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

Learning Objective: Use information from a media report of a survey to construct and interpret a confidence interval.

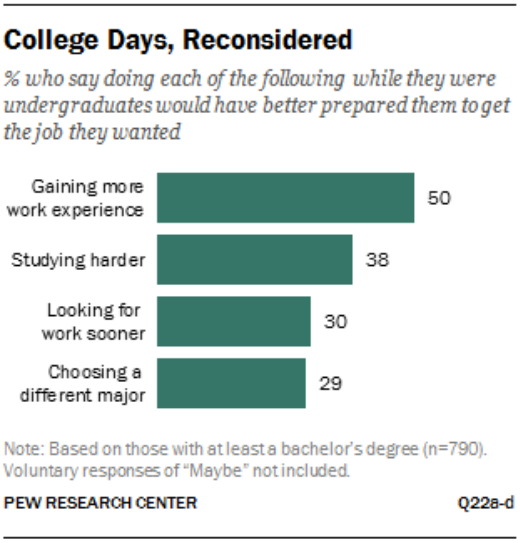
Introduction: Our goal in statistical inference is to *infer* something about a population from a sample. For example, suppose that researchers want to estimate the proportion of the population of U.S. adults that support the death penalty. It is too expensive and difficult to survey every U.S. adult. So researchers will poll a random sample of U.S. adults and use the sample proportion to draw a conclusion about the population proportion.

In this activity we will learn about one type of statistical inference called a *confidence interval.* We construct a confidence interval when our goal is to estimate a population proportion.

Example: Constructing a confidence interval from a poll.

The Pew Research Center asked college graduates what they could have done, while still in school, to be better prepared for their dream job.

We will use the results of this poll to estimate the proportion of the population of all college graduates who wished they had studied harder.
In the report, the researchers state that the poll had a margin of error of 4.3 percentage points at the 95% confidence level.



Source: http://www.pewresearch.org/methodology/u-s-survey-research/sampling/

1) What percentage of the sample of 790 college graduates said they wished they had studied harder?
2) Our estimate for the population proportion will be an interval of values, instead of a single number. Statisticians call this interval a *confidence interval.*
   The confidence interval is based on a sample proportion and a margin of error. Here is the formula: *sample proportion ± margin of error*

   Write the confidence interval to estimate the proportion of the population of all college graduates who wished they had studied harder.

3) Interpreting the confidence interval: For the population of all college graduates, we are 95% confident that the proportion that wish they had studied harder in college is between _____ and _____.

   Note: We will discuss the phrase '95% confident' in depth later.

4) In order to be confident that the sample represents the population without bias, we must have a random sample. Read the following excerpt from the Pew Research methodology for this *College Days* study.

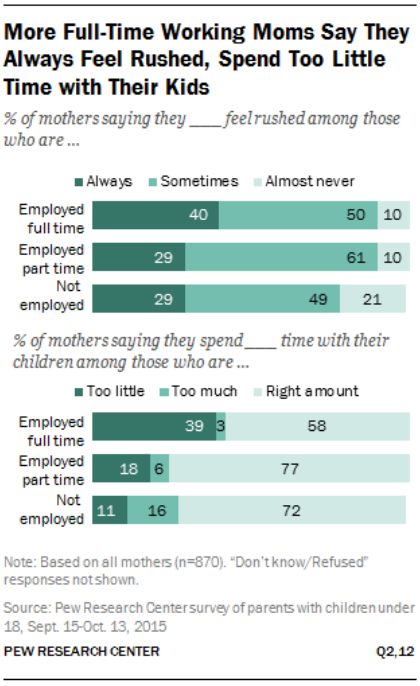
   "A majority of Pew Research Center surveys are conducted among the U.S. general public by telephone using a sampling method known as random digit dialing or "RDD." This method ensures that all telephone numbers in the U.S – whether landline or cellphone – have a known chance of being included. As a result, samples based on RDD should be unbiased, and a margin of sampling error and a confidence level can be computed for them.

   Is this a random sample? How do you know?

Group work:

1) The graph pictured here summarizes some of the results of a poll about parenting conducted by the Pew Research Center.

   Source:http://www.pewsocialtrends.org/2015/11/04/raising-kids-and-running-a-household-how-working-parents-share-the-load/



   a) The report states, "the error attributable to sampling is plus or minus 3.9% at the 95% level of confidence."

      For the population of all mothers employed full time, calculate a confidence interval to estimate the proportion of the who feel that they spend too little time with their children.

   b) Interpret the confidence interval:
      For the population of all mothers employed full time, we are 95% confident that between _____ and _____ feel they spend too little time with their children.

c) In order to be confident that the sample represents the population without bias, we must have a random sample. Read the following excerpt from the Pew Research methodology.

"The analysis in this report is based on telephone interviews conducted from Sept. 15 to Oct. 13, 2015, among a nationally representative sample of 1,807 parents, 18 years of age or older, with children under 18, living in all 50 U.S. states and the District of Columbia. A combination of landline and cell phone random digit dial (RDD) samples was used. Interviews were conducted in English and Spanish. The survey results pictured here came from the responses of 870 mothers."

Is this a random sample? How do you know?

2) Based on the McClatchy-Marist Poll summarized below, can we conclude that the majority of registered voters nationwide support a ban on the sale of assault weapons and semi-automatic weapons? Why or why not?

McClatchy-Marist Poll. July 5-9, 2016. N=1,053 registered voters nationwide. Margin of error ± 3.

"Do you think Americans are safer with more guns or fewer guns?"

	More guns %	Fewer guns %	Number is about right (vol.) %	Unsure %
ALL	45	46	3	5
Democrats	16	77	2	4
Republicans	79	16	3	3
Independents	44	44	5	7

"Do you favor or oppose a law to ban the sale of assault weapons and semi-automatic rifles?"

	Favor %	Oppose %	Unsure %
ALL	51	46	3
Democrats	74	23	3
Republicans	30	66	4
Independents	48	49	3

3) According to a Gallup poll conducted in September 2016, "64.2% of U.S. adults report that they ate healthy all day yesterday."

   Results are based on telephone interviews conducted as part of the Gallup-Healthways Well-Being Index survey, with a random sample of 2,415,499 adults, aged 18 and older, living in all 50 U.S. states and the District of Columbia. For results based on the total sample of national adults, the margin of sampling error is ± .08 percentage points at the 95 percent confidence level.

   a) Construct a confidence interval to estimate the proportion of U.S. adults who ate healthy yesterday.

   b) Interpret your confidence interval.

   c) This is a very precise confidence interval with a very, very small margin of error (0.08% is less than 1/10 of one percent). Why do you think the margin of error is so small for this poll?

4) What is the purpose of a confidence interval?