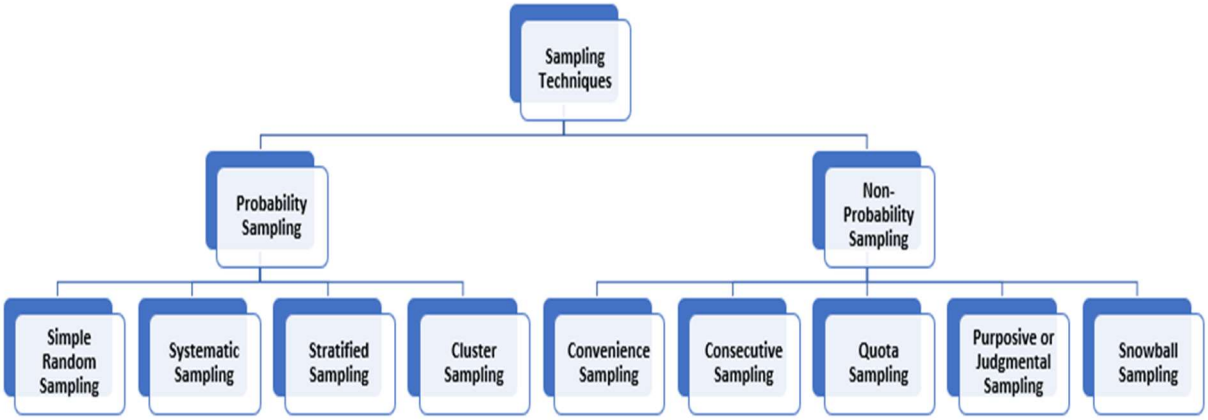
Sampling Techniques

In Statistics, the sampling method or sampling technique is the process of studying the population by gathering information and analyzing that data. It is the basis of the data where the sample space is enormous.



There are two different types of sampling methods are:

- Probability Sampling
- Non-probability Sampling

Probability Sampling:

Probability Sampling is a way of picking a group of things from a bigger group by chance. It ensures that the smaller group represents the bigger one. There are four types:

Simple Random Sampling:

Definition:

Simple random sampling involves selecting a subset of individuals from a population in such a way that every individual has an equal chance of being chosen, and each combination of individuals is equally likely.

Example:

You have a bag of candies with different colors. To do simple random sampling, close your eyes, pick one candy, and that represents your sample, showing the colors in the bag without checking each one.

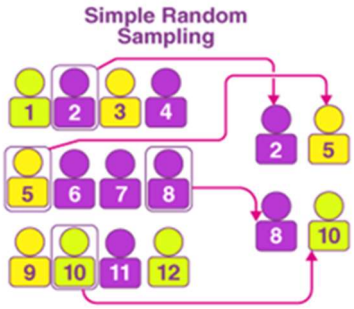


Figure: Simple Random Sampling

Advantages:

- Representativeness: Provides an unbiased and representative sample of the population.
- Simplicity: Easy to understand and implement.
- Statistical Inference: Results can be generalized to the entire population.

Disadvantages:

- **Resource-Intensive:** Can be impractical and time-consuming for large populations.
- **Not Suitable for Strata:** May not ensure representation of specific subgroups within the population.

Stratified Sampling:

Definition:

Stratified sampling involves dividing the population into subgroups or strata based on certain characteristics, and then randomly selecting samples from each stratum.

Example:

Let's say you're conducting a survey at a university with three different majors: Engineering, Business, and Arts. Instead of randomly selecting students from the entire student body, you first divide them into these three groups (strata). Then, from each group, you randomly select a proportionate number of students to represent their respective majors in the survey. This ensures that each major is adequately represented in the sample.

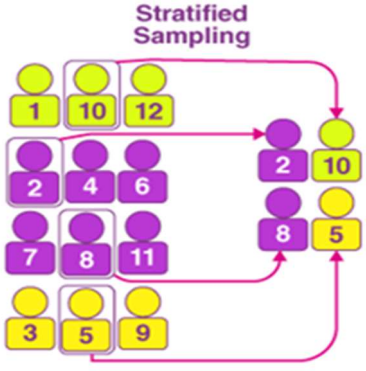


Figure: Stratified Sampling

Advantages:

- **Representativeness:** Ensures representation of various subgroups, making results more accurate.
- **Precision:** Allows for more precise analysis within each stratum.
- **Reduced Variability:** Can reduce the overall variability in the sample.

Disadvantages:

- **Complexity:** More complex to execute than simple random sampling.
- **Stratum Identification:** Requires accurate identification and classification of strata.

Systematic Sampling:

Definition:

Systematic sampling involves selecting every kth element from a list or sampling frame after selecting a random starting point.

Example:

In a class of 50 students, you select every 5th student on the attendance list to represent the entire class, ensuring a systematic sampling approach.

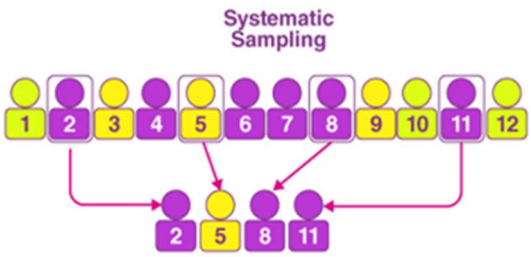


Figure: Systematic Sampling

Advantages:

- Simplicity: Easier and less time-consuming than simple random sampling.
- Uniform Representation: Ensures a level of randomness while maintaining a systematic approach.
- Efficiency: Requires a smaller sample size compared to simple random sampling for the same level of precision.

Disadvantages:

- Periodicity: If there is a pattern in the population list, it could lead to biased results.
- Potential Bias: Can introduce bias if there is an unintentional pattern in the sampling frame.

Clustered Sampling:

Definition:

Cluster sampling involves dividing the population into clusters, randomly selecting some clusters, and then including all individuals within the selected clusters in the sample.

Example:

In a city, you divide neighbourhoods into clusters and randomly select a few clusters. Then, you survey all households within the chosen clusters to represent the entire city's population in clustered sampling.

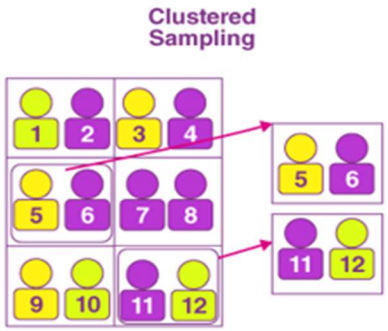


Figure: Cluster Sampling

Advantages:

- Cost-Effective: Can be more cost-effective than other methods, especially when dealing with large populations.
- Logistics: Easier to implement when the population is naturally grouped into clusters.
- Efficiency: Suitable for geographically dispersed populations.

Disadvantages:

- Less Precision: May result in less precise estimates compared to other methods.
- Homogeneity Within Clusters: If clusters are very homogeneous, it may not represent the diversity of the entire population.
- Complexity: Requires careful planning and execution to avoid biases.

(V.V.V. Important)→ Difference Between Stratified and Random Sampling

Aspect	Simple Random Sampling	Stratified Sampling	Systematic Sampling	Cluster Sampling
Purpose	Selects individuals randomly from population	Divides population into homogeneous subgroups	Selects individuals at regular intervals	Divides population into clusters, randomly selects some clusters
Method	Individuals chosen without any specific pattern	Samples from each stratum, may be random or proportional	Selects every nth individual after initial random start	Randomly selects clusters, samples all individuals within chosen clusters

Representativeness	May not ensure representation from all subgroups	Ensures representation from each subgroup	May not ensure representation from all subgroups	Ensures representation from selected clusters
Precision	May not provide precise estimates for subgroups	Often more precise estimates due to subgroup-specific analysis	May not provide precise estimates for subgroups	May provide more precise estimates for selected clusters
Complexity	Simple, requires fewer assumptions about population	More complex, requires knowledge of population characteristics	Simple, but requires a systematic approach	More complex, requires delineation of clusters and sampling within clusters

Non-Probability Sampling: (Not in Syllabus)

Non-Probability Sampling is a method where researchers choose samples based on their judgment rather than random selection, and not all population members have an equal chance of being included. There are five types:

Convenience Sampling:

Definition:

Convenience sampling involves selecting individuals who are easiest to reach or are readily available to participate in the study. This method is based on accessibility rather than random selection.

Advantages:

- Ease of Implementation: Quick and easy to execute.
- Cost-Effective: Requires fewer resources than many other sampling methods.
- Useful for Exploratory Research: Suitable for preliminary or exploratory studies.

Disadvantages:

- Bias: Prone to selection bias, as participants may not be representative of the entire population.
- Lack of Generalizability: Results may not be generalizable to the broader population.

Consecutive Sampling:

Definition:

Consecutive sampling, also known as consecutive or sequential sampling, is a non-probability sampling method where participants are selected based on their availability and willingness to participate in the study during a specific time frame. It involves recruiting individuals who meet the inclusion criteria as they become available, one after the other, until the desired sample size is reached.

Advantages:

- Efficiency: Consecutive sampling is a straightforward and efficient method, especially when the researcher has a limited timeframe for data collection.
- Practicality: It is practical for studies where it is challenging to identify or recruit participants in advance.
- Temporal Trends: Allows for the examination of temporal trends or changes over time by capturing data from a continuous and sequential process.

Disadvantages:

- Selection Bias: There is a risk of selection bias if the characteristics of participants who are available consecutively differ from those who are not.
- Limited Representativeness: The sample may not be representative of the broader population, as it depends on those who are accessible and willing to participate.
- Generalizability: Findings may have limited generalizability beyond the specific context or time period in which the data are collected.

Quota Sampling:

Definition:

Quota sampling involves dividing the population into subgroups or strata based on certain characteristics and then setting quotas for each subgroup. Participants are non-randomly selected to fill these quotas.

Advantages:

- Diversity: Ensures representation of various subgroups within the population.
- Practicality: More practical than some other sampling methods, especially when random sampling is challenging.
- Efficiency: Can be more efficient than simple random sampling for certain studies.

Disadvantages:

- Subjectivity: The lack of randomness and the subjectivity in selecting participants may introduce bias.
- Complexity: Requires accurate identification and definition of strata.

Purposive or Judgmental Sampling:

Definition:

Purposive sampling involves selecting participants based on specific characteristics or criteria that are relevant to the research question. The researcher intentionally chooses individuals who possess the desired traits.

Advantages:

- Relevance: Ensures that participants are relevant to the research objectives.
- Efficiency: Useful when specific characteristics are critical for the study.
- In-Depth Exploration: Allows for in-depth exploration of specific characteristics.

Disadvantages:

- Potential Bias: The researcher's judgment in selecting participants may introduce bias.
- Limited Representativeness: Results may not be generalizable to the broader population.

Snowball Sampling:

Definition:

Snowball sampling involves identifying a few initial participants who meet the criteria and then asking them to refer other potential participants. This method is often used when the population is hard to reach.

Advantages:

- Access to Hard-to-Reach Groups: Useful for studying populations that are difficult to identify or contact.
- Cost-Effective: Can be cost-effective, especially when researchers have limited resources.
- Network Exploration: Allows for the exploration of social networks and connections.

Disadvantages:

- Bias: May lead to biased samples, especially if participants share similar characteristics.
- Limited Representativeness: Results may not be generalizable beyond the initial participants.