

Section 8.2

Check Your Understanding, page 496:

1. The Random condition is not met because this was a convenience sample. They asked the first 100 students who arrive that day and these students are likely to be different than the rest of the population with regard to sleep time. The 10% condition is met because the sample of 100 is less than 10% of the population at a large high school. The Large Counts condition is met because they have 17 successes (those who had slept 8 hours the previous night) and 83 failures (those who hadn't). Both of these are at least 10.
2. The Random condition is met because the inspector chose an SRS of bags. The 10% condition is met because the sample of 25 is less than 10% of the thousands of bags filled in an hour. The Large Counts condition is not met because there were only 3 successes (bags with too much salt), which is less than 10. Note that there were 22 failures (bags with an appropriate amount of salt), which is at least 10, but both values must be at least 10 for this condition to be met.

Check Your Understanding, page 499:

1. The parameter of interest is p = the true proportion of all U.S. college students who are classified as frequent binge drinkers.
2. The Random condition is met because the statement says that the students were chosen randomly. The 10% condition is met because the sample of 10,904 is less than 10% of all U.S. college students. The Large Counts condition is met because there were 2486 successes (binge drinkers) and 8418 failures (non-binge drinkers), both of which are at least 10. All conditions are met.

3. Because $\frac{1-0.99}{2} = 0.005$, z^* for a 99% confidence interval can be found by looking for a left-tail area of 0.005. The closest area is 0.0051 (or 0.0049), corresponding to a critical value of $z^* = 2.57$ (or 2.58). *Using technology:* $\text{invNorm}(\text{area}: 0.005, \mu: 0, \sigma: 1) = -2.576$. Thus, the critical value is $z^* = 2.576$. For this sample, $\hat{p} = \frac{2486}{10904} = 0.228$, so the confidence interval is

$$0.228 \pm 2.576 \sqrt{\frac{0.228(1-0.228)}{10,904}} = 0.228 \pm 0.010 = (0.218, 0.238).$$

4. We are 99% confident that the interval from 0.218 to 0.238 captures the true proportion of all U.S. college students who are classified as frequent binge drinkers.

Check Your Understanding, page 503:

1. Solving $1.96 \sqrt{\frac{0.80(0.20)}{n}} \leq 0.03$ for n gives $n \geq (0.80)(0.20) \left(\frac{1.96}{0.03} \right)^2 = 682.95$. This means that we should select a sample of at least 683 customers.
2. If the company president demands 99% confidence instead, the required sample size will be larger because the critical value is larger for 99% confidence (2.576) versus 95% confidence (1.96). The company would need to select at least 1180 customers to have 99% confidence.

Exercises, page 504:

- 8.27 The Random condition is met because Latoya selected an SRS of students. The 10% condition is not met because the sample size (50) is more than 10% of the population of seniors in the dormitory (175). The Large Counts condition is met because $n\hat{p} = 14 \geq 10$ and $n(1 - \hat{p}) = 36 \geq 10$.

8.28 The Random condition was met because the sample was an SRS. The 10% condition is met because the sample size (50) is less than 10% of the population of students at his college (2400). The Large Counts condition is met because $n\hat{p} = 38 \geq 10$ and $n(1 - \hat{p}) = 12 \geq 10$.

8.29 The Random condition may not be met because we do not know if the people contacted were a random sample. The 10% condition is met because the sample size (2673) is less than 10% of the population of adult heterosexuals. The Large Counts condition is not met because $n\hat{p} = 2673(0.002) \approx 5$ is not at least 10.

8.30 The Random condition may not be met because we do not know if the whelk eggs were a random sample. The 10% condition is met because the sample size (98) is less than 10% of all whelk eggs. The Large Counts condition is not met because $n\hat{p} = 9$, which is not at least 10.

8.31 Because $\frac{1 - 0.98}{2} = 0.01$, z^* for a 98% confidence interval can be found by looking for a left-tail area of 0.01. The closest area is 0.0099, corresponding to a critical value of $z^* = 2.33$. *Using technology:* $\text{invNorm}(\text{area: } 0.01, \mu : 0, \sigma : 1) = -2.326$. Thus, the critical value is $z^* = 2.326$.

8.32 Because $\frac{1 - 0.93}{2} = 0.035$, z^* for a 93% confidence interval can be found by looking for a left-tail area of 0.035. The closest area is 0.0351, corresponding to a critical value of $z^* = 1.81$. *Using technology:* $\text{invNorm}(\text{area: } 0.035, \mu : 0, \sigma : 1) = -1.812$. Thus, the critical value is $z^* = 1.812$.

8.33 (a) The population consists of the seniors at Tonya's high school. The parameter of interest is the true proportion of all seniors who plan to attend the prom.

(b) Random: the sample is a simple random sample. 10%: The sample size (50) is less than 10% of the population size (750). Large Counts: $n\hat{p} = 36 \geq 10$ and $n(1 - \hat{p}) = 14 \geq 10$. All conditions are met.

(c) For a 90% confidence interval $z^* = 1.645$. For this sample $\hat{p} = \frac{36}{50} = 0.72$. So the confidence

interval is $0.72 \pm 1.645 \sqrt{\frac{0.72(0.28)}{50}} = 0.72 \pm 0.10 = (0.62, 0.82)$.

(d) We are 90% confident that the interval from 0.62 to 0.82 captures the true proportion of all seniors at Tonya's high school who plan to attend the prom.

8.34 (a) The population consists of the undergraduates at a large university. The parameter of interest is the true proportion of all undergraduates who would be willing to report cheating.

(b) Random: the sample was a simple random sample. 10%: the sample size (172) is less than 10% of the population of all undergraduates at a large university. Large Counts: $n\hat{p} = 19 \geq 10$ and $n(1 - \hat{p}) = 153 \geq 10$. All conditions are met.

(c) For a 99% confidence interval $z^* = 2.576$. For this sample $\hat{p} = \frac{19}{172} = 0.11$. So the

confidence interval is $0.11 \pm 2.576 \sqrt{\frac{0.11(0.89)}{172}} = 0.11 \pm 0.06 = (0.05, 0.17)$.

(d) We are 99% confident that the interval from 0.05 to 0.17 captures the true proportion of all undergraduates at this university who would be willing to report cheating.

8.35 (a) *State:* We want to estimate p = the true proportion of all full-time U.S. college students who are binge drinkers at a 99% confidence level. *Plan:* We should use a one-sample z interval for p if the conditions are met. Random: the students were selected randomly. 10%: the sample size (5914) is less than 10% of the population of all college students. Large Counts: $n\hat{p} = 2312$

≥ 10 and $n(1 - \hat{p}) = 3602 \geq 10$. *Do:* $\hat{p} = \frac{2312}{5914} = 0.391$, so the confidence interval is

$0.391 \pm 2.576 \sqrt{\frac{0.391(1 - 0.391)}{5914}} = 0.391 \pm 0.016 = (0.375, 0.407)$. *Conclude:* We are 99%

confident that the interval from 0.375 to 0.407 captures the true proportion of full-time U.S. college students who are binge drinkers.

(b) Because the value 0.45 does not appear in our 99% confidence interval, it isn't plausible that 45% of full-time U.S. college students are binge drinkers.

8.36 (a) *State:* We want to estimate p = the population proportion of all teens who would report texting with their friends every day at a 95% confidence level. *Plan:* We should use a one-sample z interval for p if the conditions are met. Random: the teens were selected randomly. 10%: the sample size (799) is less than 10% of the population of all teens. Large Counts: $n\hat{p} =$

$392 \geq 10$ and $n(1 - \hat{p}) = 407 \geq 10$. *Do:* $\hat{p} = \frac{392}{799} = 0.491$, so the confidence interval is

$0.491 \pm 1.960 \sqrt{\frac{0.491(1 - 0.491)}{799}} = 0.491 \pm 0.035 = (0.456, 0.526)$. *Conclude:* We are 95%

confident that the interval from 0.456 to 0.526 captures the true proportion of teens who would report texting with their friends every day.

(b) Because the value 0.45 does not appear in our 95% confidence interval, it isn't plausible that 45% of American teens who would report texting with their friends every day.

8.37 Answers will vary. The margin of error does not include whether or not the students told the truth in the survey. When questions about personal behavior involving alcohol or drugs are asked, there is always a concern about the answers being truthful.

8.38 Answers will vary. The margin of error does not include whether or not the students told the truth in the survey. For example, a teen without a phone may report texting every day to avoid having to admit that he or she doesn't have a phone.

8.39 *State:* We want to estimate p = the true proportion of all students retaking the SAT who receive coaching at a 99% confidence level. *Plan:* We should use a one-sample z interval for p if the conditions are met. *Random:* the students were selected randomly. 10%: the sample size (3160) is less than 10% of the population of all students taking the SAT twice. *Large Counts:* $n\hat{p} = 427 \geq 10$ and $n(1 - \hat{p}) = 2733 \geq 10$. *Do:* $\hat{p} = \frac{427}{427 + 2733} = \frac{427}{3160} = 0.135$, so the confidence interval is $0.135 \pm 2.576 \sqrt{\frac{0.135(0.865)}{3160}} = 0.135 \pm 0.016 = (0.119, 0.151)$. *Conclude:* We are 99% confident that the interval from 0.119 to 0.151 captures the true proportion of students retaking the SAT who receive coaching.

8.40 *State:* We want to estimate p = the true proportion of all U.S. adults who were satisfied with the way things were going in the United States in January 2010 at a 90% confidence level. *Plan:* We should use a one-sample z -interval for p if the conditions are satisfied. *Random:* the adults were selected randomly. 10%: the sample size (1025) is less than 10% of the population of U.S. adults. *Large Counts:* $n\hat{p} = 256 \geq 10$ and $n(1 - \hat{p}) = 769 \geq 10$. *Do:* $\hat{p} = \frac{256}{256 + 769} = \frac{256}{1025} = 0.25$, so the confidence interval is $0.25 \pm 1.645 \sqrt{\frac{0.25(0.75)}{1025}} = 0.25 \pm 0.02 = (0.23, 0.27)$. *Conclude:* We are 90% confident that the interval from 0.23 to 0.27 captures the true proportion of all U.S. adults who were satisfied with the way things were going in the United States January 2010.

8.41 (a) We do not know the sample sizes for the men and for the women.
(b) The margin of error for women alone would be greater than 0.03 because the sample size for women alone is smaller than 1019.

8.42 No, the data are not from a random sample, so the conditions for inference have not been met.

8.43 (a) Our guess is $\hat{p} = 0.75$, so we need $1.645 \sqrt{\frac{0.75(0.25)}{n}} \leq 0.04$ or

$n \geq \left(\frac{1.645}{0.04} \right)^2 (0.75)(0.25) = 317.11$. Take a sample of $n = 318$ Americans with at least one Italian grandparent.

(b) If we use $\hat{p} = 0.5$ instead, we need $1.645 \sqrt{\frac{0.5(0.5)}{n}} \leq 0.04$ or

$n \geq \left(\frac{1.645}{0.04} \right)^2 (0.5)(0.5) = 422.82$. Take a sample of $n = 423$ Americans with at least one Italian grandparent. In this case, the sample size needed is 105 people larger.

8.44 (a) Our guess is $\hat{p} = 0.44$, so we need $2.576 \sqrt{\frac{0.44(0.56)}{n}} \leq 0.03$ or

$n \geq \left(\frac{2.576}{0.03} \right)^2 (0.44)(0.56) = 1816.73$. Take a sample of $n = 1817$ adults.

(b) If we use $\hat{p} = 0.5$ instead, we need $2.576\sqrt{\frac{0.5(0.5)}{n}} \leq 0.03$ or

$n \geq \left(\frac{2.576}{0.03}\right)^2 (0.5)(0.5) = 1843.27$ or 1844 adults. The conservative approach requires 27 more adults.

8.45 Because we do not have a best guess for \hat{p} , we will need to use $\hat{p} = 0.5$. So we need

$1.96\sqrt{\frac{0.5(0.5)}{n}} \leq 0.03$ or $n \geq \left(\frac{1.96}{0.03}\right)^2 (0.5)(0.5) = 1067.11$. Take a sample of $n = 1068$ registered voters in the city.

8.46 Because we do not have a best guess for \hat{p} , we will need to use $\hat{p} = 0.5$. So we need

$1.645\sqrt{\frac{0.5(0.5)}{n}} \leq 0.04$ or $n \geq \left(\frac{1.645}{0.04}\right)^2 (0.5)(0.5) = 422.82$. Take a sample of $n = 423$ students.

8.47 (a) The margin of error is stated to be 0.03 and the sample proportion is $\hat{p} = 0.64$. We

know that the margin of error is $z^* \sqrt{\frac{\hat{p}(1-\hat{p})}{n}}$ so $0.03 = z^* \sqrt{\frac{0.64(0.36)}{1028}}$. Thus,

$z^* = \frac{0.03}{\sqrt{\frac{0.64(0.36)}{1028}}} = 2.00$. This is very close to the value 1.96 used for a 95% interval and is

probably different only because of roundoff error. The confidence level is likely 95%.

(b) Teens are hard to reach and often unwilling to participate in surveys, so nonresponse bias is a major “practical difficulty” for this type of poll.

8.48 (a) The margin of error is stated to be 0.01 and the sample proportion is

$\hat{p} = \frac{3547}{5594} = 0.6341$. We know that the margin of error is $z^* \sqrt{\frac{\hat{p}(1-\hat{p})}{n}}$ so

$0.01 = z^* \sqrt{\frac{0.6341(0.3659)}{5594}}$. Thus, $z^* = \frac{0.01}{\sqrt{\frac{0.6341(0.3659)}{5594}}} = 1.55$. The area between -1.55 and

1.55 under the standard Normal curve is 0.8788. The confidence level is likely 88%.

(b) We do not know if those who *did* respond can reliably represent those who did not.

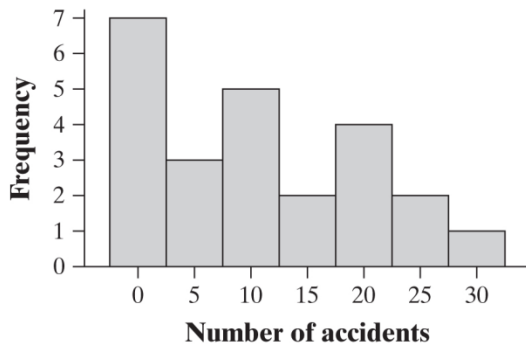
8.49 a.

8.50 d.

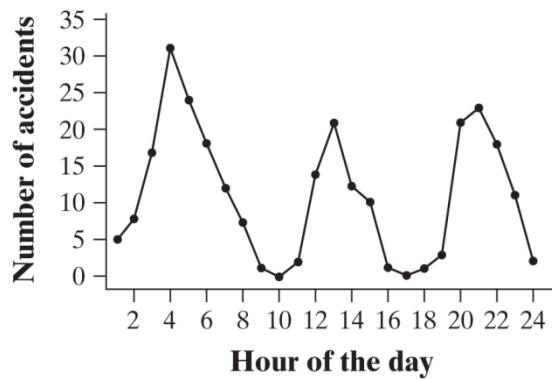
8.51 c.

8.52 a.

8.53 (a) A histogram of the number of accidents per hour is given below.



(b) A graph of the number of accidents is given below.



(c) The histogram in part (a) shows that the number of accidents has a distribution that is skewed to the right.

(d) The graph in (b) shows that there is a cyclical nature to the number of accidents. It looks like there are three shifts and that there is a peak number of accidents after the same amount of time in each shift.

8.54 The boxplots below show that both distributions are slightly skewed to the right. The median number of accidents is greater for the midnight to 8 A.M. shift (shift 1) than for the 4 P.M. to midnight shift (shift 3). In addition, the values of the minimum, Q_1 , Q_3 , and maximum are also higher for the midnight to 8 A.M. shift. The variability of the two distributions are similar, with the midnight to 8 A.M. shift having a slightly larger range and the 4 P.M. to midnight shift having a slightly larger IQR . Overall, it does seem that the number of accidents is higher during the midnight to 8 A.M. shift than in the 4 P.M. to midnight shift.

