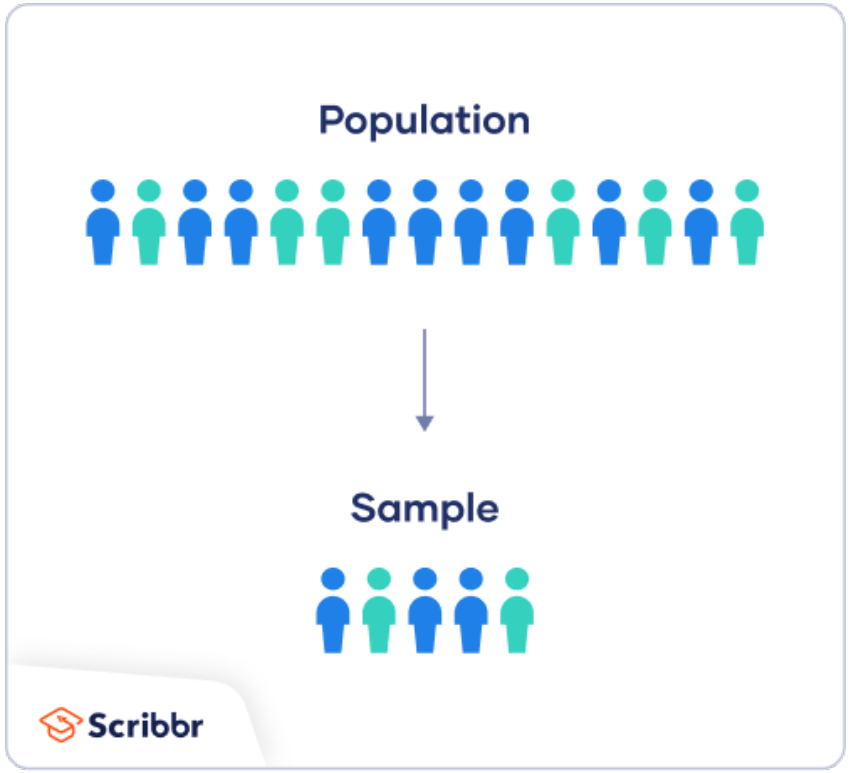
Biostatistics
Week XII

Ege Ülgen, M.D.

23 December 2021

ACIBADEM
MEHMET ALİ AYDINLAR
ÜNİVERSİTESİ

Sampling Methods



- There are 2 main types of sampling methods:
  - Probability sampling
  - Non-probability sampling

Probability Sampling

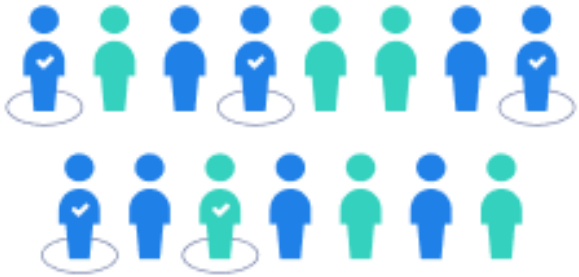
- Probability sampling involves random selection of elements in which each element has a chance of being selected.

- Four main techniques used for a probability sample:
  - Simple random
  - Systematic
  - Stratified random
  - Cluster

An introduction to sampling methods [Internet]. Scribbr. 2019 [cited 2021 Oct 9]. Available from: https://www.scribbr.com/methodology/sampling-methods/

Simple Random Sampling

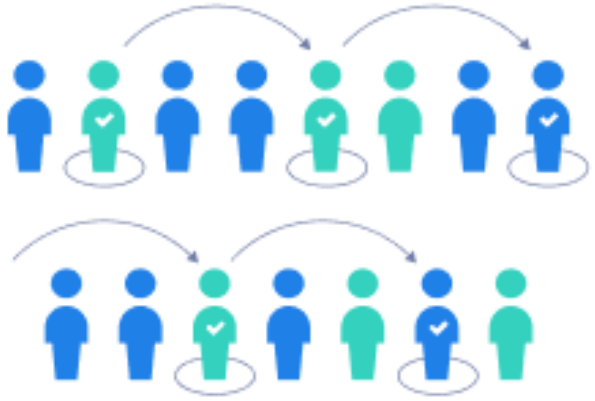
- Completely random method of selecting the sample
- This sampling method is as easy as assigning numbers to the individuals (sample) and then randomly choosing from those numbers through an automated process



An introduction to sampling methods [Internet]. Scribbr. 2019 [cited 2021 Oct 9]. Available from: https://www.scribbr.com/methodology/sampling-methods/

Systematic Sampling

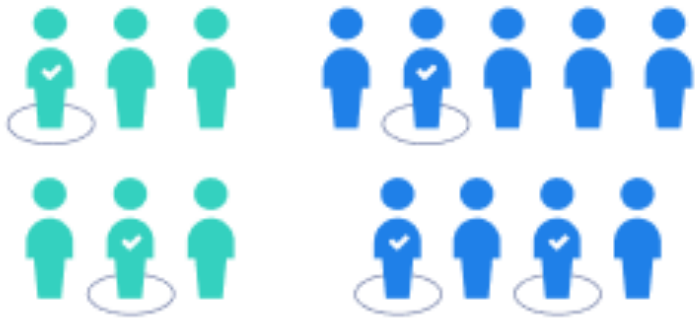
- Systematic sampling is similar to simple random sampling, but it is usually slightly easier to conduct
- Every member of the population is listed with a number, but instead of randomly generating numbers, individuals are chosen at regular intervals



An introduction to sampling methods [Internet]. Scribbr. 2019 [cited 2021 Oct 9]. Available from: https://www.scribbr.com/methodology/sampling-methods/

Stratified Random Sampling

- Stratified sampling involves dividing the population into **subpopulations** that may differ in important ways
- To use this sampling method, you divide the population into subgroups based on the relevant characteristic (e.g., gender, age range, income bracket, job role)



An introduction to sampling methods [Internet]. Scribbr. 2019 [cited 2021 Oct 9]. Available from: https://www.scribbr.com/methodology/sampling-methods/

Cluster Sampling

- Cluster sampling also involves dividing the population into subgroups, but each subgroup should have similar characteristics to the whole sample
- Instead of sampling individuals from each subgroup, you randomly select entire subgroups.

Non-Probability Sampling

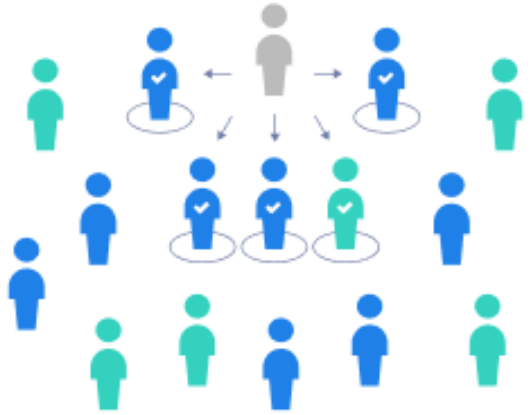
- **Non-probability sampling** involves non-random methods in the selection of elements in which not all have equal chances of being selected

- Four main techniques used for a non-probability sample:
  - Convenience
  - Purposive
  - Snowball
  - Quota

An introduction to sampling methods [Internet]. Scribbr. 2019 [cited 2021 Oct 9]. Available from: https://www.scribbr.com/methodology/sampling-methods/

Convenience Sampling

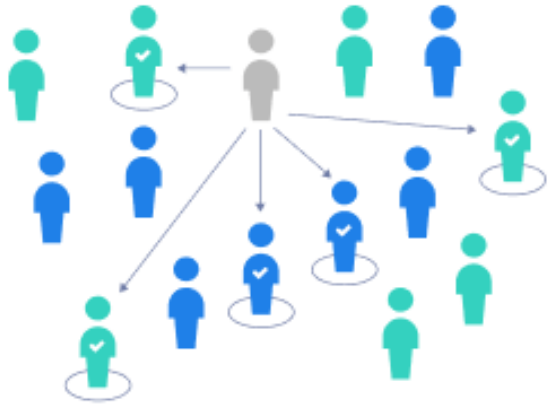
- A convenience sample simply includes the individuals who happen to be most accessible to the researcher
- This is an easy and inexpensive way to gather initial data, but there is no way to tell if the sample is representative of the population, so it can't produce generalizable results



An introduction to sampling methods [Internet]. Scribbr. 2019 [cited 2021 Oct 9]. Available from: https://www.scribbr.com/methodology/sampling-methods/

Purposive Sampling

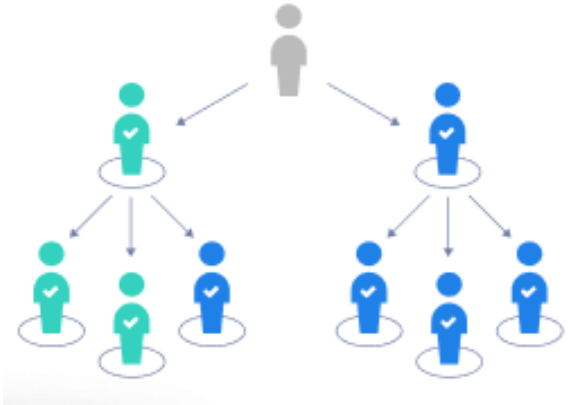
- This type of sampling, also known as judgement sampling, involves the researcher using their expertise to select a sample that is most useful to the purposes of the research



An introduction to sampling methods [Internet]. Scribbr. 2019 [cited 2021 Oct 9]. Available from: https://www.scribbr.com/methodology/sampling-methods/

Snowball Sampling

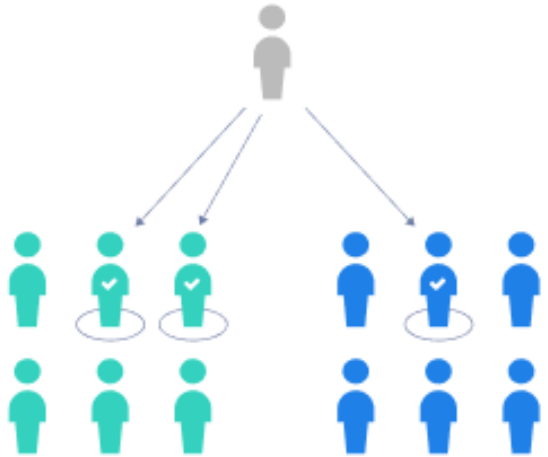
- If the population is hard to access, snowball sampling can be used to recruit participants via other participants
- The number of people you have access to “snowballs” as you contact more people



An introduction to sampling methods [Internet]. Scribbr. 2019 [cited 2021 Oct 9]. Available from: https://www.scribbr.com/methodology/sampling-methods/

Quota Sampling

- The selection of members in this sampling technique happens based on a pre-set standard
- In this case, as a sample is formed based on specific attributes, the created sample will have the same qualities found in the total population
- It is a rapid method of collecting samples.



Missing Data

- Missing data, or missing values, occur when no data value is stored for the variable in an observation

- **complications** in handling and analyzing the data
- **bias** resulting from differences between missing and complete data

Missing Data - Missing Completely at Random (MCAR)

- The missingness of the data is not associated with either any variable or outcome
- There is **nothing systematic** going on that makes some data more likely to be missing than others

- e.g., Questionnaire lost, blood tube for testing a blood level broken etc.

Missing Data - Missing at Random (MAR)

- The missingness of the data is **associated with a variable**
- e.g., Supposing men are more likely to tell their weight than women, missingness in weight is MAR

Missing Data - Missing Not at Random (MNAR)

- The missingness of the data is **related with the outcome**
- e.g., In a depression study, the depression score wasn't calculated for a participant because they committed suicide

Missing Data

- There are several strategies to cope with missing data:
  - **Try to collect the missing data** (obvious best choice)
  - **Exclude** subjects with any missing data (may reduce the power of the study)
  - **Replace** the missing data with a conservative estimate (e.g., sample mean)
  - **Estimate** the missing data from other data on the same subject (imputation)

Brief Summary

- There are 2 main sampling methods:
  - Probability sampling
  - Non-probability sampling
- There are 3 kinds of missing data:
  - MCAR: nothing systematic
  - MAR: missingness associated with a variable
  - MNAR: missingness related with the outcome