



Research Methodology

# Sample Selection Methods

## Advantages of Sample Surveys

- As in sample study few units are to be examined detailed study of the survey can be done.
- As few units are to be examined the survey work requires less time. Thus in this way sample survey saves time.
- As few units are to be examined the survey work requires less money. Thus in this way sample survey saves lots of money.
- In sample survey few persons are required for the survey work so experts can be appointed for the survey. This will increase the reliability of the survey results.

- When the test is of destructive nature, sampling is only the way out. In such cases the population survey is not possible.
- A large area can be covered under survey in the available time and money.
- Sample survey is also used to check the accuracy of the census data (population survey).
- Surveys are an efficient way of collecting information from a large number of respondents. Very large samplings are possible. Statistical techniques can be used to determine validity, reliability, and statistical significance.

# 1. Probability Sampling

- In probability (random) sampling, you start with a complete sampling frame of all eligible individuals from which you select your sample.
- In this way, all eligible individuals have a chance of being chosen for the sample, and you will be more able to generalize the results from your study.
- Probability sampling methods tend to be more time-consuming and expensive than non-probability sampling. It is further classified in following types:
  - A. Simple Random Sampling
  - B. Systematic Sampling
  - C. Probability Sampling
  - D. Cluster Sampling

## 2. Non-Probability Sampling

- In non-probability (non-random) sampling, you do not start with a complete sampling frame, so some individuals have no chance of being selected.
- Consequently, you cannot estimate the effect of sampling error and there is a significant risk of ending up with a non-representative sample which produces non-generalisable results.
- However, non-probability sampling methods tend to be cheaper and more convenient, and they are useful for exploratory research and hypothesis generation.
  - A. Convenience sampling
  - B. Quota Sampling
  - C. Purposive Sampling
  - D. Snow Ball Sampling

# Simple Random Sampling

- In this case each individual is chosen entirely by chance and each member of the population has an equal chance, or probability, of being selected.
- One way of obtaining a random sample is to give each individual in a population a number, and then use a table of random numbers to decide which individuals to include.
- For example, if you have a sampling frame of 1000 individuals, labeled 0 to 999, use groups of three digits from the random number table to pick your sample.
- So, if the first three numbers from the random number table were 094, select the individual labeled “94”, and so on.

# Simple Random Sampling

- As with all probability sampling methods, simple random sampling allows the sampling error to be calculated and reduces selection bias.
- A specific advantage is that it is the most straightforward method of probability sampling.
- A disadvantage of simple random sampling is that you may not select enough individuals with your characteristic of interest, especially if that characteristic is uncommon.
- It may also be difficult to define a complete sampling frame and inconvenient to contact them, especially if different forms of contact are required (email, phone, post) and your sample units are scattered over a wide geographical area.

# Systematic Sampling

- Individuals are selected at regular intervals from the sampling frame. The intervals are chosen to ensure an adequate sample size.
- If you need a sample size  $n$  from a population of size  $x$ , you should select every  $x/n^{\text{th}}$  individual for the sample.
- For example, if you wanted a sample size of 100 from a population of 1000, select every  $1000/100 = 10^{\text{th}}$  member of the sampling frame.
- Systematic sampling is often more convenient than simple random sampling, and it is easy to administer.



# Systematic Sampling

- However, it may also lead to bias, for example if there are underlying patterns in the order of the individuals in the sampling frame, such that the sampling technique coincides with the periodicity of the underlying pattern.
- As a hypothetical example, if a group of students were being sampled to gain their opinions on college facilities, but the Student Record Department's central list of all students was arranged such that the sex of students alternated between male and female, choosing an even interval (e.g. every 20<sup>th</sup> student) would result in a sample of all males or all females.
- Whilst in this example the bias is obvious and should be easily corrected, this may not always be the case.

# Stratified Sampling

- In this method, the population is first divided into subgroups (or strata) who all share a similar characteristic.
- It is used when we might reasonably expect the measurement of interest to vary between the different subgroups, and we want to ensure representation from all the subgroups.
- For example, in a study of stroke outcomes, we may stratify the population by sex, to ensure equal representation of men and women.
- The study sample is then obtained by taking equal sample sizes from each stratum. In stratified sampling, it may also be appropriate to choose non-equal sample sizes from each stratum.

# Stratified Sampling

- For example, in a study of the health outcomes of nursing staff in a county, if there are three hospitals each with different numbers of nursing staff (hospital A has 500 nurses, hospital B has 1000 and hospital C has 2000), then it would be appropriate to choose the sample numbers from each hospital *proportionally* (e.g. 10 from hospital A, 20 from hospital B and 40 from hospital C).
- This ensures a more realistic and accurate estimation of the health outcomes of nurses across the county., whereas simple random sampling would over-represent nurses from hospitals A and B.
- Stratified sampling improves the accuracy and representation of the results by reducing sampling bias.

# Clustered Sampling

- In a clustered sample, subgroups of the population are used as the sampling unit, rather than individuals. The population is divided into subgroups, known as clusters, which are randomly selected to be included in the study.
- Clusters are usually already defined, for example individual group practices or towns could be identified as clusters.
- In single-stage cluster sampling, all members of the chosen clusters are then included in the study.
- In two-stage cluster sampling, a selection of individuals from each cluster is then randomly selected for inclusion. Clustering should be taken into account in the analysis.

# Clustered Sampling

- The General Household survey, which is undertaken annually in England, is a good example of a (one-stage) cluster sample. All members of the selected households (clusters) are included in the survey.
- Cluster sampling can be more efficient than simple random sampling, especially where a study takes place over a wide geographical region.
- For instance, it is easier to contact lots of individuals in a few group practices than a few individuals in many different group practices.
- Disadvantages include an increased risk of bias, if the chosen clusters are not representative of the population, resulting in an increased sampling error.

# Convenience sampling

- Convenience sampling is perhaps the easiest method of sampling, because participants are selected based on availability and willingness to take part.
- Useful results can be obtained, but the results are prone to significant bias, because those who volunteer to take part may be different from those who choose not to (volunteer bias), and
- the sample may not be representative of other characteristics, such as age or sex.
- Note: volunteer bias is a risk of all non-probability sampling methods.

# Quota Sampling

- This method of sampling is often used by market researchers. Interviewers are given a quota of subjects of a specified type to attempt to recruit.
- For example, an interviewer might be told to go out and select 20 adult men, 20 adult women, 10 teenage girls and 10 teenage boys so that they could interview them about their television viewing.
- Ideally the quotas chosen would proportionally represent the characteristics of the underlying population.
- Whilst this has the advantage of being relatively straightforward and potentially representative, the chosen sample may not be representative of other characteristics that weren't considered (a consequence of the non-random nature of sampling).

# Judgmental/Selective Sampling

- Also known as subjective, sampling, this technique relies on the judgment of the researcher when choosing who to ask to participate.
- Researchers may implicitly thus choose a “representative” sample to suit their needs, or specifically approach individuals with certain characteristics. This approach is often used by the media when canvassing the public for opinions and in qualitative research.
- Judgment sampling has the advantage of being time-and cost-effective to perform whilst resulting in a range of responses (particularly useful in qualitative research).
- However, in addition to volunteer bias, it is also prone to errors of judgment by the researcher and the findings, whilst being potentially broad, will not necessarily be representative.



# Snow ball Sampling

- This method is commonly used in social sciences when investigating hard-to-reach groups. Existing subjects are asked to nominate further subjects known to them, so the sample increases in size like a rolling snowball.
- For example, when carrying out a survey of risk behaviors amongst intravenous drug users, participants may be asked to nominate other users to be interviewed.
- Snowball sampling can be effective when a sampling frame is difficult to identify.
- However, by selecting friends and acquaintances of subjects already investigated, there is a significant risk of selection bias (choosing a large number of people with similar characteristics or views)