

Sampling

Sampling

- A sample is a smaller (but hopefully representative) collection of elements from a population used to determine truths about that population
- It saves resources (time, money) and workload
- It gives results with known accuracy that can be calculated mathematically
- The sampling frame is the list from which the potential elements are drawn

Sampling process

- 1) Define the population of concern
- 2) Specify a sampling frame possible to measure
- 3) Specify a sampling method
- 4) Determine the sample size
- 5) Implement the sampling plan
- 6) Sample and collect data
- 7) Review the sampling process

Sampling methods

The sampling method is the scientific procedure of selecting those sampling units which would provide the required estimates with associated margins of uncertainty, arising from examining only a part and not the whole.

- **Probability methods** : Every element in the population has a chance (greater than zero) of being selected in the sample.
 - Simple random sampling
 - Systematic sampling
 - Stratified sampling
- **Non-probability methods** : Some elements in the population have no chance of selection or the probability of selection cannot be accurately determined.

Simple random sampling

- Create the sample by selecting randomly from the population. For instance, using random numbers by computer.
- It has two **key properties**:
 - 1) **Unbiased**: Each element has the same chance of being selected.
 - 2) **Independent**: The selection of each element is not affected by the selection of other elements.
- It can be used with or without replacement.
- The elements of the population must all be accessible and available.
- Use when the target group is sufficiently large.
- Random selection may not always be possible.

Stratified sampling

- Split the population into smaller non-overlapping sub-groups, then do a simple random sample in each sub-group.
- It reduces standard error by providing some control over variance. If you know that there are groups in the population, you can deliberately sample these in a due proportion.
- If groups are homogeneous, stratified sampling is more accurate than simple random sampling.

Cluster sampling

- Split the population into clusters of homogeneous elements, for instance, geographic contiguity.
- Sampling elements are the clusters, not the individuals. All elements in the sampled cluster are studied.
- May reduce costs, but sampling error is higher.
- Note that all strata are represented in the sample, but only a subset of clusters are in the sample.

Systematic sampling

- Order the population and select every n -th element from the population, starting from a randomly chosen one.
- It is an alternative to simple random sampling when it is easier to select every n -th element.
- It only works if the population can be sorted into a sequence in which you can ensure that the studied attribute is random.
- Sample may be biased if hidden periodicity in population coincides with that of selection.

Non-probability sampling methods

- **Quota sampling**: The sample size of each group is defined a priori with some goal in mind.
- **Purposive sampling**: The sample is selected to include elements of interest and exclude those who do not suit the purpose of the study.
- **Convenience sampling**: The elements in the sample are selected in the most convenient way, for instance, by volunteering.
- **Snowball sampling**: Find people to study. Ask them to refer you other people who fit your study requirements, then follow up with these new people. Repeat this method of requesting referrals until you have studied enough people.