Margin of Error when Estimating a Population Proportion

Student outcomes:
- Students will use data from a random sample to estimate a population proportion.
- Students will calculate and interpret margin of error in context.

Investigation: Mystery Bags
Students should recognize that the values of statistics calculated from samples selected from known populations will vary from sample to sample. For example, if you toss a coin 20 times, you might have 11 heads; if you toss the coin another 20 times you might have 7 heads. In this lesson students see how understanding variability can tell you about an unknown population. The way to start is by first understanding how samples from a population that you know all about would behave. The following question could be used to help students think about the variability inherent in different samples from the same population.
"Suppose you know that 20% of the chips in a bag are blue. Write down an estimate of the number of blue chips you are likely to see in a random sample of 30 chips from the bag. "

Before the lesson
Prepare at least nine small bags of colored chips, one bag for each pair of students. Each bag should have 20 chips with the following numbers of blue chips. If the class is large, prepare duplicate bags for the percentages.

2 blue chips – 10% of the chips are blue
4 blue chips – 20% of the chips are blue
6 blue chips – 30% of the chips are blue
8 blue chips – 40% of the chips are blue
10 blue chips- 50% of the chips are blue
12 blue chips- 60% of the chips are blue
14 blue chips- 70% of the chips are blue
16 blue chips- 80% of the chips are blue
18 blue chips – 90% of the chips are blue

Drawing a blue chip will constitute a "success". The proportion of blue chips in each bag should be clearly written on the bottom of the bag. (Several pairs of students can have the same proportion, if it the class is large.) The other chips can be any color other than blue. These chips will represent "failures".

The Mystery Bag
Prepare one "mystery bag" that has 20 chips with an unknown (to the class) proportion of blue chips. (You might put 8 blue chips in the bag for a mystery proportion of 40%. Try not to use 6 blue chips (30%) as students may confuse the percent with the sample of 30 chips that will be drawn with replacement from the bag.)

The Lesson
Show the Mystery Bag and indicate it has some proportion of blue chips: the big question is what is that proportion and how can you find out without looking in the bag.

Have one student draw a chip from the Mystery Bag, another record whether it was blue or not blue. Return the chip to the bag, mix the bag, and repeat the process until a random sample of 30 blue chips has been drawn.

Record the number of blue chips – and label it as the observed outcome. Now the question is, "What does this tell us about the actual proportion of blue chips in the bag?" Put the observed outcome aside and move to the next part of the lesson.

Working from the Known
Give each pair of students a bag with the proportion of blues marked on the bottom of the bag. Students should draw 30 chips, one at a time, replacing the chip after recording the color. They should repeat the experiment several times.

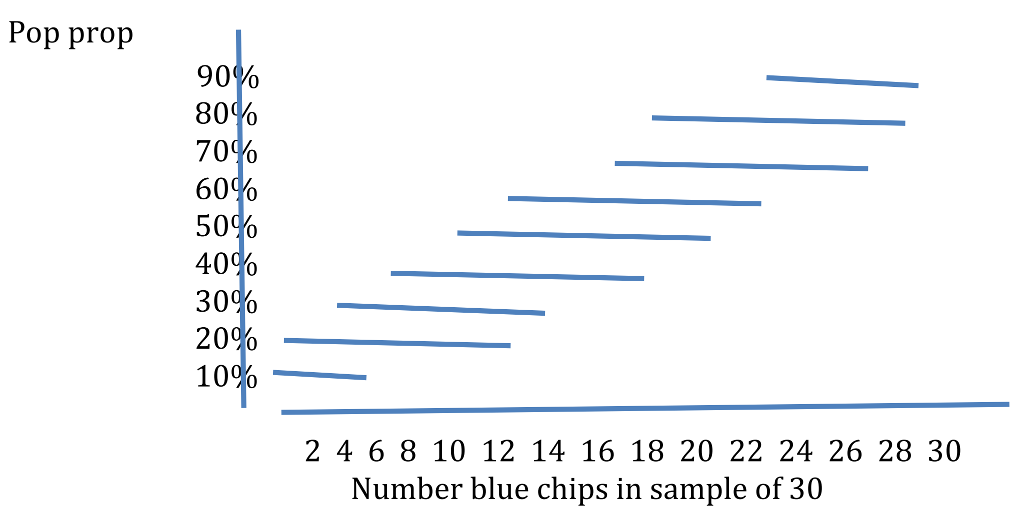
Monitor the class and when most of the students have two sets of 30 observations, bring the class together and suggest they use technology to generate their random samples to speed up the process.

Technology: Use the TI Nspire Sample Proportions tns file and activity from Building Concepts: Statistics and Probability (https://education.ti.com/en/building-concepts/activities/statistics/sequence1/sample-proportions

After most of the students have generated at least 50 random samples of size 30 from their bags (with Nspires it can be 500 samples) and recorded the number of blues in each sample in a frequency tally, bring the class together. Ask them to look at their simulated sampling distributions for the number of blue chips in their samples and to write down an interval that seems to describe the number of blues they typically got for the proportion of blue chips in their bags.

Make a public record of the known proportion of blue chips in each pair's bag and the interval that describes the number of blue chips in their samples of 30.

Example of chart with intervals representing the number of blue chips for different known population proportions



Putting it all together:

Refer back to the number of blue chips in the random sample of size 30 that was drawn from the mystery bag. The question is given the variability for sample proportions from known population proportions, which of the populations do you think would be a plausible population for our observed results?

Based on the simulated sampling distributions, do you think that the mystery bag might have had 10% blue chips? Explain your reasoning.

Based on the simulated sampling distributions, which of the percentages 10%, 20%, 30%, 40%, 50%, 60%, 70%, 80% and 90% might reasonably be the percentage of blue chips in the mystery bag?

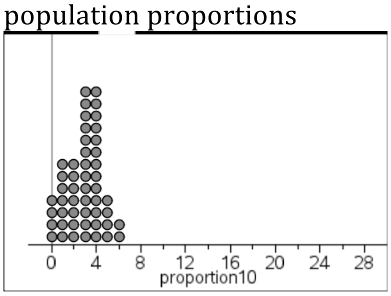
What do you think are plausible population proportions of blue chips for the chips in the mystery bag?

If the plausible populations had from 9 to 18 blue chips, that would be the same as 0.30 to 0.60 proportions of blue chips. This is sometimes written as 0.45 $\pm$ 0.15. The value 0.15 is called a "**margin of error".**

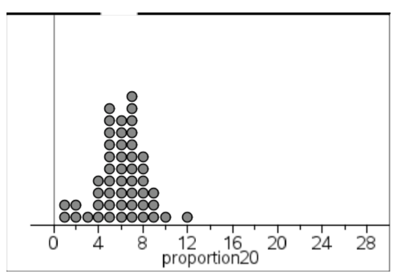
The last step is to find the margin of error for the proportion of blue chips in the mystery bag.

Sample results:

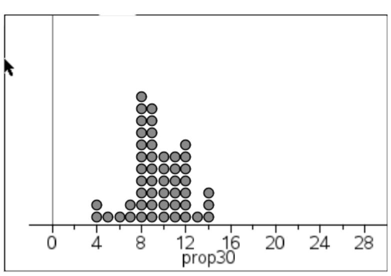
Simulated sampling distribution for number of blue chips in samples of size 30 for different population proportions



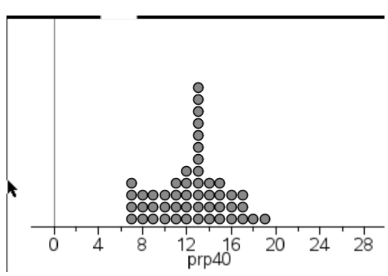
Population 10% blue chips



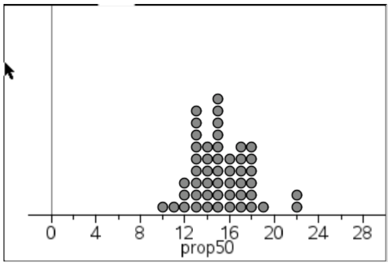
Population 20% blue chips



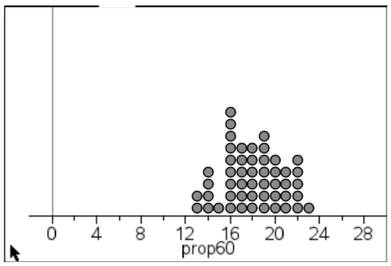
Population 30% blue chips



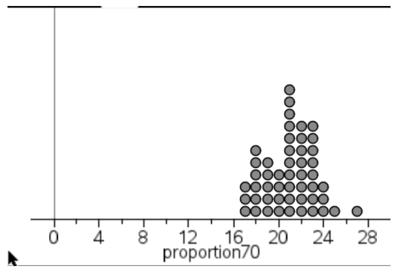
Population 40% blue chips



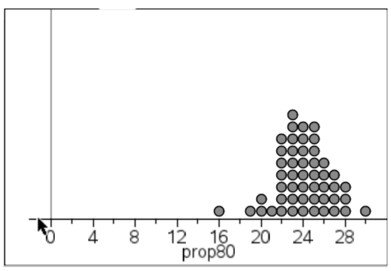
Population 50% blue chips



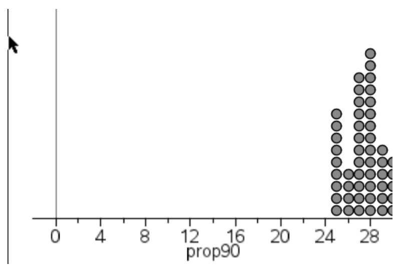
Population 60% blue chips



Population 70% blue chips



Population 80% blue chips



Population 90% blue chips