

7.5: Distribution of Sample Proportions (2 of 6)

Learning Objectives

- Describe the sampling distribution for sample proportions and use it to identify unusual (and more common) sample results.

Recall that our goal is to create a probability model that describes the long-run behavior of proportions from random samples. Previously, we used a simulation to collect a few random samples to get acquainted with making a distribution of sample proportions. We randomly selected 25 students at a time from a population of part-time college students that is 60% female. In the next example, we predict what happens in the long run when we select many, many random samples of 25 students at a time from this population. Then we watch a simulation to see if our predictions are correct.

Example

Predicting the Behavior of Sample Proportions

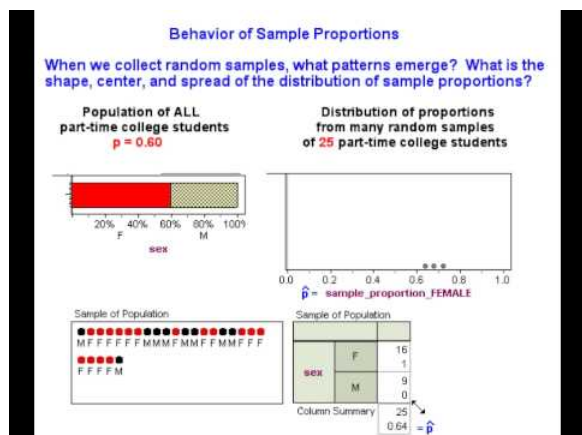
Based on our intuition and what we observed with the simulation, we might expect the following about the distribution of sample proportions that come from a population where $p = 0.60$:

Center: Some sample proportions will be on the low side – such as 0.52 or 0.56 – and others will be on the high side – such as 0.64 or 0.68. It is reasonable to expect all the sample proportions in repeated random samples to average out to the underlying population proportion, 0.6. In other words, the mean of the distribution of sample proportions should be p .

Spread: For samples of 25, we expect sample proportions of females not to stray too far from the population proportion 0.6. Sample proportions lower than 0.44 or higher than 0.72 will be unusual. Previously, we took smaller random samples of 8 and observed more variability in the sample proportions. We therefore think that sample size plays a role in the spread of the distribution of sample proportions. Smaller samples may be less accurate and more variable than larger samples.

Shape: Sample proportions closest to 0.6 will be most common, and sample proportions far from 0.6 in either direction will be progressively less likely. In other words, the shape of the distribution of sample proportions may be somewhat bell-shaped.

Now we use a simulation to collect numerous samples to see what happens in the long run. We use the simulation to check whether our intuition about the center, spread, and shape of the distribution of sample proportions is right.



A YouTube element has been excluded from this version of the text. You can view it online here: <http://pb.libretexts.org/sss/?p=272>

Learn By Doing

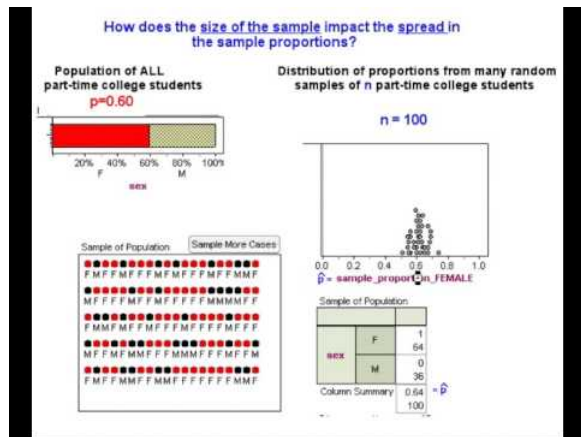
<https://assessments.lumenlearning.co...sessments/3576>

Learn By Doing

<https://assessments.lumenlearning.co...sessments/3889>

At this point, we have a good sense of what happens as we take random samples from a population. Our simulation suggests that our initial intuition about the shape and center of the distribution of sample proportions is correct.

Now we use another simulation to help us think more precisely about the variability we expect to see in the sample proportions. Our intuition tells us that larger samples will better approximate the population, so we might expect less variability in large samples. In the next walk-through, we use a simulation to investigate this idea. After that walk-through, we tie these ideas to more formal theory about the probability model for the long-run behavior of proportions from random samples.



A YouTube element has been excluded from this version of the text. You can view it online here: <http://pb.libretexts.org/sss/?p=272>

We are now ready to use what we have observed to develop a formal probability model to describe the behavior of sample proportions. First, let's return to the original question that prompted our investigation of sample proportions.

We based our investigation on the prediction that 60% of part-time college students will be female. In our investigation, we asked, *How much could the sample proportion vary from a population proportion of 0.60 for us to feel confident in our prediction?*

We don't expect a sample proportion to be exactly equal to the population proportion. But how much error seems reasonable?

We now see that the answer to this question depends on the size of the sample.

- If we select a random sample of 25 students, the distribution of sample proportions has a standard deviation of about 0.1. We can see that most sample proportions fall within 2 standard deviations of 0.60. Therefore, we might decide that $2 \times 0.10 = 0.20$ is a reasonable margin of error, so a sample proportion between 0.40 and 0.80 is not surprising if 0.60 of all part-time college students are female.
- If we select a random sample of 100 students, the distribution of sample proportions has less variability. It has a standard deviation of about 0.05. Again we see that most sample proportions fall within 2 standard deviations of 0.60, so we might decide that $2 \times 0.05 = 0.10$ is a reasonable amount of error for these larger samples. For a sample of 100 students, then, a sample proportion between 0.50 and 0.70 is not surprising if 0.60 of all part-time college students are female.

We discuss this idea further in "Introduction to Statistical Inference" in this unit.

CC licensed content, Shared previously

- Concepts in Statistics. **Provided by:** Open Learning Initiative. **Located at:** <http://oli.cmu.edu>. **License:** CC BY: Attribution

This page titled 7.5: Distribution of Sample Proportions (2 of 6) is shared under a CC BY 4.0 license and was authored, remixed, and/or curated by Bill Pelz via source content that was edited to the style and standards of the LibreTexts platform.