Lesson 27 Part 1: Introduction

Making Statistical Inferences

CCLS 7.SP.A.2

In Lesson 26 you learned about random samples. In this lesson you will look at random samples to make predictions. Take a look at this problem.

Ms. Jennings held a "mystery bag" full of marbles up in front of her class. She said, "There are 100 marbles in this bag. Some are red, and all of the others are blue. I want you to estimate how many red marbles are in the bag without looking in the bag and counting all of them." She then let Joe pick 10 marbles from the bag without looking.

Explore It

Use the math you already know to solve the problem.

- Suppose Joe had 4 red marbles in his sample. Write a ratio representing the number of red marbles in Joe's sample.

- If Joe is not allowed to pull out any more marbles, what do you think his best estimate will be for the total number of red marbles in the mystery bag? Explain.

- Ms. Jennings had Joe put his 10 marbles back in the bag and shook it up. Then she let Angela pull 10 marbles from the bag without looking. Angela had 6 red marbles in her sample. Why do you think Angela's sample was different from Joe's?

- If Angela is not allowed to draw any more samples, and she does not know about Joe's sample, what do you think her best estimate would be for the total number of red marbles in the mystery bag? Explain.

- Ms. Jennings had Angela put her marbles back in the bag, shook the bag up again, and let Isabella reach in and draw out 10 marbles without looking. Isabella got 9 reds in her sample. Why might Isabella's sample be different from Joe's and Angela's?

Find Out More

Random samples can differ from one another due to **random variation**. The amount of variation may be small, as in the difference between Joe's and Angela's samples from the mystery bag. Occasionally, the variation is much larger, as in the difference between Joe's and Isabella's samples.

In order to get a better estimate of the number of red marbles in the mystery bag, one thing that can help is to draw more random samples. That would help you judge how typical the samples from Joe, Angela, and Isabella are. If you see a number of samples clustering around the same value, you can be more confident about using that value to estimate what you would like to know about the population.

Random sampling is helpful in situations beyond the mystery bag scenario. It can be useful, for example, when conducting surveys. Suppose, for example, you want to make a prediction about who will win a school election, but you don't have the time or resources to survey everyone in school. You could draw a random sample of names from a list of students in school and administer the survey to them. The random sample could be as large as time and resources allow.

Reflect

1 Suppose Joe, Angela, and Isabella had drawn samples of 25 marbles instead of 10. How might this change their results?

Read the problem below. Then explore different ways to represent distributions of statistics from random samples.

Ms. Jennings decided to let all 20 students in her class draw random samples from the mystery bag. Each student was allowed to draw a sample of 10 marbles, count the number of reds in the sample, and then put the 10 marbles back in the bag. Angela thinks the extra samples will help the class make a better estimate of the number of reds in the mystery bag. How should students organize their results to get a good estimate?

Model It

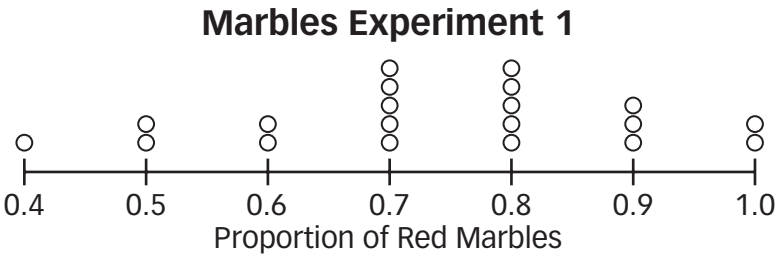
Joe suggested organizing the results in the following table.

	1	2	3	4	5	6	7	8	9	10	11	12	13	14	15	16	17	18	19	20
Number of Reds	8	10	7	10	7	5	9	8	7	9	7	6	8	8	4	7	5	9	8	6
Proportion of Reds	0.8	1.0	0.7	1.0	0.7	0.5	0.9	0.8	0.7	0.9	0.7	0.6	0.8	0.8	0.4	0.7	0.5	0.9	0.8	0.6

The table shows the proportion of reds obtained in each sample.

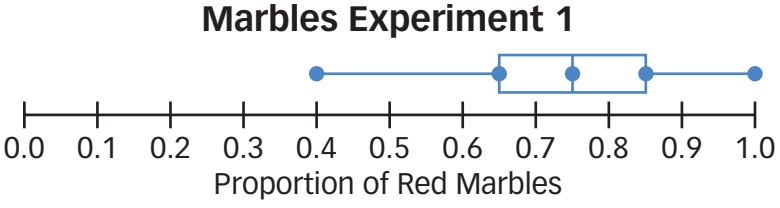
Model It

Angela thought it would be easiest to organize the data with a dot plot.



Model It

Isabella decided to construct a box plot to organize the class data.



Connect It

Now you will represent your results of taking random samples.

2 Make your own mystery bag with 100 marbles (or similar objects). Decide how many objects will be red and how many will be blue. Record the number of reds in the bag and the number of blues in the bag here. _____

3 Pretend to be each of the 20 students in Ms. Jennings' class. Repeat the experiment. Take 10 marbles out of the bag without looking. Count the number of reds and record that number in a table. Replace the marbles and repeat until you have 20 samples.

4 Use your table to make a dot plot.

5 Use your dot plot to make a box plot.

6 Imagine someone who doesn't know how many red marbles you put into the mystery bag. Which representation (table, dot plot, or box plot) would best help them estimate the number of reds in the mystery bag? Explain.

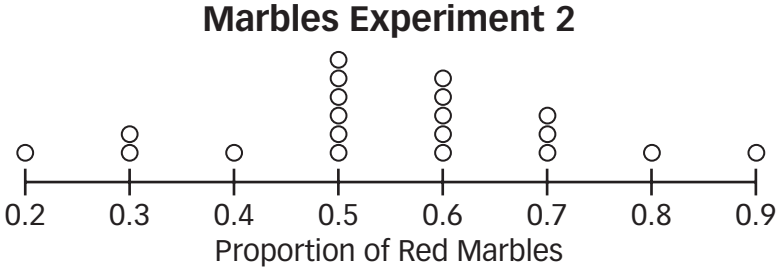
_____

_____

Try It

Use what you just learned about representing data to solve this problem.

7 Suppose one of your classmates used a different number of reds. Her dot plot is shown below. What is a good estimate for the number of reds in her bag? Explain.



_____

_____

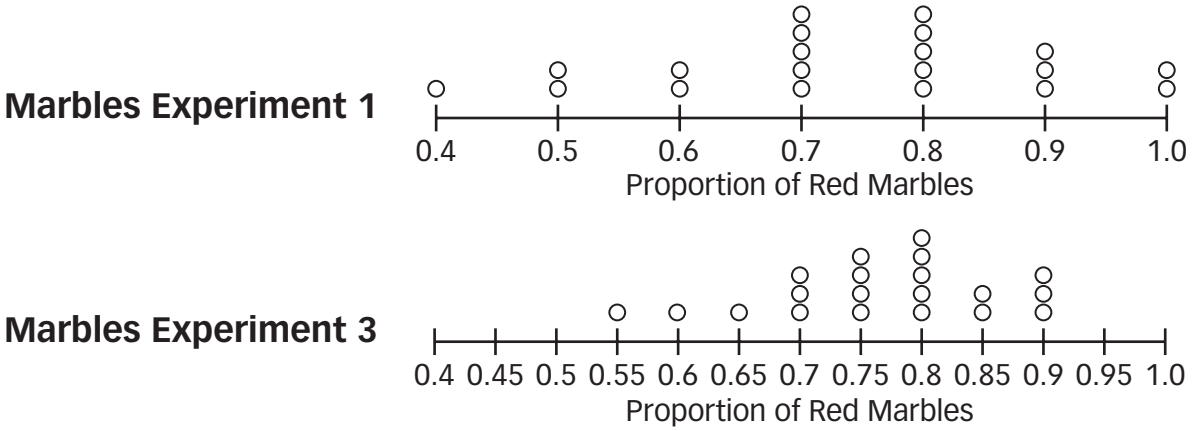
_____

Read the problem below. Then explore different ways to compare distributions of statistics from random samples when the samples are different sizes.

Angela asked Ms. Jennings if they could re-do the mystery bag activity. Angela wanted to change just one thing about it: she asked if each student could draw 20 marbles instead of just 10 marbles. Ms. Jennings agreed, and the class re-did the activity in the way Angela asked. How do the results of this new activity compare to the results of the previous one?

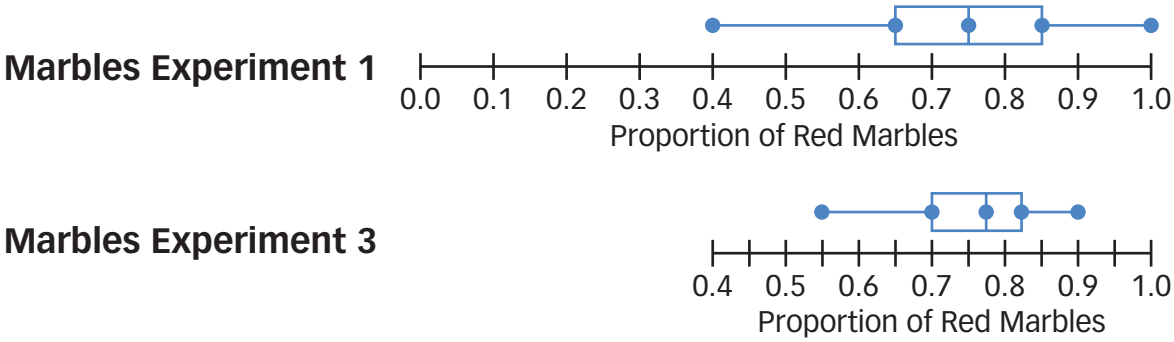
Model It

You can stack dot plots to compare the two distributions of results.



Model It

You can stack box plots to compare the two distributions of results.



Connect It

Now you will use the dot plots and box plots to compare the two data distributions.

8 Look at the models on the previous page. Compare the shapes of the two distributions of data shown in the dot plots from the two experiments.

9 Compare the centers of the two distributions to one another.

10 Compare the spreads. Which distribution is more spread out? Why?

11 If you could only take one sample from the mystery bag in order to make a prediction, would you rather draw a sample of 10 marbles or a sample of 20 marbles? Explain your choice by referring to the representations of the two distributions.

Try It

Use what you just learned about data distributions to solve this problem.

12 Perform the activity again, but this time draw 20 samples of only 5 marbles each. Record your results in a dot plot. How does the distribution compare to the dot plots you produced when you drew samples of 10?

Read the situation described below. Then solve problems 13–15.

Jonathan used the ratio of students who said they would vote for Tammy in his sample to make a prediction about the whole population.

Pair/Share

Why is it important to have random samples rather than just samples you pick out on your own students to survey?

Student Model

Tammy was running for class president. She wanted to know if she had a good chance of winning the election, so she decided to have her friends help her with surveys.

Jonathan surveyed a random sample of 10 students from the school. 7 of them said they would vote for Tammy. If there are 230 students in the school, and all of them vote, what is Jonathan's best estimate of the number of students who would vote for Tammy?

Look at how you could show your work.

$\frac{7}{10}$ of the students said they would vote for Tammy.

$\frac{7}{10} = 0.7$

$0.7 \times 230 = 161$

Solution: **161**

Sampling lets us make predictions about an entire population without surveying everyone.

Pair/Share

What are some reasons someone would have to draw a small sample rather than a large one?

13 Kimberly had more time to conduct surveys and decided to survey a random sample of 30 students from the school. In Kimberly's sample, 24 students said they would vote for Tammy. What is Kimberly's best estimate of the number of students in the school who would vote for Tammy?

Solution: _____

14 Who is likely to have a better estimate of the number of students who will vote for Tammy: Jonathan or Kimberly? Explain your choice by using what you learned in this lesson.

Whose sample size was larger?

Solution: _____

_____

_____

_____

_____

Pair/Share

How big do you think a "good" random sample has to be?

15 Which size random sample is likely to provide the most trustworthy results?

A 5

B 10

C 30

D There is no difference.

Katie chose **D** as the correct answer. How did she get that answer?

How does sample size affect results?

_____

_____

_____

_____

_____

Pair/Share

How would you help Katie understand her error?

Solve the problems.

1 Jamie wanted to estimate the mean word length in her science textbook. She did not have time to count every word and compute the average, so she took a random sample of 50 words. Which of these is the best thing to do to estimate the mean word length in the book?

A Find the average word length for the words in her sample and then multiply by 50.

B Find the average word length for the words in her sample and use it as the prediction for average word length for the entire book.

C Use a smaller sample of words to reduce the amount of spread in the distribution of averages.

D Draw a random sample of 50 words from a different science textbook and compare the new results to the original results.

2 A representative sample of 60 students from a high school is surveyed. Each student is asked which elective he or she is taking. The table shows the responses.

Elective	Number of Students
Dance	19
Music	7
Art	15
Photography	11
Electronics	8

Based on the survey results, choose True or False for each statement.

A There are many excellent dancers at the school. ☐ True ☐ False

B About 25% of the students at the high school are taking art. ☐ True ☐ False

C In a group of 30 students, it is expected that 11 of the students are taking photography. ☐ True ☐ False

D Next year, 7 out of every 60 students will be taking music. ☐ True ☐ False

E In a group of 120 students, it is expected that 16 of the students are taking electronics. ☐ True ☐ False

3 Kayla developed a study to determine the populations of fish in a lake. She took two random samples in the winter and again in the summer. She organized her data in the following table.

	Trout	Whitefish	Walleye	Total Sample
Winter	42	44	14	100
	46	42	12	100
Summer	91	84	25	200
	85	89	26	200

What valid inference can Kayla make about the entire fish population in the pond? Select all that apply.

A The total number of fish in the pond is 600.

B The walleye population comprises anywhere from 12% to 14% of the total population in both the winter and summer.

C The number of white fish in the pond is greater than the number of trout.

D The ratios of the populations of trout, whitefish, and walleye are relatively stable regardless of which season the samples were taken.

4 Suppose you flipped a fair coin 15 times. Then, 11 of your friends did the same thing. Without actually flipping a coin, make a table to show a realistic set of data for this situation. Your table should show the percentage of heads obtained by each person doing the coin flips.

Organize the data in the table you constructed using a dot plot.

Suppose you and each of your friends decided to flip the coin 50 times each instead of 15. Explain how the dot plot above would change and sketch a predicted dot plot for the situation below.

✓ Self Check *Go back and see what you can check off on the Self Check on page 247.*