Learning Goal: Interpret the confidence level associated with a confidence interval.

Learning Objective: Explain the meaning of "95% confident."

Introduction:
When we say "we are 95% confident that between 44% and 52% of Americans drink soda every day, what do we really mean by the phrase "we are 95% confident"? Let's find out.

To investigate the meaning of "95% confident", we will assume that we know the population proportion for some context and then we will examine how often a confidence interval contains this population proportion.

Simulation:
According to the Community College Research Center, 80% (or more) of community college students intend to earn at least a bachelor's degree.
Source: http://ccrc.tc.columbia.edu/media/k2/attachments/what-we-know-about-transfer.pdf

For our simulation, let's assume that 80% is a population parameter. In other words, we are going to assume that the finding in this study actually holds true for the entire population of community college students.

Based on this assumption, let's simulate collecting random samples of 100 students and calculate some confidence intervals. Our goal is to see how often a confidence interval will accurately estimate the population proportion. By this we mean that the confidence interval contains the population proportion of 0.80.

1) Verify that the distribution of sample proportions for this situation can be modeled by a normal curve.

2) Use the StatCrunch Sampling Distribution applet to select a random sample with p=0.80 and n=100. Calculate the sample proportion and the associated confidence interval.

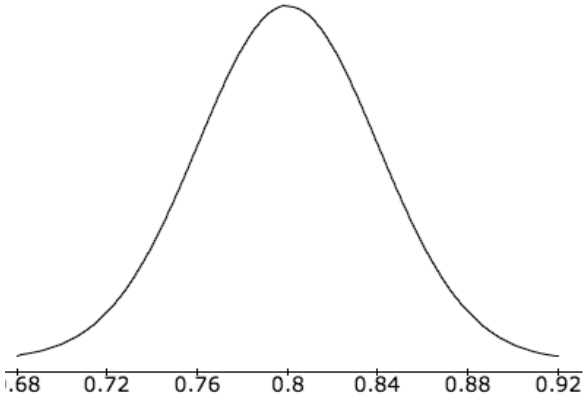
Your instructor's sample proportion and confidence interval:

Interpretation:

Your sample proportion and confidence interval:

Interpretation:

3) Here is the normal curve with mean of p=0.80 and standard error of 0.04, which models the distribution of sample proportions. We will plot a few confidence intervals below it.

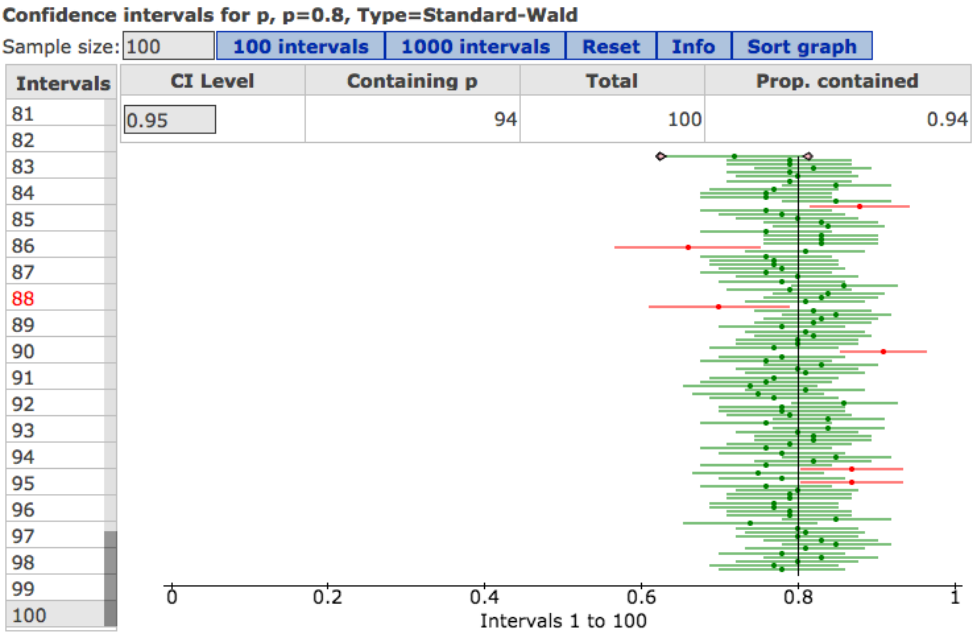


All of the confidence intervals are the same length. Why does this make sense?

4) Which of the confidence intervals that we plotted accurately estimate the population proportion? (In other words, which of the confidence intervals that we plotted contain p=0.80?)

5) What proportion of the class got a confidence interval that did NOT contain p=0.80?

6) The image below represents 100 confidence intervals with p=0.80 and n=100.



a) What does each line segment represent?

b) Why is there a dot in the middle of each line segment? What does it represent?

c) Circle a confidence interval that does not accurately estimate the population proportion. Why doesn't this interval contain p?

d) What percentage of the confidence intervals plotted here accurately estimate the population proportion? (In other words, what percentage of the confidence intervals plotted here contain p=0.80?)

e) In the long run, if we collected thousands and thousands of random samples, what percentage of the associated confidence intervals will contain p? Why do you think so?

f) What do you think "95% confident" means?

7) A student selected a random sample with a sample proportion of 0.85 and correctly calculated the associated confidence interval as 0.77 to 0.93. He also correctly interpreted the interval as "Of all community college students, we are 95% confident that the proportion who intend to get at least a bachelor's degree is between 0.77 and 0.93."

   When asked to interpret the meaning of the phrase "95% confident," the student wrote "95% confident means that 95% of the time the population proportion lies between 0.77 and 0.93." But this is incorrect.

   Use the image of the StatCrunch simulation on the previous page to explain why his interpretation of "95% confident" is incorrect.