1.1 Sampling

Table of Contents

Definitions

A population is the entire collection of outcomes, responses, measurements, or counts that are of interest.

A sample is a subset of a population.

Why do we need a sample?

It is impossible to examine every member of a huge population. Thus, we select a group of a manageable size to examine. This group is a sample.

Sampling is a process taking a sample from a larger population.

The objective of sampling is to make a sample representative of the population.

Question 1. Identify populations and samples

Answers

1. A survey of 12,082 U.S. adults found that 45.5% received an influenza vaccine for a recent flu season.
(Source: U.S. Centers for Disease Control and Prevention)

Population: Sample:

2. A survey of 202 pilots found that 20% admit that they have made a serious error due to sleepiness.
(Source: National Sleep Foundation)

Population: Sample:

Table of Contents

Six Types of Sampling

1. Simple Random Sampling
2. Samples of Convenience
3. Stratified Sampling
4. Cluster Sampling
5. Systemic Sampling
6. Voluntary Response Sampling

1. Simple Random Sampling

A simple random sample is a sample in which every possible sample of the same size has the same chance of being selected, just as in a lottery.

(image source: vectorstock.com)

Examples.

1. A student artist puts the names of all her pieces in a hat and draws out 8 names for her gallery showing.
2. An electronics store pulls all receipts of customers who purchased a computer over the past two years and uses a random number generator to select 100 of them to poll about high-speed Internet rates.
3. A manager at a family amusement park uses a random number generator to draw a sample based on the admission ticket number to ask for feedback on the new rides.

2. Samples of Convenience

A sample of convenience is a sample which consists only of members of the population that are easy to get.

Table of Contents

- This type of sample often leads to biased studies. Don't use a sample of convenience when it is possible to draw a simple random sample.
- A sample of convenience may be acceptable when it is reasonable to believe that there is no systematic difference between the sample and the population.

Examples.

1. Jennifer is a professor at the university of Wonderland. She is doing a study to determine the opinions of students at the school regarding stem cell research. She selects students who are in her biology class to survey.

2. A sports columnist asks people in a bar about the potential basketball lockout.

3. Stratified Samples

When it is important for the sample to have members from each segment of the population, you should use a stratified sample.

In using a stratified sample,

1) Members of the population are divided into groups, called *strata*, that share a similar characteristic such as age, gender, ethnicity, or even political preference.

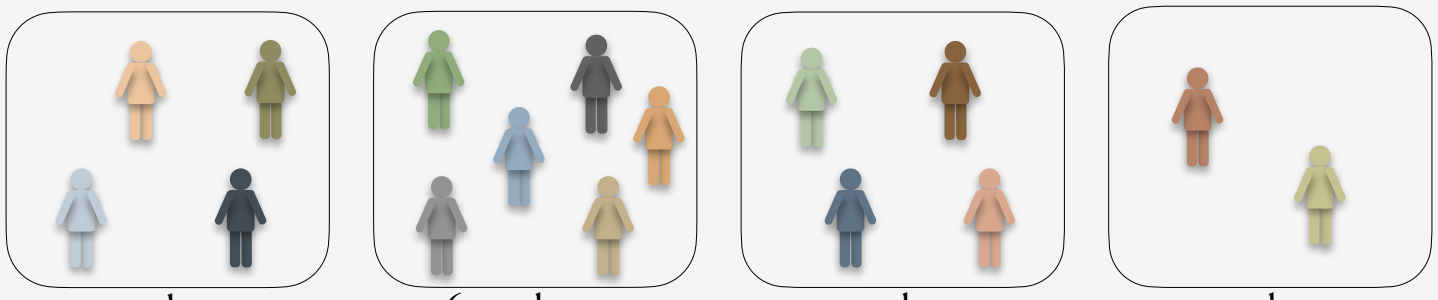
2) A simple random sample of members is selected from each of the strata.

3) Care must be taken to ensure that all strata are sampled in proportion to their actual percentages of occurrence in the population.

Examples.

1. There are 16 employees in Wonderland Cooperation. Let us collect a stratified sample of 8.

Step 1. Divide the employees into marital status as follows.



4 employees
Married

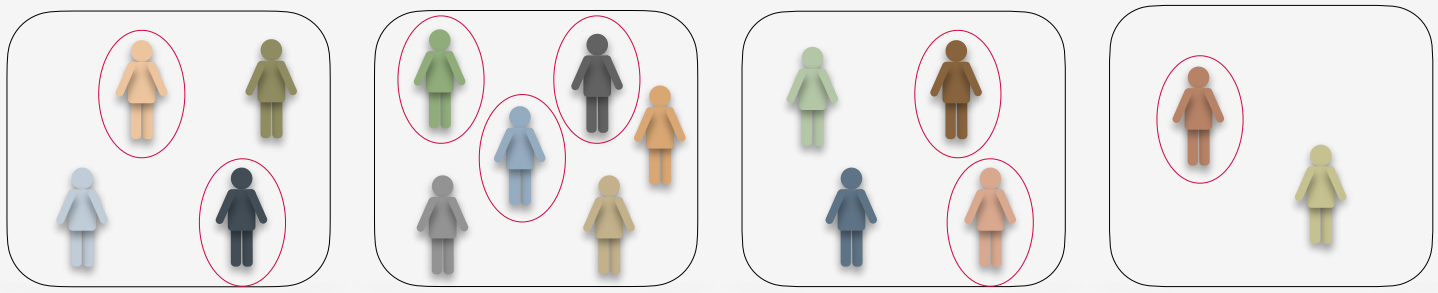
6 employees
Single

4 employees
Divorced

2 employees
Widowed

Table of Contents

Step 2. select simple random samples of of 2 employees from Married, of 3 employees from Single, of 2 employees from Divorced, and of 1 employees from Widowed.



The number of individuals chosen from each stratum is determined by the proportion or ratio of that stratum's representation in the overall population.

2. A researcher asked a simple random sample of 20 home-schooled children, a simple random sample of 20 children who attend private school, and a simple random sample of 20 children who attend public school their opinion on the new town curfew.

3. A cell phone company wants to draw a sample of 600 customers to gather opinions about potential new features on upcoming phone models. The company draws a random sample of 200 from customers with iPhones, a random sample of 100 from customers with LG phones, a random sample of 100 from customers with Samsung phones, and a random sample of 200 from customers with other phones.

4. Cluster Samples

Cluster sampling is useful when the population is too large and spread out for simple random sampling to be feasible.

In using a cluster sample,

1) Members of the population are divided into groups, called *clusters*, whose examples are different sections of the same course or different branches of a bank.

2) A simple random sample of clusters is selected from the population.

3) Every member of the selected clusters is sampled.

Table of Contents

Examples

1. There are 20 households in Wonderland Village. The population of Wonderland Village is 59.

Consider each household as a cluster.

Select a simple random sample of 8 clusters as above.

Every member of the chosen 8 clusters is sampled.

2. A football coach randomly selected the varsity or junior varsity team and then asks all players from that team their opinion on a new logo.

5. Systemic Sample

A systemic sample is a sample in which

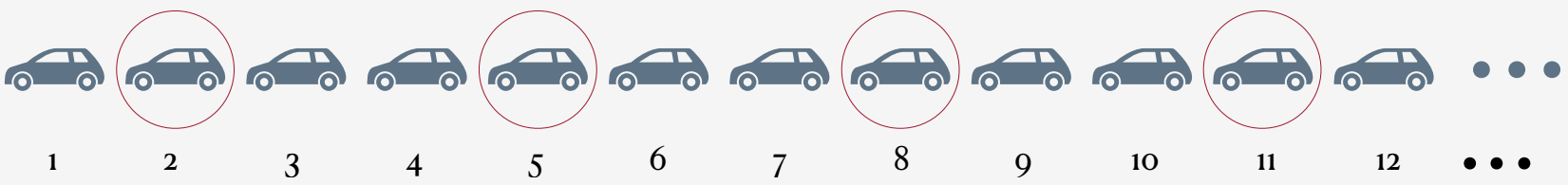
1) members of the population are ordered in some way,

2) a starting number is randomly selected,

3) sample members are selected at regular intervals from the starting number. For instance, every 3rd, 5th, or 100th member is selected.

Example.

Automobiles are coming off an assembly line. It is decided to draw a systematic sample for a detailed check of the steering system. The starting point will be the second car, then every third car after that will be sampled. Which cars will be sampled?



The sample consists of the cars numbered 2, 5, 8, 11, and so on.

Table of Contents

6. Voluntary Response Sample

A voluntary response sample is a sample made up of volunteers.

- Voluntary response samples are often used by the media to try to engage the audience. For instance, a radio talk show host invites listeners to send an email to express their opinions on an upcoming election.

- Voluntary response samples are never reliable. For example, people who send an email to the radio talk show may have strong opinions about the topic in either direction.

- Advantage: a voluntary response sample is an inexpensive way to conduct a study as data is very easy to gather.

Examples.

1. A salary committee for a Midwestern college sends an email to all employees asking them to fill out an optional online survey about the proposed benefit package.
2. A smartphone app produces a message requesting customers to click on a link to rate the app.

Definitions

A parameter is a numerical description of a population characteristic.

A statistic is a numerical description of a sample characteristic.

Statistics is the science of collecting, organizing, analyzing, and interpreting data in order to make decisions.

Question 2. Determine whether the number described is a statistic and a parameter. **Answers**

1. The average age of the employees in a certain company is 35 years.

2. In a survey of new automobile owners, 43% of the owners plan to buy another new vehicle within four years.

Table of Contents

Question 3. Identifying Sampling Techniques.

Answers

1. Using random digit dialing, researchers call 1400 people and ask what obstacles (such as childcare) keep them from exercising.

2. Chosen at random, 500 rural and 500 urban people age 65 or older are asked about their health and their experience with prescription drugs.

3. Questioning students as they leave a university library, a researcher asks 358 students about their drinking habits.

4. After a hurricane, a disaster area is divided into 200 equal grids. Thirty of the grids are selected, and every occupied household in the grid is interviewed to help focus relief efforts on what residents require the most.

5. Chosen at random, 580 customers at a car dealership are contacted and asked their opinions of the service they received.

6. Every tenth person entering a mall is asked to name his or her favorite store.

7. Soybeans are planted on a 48-acre field. The field is divided into one-acre subplots. A sample is taken from each subplot to estimate the harvest.

8. From calls made with randomly generated telephone numbers, 1012 respondents are asked if they rent or own their residences.

Table of Contents

Answers

Question 1. Identify populations and samples

1. A survey of 12,082 U.S. adults found that 45.5% received an influenza vaccine for a recent flu season.
(Source: U.S. Centers for Disease Control and Prevention)

Population: all adults in the U.S. Sample: 12,082 U.S. adults

2. A survey of 202 pilots found that 20% admit that they have made a serious error due to sleepiness.
(Source: National Sleep Foundation)

Population: all the pilots Sample: 202 pilots

Question 2. Determine whether the number described is a statistic and a parameter.

1. The average age of the employees in a certain company is 35 years.

Answer: parameter
The average 35 is based on all the employees in a certain company.

2. In a survey of new automobile owners, 43% of the owners plan to buy another new vehicle within four years.

Answer: statistic
The number 43% is a sample of new automobile owners.

Question 3. Identifying Sampling Techniques.

1. Using random digit dialing, researchers call 1400 people and ask what obstacles (such as childcare) keep them from exercising. Simple Random Sampling

2. Chosen at random, 500 rural and 500 urban people age 65 or older are asked about their health and their experience with prescription drugs. Stratified Sampling

3. Questioning students as they leave a university library, a researcher asks 358 students about their drinking habits. Convenience Sampling

4. After a hurricane, a disaster area is divided into 200 equal grids. Thirty of the grids are selected, and every occupied household in the grid is interviewed to help focus relief efforts on what residents require the most. Cluster Sampling

5. Chosen at random, 580 customers at a car dealership are contacted and asked their opinions of the service they received. Simple Random Sampling

6. Every tenth person entering a mall is asked to name his or her favorite store. Systemic Sampling

7. Soybeans are planted on a 48-acre field. The field is divided into one-acre subplots. A sample is taken from each subplot to estimate the harvest. Stratified Sampling

8. From calls made with randomly generated telephone numbers, 1012 respondents are asked if they rent or own their residences. Simple Random Sampling

Table of Contents