

Statistics

Lecture 02: | Sampling

Prof. Dr. Mostafa Elhosseini

<https://youtube.com/drmelhosseini>

| Sampling

- Statistical methods are based on the idea of **analyzing a sample drawn from a population**.
 - The sample must be chosen in an appropriate way
- Ex. Study the heights of students at the Colorado School of Mines by measuring a sample of 100 students.
- How should we choose the 100 students to measure? Some methods are obviously bad
 - Choosing the students from the rosters of the football and basketball teams
 - Would undoubtedly result in a sample that would fail to represent the height distribution of the population of students

| Sampling

- The best sampling methods involve **random sampling**.
- There are many different random sampling methods, the most basic of which is **simple random sampling**
- Imagine that **10,000 lottery tickets** have been sold and that **5 winners are to be chosen**.
- What is the fairest way to choose the winners?
- The fairest way is to **put the 10,000 tickets in a drum, mix them thoroughly, and then reach in and one by one draw 5 tickets out**
- **These**
- 5 winning tickets are a **simple random sample** from the population of 10,000 lottery tickets

| Simple Random Sample

- Each ticket is equally likely to be one of the 5 tickets drawn.
- More importantly, each collection of 5 tickets that can be formed from the 10,000 is equally likely to be the group of 5 that is drawn

| To Conclude

- A **population** is the entire collection of objects or outcomes about which information is sought.
- A **sample** is a subset of a population, containing the objects or outcomes that are actually observed.
- A **simple random sample** of size n is a sample chosen by a method in which each collection of n population items is equally likely to make up the sample, just as in a lottery.

| How to draw simple random sample?

- With a computer random number generator.
- Suppose there are N items in the population. One **assigns** to each **item** in the population an **integer between 1 and N** . Then one **generates a list of random** integers between **1 and N** and **chooses** the corresponding population items to make up the simple random sample

| Is this a simple random sample?

- A **physical education professor** wants to **study** the physical **fitness** levels of students at her university. There are **20,000 students** enrolled at the university, and she wants to draw a **sample of size 100** to take a physical fitness test.
- She obtains a list of all 20,000 students, numbered from 1 to 20,000. She **uses** a computer **random number generator** to **generate 100 random** integers **between 1 and 20,000** and then invites the 100 students corresponding to those numbers to participate in the study. Is this a simple random sample?

| Is this a simple random sample?

- Yes, this is a **simple random sample**. Note that it is **analogous to a lottery** in which each student has a ticket and 100 tickets are drawn

| Is this a simple random sample?

- A **quality engineer** wants to **inspect rolls of wallpaper** in order to obtain information on the rate at which flaws in the printing are occurring. She decides to **draw a sample of 50 rolls** of wallpaper from a day's production. Each hour for 5 hours, she takes the 10 most recently produced rolls and counts the number of flaws on each. Is this a simple random sample?

| Is this a simple random sample?

- No. **Not every subset of 50 rolls of wallpaper is equally likely to make up the sample.**
- To construct a simple random sample, the engineer would need to assign a number to each roll produced during the day and then generate random numbers to determine which rolls make up the sample