



Lesson 13: Using Sample Data to Estimate a Population Characteristic

Student Outcomes

- Students differentiate between a population and a sample.
- Students differentiate between a population characteristic and a sample statistic.
- Students recognize statistical questions that are answered by estimating a population mean or a population proportion.

Lesson Notes

This lesson reviews and extends students' previous work from Grade 7 (Module 5, Lessons 18–20). The topics covered in this lesson include the distinction between a population and a sample and between population characteristics and sample statistics. Population characteristics of interest are the mean and the proportion. Because generalizing from a sample to a population requires a random sample, selecting a random sample using a random-number table (or calculator if available) is also reviewed. Students new to the curriculum may need extra support in understanding review topics and using random-number tables.

Classwork

Example 1 (5 minutes): Population and Sample

Let students answer the questions posed in the example and then share their responses with a neighbor. Then have them answer the following in writing:

- What is a population?
 - *A population is the entire set of subjects in which there is an interest.*
- What is a sample?
 - *A sample is a part of the population from which information (data) is gathered, often for the purpose of generalizing from the sample to the population.*

When students think of a population and sample, they likely think only of people. Convey to students that populations and samples are not always just composed of people. Provide other examples of populations and samples. In biology, the subjects of interest could be plants or insects. In psychology, the subjects could be rats or mice. Television sets could be the subjects in a study to determine brand quality.

Scaffolding:

- If students have trouble differentiating between a population and a sample, use a relevant example from class.
- The principal of our school is interested in determining the average number of days students are absent.
- The population being studied is all of the students enrolled at the school.
- He could use the daily attendance recorded for our math class as a sample.

Whether a set of people or objects is a population or a sample depends on the context of the situation. For instance, if the players on a specific baseball team were studied to determine, for example, the team's most valuable player for that year, then that team's players would be considered a population. There would be no need to generalize beyond that set of players. But in a study concerning the whole league, those players could be considered a sample.

Example 1: Population and Sample

Answer the following questions, and then share your responses with a neighbor.

- a. A team of scientists wants to determine the average length and weight of fish in Lake Lucerne. Name a sample that can be used to help answer their question.

Answers will vary. The scientists could spend one day catching fish and then record the length and weight of each fish caught.

- b. Golf balls from different manufacturers are tested to determine which brand travels the farthest. What is the population being studied?

The population is all golf balls made by the manufacturers.

Exercise 1 (5-7 minutes)

Let students work independently and confirm their answers with a neighbor.

Exercise 1

For each of the following, does the group described constitute a population or a sample? Or, could it be considered to be either a population or a sample? Explain your answer.

- a. The animals that live in Yellowstone National Park.

Population. The subjects are all the animals that live in Yellowstone National Park.

- b. The first-run movies released last week that were shown at the local theater complex last weekend.

The subjects are first-run movies. If all of those released last week were acquired by our local theater management and were shown last weekend, then the first-run movies released last week would be considered a population. If our local theater management chose some of the first-run movies, then they would be a sample.

- c. People who are asked how they voted in an exit poll.

If all of the people who exited the polling place were asked how they voted, then they would be considered the population of people who exited a certain polling place on a certain day. If, however, only some of them were asked how they voted, then they would be considered a sample.

- d. Some cars on the lot of the local car dealer.

Sample. For example, a customer might only be interested in two-door models.

- e. The words of the Gettysburg Address.

Population. Note that, the number of words varies depending on what version of the address is used. For example, “can not” is two words whereas “cannot” is one.

- f. The colors of pencils available in a 36-count packet of Crayola colored pencils.

Population. The 36 colors available in the packet constitute the population of colors.

- g. The students from your school who watched your school’s soccer team play yesterday.

If the entire student body attended the game, then the students who watched the game would be considered a population. If just some of the student body attended, then they would be considered a sample of students at the school.

Example 2 (5 minutes): Representative Sample

For a sample result to be generalized to the population from which the sample was taken, the sample values need to be representative of the population.

- Ask students to explain why the scientists from Example 1 have to use a sample of fish, instead of studying the whole population.
 - *It would be impossible to measure the length and weight of every fish in the lake.*

Give students a moment to think about and respond to the question posed in the example. Ask students to share their answers for a whole-class discussion.

There is no way to guarantee that a sample exactly reflects characteristics of the population. Point out that one way that is likely to result in a representative sample is to make sure that everyone in the population has the same chance of getting into the sample. Random selection accomplishes this. Students often think of random selection as haphazard selection. So emphasize that random selection is a formal procedure that must be followed. If a population is small enough, then number each individual (person or object) in the population from 1 to as many as there are, and then use random numbers from a calculator or from a random-digit table to select the individuals for the sample.

Note that if a table is used, then the number of digits to use is the number of digits in the population size, e.g., a district with 2,446 students requires each number to have four digits. The number “1” would be “0001” and the number “240” would be “0240.” The advantage of using a calculator with a random-number generator key is that leading zeros are not used. For example, to generate 35 random numbers from a population of 2,446, the typical syntax would be random (1, 2446, 35).

If students have access to graphing calculators, demonstrate how random numbers can be generated using technology.

Calculator note: To generate random integers on a TI-84 calculator, go to the MATH menu and highlight PRB. Choose option 5: randInt. For example, to generate 30 random integers from a population of 314, use the syntax `randInt(1, 314, 30)`.

Example 2: Representative Sample

If a sample is taken for the purpose of generalizing to a population, the sample must be representative of the population. In other words, it must be similar to the population even though it is smaller than the population. For example, suppose you are the campaign manager for your friend who is running for Senior Class President. You would like to know what proportion of students would vote for her if the election were held today. The class is too big to ask everyone (314 students). What would you do?

Comment on whether or not each of the following sampling procedures should be used. Explain why or why not.

- a. Poll everyone in your friend's math class.

This may not be the best option because everyone in the class may be biased towards your friend. For example, your friend may be great at math and help a lot of students prepare for tests. Everyone may look at her favorably.

- b. Assign every student in the senior class a number from 1 to 314. Then use a random number generator to select 30 students to poll.

This is a good option because it ensures that the sample looks like the rest of the senior class.

- c. Ask each student that is going through the lunch line in the cafeteria who they will vote for.

Answers may vary. This may or may not be a good option. The students going through the lunch line might be freshman, sophomores, or juniors who do not vote for senior class president, so their opinions do not matter. However, if only seniors in the lunch line are asked how they will vote, this could be a good sample of the senior class.

Exercise 2 (3-4 minutes)**MP.5**

This exercise reviews use of a random-number table that students encountered in Grade 7. If students have access to graphing calculators or appropriate software, random numbers can be generated using technology. If time allows, students should explain both procedures. If students do not have access to a table of random numbers, display the following line of random numbers on the front board of the classroom for students to use.

54580 81507 27102 54537 55894 33006 04229 91828

Let students work in small groups to complete the exercise. Assist students who are new to the curriculum and have not used this type of table before.

Exercise 2

There is no procedure that guarantees a representative sample. But the best procedure to obtain a representative sample is one that gives every different possible sample an equal chance to be chosen. The sample resulting from such a procedure is called a random sample.

Suppose that you want to randomly select 60 employees from a group of 625 employees.

Explain how to use a random-number table or a calculator with a random-number generator to choose 60 different numbers at random and include the students with these numbers in the sample.

Random-number generator: Assign each employee a number from 1 to 625. Then generate 60 random numbers from the random-number generator. If there are any duplicate numbers, generate additional numbers as needed.

Random-number table: To do so, you could number all the employees from 001 to 625. Then you could use a random number table. Any three-digit number beyond 625 is ignored, and any duplicated three-digit number is ignored. The starting point on a random-number table does not matter. For the purposes of this example, start at the far left. For example, from the following line of random digits, the first four employees randomly selected would be those with identification numbers: 545, 150, 102, and 375.

Example 3 (3-5 minutes): Population Characteristics and Sample Statistics**MP.1**

A statistical study begins with asking a question that can be answered by data. Students should be familiar with this since it was the first Common Core Statistics Standard in Grade 6.

Let students work independently on the example and then share their responses with a neighbor. Then ask students to summarize the following summary measures in their own words:

- What is a population characteristic?
 - *A summary measure calculated using all the individuals in a population is called a population characteristic. A population proportion and a population mean are two examples of population characteristics.*
- What is a sample statistic?
 - *If the summary measure is calculated using data from a random sample, it is called a sample statistic. For example, a sample proportion or a sample mean are sample statistics.*

A summary measure is used to analyze the data and arrive at a conclusion. Population summary measures are called population characteristics (or parameters), and sample summary measures are called statistics. Examples include a total (e.g., gross movie receipts), a mean (e.g., average time to run a mile), a proportion (e.g., the proportion of voters who favor a particular candidate), a correlation coefficient (e.g., the correlation coefficient for the amount of vitamin C taken and the number of colds contracted).

Example 3: Population Characteristics and Sample Statistics

A statistical study begins with a question of interest that can be answered by data. Depending on the study, data could be collected from all individuals in the population or from a random sample of individuals selected from the population. Read through the following and identify which of the summary measures represents a *population characteristic* and which represents a *sample statistic*. Explain your reasoning for each.

Suppose the population of interest is the words of the Gettysburg Address. There are 269 of them (depending on the version).

- a. The proportion of nouns in all words of the Gettysburg Address.
It is a population characteristic because it is determined from all of the words (the entire population) of the Gettysburg Address.
- b. The proportion of nouns or the proportion of words containing the letter “e” in a random sample of words taken from the Gettysburg Address.
It is a sample statistic because it is determined from a sample of the words (the entire population) of the Gettysburg Address.
- c. The mean length of the words in a random sample of words taken from the Gettysburg Address.
It is a sample statistic because it is determined from a sample of the words (the entire population) of the Gettysburg Address.

- d. The proportion of all words in the Gettysburg Address that contain the letter “e.”

It is a population characteristic because it is determined from all of the words (the entire population) of the Gettysburg Address.

- e. The mean length of all words in the Gettysburg Address.

It is a population characteristic because it is determined from all of the words (the entire population) of the Gettysburg Address.

Exercise 3 (7–9 minutes)

Let students work in small groups. If time allows, let each group present one of their answers.

Exercise 3

For the following items of interest, describe an appropriate population, population characteristic, sample, and sample statistic. Explain your answer.

Answers will vary. The following are examples of sample responses.

- a. Time it takes students to run a quarter-mile.

Population: All students in the eleventh grade at your school.

Population characteristic: Mean time to run a quarter-mile for all students in the eleventh grade at your school.

Sample: A random selection of students in the eleventh grade at your school.

Sample statistic: Mean time to run a quarter-mile for the sampled students in the eleventh grade at your school.

- b. National forests that contain bald eagle nests.

Population: All national forests in the United States.

Population characteristic: The proportion of all national forests in the U.S. that contain bald eagle nests.

Sample: A random selection of national forests in the U.S..

Sample statistic: The proportion of the sampled national forests in the U.S. that contain bald eagle nests.

- c. Curfew time of boys compared to girls.

Population: All boys and girls in the eleventh grade at your school.

Population characteristic: The difference in mean curfew time of all eleventh grade boys at your school, and the mean curfew time of all eleventh grade girls at your school.

Sample: A random selection of some boys and girls sampled from the eleventh grade at your school.

Sample statistic: The difference in mean curfew time of the sampled eleventh grade boys at your school and the mean curfew time of the sampled eleventh grade girls in at your school.

Scaffolding:

If students struggle with the exercise, provide the following on the front board as a sample response to use as a guide.

For example, an item of interest is SAT scores.

- The population might be all students in your school who take the SAT exam in mathematics this year.
- One population characteristic is the proportion of all students who take the mathematics exam this year who score 700 or higher, and another is the mean score of all students who take the mathematics exam this year.
- A sample of 40 students might be selected from those who take the SAT exam this year.
- Two examples of sample statistics are the proportion of the 40 students who score 700 or higher and the mean score for the 40 students who take the SAT exam in mathematics this year.

- d. Efficiency of electric cars.

Population: All electric cars currently being marketed.

Population characteristic: The mean number of kilowatt-hours to drive 100 miles for all electric cars currently being marketed.

Sample: A random selection of electric cars currently being marketed.

Sample statistic: The mean number of kilowatt-hours to drive 100 miles for the sampled electric cars currently being marketed.

Exercise 4 (5–7 minutes)

Let students continue to work in small groups. Several statistical questions about the population of students at your school or all students in a particular grade (e.g., the eleventh grade) are suggested. Discuss the questions with the whole class. You might consider asking students to add questions to this list. Ask each group to select a question and write a paragraph as directed in the exercise. After 3 to 5 minutes, allow as many groups as possible to discuss their paragraphs. Focus on each group's responses that indicate how the sample will be obtained and how the sample mean or proportion is used to generalize to the students in the population (e.g., the whole school or the students in the eleventh grade).

Exercise 4

Consider the following questions:

- What proportion of eleventh graders at our high school are taking at least one advanced placement course?
- What proportion of eleventh graders at our high school have a part-time job?
- What is the typical number of hours an eleventh grader at our high school studies outside of school hours on a weekday (Monday, Tuesday, Wednesday, or Thursday)?
- What is the typical time (in minutes) that students at our high school spend getting to school?
- What is the proportion of students at our high school who plan to attend a college or technical school after graduation?
- What is the typical amount of time (in hours per week) that students at our high school are involved in community service?

Select one of these questions (or a different statistical question that has been approved by your teacher). Working with your group, write a paragraph that:

- states the statistical question of interest pertaining to the students in the population for the statistical question selected.
- identifies a population characteristic of interest.
- identifies the appropriate statistic based on a sample of 40 students.
- states what property your sample must have for you to be able to use its results to generalize to all students in your high school.
- includes the details on how you would select your sample.

Possible responses for two of the above questions:

We are interested in answering the statistical question: “What proportion of eleventh graders at our high school are taking at least one advanced placement course?” We would go to our high school counselors and obtain a list of the 403 eleventh graders at our high school. We would number the students from 001 to 403. We decided to randomly select 40 students for our sample. To do so, we would find a random number table, choose a place to start, and then generate 40 different random numbers between 001 and 403. If we find that 14 of the 40 students in the sample are taking at least one AP course this year, we would use that information to calculate the sample proportion of students who are taking at least one AP course. Since our sample was a random sample, we can generalize our estimate to the population of eleventh graders at our school and estimate that the proportion of all eleventh graders who are taking at least one AP course this year is 0.35.

Also consider the following paragraph:

We are interested in the question about community service. The statistical question would be “What is the typical amount of time (in hours per week) that students at our school are involved in community service?” We would go to our school’s guidance department and ask them to help us. The guidance counselor has an alphabetical list of the 1,245 students enrolled in our school. Listed students are numbered from 0001 to 1245. We decided we would select 40 students at random for our study. Using our graphing calculator, we would obtain 40 different random numbers between 1 and 1245. We would ask the guidance counselor to identify the student associated with each of the 40 numbers. We would also ask for the location of the student during the last period of the day, so we could quickly ask them a question before leaving school. We would ask “How many hours of community service did you complete last week?” We would find the mean of the 40 numbers reported by the students in our sample. The mean is our statistic for this study. If our sample was randomly selected, we can generalize to the population of all students at the school, and we would use the sample mean to estimate the mean number of hours of community service per week for students at our school. For example, if the mean of the 40 randomly selected students is 14.5 hours, we would say that our best estimate of the mean number of hours students in our school are involved in community service is 14.5 hours per week.

Closing (2 minutes)

MP.3

- Present students with a scenario, and ask them to critique whether or not the sample chosen will be representative of the population and provide a good estimate for a population characteristic. Allow for multiple responses and arguments. For example, suppose that an elementary school principal is interested in how many hours of sleep students get each night. He decides to use Mr. Ross’s fifth-grade class as a sample. Do you think that this sample will provide a good estimate of the population characteristic (average number of hours of sleep for all students in the school)? Explain.
 - *Answers will vary. This may not be a representative sample. Older students may require less sleep than younger students. The estimate may be too low and not represent the sleep times of students in the lower grades.*
- Ask students to summarize the main ideas of the lesson in writing or with a neighbor. Use this as an opportunity to informally assess comprehension of the lesson. The Lesson Summary below offers some important ideas that should be included.

Lesson Summary

We refer to summary measures calculated using data from an entire population as population characteristics. We refer to summary measures calculated using data from a sample as sample statistics. To generalize from a sample to the corresponding population, it is important that the sample be a random sample from the population. A random sample is one that is selected in a way that gives every different possible sample an equal chance of being chosen.

Exit Ticket (5 minutes)

Name _____

Date _____

Lesson 13: Using Sample Data to Estimate a Population Characteristic

Exit Ticket

Indicate whether each of the following is a summary measure from a population or from a sample. Choose the one that is more realistically the case. If it is from a population, identify the population characteristic. If it is from a sample, identify the sample statistic. Explain your reasoning.

- a. 88% of the more than 300-million automobile tires discarded per year are recycled or used for fuel.
- b. The mean number of words that contain the letter “e” in the Gettysburg Address.
- c. 64% of respondents in a recent poll indicated that they favored building a proposed highway in their town.

Exit Ticket Sample Solutions

Indicate whether each of the following is a summary measure from a population or from a sample. Choose the one that is more realistically the case. If it is from a population, identify the population characteristic. If it is from a sample, identify the sample statistic. Explain your reasoning.

- a. 88% of the more than 300-million automobile discarded tires per year are recycled or used for fuel.

Population. The statement concerns all discarded tires in a year.

Characteristic: % of all discarded tires that are recycled or used for fuel.

- b. The mean number of words that contain the letter “e” in the Gettysburg Address.

Population. The statement refers to the entire document.

Characteristic: mean number of words that contain the letter “e.”

- c. 64% of respondents in a recent poll indicated that they favored building a proposed highway in their town.

Sample. Not everyone in the town contributed.

Characteristic: % of those who favor the highway construction.

Problem Set Sample Solutions

Use this space to describe specific details about the problem set for teacher reference.

1. In the following, identify whether the subjects being measured are the sample or the population. In some cases, they could be considered a sample or a population. Explain each answer.

Subjects	What is being measured	Sample or Population? Explain
Some students in your class	Number of books in backpack	<i>Sample</i> <i>Only some of the students in the class were in the study.</i>
AA batteries of a certain brand	Lifetime	<i>Sample</i> <i>Realistically a population doesn't make sense since if all the batteries are tested, there would be none left.</i>
Birds in Glacier National Park	Number of species	<i>Population</i> <i>It appears that all of the birds in the park are to be catalogued.</i>
Students in your school	Number absent or present today	<i>Population</i> <i>All students in your school are being recorded as being present or being absent today.</i>
Words in the Constitution of the U.S.	Whether a noun or not	<i>Population</i> <i>All the words in the Constitution are being considered.</i>
Americans of voting age	Opinion on an issue	<i>Sample</i> <i>It is unrealistic to think that all Americans of voting age could be asked their opinion.</i>

2. For the following items of interest, describe an appropriate population, a population characteristic, sample, and sample statistic.

- a. Whether or not a driver is speeding in your school zone during school hours in a day.

Answers will vary. A sample response is given.

Population: all drivers who drive through your school zone during school hours on a certain day.

Population characteristic: proportion of all drivers who exceed the speed limit in your school zone during school hours on a certain day.

Sample: some drivers who drive in your school zone from noon to 1:00 p.m. on a certain day.

Sample statistic: proportion of the sampled drivers who exceed the speed limit in your school zone from noon to 1:00 p.m. on a certain day.

- b. Seatbelt usage of men compared to women.

Answers will vary. A sample response is given.

Population: all men and women drivers in the state of New York.

Population characteristic: the difference in the proportion of all male drivers in the state of New York who always wear a seatbelt and the proportion of all female drivers in the state of New York who always wear a seatbelt.

Sample: some men and women drivers in the state of New York.

Sample statistic: the difference in the proportion of the male drivers sampled who always wear a seatbelt and the proportion of the female drivers sampled who always wear a seatbelt.

- c. Impact of a new antidepressant on people with severe headaches.

Answers will vary. A sample response is given.

Population: all people in New York with severe headaches.

Population characteristic: the mean time it takes the new antidepressant to eliminate the headaches for all people in New York with severe headaches.

Sample: some people in New York with severe headaches.

Sample statistic: the mean time it takes the new antidepressant to eliminate the headaches for the sampled people in New York with severe headaches.

3. What are the identification numbers for ten students chosen at random from a population of 78 students based on the following string of random digits? Start at the left.

27816 78416 01822 73521 37741 016312 68000 53645 56644 97892 63408 77919 44575

27 67 16 01 27 35 21 37 74 10

Table of Random Digits

Row																				
1	6	6	7	2	8	0	0	8	4	0	0	4	6	0	3	2	2	4	6	8
2	8	0	3	1	1	1	1	2	7	0	1	9	1	2	7	1	3	3	5	3
3	5	3	5	7	3	6	3	1	7	2	5	5	1	4	7	1	6	5	6	5
4	9	1	1	9	2	8	3	0	3	6	7	7	4	7	5	9	8	1	8	3
5	9	0	2	9	9	7	4	6	3	6	6	3	7	4	2	7	0	0	1	9
6	8	1	4	6	4	6	8	2	8	9	5	5	2	9	6	2	5	3	0	3
7	4	1	1	9	7	0	7	2	9	0	9	7	0	4	6	2	3	1	0	9
8	9	9	2	7	1	3	2	9	0	3	9	0	7	5	6	7	1	7	8	7
9	3	4	2	2	9	1	9	0	7	8	1	6	2	5	3	9	0	9	1	0
10	2	7	3	9	5	9	9	3	2	9	3	9	1	9	0	5	5	1	4	2
11	0	2	5	4	0	8	1	7	0	7	1	3	0	4	3	0	6	4	4	4
12	8	6	0	5	4	8	8	2	7	7	0	1	0	1	7	1	3	5	3	4
13	4	2	6	4	5	2	4	2	6	1	7	5	6	6	4	0	8	4	1	2
14	4	4	9	8	7	3	4	3	8	2	9	1	5	3	5	9	8	9	2	9
15	6	4	8	0	0	0	4	2	3	8	1	8	4	0	9	5	0	9	0	4
16	3	2	3	8	4	8	8	6	2	9	1	0	1	9	9	3	0	7	3	5
17	6	6	7	2	8	0	0	8	4	0	0	4	6	0	3	2	2	4	6	8
18	8	0	3	1	1	1	1	2	7	0	1	9	1	2	7	1	3	3	5	3
19	5	3	5	7	3	6	3	1	7	2	5	5	1	4	7	1	6	5	6	5
20	9	1	1	9	2	8	3	0	3	6	7	7	4	7	5	9	8	1	8	3
21	9	0	2	9	9	7	4	6	3	6	6	3	7	4	2	7	0	0	1	9
22	8	1	4	6	4	6	8	2	8	9	5	5	2	9	6	2	5	3	0	3
23	4	1	1	9	7	0	7	2	9	0	9	7	0	4	6	2	3	1	0	9
24	9	9	2	7	1	3	2	9	0	3	9	0	7	5	6	7	1	7	8	7
25	3	4	2	2	9	1	9	0	7	8	1	6	2	5	3	9	0	9	1	0
26	2	7	3	9	5	9	9	3	2	9	3	9	1	9	0	5	5	1	4	2
27	0	2	5	4	0	8	1	7	0	7	1	3	0	4	3	0	6	4	4	4
28	8	6	0	5	4	8	8	2	7	7	0	1	0	1	7	1	3	5	3	4
29	4	2	6	4	5	2	4	2	6	1	7	5	6	6	4	0	8	4	1	2
30	4	4	9	8	7	3	4	3	8	2	9	1	5	3	5	9	8	9	2	9
31	6	4	8	0	0	0	4	2	3	8	1	8	4	0	9	5	0	9	0	4
32	3	2	3	8	4	8	8	6	2	9	1	0	1	9	9	3	0	7	3	5
33	6	6	7	2	8	0	0	8	4	0	0	4	6	0	3	2	2	4	6	8
34	8	0	3	1	1	1	1	2	7	0	1	9	1	2	7	1	3	3	5	3
35	5	3	5	7	3	6	3	1	7	2	5	5	1	4	7	1	6	5	6	5
36	9	1	1	9	2	8	3	0	3	6	7	7	4	7	5	9	8	1	8	3
37	9	0	2	9	9	7	4	6	3	6	6	3	7	4	2	7	0	0	1	9
38	8	1	4	6	4	6	8	2	8	9	5	5	2	9	6	2	5	3	0	3
39	4	1	1	9	7	0	7	2	9	0	9	7	0	4	6	2	3	1	0	9
40	9	9	2	7	1	3	2	9	0	3	9	0	7	5	6	7	1	7	8	7