SAMPLE DESIGNING

PROF. RANJEET KUMAR CHOUDHARY
PROFESSOR & HEAD, UNIVERSITY LIBRARIAN I/C
DEPARTMENT OF LIBRARY & INFORMATION SCIENCE
MAHATMA GANDHI CENTRAL UNIVERSITY, MOTIHARI-845401 (BIHAR) INDIA
MOBILE NO. +91-9935787978, E-MAIL: RKCHOUDHARY@MGCUB.AC.IN, RKCLIS@YAHOO.CO.IN

INTRODUCTION

SAMPLE

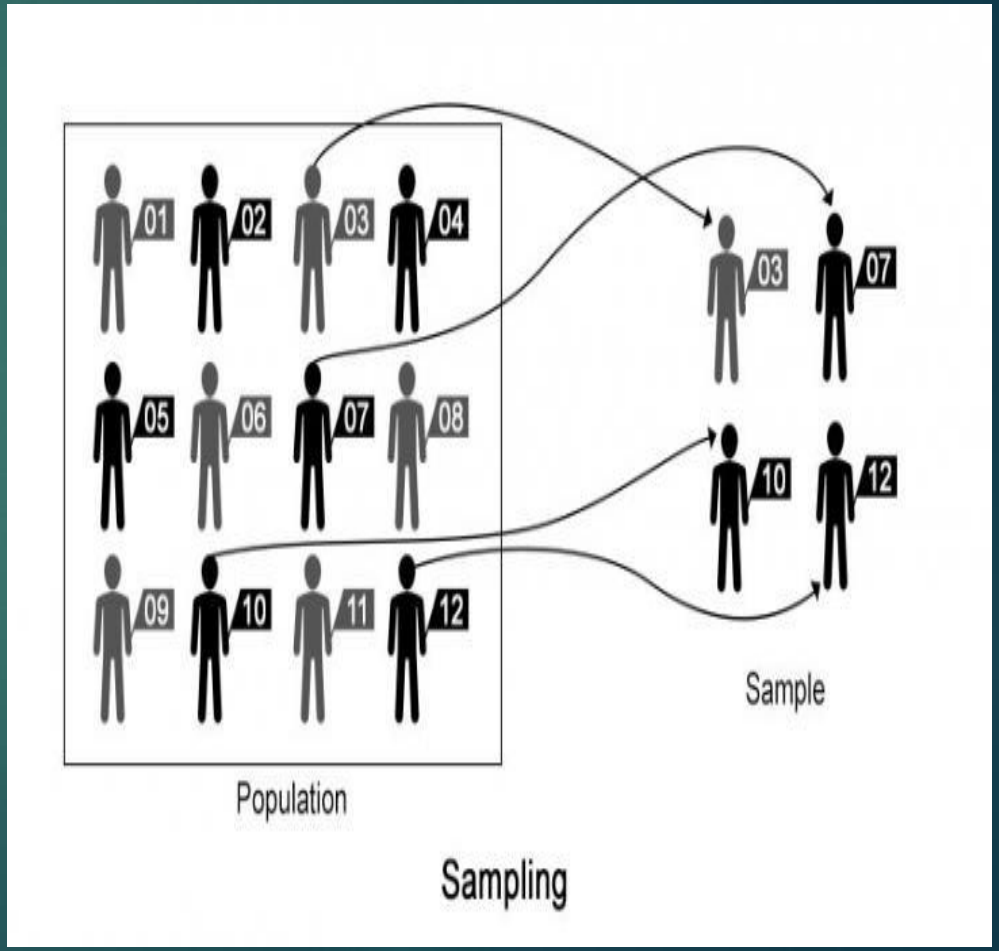
A sample is group of people, objects or items that are taken from a large population for a measurement. The sample should be representative of the population to ensure that we can generalize the findings from the research sample to the population as a whole.

SAMPLING

Sampling is the act, process, or technique of selecting a suitable sample, or a representative part of a population for the purpose of determining parameters or characteristics of the whole population.

PURPOSE OF SAMPLING

- The large size of many populations.
- The time factor- a sample may provide needed information quickly.
- Inaccessibility (associated with cost or time or just access) of some of the population e.g. prisoners, people with severe mental illness, disaster survivors etc.
- Destructiveness of the observation e.g. to determine the quality of a fuse and whether it is defective, it must be destroyed. Therefore, if you tested all the fuses, all would be destroyed.
- Accuracy and sampling i.e. a sample may be more accurate.



CHARACTERISTICS

- Much cheaper.
- Saves time.
- Much reliable.
- Very suitable for carrying out different surveys.
- Scientific in nature.

ADVANTAGES

- Very accurate.
- Economical in nature.
- Very reliable.
- High suitability ratio towards the different surveys.
- Takes less time.
- In cases, when the universe is very large, then the sampling method is the only practical method for collecting the data.

DISADVANTAGES

- Inadequacy of the samples.
- Chances for bias.
- Problems of accuracy.
- Difficulty of getting the representative sample.
- Untrained manpower.
- Absence of the informants.
- Chances of committing the errors in sampling.

STAGES OF SAMPLING DESIGN PROCESS

- Define the population.
- Specifying the sampling frame.
- Specifying the sampling unit.
- Selection of the sampling method.
- Determination of sample size.
- **Specifying the Sampling Plan**.
- **Selecting the Sample.**

STEP – I
Define the Population

- Population must be defined in terms of elements, sampling units, extent and time. Because there is very rarely enough time or money to gather information from everyone or everything in a population, the goal becomes finding a representative sample (or subset) of that population.
- For ex, if a kitchen appliances firm wants to conduct a survey to ascertain the demand for its micro ovens, it may define the population as ‘all women above the age of 20 who cook (assuming that very few men cook)’. However this definition is too broad and will include every household in the country, in the population that is to be covered by the survey. Therefore the definition can be further refined and defined at the sampling unit level, that, all women above the age 20, who cook and whose monthly household income exceeds Rs.20,000. This reduces the target population size and makes the research more focused. The population definition can be refined further by specifying the area from where the researcher has to draw his sample, that is, households located in Hyderabad.

STEP- 2
Specifying the Sampling Frame

Once the definition of the population is clear a researcher should decide on the sampling frame. A sampling frame is the list of elements from which the sample may be drawn. Continuing with the micro oven ex, an ideal sampling frame would be a database that contains all the households that have a monthly income above Rs.20,000. However, in practice it is difficult to get an exhaustive sampling frame that exactly fits the requirements of a particular research. In general, researchers use easily available sampling frames like telephone directories and lists of credit card and mobile phone users. Various private players provide databases developed along various demographic and economic variables. Sometimes, maps and aerial pictures are also used as sampling frames. Whatever may be the case, an ideal sampling frame is one that entire population and lists the names of its elements only once.

STEP – 3
Specifying the Sampling Unit

A sampling unit is a basic unit that contains a single element or a group of elements of the population to be sampled. In this case, a household becomes a sampling unit and all women above the age of 20 years living in that particular house become the sampling elements. If it is possible to identify the exact target audience of the business research

STEP – 4
Selection of the Sampling Method

The sampling method outlines the way in which the sample units are to be selected. The choice of the sampling method is influenced by the objectives of the business research, availability of financial resources, time constraints, and the nature of the problem to be investigated.

STEP – 5
Determination of Sample Size

The sample size plays a crucial role in the sampling process. There are various ways of classifying the techniques used in determining the sample size. A couple those hold primary importance and are worth mentioning are whether the technique deals with fixed or sequential sampling and whether its logic is based.

STEP – 6
Specifying the Sampling Plan

In this step, the specifications and decisions regarding the implementation of the research process are outlined. Suppose, blocks in a city are the sampling units and the households are the sampling elements.

STEP – 7
Selecting the Sample

This is the final step in the sampling process, where the actual selection of the sample elements is carried out. At this stage, it is necessary that the interviewers stick to the rules outlined for the smooth implementation of the business research. This step involves implementing the sampling plan to select the sampling plan to select a sample required for the survey.

Types of Sampling

Probability Sampling

- Simple Random Sampling
- Cluster Sampling
- Systematic Sampling
- Stratified Random Sampling

Non-probability Sampling

- Convenience sampling
- Judgmental or Purposive Sampling
- Snowball sampling
- Quota sampling

Probability Sampling

- Probability Sampling is a sampling technique in which sample from a larger population are chosen using a method based on the theory of probability.
- For a participant to be considered as a probability sample, he/she must be selected using a random selection.
- So that probability sampling uses statistical theory to select randomly, a small group of people (sample) from an existing large population and then predict that all their responses together will match the overall population.

Types of Probability Sampling

- **Simple Random Sampling:** This is a sampling technique where every item in the population has an even chance and likelihood of being selected in the sample.
- **Cluster Sampling:** The researchers divide the entire population into sections or clusters that represent a population.
- **Systematic Sampling:** The members of a sample are chosen at regular intervals of a population. It requires selection of a starting point for the sample and sample size that can be repeated at regular intervals. This type of sampling method has a predefined interval and hence this sampling technique is the least time-consuming.
- **Stratified Random Sampling:** The population can be divided into smaller groups, that don't overlap but represent the entire population together. While sampling, these groups can be organized and then draw a sample from each group separately.

Non-probability Sampling Methods

The non-probability method is a sampling method that involves a collection of feedback on the basis of a researcher or statistician’s sample selection capabilities and not on a fixed selection process. In most situations, output of a survey conducted with a non-probable sample leads to skewed results, which may not totally represent the desired target population.

Types of Non-probability Sampling Methods

- **Convenience Sampling:** This method is dependent on the ease of access to subjects such as surveying customers at a mall or passers-by on a busy street. It is usually termed as convenience sampling, as it's carried out on the basis of how easy is it for a researcher to get in touch with the subjects.
- **Judgmental or Purposive Sampling:** In judgemental or purposive sampling, the sample is formed by the discretion of the judge purely considering the purpose of study along with the understanding of target audience.
- **Snowball Sampling**: This sampling method involves a primary data source nominating other potential data sources that will be able to participate in the research studies. Snowball sampling method is purely based on referrals and that is how a researcher is able to generate a sample. Therefore this method is also called the chain-referral sampling method.
- **Quota Sampling:** In Quota sampling, selection of members in this sampling technique happens on basis of a pre-set standard. In this case, as a sample is formed on basis of specific attributes, the created sample will have the same attributes that are found in the total population. It is an extremely quick method of collecting samples.

Use of the Non-Probability Sampling Method

- **Create a hypothesis:** The non-probability sampling method is used to create a hypothesis when limited to no prior information is available. This method helps with immediate return of data and helps to build a base for any further research.
- **Exploratory research:** This sampling technique is widely used when researchers aim at conducting qualitative research, pilot studies or exploratory research.
- **Budget and time constraints:** The non-probability method when there are budget and time constraints and some preliminary data has to be collected. Since the survey design is not rigid, it is easier to pick respondents at random and have them take the survey or questionnaire.

SAMPLING ERROR AND SURVEY BIAS

ERROR

- Population Specification Error
- Sample Frame Error
- Selection Error
- Non-Response
- Sampling Errors

BIAS

Bias often occurs when the survey sample does not accurately represent the population. The bias that results from an unrepresentative sample is called **selection bias**.

Determination of Sample Size

- The sample size computation must be done appropriately because if the sample size is not appropriate for a particular study then the inference drawn from the sample will not be authentic and it might lead to some wrong conclusions.
- A sample with the smallest sampling error will always be considered a good representative of the population. Bigger samples have lesser sampling errors. When the sample survey becomes the census survey, the sampling error becomes zero.
- There are different formulae given by different educationists for the determination of appropriate sample sizes. The researcher should choose the formula according to their needs and convenience.

Size of Your Sample

- **Population** : This is the complete set of data points, for example, all Americans.
- **Target Population** : This is the complete group for which you are studying; your data will have specific characteristics (demographics, clinical characteristics) that you are interested in for example, Americans over the age of 65, who live at home and have had a stroke in the past 6 months.
- **Sample** :A subset of the target population that represents the target population.
- **Margin of Error :** The margin of error is about a degree of uncertainty in statistics. How much error will you allow? We would like the mean of our data to represent the mean of the target population; however, this is generally not going to happen. The margin of error tells us how much higher or lower than the true value will we let our sample mean fall. In articles, you usually see a +/-5% or +/-3% margin of error.
- **Confidence Interval :** Confidence interval (CI) is usually set at 90%, 95% or 99%. It tells us how confident we are that if the study was repeated again and again, we would get the same results. If confidence level is 95%, we would get the same results in 95% of the cases.
- **Standard Deviation :** Standard deviation tells us the variation in the data from your sample.
- **Power** : This refers to the chance of missing a real difference (‘false negatives’). Usually, studies have a power of around 80%, which means that you accept the possibility that in 20% of the cases, the real difference was missed (you concluded there was no effect when there was one). Larger samples generally yield higher statistical power.

Calculate a Sample Size

FORMULA

$(\text{Z value})^2 \times \text{standard deviation (1-standard deviation)}/(\text{margin of error})^2 = n$

(This formula, however, can only be used for large populations or unknown population sizes.)

The Z-value or Z-score corresponds with your chosen confidence level. There are usually Z tables available that tell you the Z-score. You can then insert that value into the formula. Below are values for the most commonly used confidence intervals.

Software for Calculating Your Sample Size

Various software programs and websites that will calculate the size for you. An example of such a site is The Survey System, which offers a free online sample size calculator. Another option is Survey Monkey's sample size calculator, which can also be accessed online. These calculators usually ask you to enter your target population size, confidence level and margin of error.

THANK YOU