BRM: Unit 4
Sampling

DR. ROHAN DAHIVALE
RAJGAD INSTITUTE OF MANAGEMENT RESEARCH & DEVELOPMENT, PUNE
• **4.1 Basic Concepts:** Defining the Universe, Concepts of Statistical Population, Sample, Characteristics of a good sample. Sampling Frame (practical approach for determining the sample frame expected), Sampling errors, Non Sampling errors, Methods to reduce the errors, Sample Size constraints, Non Response.

• **4.2 Probability Sample:** Simple Random Sample, Systematic Sample, Stratified Random Sample, Area Sampling & Cluster Sampling.

• **4.3 Non Probability Sample:** Judgment Sampling, Convenience Sampling, Purposive Sampling, Quota Sampling & Snowballing Sampling methods.

• Determining size of the sample - Practical considerations in sampling and sample size
**Sampling** is a process used in statistical analysis in which a predetermined number of observations are taken from a larger population.
<table>
<thead>
<tr>
<th>Term</th>
<th>Definition</th>
</tr>
</thead>
<tbody>
<tr>
<td>Universe or Population</td>
<td>Group of people or objects under study or investigation</td>
</tr>
<tr>
<td></td>
<td>Example: In a study of worker absenteeism in a factory, the population is all workers in that factory</td>
</tr>
<tr>
<td>Sampling Frame</td>
<td>All elements in the population with their identification</td>
</tr>
<tr>
<td></td>
<td>Example: all workers with identification like employee ID</td>
</tr>
<tr>
<td>Sample</td>
<td>Some elements (subset of population) in population under study</td>
</tr>
<tr>
<td></td>
<td>Example: A group of workers above age 40 yrs or working in logistics department only are selected for the study.</td>
</tr>
<tr>
<td>Term</td>
<td>Definition</td>
</tr>
<tr>
<td>-----------------</td>
<td>---------------------------------------------------------------------------</td>
</tr>
<tr>
<td><strong>Sampling Unit</strong></td>
<td>A single member of a sample</td>
</tr>
<tr>
<td></td>
<td>Example: Each &amp; every individual worker in a factory is a sampling Unit</td>
</tr>
<tr>
<td><strong>Sampling</strong></td>
<td>A process of selecting adequate number of elements from population is called as sampling</td>
</tr>
<tr>
<td><strong>Census</strong></td>
<td>When researcher examines each and every element of the population is called as Census</td>
</tr>
</tbody>
</table>
Sample Design

Population

Sampling frame

Sample
Characteristics of a good sample

**It should be random**
- A random sample is one in which each item in the sample has an equal chance of being selected. Statistical inferences may not be valid unless the sample is random.

**It should be representative**
- The sample should be representative of the differing items in the whole population.
- For example it should contain a similar proportion of high and low value items of the population.

**Protective**
- Protective towards the researcher. More intensive auditing should occur on high value items known to be high risk.

**Unpredictable**
- Client should not be able to know which items will be examined.
Sampling Frame

- A sampling frame is the source material or device from which a sample is drawn.
- It is a list of all those within a population who can be sampled, and may include individuals, households or institutions.
Sampling Errors

- Sampling error is incurred when the statistical characteristics of a population are estimated from a subset, or sample, of that population.
- Since the sample does not include all members of the population, statistics on the sample, such as means, generally differ from the characteristics of the entire population, which are known as parameters.
- **Sampling Error = (Frame Error) + (Chance Error) + (Response Error)**
Errors

Sampling
  Non Sampling

Biased
  Unbiased
    Data Analysis
      Non Response

Frame
  Measurement
<table>
<thead>
<tr>
<th>Errors</th>
<th>Description</th>
</tr>
</thead>
<tbody>
<tr>
<td>Sampling Errors</td>
<td>Originate at the time of collecting samples</td>
</tr>
<tr>
<td>Biased Sampling Errors</td>
<td>Originates from bias selection of a respondent</td>
</tr>
<tr>
<td>Unbiased Sampling Errors</td>
<td>Originates when there are few restrictions in sample selection</td>
</tr>
<tr>
<td>Non Sampling Errors</td>
<td>Occurs at the time of observing or processing the data</td>
</tr>
<tr>
<td>Data Analysis Error</td>
<td>Arithmetic or calculation error while data analysis</td>
</tr>
<tr>
<td></td>
<td>This error mostly occurs at the execution stage</td>
</tr>
<tr>
<td>Errors</td>
<td>Description</td>
</tr>
<tr>
<td>-------------------</td>
<td>-----------------------------------------------------------------------------</td>
</tr>
<tr>
<td>Non Response Error</td>
<td>Occurs after respondents refuse to give information that needed</td>
</tr>
<tr>
<td>Frame Error</td>
<td>Occurs if sampling frame is incomplete or not accurate</td>
</tr>
<tr>
<td>Measurement Error</td>
<td>Occurs when there is a difference between information gathered and information sought</td>
</tr>
</tbody>
</table>
Reducing Sampling Error

There are mainly two ways by which this sampling error can be reduced. The ways are:

- Increasing sample size
- Stratification
Reducing Non Sampling Error (Non Response)

1. Appropriate time of the survey is selected so as to reduce non response of the respondents
2. Design of questionnaire should be done with care and in such a way that respondents responds it properly.
3. Sampling frame error is avoided
4. Proper follow up of the respondents
A process of selecting adequate number of elements from population is called as sampling.

On the basis of sample selection, there are two types of Sampling Methods:

- **Probability Sampling**
- **Non-Probability Sampling**
A probability sampling method is any method of sampling that utilizes some form of random selection. In order to have a random selection method, you must set up some process or procedure that assures that the different units in your population have equal probabilities of being chosen.

Probability sampling can be further achieved by following various methods.

a) Simple Random Sample
b) Systematic Sample
c) Stratified Random Sample
d) Area Sampling
e) Cluster Sampling
a) Simple Random Sampling

- Simplest method
- Each & every element in the sample frame has equal chance of getting selected
- Respondent is randomly chosen from the sample frame
- **Example**: In a survey conducted in school, about reading habits of the students, any student of the school can get selected for the study.
- **Advantages**: Free from bias opinion & ease of sampling
- **Disadvantages**: Costly if large population & administrative difficulties
b) Systematic Sampling

- Sample Interval is calculated =
  \[
  \text{Sample Interval} = \frac{\text{Population Size}}{\text{Sample Size}}
  \]

- First respondent is selected randomly and then next respondents are taken at regular intervals (as per calculated sample interval).

- **Example:** For the above same example, if Sample Interval = 15, then First student is selected randomly and then next students are selected at interval of 15.

- **Advantages:** Simple, very low bias & convenient

- **Disadvantages:** Not suitable for large population & high sample error
c) Stratified Random Sample

- Random selection from the heterogeneous universe as a whole but homogeneous as a strata of the universe.
- Better than simple random sampling as representative of sampling increases.
- Universe is divided into homogeneous group which is called as strata, then simple random sampling is done within each strata.

**Example:** For the above same example, strata can be done as male students & female students. OR
- strata can be done as 5th, 6th, 7th up to 10th division.

**Advantages:** More precise & more representative

**Disadvantages:** Complicated & time consuming
d) Area Sampling

- Maps are used as sampling frame
- A method in which an area to be sampled is subdivided into smaller blocks that are then selected at random and then again sub-sampled or fully surveyed. This method is typically used when a complete frame of reference is not available to be used.

**Example:** Divide population according to geographical boundaries.

**Advantages:** Feasible & convenient

**Disadvantages:** Costly
e) Cluster Sampling

- Suitable for smaller population.
- A method in which an area to be sampled is sub-divided into smaller blocks that are then selected at random and fully surveyed.
- These groups are **heterogeneous within** sub groups.
- These groups are **homogeneous between** sub groups.
- **Example:** In a survey conducted in a company, Logistics department is selected and fully surveyed.
- **Advantages:** Fast to apply & cost effective
- **Disadvantages:** High sample error & less representative
Non probability sampling is a sampling technique where the samples are gathered in a process that does not give all the individuals in the population equal chances of being selected.

Non Probability sampling can be further achieved by following various methods.

i. Judgment Sampling
ii. Convenience Sampling
iii. Purposive Sampling
iv. Quota Sampling
v. Snowballing Sampling
i) Convenience Sampling

- Researcher choose samples as per his/her convenience
- Suitable for pre test questionnaire test
- Example: Survey of those students who are coming out of the Institute.
- Advantages: Easy & less costly
- Disadvantages: Personal bias & not suitable for large population
ii) Purposive Sampling

- Purposive sampling is a non-probability sample that is selected based on characteristics of a population and the objective of the study.

Further can be classified as
1. Judgment Sampling
2. Quota sampling
iii) Judgmental Sampling

- **Sample** is selected based on the opinion of an expert.
- Results obtained from a *judgment sample* are subject to some degree of bias, due to the frame and population not being identical.
- **Example:** In a survey to be conducted about the malls in the city, Expert feels to go to a particular mall only for conducting the survey than other malls.
- **Advantages:** Suitable for small surveys
- **Disadvantages:** Bias & non scientific
iv) Quota Sampling

- A **sampling** method of gathering representative data from a group.
- As opposed to random **sampling**, **quota sampling** requires that representative individuals are chosen out of a specific subgroup.
- **Example**: A researcher might ask for a **sample** of 100 females, or 100 individuals between the ages of 20-30.
- **Advantages**: Economic & Convenient
- **Disadvantages**: Difficult to calculate standard error
v) Snowballing Sampling

- Snowball sampling (or chain sampling) is a technique where existing study subjects recruit future subjects from among their acquaintances.
- Thus the sample group is said to grow like a rolling snowball. As the sample builds up, enough data are gathered to be useful for research.
- **Example:** Survey of deep ocean swimmers, one swimmer is selected, and then next respondents are selected by reference of the previous one.
- **Advantages:** Less costly & easy to plan
- **Disadvantages:** Biased & Limited control
Thank You