Estimate a Population Proportion

MATERIALS • graphing calculator

MATHEMATICAL PRACTICES

Reason abstractly and quantitatively.

**QUESTION** **How can you use a sample proportion to estimate a population proportion?**

You can use statistics to make reasonable predictions, or *inferences*, about an entire population from a sample of the population. A *population proportion* is the ratio of members of a population with a particular characteristic to total members of the population. A *sample proportion* is the ratio of members of a sample of the population with a particular characteristic to total members of the sample.

EXPLORE 1 **Compare sample proportions from a population with known proportion**

In the "population" of randomly-generated decimals between 0 and 1, the proportion of decimals less than 0.6 is 0.6. Simulate random samples of decimals from this population, and compare the sample proportions with the population proportion.

STEP 1 *Perform a simulation*

Use a graphing calculator to generate 40 random decimals between 0 and 1. Press MATH , choose the "PRB" menu, select "rand," and press ENTER . The display will be similar to the one at the right.

```
rand
        .8999652259
```
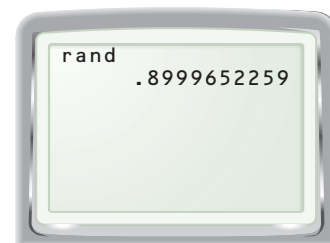
Press ENTER 40 times to generate 40 random numbers. Keep a tally of the number of decimals below 0.6. Then find your sample proportion by dividing the number of decimals below 0.6—the "successes"—by the number of trials, which is 40.

Repeat the simulation to find another sample proportion. Each group will need to do enough simulations so that there are 40 in all for the class.

STEP 2 *Display sample proportions*

Collect all 40 sample proportions. Make a histogram of the sample proportions using intervals of 0.05, for example, [0.55, 0.6). The histogram for one set of 40 sample proportions is shown at the right.
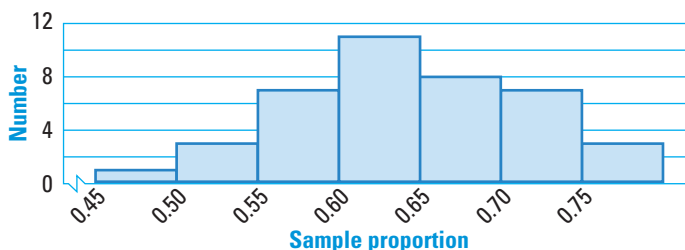
STEP 3 *Investigate sample proportions*

Exclude the single smallest and single greatest of the 40 sample proportions. Then find the interval that contains the remaining 38 sample proportions.

Plot the interval containing the middle 38 sample proportions as shown. The blue horizontal segment represents the middle 38 of 40, or 95%, of the sample proportions. So, 95% of the time, the sample proportion from the simulation fell within the interval graphed.
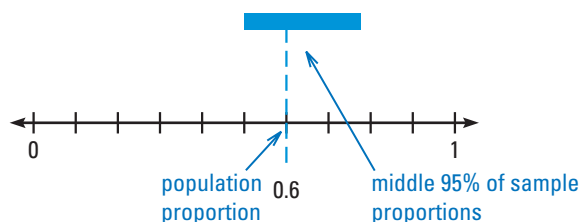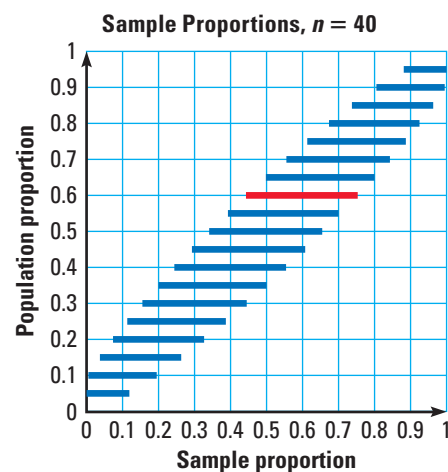
EXPLORE 2 **Estimate a population proportion from a sample proportion**

Sample Proportions, *n* = 40

STEP 1 *Investigate sample proportions*

There are formulas for calculating intervals expected to contain 95% of the sample proportions given any population proportion and sample size. The chart at the right shows the 95% sample proportion intervals for population proportions of 0.05, 0.1, 0.15, ..., 0.9, 0.95 where the sample size is $n = 40$.

The interval in red is the calculated 95% sample proportion interval for a population proportion of 0.6. Compare this interval with the 95% interval your class created for the simulation in Explore 1.

STEP 2 *Use a sample proportion*

It's not usually possible to collect data for entire populations. So, data from samples are used to estimate unknown population proportions. Suppose you want to estimate a population proportion using a sample of size 40, and your sample has a sample proportion of 0.6. Which populations are most likely to have produced a sample like yours?

The green vertical line shown at right highlights a sample proportion of 0.6. Any horizontal bar that this line touches corresponds to a population that will produce a sample whose sample proportion lies within that bar 95% of the time. The green line touches the bars for population proportions of 0.45 up to 0.75. So, *if we have a sample of size 40 with a sample proportion of 0.6, then, 95% of the time, the population proportion should lie in the interval from 0.45 to 0.75.*
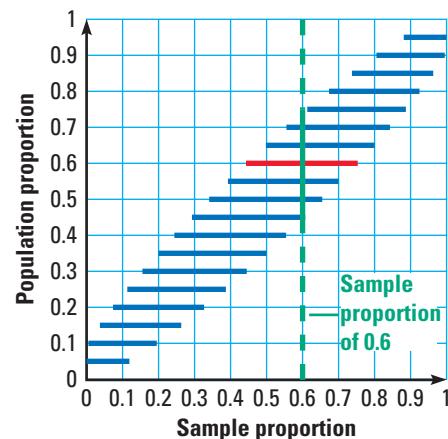
STEP 3 *Estimate a population proportion*

For a sample of size 40 with sample proportion 0.3, use the chart to estimate the interval that contains the corresponding population proportion 95% of the time.

DRAW CONCLUSIONS **Use your observations to complete these exercises**

1. As a class, repeat Steps 1 and 2 of Explore 1, but with each trial count the number of decimals less than 0.6 out of *80* random decimals. *Describe* how increasing the sample size in each trial affects the shape of the histogram.

2. You have used the *margin of error* $\pm \frac{1}{\sqrt{n}}$ for a sample of size *n* to estimate an interval containing a population proportion. More precisely, this expression is $\pm 2\sqrt{\frac{\hat{p}(1 - \hat{p})}{n}}$, where the symbol $\hat{p}$ (read as "*p*-hat") is the sample proportion. Then $\hat{p} - 2\sqrt{\frac{\hat{p}(1 - \hat{p})}{n}} < p < \hat{p} + 2\sqrt{\frac{\hat{p}(1 - \hat{p})}{n}}$ is a 95% *confidence interval* for the population proportion *p*. We can expect the population proportion to be in this interval 95% of the time. Find a 95% confidence interval for the population proportion for a sample of size 150 with a sample proportion of 0.8.