Lesson 7.2.2

7-40. a. Assuming there are more than 500 hydrangea bushes in Nashville, the conditions are satisfied because it is a random sample, the population is large compared to the sample size, and the number of "blue" (32) and "not blue" (18) are both larger than 10 allowing the normal approximation.

b. Yes for pink, no for the other two as the counts are not large enough to use the normal approximation.

c. blue: 0.64 ± 0.136; pink: 0.28 ± 0.127

7-41. The population proportion is a fixed value; no matter what interval you find, it either contains the population proportion or it does not. So there is not a 95% chance that it contains it—there is either a 0% or a 100% chance that it contains it! The word "confident" is better because it does not connect directly to probability.

7-42. a. A bunch of lines appear. Most of them are black, some of them are red.

b. Answer varies. There are usually between 2 and 8 red lines, though there could be as few as 0 and as many as 10 or 11.

c. They represent the samples whose confidence interval does not contain the true proportion—they "miss."

d. It represents the proportion of black lines—the proportion of confidence intervals that successfully capture the true proportion p.

e. The lines get narrower, and some of the black ones turn red so the capture proportion decreases.

f. The total proportion captured is about 0.95, maybe a little more or little less.

g. …the confidence interval will contain the true proportion about 95% of the time.

h. Rather than using $\hat{p}$ to create the confidence interval, the calculation used p. This is not possible in real-world situations because why would you calculate a confidence interval for p if you already knew it?

7-43. a. 1.645

b. Instead of using 2 standard deviations for the margin of error, use 1.645 standard deviations to capture 90% of the data. This gives a 90% confidence interval of 0.64 ± 0.112 for the proportion of blue hydrangea bushes in Nashville.

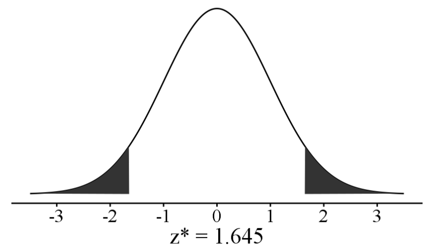
7-44. Because the inverse normal of 0.025 gives a z-score of $1.96 \approx 2$. Using an inverse normal function or standard normal probability table, determine the critical value (z^*) for the confidence level—how many standard deviations you need to go from the mean to capture that percent of the data in a normal curve. Multiply the SE times the critical value to get your margin of error.

7-45. Answers will vary based on $\hat{p}$. Responses should sound like: "I am 90% confident that the interval from [*minimum value*] to [*maximum value*] contains the true value of my team's long term trashketball shot percentage."

7-46. 80%: 1.282, 90%: 1.645, 95%: 1.960, 98%: 2.326, 99%: 2.579

7-47. Identify: This is a 90% confidence interval for p, the proportion of all U.S. citizens who know what a Hoosier is.
Check conditions: Random, to avoid bias: says "SRS." Independent trials, for an accurate σ: 305 citizens· 10 < U.S. population. Large counts, so the sampling distribution ≈ normal: $n\hat{p}$ = successes = 85 citizens, $n(1 - \hat{p})$ = failures = 220 citizens, both >10.

standard normal distribution



Calculate: $\hat{p} = \frac{85}{305} \approx 0.2787$, SE ≈ 0.02567 citizens, z* = 1.645, margin of error = 1.645(0.02567) = 0.04223 citizens, confidence interval = 0.2787 ± 0.04223, {0.2365, 0.3209} citizens. See graph above right.
Conclude: I am 90% confident that the interval from 0.237 to 0.321 contains the true population proportion of U.S. citizens who know what a "Hoosier" is.

7-48. a. The population is not more than 10 times the sample size.

b. The confidence interval will be predictably too wide, meaning intervals made with this technique will capture the parameter more often than expected.

c. The calculated interval is 0.11 ± 0.059. We can be (somewhat more than) 94% confident that the true proportion of ripe strawberries in the patch is between 5.1% and 16.9%.

7-49. a. An observational study would be best since you simply need to observe whether or not students are reading for pleasure during summer vacation.

b. A simple random sample of students from the school would be given a survey about their summer reading habits. The survey could be distributed in the fall when the new school year begins. The teachers could follow up with students who did not return their surveys.

7-50. See diagrams below. Mean and standard deviation are appropriate statistics because both distributions are fairly symmetric with no outliers. Both types of toads lay a mean of between 9000 and 9100 eggs. Both distributions are single-peaked and symmetric with no apparent outliers. However there is much greater variability in the number of eggs that American toads lay. The standard deviation for the sample of American toads is about 654 eggs while the standard deviation for the sample of Fowler toads is only half as much, about 324 eggs.