



SAMPLING VS. CENSUS

Lexile Measure: 1210L

FIELDS OF STUDY

Sampling Design; Sampling Techniques; Statistical Analysis

ABSTRACT

Censuses and sample surveys are two common techniques used to gather data. Conducting a census requires data to be collected from every member of the population being surveyed. Sample surveying allows data to be collected from a subset of the population.

PRINCIPAL TERMS

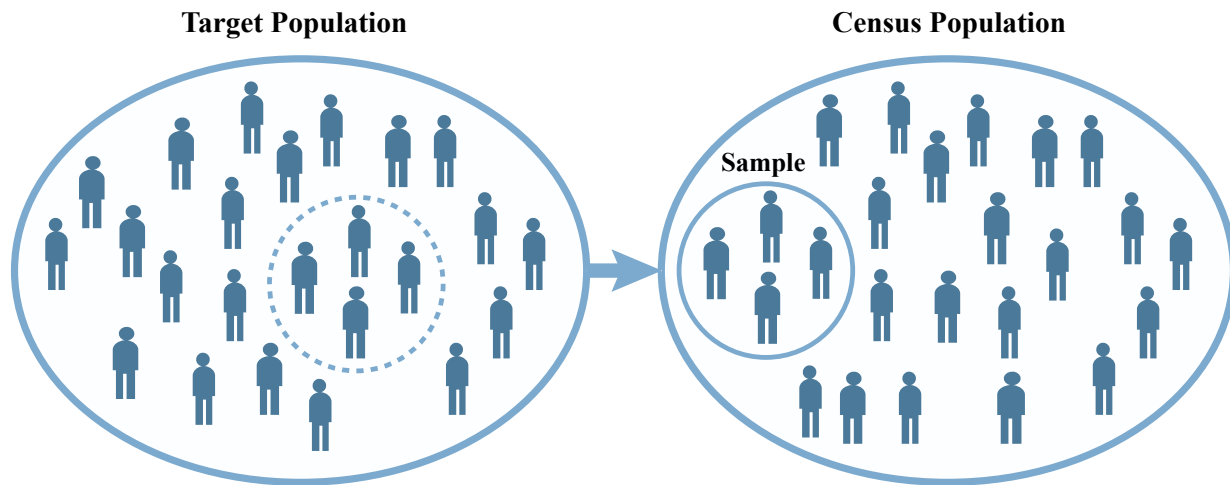
- **convenience sampling:** a non-probabilistic sampling method in which members of the sample group are drawn from a population based on their availability and ease of access; also called accidental sampling or haphazard sampling.
- **simple random sampling:** a probabilistic sampling method in which members of the sample group are drawn from the population at random, so that each member of the population has an equal chance of being included.
- **voluntary response sampling:** a non-probabilistic sampling method in which members of the population of interest are made aware of a survey or study and can choose whether or not to participate.

Overview of Censuses and Sampling

A population is a set of elements that have at least one common trait and are studied for statistical purposes. A census compiles target statistics from each individual in the population. A sample only collects data from a subset of the population.

Since ancient times, censuses have been used by leaders to understand important statistics about their people. The first census in the United States was conducted in 1790 and resulted in a population count of 3,929,214 people. Since then, the US government has conducted a census every ten years to get an official count of the number of people in the country, as well as descriptive information about them. The government also collects more detailed data on a sample of residents. It began conducting the Current Population Survey in 1940 and the American Community Survey in 2005.

There are two basic ways to collect data about a population. One way is to compile target statistics from every single member of the population. This process is known as a census. A census is defined by its comprehensiveness. The other way is to collect data from a subset of the population, known as a sample. Findings can then be extrapolated to the larger population. A sample survey is more practical than a census, because it does not require data from every member in the population. However, it has a greater margin of error.



When data is collected from the entire target population, it is a census. When data is collected only from a subset of the target population, it is a sample.

There are many ways to select a sample from a population. Sampling methods can be broadly characterized as either probabilistic or non-probabilistic. A probabilistic sampling method is one that involves an element of random chance. Every member of the population has a probability of being selected, and that probability can be calculated for any given individual. A non-probabilistic sampling method is one in which either some members of the population have no probability of being selected, or the probability cannot be calculated for every individual.

One of the easiest sampling methods is **convenience sampling**. Convenience sampling takes little effort from the sampler, because it involves collecting data from whoever happens to be nearby or otherwise readily available at the time of sampling. It is a non-probabilistic method because it is impossible to calculate the odds of any given member of the population being selected. Another method is **simple random sampling**. This is the most basic form of probabilistic sampling. In this method, a subset of individuals is chosen at random from the entire population in such a way that every individual has an equal chance of being selected. A simple random sample is much more likely than a convenience sample to be representative of the population as a whole.

Advantages and Disadvantages

Censuses and sample surveys each have advantages and disadvantages. Because a census collects data from every individual in a population, it provides greater accuracy than a sample survey when done correctly. Unfortunately, it is often impossible to collect data from every member of a population because of constraining factors such as population size, time, and cost.

When a census is not possible, researchers use sampling instead. If the sample is chosen in such a way that it is demographically representative of the larger population, researchers can use the data gathered to make estimates about the whole population with a fair degree of accuracy. However, sampling is less effective when studying smaller populations, because larger samples are required to obtain statistically meaningful results. For

example, if a researcher were studying a population of one hundred people, they would need a sample size of eighty in order to achieve a standard 95 percent confidence level and 5 percent margin of error.

Another challenge of sampling is avoiding sampling bias. Sampling bias occurs when not all members of the population have an equal chance of being chosen for the sample, and so it is not truly representative. A common type of sampling bias is self-selection bias. This arises when individuals can choose whether or not to be part of the sample, as in **voluntary response sampling**. Self-selection bias can lead to inaccurate results in any type of study, but it is particularly problematic in opinion polls and surveys, because those who choose to participate tend to be the ones who feel most strongly about a topic, and those with less firmly held viewpoints are underrepresented.

Examples of Census and Sample Surveying

If a basketball coach wanted to know the average shoe size of the team, it would be best to take a census. Typically, a basketball team has around twelve players at any given time. It would take minimal time and effort to collect information about the shoe sizes of all twelve players. A sample would not be the best choice, because in such a small population, the margin of error would be so great that the results would be all but meaningless.

Alternatively, for a scientist researching how a specific bug spray affects mosquitoes, a sample would be more appropriate. The scientist could gather a group of mosquitoes to observe how the spray affects each of them. A census would be impossible, as the scientist would have to observe how all mosquitoes in the world are affected by the spray.

SAMPLE PROBLEM

A statistics teacher wants to compare the average SAT score of the students in their class to the national average SAT score. There are thirty students in the statistics class. Should the teacher conduct a census or a sample survey of the students in order to calculate their average SAT score? Assume that the national average can be looked up online and that all of the students in the class have already taken the SAT.

Answer:

The teacher has direct access to the students in the class, and the number of students is not prohibitively large, so it would not be very difficult to gather information from every student. In addition, because the population is small—only thirty individuals—any sample would have a large margin of error. Thus, a census is the best choice.

Real-World Application

Censuses and sample surveys are utilized by a variety of professionals. For example, college professors often want feedback from students about a particular course. They may ask a small sample of students for their opinions, or they may have the entire class submit feedback. Once the professor has performed either the sample survey or the census, they can then use the feedback to improve the course.

Knowing when to use a census and when to use a sample survey is important. Both techniques have distinct advantages and disadvantages. Using the appropriate technique enables researchers to make the best possible predictions about the populations they are studying.

— Brandon Chupp, Austin Huff, and Daniel Showalter, PhD

Bibliography

- Bennett, Jeffrey O., et al. *Statistical Reasoning for Everyday Life*. 4th ed., Pearson Education, 2014.
- Bryan, Thomas. "Basic Sources of Statistics." *The Methods and Materials of Demography*, edited by Jacob S. Siegel and David A. Swanson, 2nd ed., Elsevier Academic Press, 2004, pp. 9–41.
- Durrett, Rick. *Elementary Probability for Applications*. Cambridge UP, 2009.
- Olofsson, Peter. *Probabilities: The Little Numbers That Rule Our Lives*. 2nd ed., John Wiley & Sons, 2015.
- "Statistical Language: Census and Sample." *Australian Bureau of Statistics*, Commonwealth of Australia, 3 July 2013, www.abs.gov.au/websitedbs/a3121120.nsf/home/statistical+language+-+census+and+sample. Accessed 11 Apr. 2017.
- Triola, Mario F. *Elementary Statistics*. 12th ed., Pearson Education, 2014.