Learning Goal: Find a confidence interval to estimate a population proportion when conditions are met. Interpret the confidence interval in context.
Specific Learning Objectives:

- Construct a 95% confidence interval from scratch.
- Check conditions that allow the use of a 95% confidence interval.

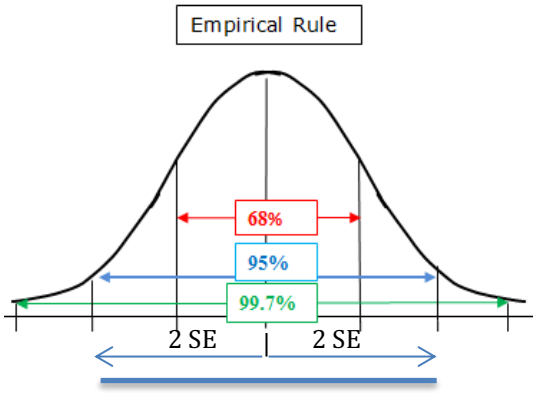
Introduction

In this activity we will learn how to construct a confidence interval from scratch.

Recall that the confidence interval is *sample proportion ± margin of error.*
For a 95% confidence interval the margin of error is two standard errors.

sample proportion ± margin of error
sample proportion ± 2 SE

This makes sense because of the Empirical Rule. If we randomly select a sample, there is a 95% chance that the sample proportion lies within two standard errors of the population proportion. Of course, this is only true if the distribution of sample proportions is normal in shape.



We learned previously that the standard error (standard deviation) for the distribution of sample proportions is $\sqrt{\frac{p(1-p)}{n}}$, where p is the population proportion and n is the sample size.

Putting the pieces together gives the formula for the 95% confidence interval:

Sample proportion ± margin of error
Sample proportion ± 2 standard errors

$$\textit{Sample proportion} \pm 2\sqrt{\frac{p(1-p)}{n}}$$

IMPORTANT NOTE: This formula can only be used if a normal model is a good fit for the distribution of sample proportions.

You may realize that the formula for the confidence interval is a bit odd, since our goal in calculating the confidence interval is to estimate the population proportion, p. Yet the formula requires that we know p in order to calculate the standard error. **For now, we use an estimate for *p* from a previous study when calculating the confidence interval, as we have done before.** This is not the usual way statisticians estimate the standard error, but it captures the main idea and allows us to practice finding and interpreting confidence intervals. Later, we explore a different way to estimate standard error that is commonly used in statistical practice.

Example: The National Health Interview Survey for 2011-2014 estimates that 37% of 18-24 year olds are not meeting the Physical Activity Guidelines for Americans. These guidelines include recommendations for both aerobic and muscle-building exercise.

Suppose that we survey 30 randomly selected LMC students in this age category and find that 45% are not meeting the Physical Activity Guidelines for Americans. Find the 95% confidence interval and interpret it.

- Verify that the conditions for the use of the normal model are met based on the NHIS study.

- Calculate the standard error based on the NHIS study.

- Find the 95% confidence interval and represent it as a line segment.

- Interpret the interval.

Here is a link if you want to read about the Physical Activity Guidelines later: https://health.gov/paguidelines/

Group Work:

1) Using the data from the American Community Survey (IPUMS), the Pew Research Center reported that 32.1% of 18- to 34-year-olds was living with their parents in their parents' home in 2014. For the first time in recorded history, this was the most common living arrangement for this age group.

   What percentage of 18- to 34-year-olds is living at home with their parents this year?

   To answer this question, suppose that this year researchers conduct another survey with a random sample of 1,100 adults in this age group. Suppose that 35% are living at home with their parents. Assuming that the variability in random samples will be the same as in the 2014 study, find the 95% confidence interval for this year and interpret it.

   a) Verify that the conditions for the use of the normal model are met based on the Pew Research Center study.

   b) Calculate the standard error based on the 2014 study.

   c) Find the 95% confidence interval for this year and represent it as a line segment.

   d) Interpret your 95% confidence interval.

2) What percentage of Contra Costa County adults supports a ban on assault-style weapons?

   Suppose that we survey a random sample of 100 adults living in Contra Costa County and find that 62% support a ban on assault-style weapons.

   According to a 2015 study by the Pew Research Center, 57% of U.S. adults favor a ban on assault-style weapons. Assuming that the variability in random samples

will be the same in Contra Costa County as in the U.S., find the 95% confidence interval to estimate the proportion of CCC adults that favor a ban on assault-style weapons.

a) Verify that the conditions for the use of the normal model are met using the 2015 study.

b) Why do we check conditions for use of the normal model?

c) Calculate the standard error based on the 2015 study. What does this number tell us?

d) Find the 95% confidence interval for Contra Costa County.

e) Write a sentence to interpret the interval.

f) Are we confident that the majority of Contra Costa County residents support a ban on assault-style weapons? Why or why not?

g) Are we confident that the percentage of Contra Costa County residents that supports a ban is greater than the percentage nationwide as reported by the Pew Research Center? Why or why not?