual sampling distribution of the sample proportion for p = 0.55.

The larger sample size results in a smaller standard error for the sampling distribution of the sample proportion and the distribution is clustered more closely to the mean, as shown in the second graph. There is less overlap of the two distributions in the second graph. The region indicated by the striped shading shows the portion of that normal distribution that lies in the rejection region for the hypothesized value p_0. This region is larger in the bottom graph indicating a greater likelihood of rejecting the null hypothesis and thus, greater power.



P37. The power is greater when the true population proportion is farther from p_0. In the graphs below, the hypothesized value p_0 is 0.06 and the unknown population proportion p is 0.08 in the first graph and 0.10 in the second graph. The solid shaded region shows the reasonably likely values of the normal approximation for the binomial distribution if the null hypotheses were true. Both graphs have the middle 95% shaded (α = 0.05). The normal distribution to the right in each graph is the normal approximation for the actual sampling distribution of the sample proportion for $p = 0.08$ in the first graph and $p = 0.10$ in the second. Notice that there is less overlap of the two normal distributions in the second graph. The region indicated by the striped shading shows the portion of that normal distribution that lies in the rejection region for the hypothesized value p_0. This region is larger in the bottom graph indicating a greater likelihood of rejecting the null hypothesis and thus, greater power.



Notice that there is less overlap of the two normal distributions in the lower graph.

Exercise Solutions

E1. a. Using $n = 1100,\ \hat{p} = 0.19,\ z^* = 1.96,$ we see that the 95% confidence level is

$$0.19 \pm 1.96\sqrt{\frac{(0.19)(0.81)}{1100}}$$

or about 0.167 to 0.213.

b. The margin of error is $E = 1.96\sqrt{\frac{(0.19)(0.81)}{1100}} \approx 0.023184$.

E2. a. Using $n = 1009,\ \hat{p} = 0.92,\ z^* = 1.96,$ we see that the 95% confidence level is

$$0.92 \pm 1.96\sqrt{\frac{(0.92)(0.08)}{1009}}$$

or about 0.903 to 0.937.

b. The margin of error is $E = 1.96\sqrt{\frac{(0.92)(0.08)}{1009}} \approx 0.016740$.

E3. a. The population is all college students graduating from 4-year colleges; the parameter being estimated is the proportion who are proficient in quantitative literacy; $\hat{p} = 0.34$ since it is measured from the sample.

b. Using n = 1000 and $\hat{p} = 0.34$, observe that $np \geq 10,\quad n(1-p) \geq 10$. The sample is taken from a binomial population (since there are only two responses allowed) and the population size certainly exceeds 10,000. So, the conditions are met.

c. Using $n = 1000$, $\hat{p} = 0.34$, $z^* = 1.96$, we see that the 95% confidence level is

$$0.34 \pm 1.96\sqrt{\frac{(0.34)(0.66)}{1000}}$$

or about 0.311 to 0.369. This means that among the college students graduating from 4-year colleges from which this sample was selected, you are 95% confident that the percentage who would test proficient in quantitative literacy is between 31% and 37%.

d. Using $n = 800$, $\hat{p} = 0.18$, $z^* = 1.96$, we see that the 95% confidence level is

$$0.18 \pm 1.96\sqrt{\frac{(0.18)(0.82)}{800}}$$

or about 0.153 to 0.207. This means that among the college students graduating from two-year colleges from which this sample was selected, you are 95% confident that the percentage who would test proficient in quantitative literacy is between 15% and 21%.

e. Yes, because the confidence intervals do not overlap.

E4. The margin of error for 2-year college students is

$$E = 1.96\sqrt{\frac{(0.18)(0.82)}{800}} \approx 0.026623.$$

The margin of error for 4-year college students is

$$E = 1.96\sqrt{\frac{(0.34)(0.66)}{1000}} \approx 0.029361.$$

The sample proportion and sample size account for the differences in the widths.

E5. **a.** The population is men of the Oxford cohort of patients; the parameter being estimated is proportion of men classified as overweight.

b. Using n = 279 and $\hat{p} = 0.409$, observe that $np \geq 10$, $n(1-p) \geq 10$. The sample is taken from a binomial population (since there are only two responses allowed) and the population size certainly exceeds 2,790. So, the conditions are met.

c. Using $n = 279$, $\hat{p} = 0.409$, $z^* = 1.96$, we see that the 95% confidence level is

$$0.409 \pm 1.96\sqrt{\frac{(0.409)(0.591)}{279}}$$

or about 0.35 to 0.47.

d. Using $n = 279$, $\hat{p} = 0.409$, $z^* = 1.645$, we see that the 90% confidence level is

$$0.409 \pm 1.645\sqrt{\frac{(0.409)(0.591)}{279}}$$

or about 0.36 to 0.46.

e. The 95% interval because to have a greater chance of capturing the true parameter value, the interval must be wider.

E6. a. The population is women of the Oxford cohort of patients; the parameter being estimated is proportion of women classified as overweight.

b. Using $n = 433$ and $\hat{p} = 0.27$, observe that $np \geq 10$, $n(1-p) \geq 10$. The sample is taken from a binomial population (since there are only two responses allowed) and the population size certainly exceeds 4,330. So, the conditions are met.

c. Using $n = 433,\ \hat{p} = 0.27,\ z^* = 1.96,$ we see that the 95% confidence level is

$$0.27 \pm 1.96\sqrt{\tfrac{(0.27)(0.73)}{433}}$$

or about 0.228 to 0.312.

d. The 95% margin of error for women is 0.041817 and for men it is 0.057691. Hence, the margin error is less for women than for men.

e. Yes, because the confidence intervals do not overlap.

E7. a. Using $n = 600$ and $\hat{p} = 0.65$, observe that $np \geq 10$, $n(1-p) \geq 10$. The sample is taken from a binomial population (since there are only two responses allowed) and the population size certainly exceeds 6,000. So, the conditions are met.

b. Using $n = 600,\ \hat{p} = 0.65,\ z^* = 1.96,$ we see that the 95% confidence level is

$$0.65 \pm 1.96\sqrt{\tfrac{(0.65)(0.35)}{600}}$$

or about 0.61 to 0.69. Note that the margin of error is $1.96\sqrt{\tfrac{(0.65)(0.35)}{600}}$, or about 0.04.

c. Using $n = 600,\ \hat{p} = 0.65,\ z^* = 1.645,$ we see that the 90% confidence level is

$$0.65 \pm 1.645\sqrt{\tfrac{(0.65)(0.35)}{600}}$$

or about 0.62 to 0.68. Note that the margin of error is $1.645\sqrt{\tfrac{(0.65)(0.35)}{600}}$, or about 0.03.

d. The 95% interval because to have a greater chance of capturing the true parameter value, the interval must be wider.

E8. a. Using $n = 600$ and $\hat{p} = 0.04$, observe that $np \geq 10$, $n(1-p) \geq 10$. The sample is taken from a binomial population (since there are only two responses allowed) and the population size certainly exceeds 6,000. So, the conditions are met.

b. Using $n = 600,\ \hat{p} = 0.04,\ z^* = 1.96,$ we see that the 95% confidence level is

$$0.04 \pm 1.96\sqrt{\tfrac{(0.04)(0.96)}{600}}$$

or about 0.024 to 0.056.

c. This interval not only does not overlap the confidence interval in E7b, the values are considerably less than those in E7b. This is because 65% of students claimed school helped them, so the corresponding grades that they would assign would likely be better than a C, and specifically not D.

E9. a. Using $n = 4884,\ \hat{p} = 0.0108,\ z^* = 1.96,$ we see that the 95% confidence level is

$$0.0108 \pm 1.96\sqrt{\tfrac{(0.0108)(0.9892)}{4884}}$$

or about 0.008 to 0.014.

b. Using $n = 4884,\ \hat{p} = 0.0613,\ z^* = 1.96,$ we see that the 95% confidence level is

$$0.0613 \pm 1.96\sqrt{\tfrac{(0.0613)(0.9387)}{4884}}$$

or about 0.054 to 0.068.

c. Yes, because the intervals do not overlap.

E10. a. Using $n = 4675,\ \hat{p} = 0.47,\ z^* = 1.96,$ we see that the 95% confidence level is

$$0.47 \pm 1.96\sqrt{\tfrac{(0.47)(0.53)}{4675}}$$

or about 0.456 to 0.484.

b. Using $n = 4675,\ \hat{p} = 0.39,\ z^* = 1.96,$ we see that the 95% confidence level is

$$0.39 \pm 1.96\sqrt{\tfrac{(0.39)(0.61)}{4675}}$$

or about 0.376 to 0.404.

c. Yes, because the intervals do not overlap.

E11. a. The 95% confidence intervals are as follows:

<u>Black</u>: Using $n = 634,\ \hat{p} = 0.74,\ z^* = 1.96,$ we see that the 95% confidence level is

$$0.74 \pm 1.96\sqrt{\tfrac{(0.74)(0.26)}{634}}$$

or about 0.71 to 0.77.

<u>White</u>: Using $n = 567,\ \hat{p} = 0.70,\ z^* = 1.96,$ we see that the 95% confidence level is

$$0.70 \pm 1.96\sqrt{\tfrac{(0.70)(0.30)}{567}}$$

or about 0.66 to 0.737.

<u>Hispanic</u>: Using $n = 314,\ \hat{p} = 0.66,\ z^* = 1.96,$ we see that the 95% confidence level is

$$0.66 \pm 1.96\sqrt{\tfrac{(0.66)(0.34)}{314}}$$

or about 0.61 to 0.71.

b. Black and Hispanic may differ because their intervals barely overlap.

c. For the population of black youth from which this sample was selected, you are 95% confident that the percentage who believes they have the skills to participate in politics is between 71% and 77%.

d. Only students with telephones could have been selected; payment for services may have attracted students from lower income brackets.

E12. a. You should wonder whether the sample was selected randomly from all teens and adults in the US to see if conditions for a confidence interval are met.

b. As stated in part a, the simple random sample condition may or may not have been met. There are several sample proportions listed. The most extreme (farthest from 0.5) proportion listed for all teens is 9%. Since $n\,\hat{p} = 500 \bullet 0.09 = 45$ and $n(1 - \hat{p}) = 500 \bullet 0.91 = 455$ are both at least 10, and there are more than $500 \bullet 10 = 5000$ teens in the US, the other two conditions for a confidence interval for each of the proportions of all teens are met.

For the statement that 4% of girls picked engineering as the field that most interested them, we are unable to determine whether $n\hat{p}$ is greater than 10 because we don't know how many of the sampled teens were girls. However, the population is clearly more then ten times the sample size.

For adults, $n\hat{p}$ and $n(1 - \hat{p})$ are both at least 10 because the sample size is bigger than for teens and all the proportions are larger than 9%. Also, there are more than $10 \bullet 1030$ adults in the US.

c. The largest sample proportion given for the teens was 33%, so the margin of error for teens would be at most $1.96\sqrt{\dfrac{0.33 \cdot 0.67}{500}} \approx 0.0412$. Giving the margin of error as 4% is slightly understating it. For adults, the largest sample proportion given was 45%, the margin of error would be $1.96\sqrt{\dfrac{0.45 \cdot 0.55}{1{,}030}} \approx 0.0304$, which is very close to 3%.

d. You are 95% confident that the proportion of all teens in the US who think engineering is an attractive career choice is between 0.10 and 0.18.

e. Suppose we could take 100 random samples from this population and construct the 100 resulting confidence intervals. We'd expect that the actual proportion of all U.S. teens who think engineering is an attractive career choice would be in about 95 of these intervals.

E13. a. The text doesn't say whether or not this is a random sample of adults, but, given that it was done by *U.S. News & World Report,* it is likely some randomization was involved. Therefore, our formulas give a reasonable approximation to the margin of error.

Both $n\hat{p}$ = 0.81(1000) = 810 and $n(1 - \hat{p})$ = 0.19(1000) = 190 are at least 10. Finally, the number of adults in the United States is greater than 10(1000). We are 95% confident that if we were to ask *all* adults from the general public whether they thought TV contributed to a decline in family values, the percentage would be between 78.6% and 83.4%. The computations follow: $\hat{p} \pm z^* \cdot \sqrt{\frac{\hat{p}(1-\hat{p})}{n}} = 0.81 \pm 1.96\sqrt{\frac{0.81(1-0.81)}{1000}} \approx 0.81 \pm 0.243$ **b.** No, the low response rate of 9.4% could seriously bias the results.
E14. a. This is a multistage sampling plan with no apparent randomization. The Epilepsy Foundation selected 20 affiliates (assuming they have more than that), each of which selected schools, presumably in their local area. The surveys were passed out to (all, some?) of the students in these schools. We probably should consider this a nationwide convenience sample. The lack of randomization makes it impossible to draw reliable conclusions from the sample. (When it says the results were "weighted" by age and region. This means that, for example, if 2% of the teens in their sample were 14-year-olds from the South, but 4% of the teens nationwide are 14-year-olds from the South, they would double-count the responses of each 14-year-old from the South in the sample.) **b.** No, there is no indication of a random sample. However, $n\hat{p}$ = 19,441 • 0.51 = 9914.91 and $n(1 - \hat{p})$ = 19,441 • 0.49 = 9526.09 are both greater than 10, and there are more than 19,441 • 10 = 194,441 teens in the US, so the other two conditions have been met. **c.** Yes. $E = z^* \cdot \sqrt{\frac{\hat{p}(1-\hat{p})}{n}} = 1.96\sqrt{\frac{0.51(1-0.51)}{19,441}} \approx 0.0070 \approx 0.01$ **d.** Suppose we could take 100 random samples from this population and construct the 100 resulting confidence intervals. We'd expect that the actual proportion of all U.S. teens who know that epilepsy isn't contagious would be in about 95 of these intervals.
E15. D, F, and H are the correct interpretations.
E16. D
E17. You should have used a sample size 9 times as big, or 9*n*. The margin of error is given by the formula $z^* \cdot \sqrt{\frac{p(1-p)}{n}}$ If you want this to be $\frac{1}{3}$ as large as it was before, you must solve for the new sample size, *m,* in the equation

$$\frac{1}{3}\cdot z^*\cdot\sqrt{\frac{p(1-p)}{n}} = z^*\cdot\sqrt{\frac{p(1-p)}{m}}$$

which will give the result $m = 9n$.

Hence, using $n = 200$, we would use a sample of size 1800.

E18. Increase the sample size by a factor of 16. Because quadrupling the sample size cuts the margin of error in half, quadrupling it again cuts it by one-fourth. To see this algebraically, suppose that a sample of size n gives a margin of error E. Then to get an error of $\frac{E}{4}$, the new sample size would have to be

$$z^2\cdot\frac{p(1-p)}{(E/4)^2} = 16z^2\cdot\frac{p(1-p)}{E^2} = 16n$$

E19. You would expect 0.90(80) = 72 of these intervals to include the proportion 0.60.

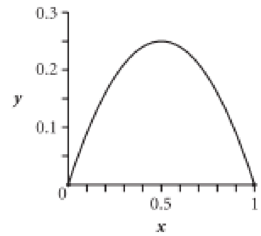
E20. One. Assuming that the sample size is at least 400 so that np and $n(1-p)$ are both at least 10, and that more than 4000 mice are produced in a week so that the sample size is less than 10% of the sample size, we would expect 95% of the intervals constructed to capture the population proportion of 0.04, meaning that 5% of the 20 samples, or one sample, would not capture this proportion.

E21. The symbol p is used for the proportion of successes in the population from which we are drawing a sample. This is the unknown parameter—the value that we are trying to estimate. The symbol $\hat{p}$ is used for the proportion of successes in a sample drawn from the population with proportion of successes p. The value of $\hat{p}$ varies from sample to sample. When constructing a confidence interval, the value of $\hat{p}$ from the sample is at the center of the confidence interval and so is always in it. The value of p may or may not be in the confidence interval.

E22. a. The sample size n appears in the denominator of the fraction in the formula for the margin of error. Thus as n gets larger, that fraction gets smaller, and so the length of the confidence interval gets smaller.

b. The length of the confidence interval increases. If you want to have more confidence that the interval captures the true population percentage p, you have to have a longer interval. This is seen in the formula, as z^* must be larger to have a larger probability of having $\hat{p}$ in the interval around p.

E23. a. The graph is the graph of the parabola $y = x - x^2$. This parabola opens down as shown here:



b. Because $x(1 - x)$ is of the form of $p(1 - p)$, which we are trying to maximize. Using y and x allows you to graph the function on a graphing calculator as you are asked to do in part b. The domain of x is restricted because a probability can be at most 1 and must be at least 0.

c. The maximum y-value occurs at the vertex. The vertex for a parabola $y = ax^2 + bx + c$ occurs at $x = \frac{-b}{(2a)}$. Here, $a = -1$ and $b = 1$, so the vertex is $x = \frac{-1}{2(-1)} = \frac{1}{2}$. The value of y at $x = \frac{1}{2}$ is $y = \frac{1}{2}(1 - \frac{1}{2}) = \frac{1}{4}$. If you graph this function on a graphing calculator you could also locate the coordinates of the vertex using the **maximum** function.

d. The standard deviation of $\hat{p}$ is maximized at $p = 0.5$.

E24. The method used in this section is based on the idea that middle 95% of sample proportions will be within about 1.96 standard errors of the population proportion. That 1.96 comes from a normal model, so this idea depends on the distribution of the sample proportion being approximately normal. As you learned earlier, larger samples give the distribution of the sample proportion a more normal shape, and when the sample is large enough so that np and $n(1 - p)$ are both at least 10, then the normal approximation to the shape of the sampling distribution gives reasonably accurate results.

E25. a. No, because the sample size 20 is greater than 0.10(135).

b. Using $n = 20,\ \hat{p} = 0.5,\ z^* = 1.96,$ we see that the 95% confidence level is

$$0.5 \pm 1.96\sqrt{\frac{(0.5)(0.5)}{20}}$$

or about 0.281 to 0.719.

c. Using $N = 135,\ n = 20,\ \hat{p} = 0.5,\ z^* = 1.96,$ we see that the 95% confidence level is

$$0.5 \pm 1.96\sqrt{\frac{(0.5)(0.5)}{20}}\sqrt{\frac{135-20}{135-1}}$$

or about 0.297 to 0.703.

d. The interval with the finite population correction is shorter; you should be able to estimate more accurately if the sample consists of a high percentage of the population values.

e. In such case, $\frac{N-n}{N-1} \sim 1$, and so $\sqrt{\frac{N-n}{N-1}} \sim 1$.

E26. a. No, because the sample size 235 is greater than 0.10(389).

b. Using $n = 235,\ \hat{p} = 0.834,\ z^* = 1.96,$ we see that the 95% confidence level is

$$0.834 \pm 1.96\sqrt{\frac{(0.834)(0.166)}{235}}$$

or about 0.786 to 0.882.

c. Using $N = 389,\ n = 235,\ \hat{p} = 0.834,\ z^* = 1.96,$ we see that the 95% confidence level is

$$0.834 \pm 1.96\sqrt{\frac{(0.834)(0.166)}{235}}\sqrt{\frac{389-235}{389-1}}$$

or about 0.804 to 0.864.

d. The interval with the finite population correction is shorter; you should be able to estimate more accurately if the sample consists of a high percentage of the population values.

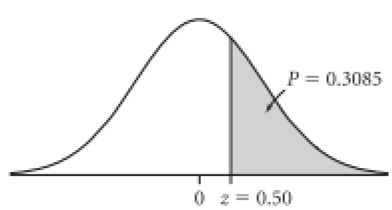
E27. a. $\frac{2}{3}$. You can explain this in two ways. The first way is to have the first student sit down anywhere. Then the probability the second student sits corner-to-corner is $\frac{2}{3}$ because, of the three seats that are left, two are adjacent seats. The other explanation is to list all 12 of the ways two students, A and B, can sit in chairs 1, 2, 3, and 4.

b. The null hypothesis is that the proportion of pairs of students who are seated on adjacent sides of a table is $\frac{2}{3}$.

c. The test should be one-sided. So, the alternate hypothesis is that the proportion of pairs is greater than $\frac{2}{3}$. (This indicates a preference for sitting on adjacent sides.)

d. The test statistic is

$$z = \frac{\hat{p} - p_0}{\sqrt{\frac{p_0(1-p_0)}{n}}} = \frac{0.70 - 2/3}{\sqrt{\frac{(2/3)(1-2/3)}{50}}} \approx 0.5$$



e. The p-value for this one-sided test is about 0.3085. We don't reject the null hypothesis.

f. The probability that two students will sit on adjacent sides of a table just by chance is $\frac{2}{3}$. In the psychologist's sample of size 50, 70% of the students were sitting on adjacent sides. This is just about what you would expect from chance variation alone. There is no evidence that students prefer sitting on adjacent sides. To establish that, the psychologist would need a larger sample size or a sample proportion quite a bit higher than 0.7.

E28. a. The null hypothesis is that the proportion of pairs of students who are seated on adjacent sides of a table is $\frac{1}{2}$.

b. The test should be one-sided. So, the alternate hypothesis is that the proportion of pairs is greater than $\frac{1}{2}$. (This indicates a preference for sitting on adjacent sides.)

c. The test statistic is

$$z = \frac{\hat{p} - p_0}{\sqrt{\frac{p_0 \cdot (1-p_0)}{n}}} = \frac{0.70 - \frac{1}{2}}{\sqrt{\frac{\left(\frac{1}{2}\right)\left(\frac{1}{2}\right)}{50}}} \approx 2.828\,.$$

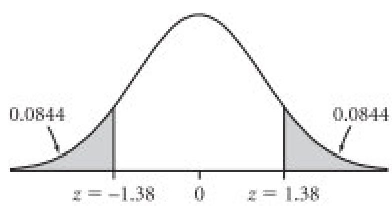
d. The p-value is $P(Z > 2.828) = 0.0023$.

e. There is sufficient evidence at the 5% level to suggest that the adjacent seating is a preference and not due to chance alone.

E29. a. H_0: $p = 0.69$, where p is the proportion of houses in your community that are occupied by their owners.
H_A: $p \neq 0.69$

b.

The z-value from the Minitab output is -1.38.



c. The *P*-value is 0.169. If the proportion of houses in your community that are occupied by their owners is indeed 69%, then the probability of getting 30 or fewer owner-occupied homes or 39 or more owner-occupied homes in a random sample of 50 homes is 16.9%.

d. Because this *P*-value is relatively high (17%), you do not have sufficient evidence to reject the null hypothesis. There is insufficient evidence that the proportion of houses occupied by their owners in your community is different from that of the country as a whole.

E30. a. H_0: $p = 0.50$, where p is the proportion of black youths who feel that black youths receive a poorer education than white youths.
H_A: $p > 0.50$

b. The test statistic, as given by the output, is 1.99.

c. The p-value, P(Z>1.99) = 0.024. This measures the strength of evidence to reject the null hypothesis in favor of the alternative.

d. There is sufficient evidence at the 5% level that black youths indeed feel that they receive a poorer education than white youths.

E31. a. $p_0 = 0.50$

b. The test should be one-sided. The null and alternative hypotheses are:
H_o: $p = 0.50$ H_a: $p < 0.50$, where p is the probability an astrologer can chose the correct CPI test result.

c. $\hat{p}$ = (40/116) = 0.345

d. The test statistic is

$$z = \frac{0.345 - 0.50}{\sqrt{\frac{(0.50)(0.50)}{116}}} = -3.339\,.$$

The p-value is P(Z < -3.339) = 0.0004.

e. There is strong evidence against the null hypothesis in favor of the claim that an astrologer chooses correct CPI test results less than 50% of the time.

E32. a. $p_0 = 0.50$

b. The test should be two-sided. The null and alternative hypotheses are:
H_o: $p = 0.50$ H_a: $p \neq 0.50$, where p is the probability an astrologer can chose the correct CPI test result.

c. $\hat{p}$ = (40/116) = 0.345

d. The test statistic is

$$z = \frac{0.34 - 0.50}{\sqrt{\frac{(0.50)(0.50)}{116}}} = -3.339\,.$$

The p-value is 2P(Z < -3.339) = 0.0008.

e. There is still strong evidence against the null hypothesis in favor of the claim that an astrologer chooses correct CPI test results a percentage of the time different from 50%.

E33. We test:
H_o: $p = 0.50$ H_a: $p < 0.50$, where p is the probability that a U.S. adult says he believes in ghosts.

The test statistic is

$$z = \frac{0.345 - 0.50}{\sqrt{\frac{(0.50)(0.50)}{1013}}} = -10.185\,.$$

The p-value is P(Z < -10.185), which is nearly zero. So, there is very strong evidence to reject the hypothesis that 50% (or more) believe in ghosts in favor of the alternative that the proportion is less.

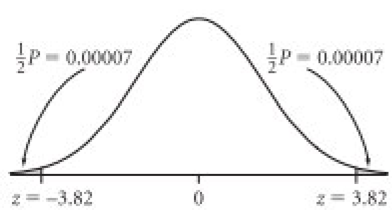
E34. Conduct a one-sided test of significance.

Check conditions. The question does not state that the USA TodayCNN/Gallup poll uses a random sample, however, the Gallup poll uses what you can consider a simple random sample of adults in the United States. Both $np_0 = 1002(0.55) = 551.1$ and $n(1 - p_0) = 1002(0.45) = 450.9$ are at least 10. The population of all adults in the United States is much larger than 10(1002) = 10,020. Thus, you can use a test for the significance of a proportion. This will be a one-sided test.

State your hypotheses.
H_0: $p = 0.55$, where p is the proportion of all adults residents of the United States who would say "yes" if asked whether a 16-year-old is too young to have a driver's license.
H_a: $p > 0.55$

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

$$z = \frac{\hat{p} - p_0}{\sqrt{\frac{p_0(1-p_0)}{n}}} = \frac{0.61 - 0.55}{\sqrt{\frac{0.55 \cdot 0.45}{1002}}} \approx 3.82$$



P-value. If the null hypothesis is true, the probability of getting a value of the test statistic that is larger than 3.82 is only 0.000067. Using the test on the calculator with $x = 611$, you get $z = 3.8037$ and a P-value of about 0.00007.

Write a conclusion in context. If you were to take a random sample from a population with 55% successes, there is almost no chance of getting a result as extreme as or more extreme than the one USA Today got in this sample. If the null hypothesis were true, USA Today would have a very unlikely result indeed. Thus, it is not plausible that only 55% of all adults would say that a 16-year-old is too young to get a driver's license and there is sufficient evidence to support the claim that the true proportion of adults who would say that a 16-year-old is too young to have a driver's license is more than 55%.

E35. a. $p_o = 0.50$

b. The test should be one-sided. The null and alternative hypotheses are:
H_o: $p = 0.50$ H_a: $p < 0.50$, where p is the probability that a spun penny lands heads up.

c. $\hat{p}=(10/40) = 0.25$

d. The test statistic is

$$z = \frac{0.25-0.50}{\sqrt{\frac{(0.50)(0.50)}{40}}} = -3.162 .$$

The p-value is P(Z < -3.162) = 0.0008.

e. There is strong evidence against the null hypothesis in favor of the claim that a spun penny lands heads-up less than 50% of the time.

E36. a. $p_o = 0.50$

b. The test should be one-sided. The null and alternative hypotheses are:
H_o: $p = 0.50$ H_a: $p > 0.50$, where p is the proportion who choose their own horoscope rather than for a different sign of the zodiac.

c. $\hat{p}=(23/40) = 0.575$

d. The test statistic is

$$z = \frac{0.575-0.50}{\sqrt{\frac{(0.50)(0.50)}{40}}} = 0.9487 .$$

The p-value is P(Z > 0.9487) = 0.1714.

e. There is insufficient evidence, even at the 10% level, to reject the null hypothesis.

E37. a. The null hypothesis is:
H_o: $p = 0.60$, where p is the proportion of all students on campus who carry a backpack to class; that is, 60% of the students on campus carry backpacks to class.

b. Note that $\hat{p}=(28/50) = 0.56$. So, the test statistic is

$$z = \frac{0.56-0.60}{\sqrt{\frac{(0.60)(0.40)}{50}}} = -0.577 .$$

So, the p-value is 2P(Z < -0.577) = 0.5639. This means that under the hypothesis that 60% of students carry backpacks, in a sample of 50 the chance of at most 28, or at least 32, carrying backpacks is 0.5639.

c. There is no evidence against the null hypothesis that 60% of the students carry backpacks.

E38. Answers will vary depending on the samples taken. The setup is identical to the one used in E37.
E39. B
E40. B is the best answer, although A is also true.
E41. We test: H_o: $p = 0.5$ H_a: $p > 0.5$, where p is the proportion of Americans who worry a great deal about the pollution of drinking water The test statistic is $$z = \frac{0.59 - 0.50}{\sqrt{\frac{(0.50)(0.50)}{1012}}} = 5.726.$$ The p-value is P(Z > 5.726), which is nearly zero. Hence, there is strong evidence to reject the null hypothesis in favor of the alternative that more than half of Americans worry a great deal about the pollution of drinking water.

E42. This will be a two-sided significance test for a proportion.

Check conditions. The problem does not state whether this was a simple random sample. However, Harris is a reputable polling organization and uses what can be considered a simple random sample. Both $np_0 = 1{,}217 \bullet (2/3) \approx 811.33$ and $n(1 - p_0) = 1{,}217 \bullet (1/3) \approx 405.67$ are at least ten, and there are more than 1,217(10) = 12,170 adults nationwide. The conditions are met for a significance test.

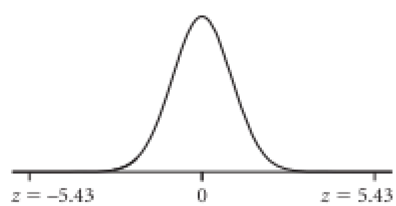
State hypotheses. H_0: The proportion p of all adults nationwide who agree with the statement, "Protecting the environment is so important that requirements and standards cannot be too high, and continuing environmental Improvements must be made regardless of cost." is 2/3.

H_a: $p \neq 2/3$

Calculate test statistic and draw a sketch.

The test statistic is

$$z = \frac{0.74 - (2/3)}{\sqrt{\frac{(2/3)\cdot(1/3)}{1217}}} \approx 5.427$$



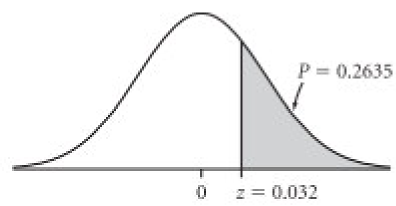
The P-value is $2(2.87 \bullet 10^{-8}) = 5.57 \bullet 10^{-8}$. Using the 1-Prop Z Test on the calculator with $x = 901$, z = 5.45 and $P = 4.98 \bullet 10^{-8}$. If you were to take a random sample from a population with two-thirds successes, there is almost no chance of getting a result as extreme as or more extreme than the one Gallup got in this sample. If the null hypothesis were true, Gallup would have a very unlikely result indeed.

Write a conclusion in context. Because the P-value is close to 0, you reject the null hypothesis that the proportion of all adults nationwide that would agree with the statement, "Protecting the environment is so important that requirements and standards

cannot be too high, and continuing environmental Improvements must be made regardless of cost." is 2/3. There is sufficient evidence to support the claim that this proportion differs from two-thirds.
E43. The confidence interval does not contain 0.5, so 0.5 is a not plausible value for the proportion of all Americans who worry a great deal about the pollution of drinking water. As such, we conclude that $p > 0.5$.
E44. The confidence interval does not contain $\frac{2}{3}$, so $\frac{2}{3}$ is not a plausible value for the proportion of the nation's adults who agree with the statement. We conclude that $p \neq \frac{2}{3}$.
E45. a. The null and alternative hypotheses are: H_0: The proportion p of all adults in the United States who would say they are satisfied with the quality of K–12 education in the country is 0.5. H_a: $p < 0.5$. The question asks whether the poll results imply that less than half of adults are satisfied with the quality of education. This implies a one-sided alternative hypothesis. **b.** The conditions are met for doing a test of significance for a proportion because you have a random sample, both $np_0 = 1000(0.5) = 500$ and $n(1 - p_0) = 1000(1 - 0.5) = 500$ are at least 10, and the number of adults in the United States is at least 10(1000) = 10,000. **c.** The test statistic is: $z = \frac{\hat{p} - p_0}{\sqrt{\frac{p_0(1-p_0)}{n}}} = \frac{0.49 - 0.5}{\sqrt{\frac{0.5 \cdot 0.5}{1000}}} \approx -0.632$. The p-value is about 0.2635. **d.** There is insufficient evidence to reject the null hypothesis that at least half of the adult residents are satisfied with the quality of K-12 education.
E46. a. The null and alternative hypotheses are: H_0: The proportion p of all adults in the United States who would say they are dissatisfied with the quality of K-12 education in the country is 0.5. H_a: $p > 0.5$. The question asks whether the poll results imply that more than half of adults are dissatisfied with the quality of education. This implies a one-sided alternative hypothesis. **b.** The conditions are met for doing a test of significance for a proportion because you have a random sample, both $np_0 = 1000(0.5) = 500$ and $n(1 - p_0) = 1000(1 - 0.5) = 500$ are at least 10, and the number of adults in the United States is at least 10(1000) = 10,000.

c. The test statistic is:

$$z = \frac{\hat{p} - p_0}{\sqrt{\frac{p_0(1-p_0)}{n}}} = \frac{0.51 - 0.5}{\sqrt{\frac{0.5 \cdot 0.5}{1000}}} \approx 0.632$$



The *P*-value is about 0.2635.

d. No significance level was given, so you may use α = 0.05. Since the *P*-value of 0.2635 is more than α = 0.05 you would not reject the null hypotheses. If the proportion of adults in the US who would say that they are dissatisfied with the quality of K-12 education is 0.5 it is reasonably likely to get a random sample of 1000 adults in which 51% say they are dissatisfied with the quality of education. In fact, if the true proportion is indeed 50%, then there is 26.35% chance of getting a result as extreme as or more extreme than the one in this sample. So, there is insufficient evidence to support the claim that the true proportion of adults who would say they are dissatisfied with the quality of K-12 education today in the U.S. is more than 50%.

E47. We test:
H_o: $p = 0.5$ versus H_a: $p < 0.5$, where p is the percentage of all Americans approving.

The test statistic is

$$z = \frac{0.45 - 0.50}{\sqrt{\frac{(0.50)(0.50)}{1004}}} = -3.169\,.$$

The p-value is P(Z < -3.169) = 0.001. Yes, the statement is fair. So, it's not plausible that half (or more) of all Americans give him positive marks.

E48. We test:
H_o: $p = 0.5$ versus H_a: $p > 0.5$, where p is the proportion of adults who believe friends are important for success.

The test statistic is

$$z = \frac{0.53 - 0.50}{\sqrt{\frac{(0.50)(0.50)}{1027}}} = 1.923\,.$$

The p-value is P(Z > 1.923) = 0.0271. So, at the 5% level, there is sufficient evidence to support the claim that at least 50% of adults believe that friends are important for success. So, yes, go ahead and use the headline.

E49. Neither C nor D is true. For C, this is an incorrect interpretation of a confidence interval and D is an incorrect interpretation of margin of error.

E50. D

E51. a. The test is one-sided.
b. The expected proportion of deaths is $p = (11.3/196) = 0.058$, so the null and alternative hypotheses are:

H_o: $p = 0.058$ and H_a: $p > 0.058$

The test statistic is

$$z = \frac{\frac{39}{196} - 0.058}{\sqrt{\frac{(0.058)(0.942)}{196}}} = 8.44 .$$

The p-value is near zero.

c. No, the result cannot reasonably be attributed to chance since the p-value is so small.

E52. a. Answers will vary here and should involve some exploratory analysis first. Students might make back-to-back stem-and-leaf plots of the heights of the winners and the heights of the losers. However, because the data come paired (by election), a complete exploratory analysis should take that into account. For example,you might examine a plot of the differences of winner's height minus loser's height. You also might make a scatterplot of (winner's height, loser's height) or make a plot of the differences over time.

b. When conducting a test of significance, the null hypothesis is that height doesn't matter. Specifically, the probability that the taller person will be elected is 0.50. You could test the proportion of pairs where the taller person was elected against this standard. (You first should decide how to deal with the case where the two heights are the same.)

c. In the 22 elections where the heights are known and not equal, the taller candidate won 15 times. This results in the test statistic

$$z = \frac{\frac{15}{22} - 0.5}{\sqrt{\frac{(0.5)(0.5)}{22}}} = 1.706 .$$

The p-value is $P(Z > 1.706) = 0.0446$. So, there is sufficient evidence that the results are not attributed to chance alone.

A Note about E52 and Other Situations Where Your "Sample" Is the Entire Population: Purists would say that we should not use a test of significance to determine whether the taller person is more likely to be elected president. They have two reasons. The first is that the numbers given are not a random sample from any population. We have all of the information for presidential elections from 1900 through 2000. And in fact, the taller candidate *did* win more than half of the elections.

The second reason is that this is a classic example of "data snooping." We probably could think of a hundred characteristics of the two candidates that we could have considered: older, heavier, more experienced, sharper dresser, and so on. The winner and the loser will not match exactly on all of these. By definition, if we check 100 characteristics, we expect that 5 will result in a rare event. Height simply may be one of those 5 characteristics and that's why people have noticed it.

However, you might use this example with your students as an example of how a test of significance can be used profitably even though we have the entire population. The question becomes, Can the fact that the taller man has won more often reasonably be attributed to chance? If the answer is yes, we have no reason to continue this investigation. If the answer is no, 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, that explanation might be that being taller helps and future voters will also tend to favor the taller candidate.

So in situations like these, the mechanics of a significance test simply let us know whether or not there is something that needs to be investigated. The data first must pass the test that the results can't reasonably be attributed to chance before we take any further steps. Note that these comments apply to the Westvaco case as well, where we had information about the entire population of employees.

E53. a. You would expect 0.05(240) = 12 of these people to make a Type-I error.
b. You cannot tell from the given information.

E54. a. You would expect that about 0.10(50) or 5 laboratories would reject the null hypothesis if it is true.

b. You cannot answer this question because the power of the test (the probability of rejecting the null hypothesis) depends on the sample size and on the difference between p_0 and p when the null hypothesis is false.

E55. a. Yes, it's possible for both to make a Type I error if they reject this true null hypothesis. They are equally likely to do so because they have the same value of α.

b. No, it isn't possible for them to fail to reject a false null hypothesis because the null hypothesis isn't false.

E56. a. Yes, it's possible for both to make a Type I error if they reject this true null hypothesis. Taline is more likely to do so because she has a larger value of α.

b. No, it isn't possible for either of them to fail to reject a false null hypothesis because the null hypothesis isn't false.

E57. a. No, it isn't possible for them to reject a true null hypothesis because the null hypothesis isn't true.

b. Yes, it's possible for both to make the Type II error of failing to reject this false null hypothesis. Jeffrey is more likely to do so because he has the smaller sample size.

E58. a. No, it isn't possible for either of them to reject a true null hypothesis because the null hypothesis isn't true.

b. Yes, it's possible for both of them to make the Type II error of failing to reject this false null hypothesis. Jeffrey is more likely to make a Type II error because he has the smaller value of α. Jeffrey thus needs a larger difference between $\hat{p}$ and p_0 before he can reject the null hypothesis.

E59. a. $\hat{p} = 0.668$, $p_0 = 0.5$.

b. The test statistic is

$$z = \frac{\hat{p} - p_0}{\sqrt{\frac{p_0(1-p_0)}{n}}} = \frac{0.668 - 0.5}{\sqrt{\frac{0.5 \cdot 0.5}{600}}} \approx 8.23.$$

The difference is due to rounding error. To see this note that the reported $\hat{p}$ is 66.8% and the sample size is 600. 66.8% of 600 is 400.8 which means the number of correct guesses in the sample was 401. $\frac{401}{600} = 0.668\overline{3}$, and the test statistic calculated using this value for $\hat{p}$ gives a result of 8.2466, rounding to 8.25.

c. Yes. A *P*-value of 0.001 corresponds to a *z*-score of around 3. A *z*-score of 8 is *much* more extreme than this, in fact it is $8.01 \bullet 10^{-17}$.

d. Not really. While the sample size always affects power, a smaller sample would have been sufficient in this case because 66.8% is so far from 50%. Even if the sample size had been only 60, with 40 successes, the result would have been statistically significant with a *P*-value of 0.01.

e. Since you are rejecting the null hypothesis, you could be making a Type-I error. The error may be that the null hypothesis is actually true. Specifically, the consequence would be believing that the winning candidate is more likely to be perceived as competent when this is not the case among the population from which the subjects were taken.

E60. a. The test is one-sided. The null and alternative hypotheses are:
H_o: $p = 0.5$ versus H_a: $p > 0.5$, where p is the proportion of those candidates who were perceived as more competent won the election.

b. $\hat{p} = \frac{22}{32} = 0.6875$ and $p_0 = 0.50$.

c. The test statistic is

$$z = \frac{0.6875 - 0.50}{\sqrt{\frac{(0.50)(0.50)}{32}}} = 2.121.$$

The p-value is $P(Z > 2.121) = 0.017$.

d. Yes – see (c).

e. Yes, the sample size plays an important role in the power of any test.

f. Since he is rejecting the null hypothesis, he might be making a Type-I error. The consequence is believing candidates who are perceived as more competent more often win the election when this is not the case.

E61. a. The test statistic is

$$z = \frac{\frac{23}{45} - 0.50}{\sqrt{\frac{(0.50)(0.50)}{45}}} = 0.149 .$$

The p-value for the one-sided test is P(Z>0.149) = 0.4404. So, we do not reject the null hypothesis (in a big way!).

b. We could be making a Type-II error. The consequence is believing that the judge cannot select the correct dog at better than chance level when in fact he or she can.

c. Use a larger sample size.

E62. a. The test statistic for this test is

$$z = \frac{\frac{42}{80} - 0.55}{\sqrt{\frac{0.55(1-0.55)}{80}}} \approx -0.449$$

This gives a *P*-value for this test of about 0.65. We can not reject the hypothesis that 55% of students prefer hamburgers.

b. The test statistic is $z = -1.35$. The *P*-value for this test is about 0.18. We can not reject the hypothesis that 55% of students prefer hot dogs.

c. The results of parts a and b may seem paradoxical to students. It looks like we are accepting the hypothesis that a majority prefer hamburgers and the hypothesis that a majority prefer hot dogs. This is precisely why we don't "accept" the null hypothesis. In this case, we don't have enough evidence to reject it, but we know that we have made a Type II error in at least one of these tests.

d. The test statistic is $z = -1.42$. Then the *P*-value for this test is about 0.155. We can not reject the hypothesis that 55% of the students prefer hamburgers.

e. The test statistic is $z = -4.264$. Then, the *P*-value for this test is 0.00002. We can reject the hypothesis that 55% of the students prefer hot dogs.

f. The results of parts d and e will look a lot better to students. The larger sample size makes it clear that more than half the students prefer hamburgers. Note that having students compare the four confidence intervals that can be constructed from this exercise would be a good exploration. Thanks to Dan Johnson, of Silver Creek High School in San Jose, California, for this idea.

E63. a. Since 82% of studies had a Type-II error, they failed to reject the null hypothesis.

b. C

E64. a. The size of each group is too small or the variability is too large.

b. We fail to reject the null hypothesis since the p-value is greater than 0.05.

c. C

E65. a.

***x*, the number of heads**	***p*(*x*)**
0	$\binom{20}{0} \cdot 0.5^{20} \approx 9.54 \cdot 10^{-7}$
1	$\binom{20}{1} \cdot 0.5^{20} \approx 1.91 \cdot 10^{-5}$
2	$\binom{20}{2} \cdot 0.5^{20} \approx 1.81 \cdot 10^{-4}$
3	$\binom{20}{3} \cdot 0.5^{20} \approx 0.00109$
4	$\binom{20}{4} \cdot 0.5^{20} \approx 0.00462$
5	$\binom{20}{5} \cdot 0.5^{20} \approx 0.0148$
6	$\binom{20}{6} \cdot 0.5^{20} \approx 0.0370$
7	$\binom{20}{7} \cdot 0.5^{20} \approx 0.0739$
8	$\binom{20}{8} \cdot 0.5^{20} \approx 0.1201$
9	$\binom{20}{9} \cdot 0.5^{20} \approx 0.1602$
10	$\binom{20}{10} \cdot 0.5^{20} \approx 0.1762$
11	$\binom{20}{11} \cdot 0.5^{20} \approx 0.1602$

12	$\binom{20}{12}\cdot 0.5^{20} \approx 0.1201$
13	$\binom{20}{13}\cdot 0.5^{20} \approx 0.0739$
14	$\binom{20}{14}\cdot 0.5^{20} \approx 0.0370$
15	$\binom{20}{15}\cdot 0.5^{20} \approx 0.0148$
16	$\binom{20}{16}\cdot 0.5^{20} \approx 0.00462$
17	$\binom{20}{17}\cdot 0.5^{20} \approx 0.00109$
18	$\binom{20}{18}\cdot 0.5^{20} \approx 1.81\cdot 10^{-4}$
19	$\binom{20}{19}\cdot 0.5^{20} \approx 1.91\cdot 10^{-5}$
20	$\binom{20}{20}\cdot 0.5^{20} \approx 9.54\cdot 10^{-7}$

b. The outer 95% can be found by adding the probabilities, starting with the probability of 0 successes and working your way up, until you get a sum of 0.025. Since the distribution is symmetric, take the same number of outcomes starting from 20 successes and working your way down. The closest you can come to the outer 5% is to add up the probabilities of 0 through 5 successes and 15 to 20 successes. This sum is approximately 0.04. You would reject the null hypothesis if you have 5 or fewer heads or 15 or more heads.

c. Here we need the part of the probability distribution for 40% successes that represents 0 to 5 outcomes and 15 to 20 outcomes.

***X*, the number of successes**	***p*(*x*)**
0	$\binom{20}{0}0.4^{0}\cdot 0.6^{20} \approx 3.656\cdot 10^{-5}$
1	$\binom{20}{1}0.4^{1}\cdot 0.6^{19} \approx 4.875\cdot 10^{-4}$
2	$\binom{20}{2}0.4^{2}\cdot 0.6^{18} \approx 3.087\cdot 10^{-3}$

3	$\binom{20}{3}0.4^3 \cdot 0.6^{17} \approx 0.0123$
4	$\binom{20}{4}0.4^4 \cdot 0.6^{16} \approx 0.0350$
5	$\binom{20}{5}0.4^5 \cdot 0.6^{15} \approx 0.0746$
15	$\binom{20}{15}0.4^{15} \cdot 0.6^5 \approx 0.00129$
16	$\binom{20}{16}0.4^{16} \cdot 0.6^4 \approx 2.697 \cdot 10^{-4}$
17	$\binom{20}{17}0.4^{17} \cdot 0.6^3 \approx 4.230 \cdot 10^{-5}$
18	$\binom{20}{18}0.4^{18} \cdot 0.6^2 \approx 4.700 \cdot 10^{-6}$
19	$\binom{20}{19}0.4^{19} \cdot 0.6^1 \approx 3.299 \cdot 10^{-7}$
20	$\binom{20}{20}0.4^{20} \cdot 0.6^0 \approx 1.0995 \cdot 10^{-8}$
Total	0.1271 or 12.7%

There is a 12.7% chance that the null hypothesis will be rejected if the true proportion of getting heads when spinning a penny is 40%.

d. To increase the power of the test your friend should increase the sample size.

E66. a.

***x*, the number of invoices not paid on time**	***P*(*x*)**
0	$\binom{10}{0}0.3^0 \cdot 0.7^{10} \approx 0.0282$
1	$\binom{10}{1}0.3^1 \cdot 0.7^9 \approx 0.1211$
2	$\binom{10}{2}0.3^2 \cdot 0.7^8 \approx 0.2335$
3	$\binom{10}{3}0.3^3 \cdot 0.7^7 \approx 0.2668$
4	$\binom{10}{4}0.3^4 \cdot 0.7^6 \approx 0.2001$

5	$\binom{10}{5} 0.3^5 \cdot 0.7^5 \approx 0.1029$
6	$\binom{10}{6} 0.3^6 \cdot 0.7^4 \approx 0.0368$
7	$\binom{10}{7} 0.3^7 \cdot 0.7^3 \approx 0.0090$
8	$\binom{10}{8} 0.3^8 \cdot 0.7^2 \approx 0.0014$
9	$\binom{10}{9} 0.3^9 \cdot 0.7^1 \approx 0.000138$
10	$\binom{10}{10} 0.3^{10} \cdot 0.7^0 \approx 5.905 \cdot 10^{-6}$

The probability of 6 or more heads is 0.047, so this corresponds to a 4.7% significance level.

b. You need the probabilities of 6 or more heads with a population proportion of 40% late invoices. The relevant part of the distribution is shown below.

***x*, the number of invoices not paid on time**	***p*(*x*)**
6	0.1115
7	0.0425
8	0.0106
9	0.0016
10	0.000105
Total	0.1663

There is 0.166, or about a 16.6% chance of rejecting this null hypothesis. So the power of the test is 0.166.

E67. a. *Blinded* means that a TT practitioner could not tell whether the investigator's hand was placed above their left hand or above their right hand. The way it was done was to have the TT practitioner rest their hands, palms up, on a flat surface. A tall screen with cutouts on its base was placed over the TT practitioner's arms so that they couldn't see the investigator's hand on the other side of the screen. Double-blinding would mean that the person who placed his or her hand above the TT practitioner's hands would not hear what the person's response was. Although this might have made the experiment a bit better, there was no judgment on the part of the investigator in evaluating the response from the TT practitioner. (The response was either right or wrong.)

b. H_0: The TT practitioners did no better than chance in identifying the correct hand. That is, $p = 0.5$, where p is the proportion of times the TT practitioners identified the correct hand.

H_a: The proportion of hands correctly identified was greater than 0.5.

c. No. The alternative hypothesis was that the TT practitioners should be able to identify the correct hand more often than random guessing. Because they actually performed more poorly than random guessing, the *P*-value will be more than 0.5 and the null hypothesis would be rejected.

d. This means that the sample size was large enough so that if TT practitioners could identify the correct hand with any consistency, the null hypothesis would have been rejected. Specifically, to reject the null hypothesis that the probability that they select the correct hand is 0.5, the value of z for a one-sided test would have to be 1.645. With a sample size of 280, the practitioners would only have to get 54.9% correct:

$$z = \frac{0.549 - 0.5}{\sqrt{\frac{0.5(1-0.5)}{280}}} \approx 1.645$$

E68. a. ***Check conditions.*** The problem states that the teenagers were randomly selected. Both $n_1\hat{p}_1 = 971 \bullet 0.78 \approx 757$ and $n_1(1 - \hat{p}_1) = 971 \bullet 0.22 \approx 214$ are at least 10. There are more than 971 • 10 = 9710 teens with internet access. So the conditions are met for constructing a confidence interval.

Calculate the confidence interval.

$$\hat{p} \pm z^* \sqrt{\frac{\hat{p}(1-\hat{p})}{n}} = 0.78 \pm 1.96\sqrt{\frac{0.78 \cdot 0.22}{971}} \approx 0.78 \pm 0.026$$

or (0.754, 0.806).

b. You are 95% confident that if you were able to ask *all* teenagers with Internet access, between 75.4% and 80.6% would agree that most teens are not careful enough about the information they give out about themselves online.

c. Suppose you could repeat this survey with 100 different random samples of teens with internet access, each of size 971. Then you would expect the (unknown) proportion of all teens with Internet access who say they agree that most teens are not careful enough about the information they give out about themselves online to be in about 95 of the resulting confidence intervals. (But each of the 100 surveys probably would have a different interval.)

E69. a. (0.05)(200) = 10.
b. $(0.95)^{200} = 0.000035$

E70. The method used for constructing 90% confidence intervals captures the true population percentage 90% of the time. Once again, the reasoning is rather subtle. See the explanation in the text.

E71. C

E72. There is one glaring error in the interpretation of a confidence interval: "To be more specific, the laws of probability say that if we were to conduct the same survey 100 times, asking people in each survey to rate the job Bill Clinton is doing as president, in 95 out of those 100 polls, we would find his rating to be between 47% and 53%."

This should say: "To be more specific, the laws of probability say that if we were to conduct the same survey 100 times, asking people in each survey to rate the job Bill Clinton is doing as president, *we expect that the proportion of the entire population that approves* would fall within 95 of the 100 confidence intervals."

There are many other sentences that students may be able to write more clearly or that students may not agree with. For example, "If Gallup were to—quite expensively—use a sample of 4,000 randomly selected adults each time it did its poll, the increase in accuracy over and beyond a well-done sample of 1,000 would be minimal and, generally speaking, would not justify the increase in cost."
Students may feel that cutting the margin of error in half from about 3% to about 1.5% might justify the increase in cost in a close election.

E73. You should include some explanation of a margin of error and how it is relatively small (around 3%) even with a sample size as small as 1000. The part that people tend to have the hardest time understanding is that, for a fixed $\hat{p}$, the margin of error depends almost entirely on the sample size *n* and not on how large the population is. That is, a random sample of size 1000 from the residents of Seattle has about the same margin of error as a random sample of size 1000 from the residents of the United States. If the sample size *n* is large relative to the size *N* of the population (more than about 10% of the population size), then you should use a correction factor for the formula for the margin of error that makes it smaller. Specifically, an approximate confidence interval is

$$\hat{p} = z^* \cdot \sqrt{\frac{\hat{p}(1-\hat{p})}{n-1}}\sqrt{\frac{N-n}{N}}$$

where *N* is the size of the population.

See also the explanation from the Gallup Organization in E72.

E74. The farther apart the hypothesized proportion and the true proportion, the greater the power. Also, a one-sided test makes it easier to reject the null hypothesis if you select the correct direction. A higher level of significance makes a larger rejection region, also increasing the power of the test. (Note that this also increases the chance of a Type I error.) A larger sample size decreases the standard error of the sampling distribution, making it easier to reject a false null hypothesis, thus increasing power.

E75. **a.** Using $n = 54,461$, $\hat{p} = 0.76$, $z^* = 1.96$, we see that the 95% confidence level is

$$0.76 \pm 1.96\sqrt{\frac{(0.76)(0.24)}{54,461}}$$

or about 0.756 to 0.764.

b. The margin of error is $1.96\sqrt{\frac{(0.76)(0.24)}{54{,}461}} \approx 0.004$. **c.** Observe that $1.96\sqrt{\frac{(0.695)(0.305)}{1000}} \approx 0.028$.
E76. a. To test whether it is plausible that $p \le p_0$, you would use the value of p within that range that would be the most plausible, or closest to $\hat{p}$ which is p_0. The test statistic would be the same with either null hypothesis. **b.** No, not if α had remained the same. A one-sided test puts the entire rejection region on one side, so a smaller test statistic is adequate to reject the null hypothesis.
E77. a. False. There should be a theoretical basis for the formulation of hypotheses. Using the data to form them would be "stacking the deck." **b.** False. A two-tailed test p-value is 2 times the p-value of a corresponding one-sided test. **c.** False. The p-value is the likelihood of rejecting a true null hypothesis. **d.** True.

Concept Review Solutions

C1. E. When the sample is selected at random, the p-value is the probability of getting a sample proportion as extreme as or more extreme than that actually observed, given that the null hypothesis is true. In this case, Ms. Chang can only ask whether the proportion of her students who get the question correct is consistent with what would be expected from a random sample of students nationwide.
C2. B. The margin of error is inversely proportional to the square root of n, so if four times as many people were surveyed, the margin of error would be half as large.
C3. C. The 95% confidence interval is (0.45579, 0.59421). This means that it is plausible that the percentage of all alumni who favor abolishing the school dress code is 50% or less.
C4. B. By definition of Type I error.
C5. C. Choice E is close, but technically incorrect.
C6. E
C7. a. 45 **b.** Forty intervals, or 80%, actually captured the population proportion of 0.20.

c. Most of the intervals are too close to 0 and shorter than average. A sample proportion that underestimates the value of p will also underestimate the standard error, and the resulting interval will be too close to 0 and too short.

d. You are using a continuous distribution to approximate a discrete one. You constructed the horizontal lines earlier in the chapter using the fact that 95% of all values in a normal distribution lie within 1.96 standard deviations of the mean; that is, the middle 95% of the binomial distribution was used as the middle 95% of the binomial distribution centered in the same spot. For most binomial distributions, it is not possible to get an exact 95% middle. Further, all binomial distributions, except those where p = 0.5, are skewed. The normal approximation is symmetric. Finally, the sample proportion was used to estimate the population proportion in the formula for the standard error. Thus, two quantities are being estimated (center and spread) by the same sample proportion. All of these approximations can cause the capture rate to differ from its nominal value.

C8. a. The standard 90% confidence interval is

$$\hat{p} \pm 1.645\sqrt{\frac{\hat{p}(1-\hat{p})}{n}} = 0.1 \pm 1.645\sqrt{\frac{(0.1)(0.9)}{50}}$$

or about 0.030 to 0.170. The Plus 4 confidence interval using $\hat{p} = \frac{x+2}{n+4} = 0.130$ is

$$\hat{p} \pm 1.645\sqrt{\frac{\hat{p}(1-\hat{p})}{n+4}} = 0.130 \pm 1.645\sqrt{\frac{(0.130)(0.870)}{54}}$$

or about 0.055 to 0.205.

The center changed from 0.10 to 0.13, which is a move toward 0.5. The length of the interval increased from about 0.14 to 0.15.

b. The standard interval is $0.4 \pm 1.645\sqrt{\frac{(0.4)(0.6)}{50}}$, or about 0.286 to 0.514. The Plus 4 interval is $0.407 \pm 1.645\sqrt{\frac{(0.407)(0.593)}{54}}$, or about 0.297 to 0.452.

The center moved toward 0.5 (from 0.4 to 0.407), but the length of the interval decreased from 0.217 to 0.155. Both changes were smaller for x = 20 than for x = 5. The impact was greater on the interval based on the smaller sample proportion.

c. Forty-five of the 50 intervals, or 90%, captured the population proportion.

d. The intervals that do not capture the true value of p are more balanced on either side of p, the intervals themselves are of more uniform length, and the capture rate is about what you would expect for 90% confidence intervals. Specifically, when $\hat{p}$ is close to zero (or close to 1), the Plus 4 interval tends to be longer than the standard interval. When $\hat{p}$ is close to 0.5, the Plus 4 interval tends to be about the same length as the standard interval.