Lesson 11: Reducing Margin of Error

- Let's estimate population proportions and explore margin of error.

11.1: Notice and Wonder: Female Leads

Five students wanted to see how many children's movies have female lead characters. They each took a sample of children's movies, found the proportion of movies that had female lead characters, then used their results to simulate 100 additional samples. The table shows some of the findings based on the original sample and the simulations.

What do you notice? What do you wonder?

student	number of movies used in the original random sample	estimated proportion	margin of error
A	20	0.273	0.204
B	20	0.205	0.189
C	30	0.280	0.180
D	30	0.232	0.157
E	50	0.205	0.115

11.2: Finding a Job

Elena and Clare are each working on a project about how high school students are having trouble finding jobs. They each find the proportion of students without jobs from a random sample, then use a computer to do 1,000 simulations using the proportion they found and report the results.

Elena says, "The proportion of high school students without jobs is about 0.70 with a margin of error of 0.280."

Clare says, "The proportion of high school students without jobs is about 0.74 with a margin of error of 0.138."

1. Both students reported the margin of error based on 2 standard deviations from their simulations. What are the mean and standard deviation each student found? For at least one student, show your reasoning.

2. Clare and Elena try to figure out why Clare had such a smaller range of values in her report.

   a. First they consider the proportion they used in the simulations. Elena says, "My simulation used 0.7 as the proportion since I found that proportion in my original sample." Clare says, "My simulation used 0.75 as the proportion since I found that proportion in my original sample." The students used different proportions in their simulations. Do you think this is why Clare has a smaller margin of error? Explain your reasoning.

   b. They look for more differences in their initial sample and discover than Elena surveyed 10 people in her initial sample and Clare surveyed 40 people. Do you think this is why Clare has a smaller margin of error? Explain your reasoning.

11.3: Exercised

Cut a sheet of paper into enough slips for each student in the class to get one. On each of the slips you cut, write "Yes" if you spend at least 5 hours intentionally exercising each week, otherwise write "No" on each of the slips. Put one of your paper slips in each student's bag, including your own.

After all the slips are distributed for all the students, return to your bag. Your teacher gave your group a number of slips to draw for each sample. Draw a sample and record the proportion of the slips that say "Yes." Return the slips to the bag and repeat the process until you have 10 sample proportions from 10 samples. Share your results with the group so that each person has 50 sample proportions to work with.

1. Use your 50 sample proportions to report an estimate and associated margin of error for the class. Explain or show your reasoning.

2. Compare the standard deviations of the 50 sample proportions for each of the different groups. Is there a connection to the number of slips chosen in each sample?

Are you ready for more?

Here is a table of standard deviations that were obtained when repeatedly sampling from a population where the proportion of "Yes" slips was 0.4.

sample size	standard deviation
5	.219
20	.110
45	.073
80	.055
125	.044
500	

Estimate the standard deviation for a sample size of 500.

Lesson 11 Summary

Estimating a population characteristic from a random sample will always have some room for error since the sample is only a subset of the population, so it does not provide complete information about the population. One way to reduce the error is to take larger random samples. Not only will this include more of the population in the sample, so a greater percentage of the total information is being recorded, but the standard deviation of a sample statistic from simulations using larger samples also tends to be smaller. With a smaller standard deviation, the difference between the sample estimate and the actual value of the population characteristic being estimated tends to be smaller.

For example, a group goes to an island and collects a random sample of 10 lizards, finding that 5 of them are male. This random sample has a proportion of 0.5 males in the group. How close is this likely to be to the actual proportion of lizards that are male on the island? To investigate, we can simulate additional samples from a population in which the proportion is 0.5 and see how far away from 0.5 the sample proportions tend to be. The distribution of simulated samples will give us an idea of how far off our sample estimate of 0.5 might be from the actual population value.

Suppose we simulate taking 30 random samples of 10 lizards from a population with a 0.5 probability of each one being male, and this results in sample proportions that have a mean of 0.503 and a standard deviation of 0.145. The dot plot of the sample proportions is approximately normal in shape, so it is reasonable to think that simulated proportions from the population should be within about 2 standard deviations, or 0.290 ($2 \cdot 0.145 = 0.290$) of the actual population mean.

Based on the simulations and analysis, we expect that the original estimate of 0.5 for the proportion of the population that is male is likely to be within 0.290 of the actual value of the population proportion of lizards that are male. The researchers should report an estimate of 0.5 for the population proportion with an associated margin of error of 0.290.

Later, another group goes to the island and collects a sample of 40 lizards, finding that 20 of them are male. After simulating 30 samples of 40 lizards with a 0.5 probability of each one being male, the mean proportion that is male is found to be 0.503 again, but the standard deviation is 0.062. This group should report an estimated proportion of lizards on the island that are male of 0.503 with a margin of error of 0.124 ($2 \cdot 0.062 = 0.124$). This means that they believe their estimate of 0.503 is within 0.124 of the actual population proportion.

samples of 40

samples of 10

0   0.1   0.2   0.3   0.4   0.5   0.6   0.7   0.8   0.9
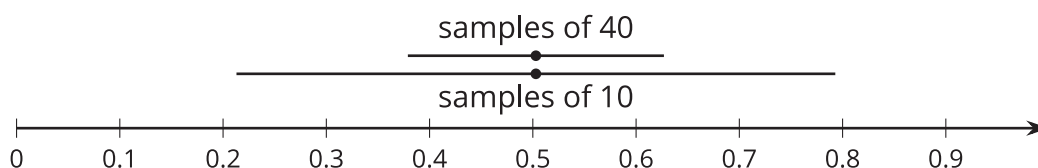
Although the means are the same in this case, the standard deviation is much less with the larger samples, so the margin of error reported was smaller.