Sampling Reese's Pieces

Adapted by: Trey Cox, Ph.D., Chandler-Gilbert Community College, trey.cox@cgc.edu

* This activity is an adaptation of an activity from Rossman and Chance (2000), *Workshop Statistics: Discovery with Data,* 2nd Edition.

Overview of Lesson

This activity uses simulation to help students understand sampling variability and reason about whether a particular sample result is unusual, given a particular hypothesis. By using first candies, a web applet, then a calculator, and varying sample size, students learn that larger samples give more stable and better estimates of a population parameter and develop an appreciation for factors affecting sampling variability.

First, students estimate the proportion of orange Reese's pieces in a random sample of 25 candies. Students then use an actual sample of Reese's pieces candy to calculate a sample proportion, and then compare results for different samples, taken by each student in the class. Next, students will use the Web Applet Reese's Pieces (at rossmanchance.com) to gather information on the sample proportions of orange candies in random samples of 25. During this stage of the activity students will compare their group and the class results to the actual parameters. In the final part of the activity, the students will use the Applet to investigate the effect of sample size on correctly estimating parameters.

CCSS

- CCSS.MATH.CONTENT.7.SP.A.2. Use data from a random sample to draw inferences about a population with an unknown characteristic of interest. Generate multiple samples (or simulated samples) of the same size to gauge the variation in estimates or predictions.

Prerequisites

Students should have an understanding of the concepts of randomization, sample, variability, population, parameter and statistic.

Learning Targets

- To understand variability between samples
- To build and describe distributions of sample proportions
- To understand the effect of sample size on how well a sample resembles a population
- To develop an understanding of a sample distribution
- To develop an understanding of a sampling distribution
- To develop an understanding of a population distribution
- To develop an understanding of the differences between sample, sampling, and population distributions

Time Required

Approximately 1 hour

Materials Required

- One Dixie cup per group (each with 25 Reeses Pieces in it)
- Reeses Pieces (enough for 25 per group)
- Reeses Pieces Web Applet - URL: ***http://www.rossmanchance.com/applets/OneProp/OneProp.htm?candy=1***

Teacher Notes

- It is helpful for the teacher to go through the entire activity using the web applet themselves, before having students do the activity.
- Be sure to let students make and discuss their conjectures before looking at real or simulated data.
- Be sure to have students turn off the animation on the web applet after they have taken a few samples.
- Be sure to ask students to compare their results and discuss the amount of sampling variability.

Sampling Reese's Pieces – Student Activity

When Reese's Pieces were first manufactured, designers knew that they wanted a peanut-flavored candy but had some problems with the filling. Original plans called for filling the candy shells with peanut butter, but the oil leaked out into the shell, leaving it soft, rather than crunchy.

The developer of the project turned the problem over to a team of outside scientists, who created a peanut-flavored filling. More experimentation was needed to determine the correct thickness of the shell. Finally, the colors of the candy coating were designed to coordinate with the color of the Reese's package. According to the Hershey Corporation who makes Reese's Pieces the goal color distribution is __% orange, __% brown, and __% yellow. Let's do some tests (no, not taste tests ☺) to see if we can determine the breakdown of percentages by color.

I. Studying the variability of orange candies

1. As we have learned, Reese's Pieces candies have three colors: orange, brown, and yellow. Which color do you think has more candies in a package: orange, brown or yellow?

2. Guess the proportion of each color in a bag/box:

   Orange____ Brown____ Yellow_____

3. If each student in the class takes a sample of 25 Reese's pieces, would you expect every student to have the same number of orange candies in their sample? Explain.

4. Pretend that 10 students each took samples of 25 Reese's pieces. Write down the number of orange candies you might expect for these 10 samples:

   ___ ___ ___ ___ ___ ___ ___ ___ ___ ___

*These numbers represent the **variability** you would expect to see in the number of orange candies in 10 **samples** of 25 pieces.*

*You will now be given a cup that holds a **random sample** of 25 Reese's pieces.*

5. Count the colors for your sample and fill in the chart below:

	Orange	Yellow	Brown
Number of candies	_____	_____	_____
Proportion of candies	_____	_____	_____

6. Write the number AND the proportion of orange candies in your sample on the board.

7. *Mark where each value should be on the two dotplots your teacher constructs (one for number of oranges, one for proportion of oranges).*

II. **Class Discussion:** The proportions of orange candies are ***sample statistics***.

a. How do you think the sample statistics will compare to the population ***parameter*** (the actual proportion of all orange Reese's Pieces produced by Hershey Co.)?

b. Do you know the value of the parameter? The values of the statistics?

c. Does the value of the ***parameter*** change each time you take a sample? Does the value of the ***statistic*** change each time you take a sample?

d. Did every group have the same proportion of orange in their samples?

e. Describe the ***distribution of sample proportions*** on the board in terms of its SOCS (shape, outlier, center, and spread).

f. Based on the distribution of sample proportions we created (on the board), what would you ESTIMATE the population parameter to be – the proportion of orange Reese's pieces actually produced by Hershey?

III. The Reese's Pieces Applet

Instead of trying this activity again with more samples of candies, we will simulate the activity using a web applet. Go to ***http://www.rossmanchance.com/applets/OneProp/OneProp.htm?candy=1***. You will see a big container of colored candies: that represents the POPULATION.

1. How many orange candies are in the population?

You will see that the proportion (probability) of orange is set to 0.50. This is the actual **population parameter** of orange candies Reese's produces.

Set the number of candies option to 25 and the number of samples to 1. Note you can choose to have your result in the number of orange candies or the proportion of orange candies. Choose "proportion". Then click "Draw Samples". Watch as the candy jar unloads 25 candies into the dish and separates the orange Reese's pieces from the brown and yellow. Also notice that a dot plot is being created with a single sample proportion plotted. This is the beginnings of a distribution of sample proportions of size 25.

2. How does the population parameter of 0.50 compare to the sample statistic of orange candies in your sample?

3. How does the population parameter of 0.50 compare to the center of the class' distribution we created on the board?

Click on the "draw samples" button again. Another sample of 25 candies will be taken and the proportion of orange candies for this sample is plotted on the graph. Repeat this again three more times.

4. Do you get the same or different values for each sample?

5. How do these numbers compare to the ones our class obtained?

6. How close is each **sample statistic** (proportion) to the **POPULATION PARAMETER**?

7. Check the "Summary Statistics" box and explain what the mean and standard deviation tells us about the distribution of sample proportions.

Turn off the animation and change the number of samples to 100.
Click on draw samples, and see the distribution of sample statistics that is built.

8. Describe its SOCS (**shape, outliers, center, and spread**.)

9. How does this compare to the distribution our class had constructed on the board?

10. When we generate sample statistics and graph them we are generating a SAMPLING DISTRIBUTION, or a distribution of the sample statistics. It looks like other distributions we have seen of raw data. What's different about a sampling distribution?

IV. Studying the Effect of Sample Size

*** Optional Extension Exercises – goes beyond 7th grade CCSS standards**

What happens to this distribution of sample statistics as we change the number of candies in each sample (sample size)? Let's find out. First, change the sample size to 10 and draw 100 samples.

1. How close is each **sample statistic** (proportion) to the **POPULATION PARAMETER**?

Next, change the sample size to 100 and draw 100 samples.

2. How close is each **sample statistic** (proportion) to the **POPULATION PARAMETER**?

3. As the sample size increases, what happens to how well the sample statistics resemble the population parameter?

4. Describe the effect of sample size on the distributions of sample statistics.

Discussion Questions/Homework:

Here are some questions that can be given to students to discuss or write answers to as homework:

1. Distinguish between how samples vary from each other, and variability of data WITHIN one sample of data.
2. Why is that taking more and more samples helps us in estimating an unknown population parameter?
3. Based on the class activity, would you be surprised to get a sample that had 5 orange candies from 25 candies drawn? Why or why not?
4. What effect does generating multiple samples (or simulated samples) of the same size on the variation in estimates or predictions?
5. Why does the size of the population (total # of Reese's Pieces produced in this case) have little influence on the behavior of statistics from random samples?

Optional Question from Extension Exercises:

6. Why is that larger samples better represent the population from which they were sampled than small samples?