Lesson 7.2.1

7-31. a. Each shot is an individual. The variable is "made it?" and is categorical. The population of interest is all possible shots, forever.

b. The statistic you will calculate is $\hat{p}$, the proportion of shots made in the sample, and it will be used to estimate p, the proportion of all shots forever that the team would make.

c. Random, to avoid bias: the problem states the first 100 shots will count as a random sample. Independent trials, for an accurate σ: the problem states that the shots are independent. Large counts, so the sampling distribution ≈ normal: assuming the team makes and misses at least 10 shots each, this will be satisfied.

7-32. $\hat{p}$ is the sample shot percentage and can be calculated (value depends on data). p is the long-term shot percentage and cannot be calculated but can be estimated using the value of $\hat{p}$. The rest of the solutions will be written using the assumption that the team made $\frac{73}{100}$ of their shots, in which case $\hat{p} = 0.73$.

7-33. a. Value depends on results from data collection. Using 0.73 as $\hat{p}$ gives an estimate for the standard deviation of $\sigma_{\hat{p}} \approx \sqrt{\frac{\hat{p}(1-\hat{p})}{n}} = \sqrt{\frac{0.73(0.27)}{100}} \approx 0.044$.

b. Using 0.73 as the original $\hat{p}$ creates new estimates of $\sqrt{\frac{0.83(0.17)}{100}} \approx 0.038$ and $\sqrt{\frac{0.63(0.37)}{100}} \approx 0.048$. These spreads differ from the original estimate but are still reasonably close.

c. The standard error of a statistic is an estimate for the standard deviation of its sampling distribution, and based on part (b), a fairly accurate one.

7-34. Answers vary based on class data.

7-35. a. $n = 642$ voters, $\hat{p} = 0.41$. $SE_{\hat{p}} = \sqrt{\frac{\hat{p}(1-\hat{p})}{n}} \approx 0.02$, so the 95% confidence interval is found by adding two standard errors above and below the sample proportion, leading to a 95% confidence interval of 37% to 45% of the vote.

b. It must be assumed that the poll is a truly random sample of U.S. voters.

c. 0.41 ± 0.04

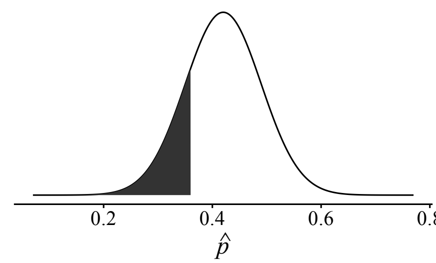
d. No. The margin of error accounts only for sample variability, NOT bias.

7-36. a. Random, to avoid bias: the problem states it is a random sample. Independent trials, for an accurate σ: $698 > 10 \cdot 50$. Large counts, sampling distribution $\approx$ normal: successes and failures are $n\hat{p} = 18$ and $n(1-\hat{p}) = 32$, both > 10.

b. The $\hat{p}$ was $\frac{18}{50} = 36\%$. Using $\sigma_{\hat{p}} = \sqrt{\frac{p(1-p)}{n}} = \sqrt{\frac{0.36(64)}{50}} \approx 0.068$ and 2σ as the margin of error yields a confidence interval of 0.36 ± 0.136. Students are 95% confident that between 22.4% and 49.6% of Hallways School girls will say they get enough sleep.

c. No to both questions. 25% is a reasonable hypothesis, as it is in the confidence interval, however it is no more likely than any other value in the confidence interval and actually significantly less likely than some values.

d. Using normal curve with $\mu = 0.42$ and $\sigma_{\hat{p}} = \sqrt{\frac{p(1-p)}{n}} = \sqrt{\frac{0.42(0.58)}{50}} \approx 0.070$, $P(\hat{p} \le 0.36) = 0.196$. There is about a 20% chance of getting a sample this low or lower. See diagram at right.



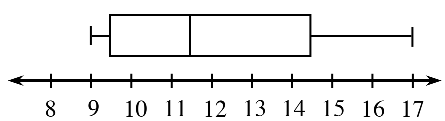
7-37. a. The independent trials condition is not met.

b. Using the formula, $SE_{\hat{p}} = \sqrt{\frac{\hat{p}(1-\hat{p})}{n}} = \sqrt{\frac{0.8(0.2)}{100}} = 0.04$. This is significantly larger than the estimated value!

c. The confidence interval would be 0.8 ± 0.08, which would be significantly wider than the one estimated.

7-38. a. Q1 = 9.5 in, median = 11.5 in, Q3 = 14.5 in

b.



c. $14.5 - 9.5 = 5$

7-39. a. $r = 0.9046$, indicating a strong positive linear association between the number of vehicles/hour versus truck traffic at intersections, which is confirmed by an inspection of the scatterplot.

b. $r^2 = 0.8183$, meaning that 82% of the variation observed in number of vehicles/hour is explained by a linear relationship with truck traffic. In other words, knowing the truck traffic associated with each number of vehicles/hour reduces the error in predicting number of vehicles/hour by 82% over using just the mean number of vehicles/hour as a predictor.

