NONPROBABILITY SAMPLING

L-2 Part 2

Prof Dr Najlaa Fawzi

- Any sampling method where some elements of population have no chance of selection (these are sometimes referred to as 'out of coverage'/'under covered'), or where the probability of selection can't be accurately determined.

- It involves the selection of elements based on assumptions regarding the population of interest, which forms the criteria for selection.

-

- Hence, because the selection of elements is nonrandom, nonprobability sampling not allows the estimation of sampling errors..

Non random sampling (non-probability sampling) technique is used when it cannot be ensured that each item has an equal chance of being selected , or when selection is based on expert knowledge of the population.

- Nonprobability Sampling includes: Accidental Sampling, Quota Sampling and Purposive Sampling.

- In addition, nonresponse effects may turn *any* probability design into a nonprobability design if the characteristics of nonresponse are not well understood, since nonresponse effectively modifies each element's probability of being sampled.

02. Non Probability Samples

A non probability sample relies on the researcher selecting the respondents.

They are considered to be:

- Interpretive
- Subjective
- Not scientific
- Qualitative
- Unrepresentative

The purpose of this method is to make an clear choice on researcher own judgment about exactly whom to include in the sample.

Despite various limitations and criticisms, there are numerous advantages of non-probability methods:

- Cheaper
- Used when sampling frame is not available

- Useful when population is so widely dispersed that cluster sampling would not be efficient.

- Often used in exploratory studies, e.g. for hypothesis generation
- Some research not interested in working out what proportion of population gives a particular response but rather in obtaining an idea of the range of responses on ideas that people have.

QUOTA SAMPLING

- The population is first segmented into mutually exclusive sub-groups, just as in stratified sampling.
-
- Then judgment used to select subjects or units from each segment based on a specified proportion.
- For example, an interviewer may be told to sample 200 females and 300 males between the age of 45 and 60.
- It is this second step which makes the technique one of non-probability sampling.
- In quota sampling the selection of the sample is non-random.

For example interviewers might be tempted to interview those who look most helpful.

The problem is that these samples may be biased because not everyone gets a chance of selection.

This random element is its greatest weakness and quota versus probability has been a matter of controversy for many years

Stratified versus Quota Sample

Similarities:

- Population is divided into segments (strata).
- Elements are selected from each segment.

Key Difference:

- Stratified sampling uses probability methods.
- Quota samples are based on a researcher's judgment.
- Therefore, stratified sampling allows the establishment of the sampling distribution, confidence intervals and statistical tests.

CONVENIENCE SAMPLING

- Sometimes known as grab or opportunity sampling or accidental or haphazard sampling.
- A type of non probability sampling which involves the sample being drawn from that part of the population which is close to hand. That is, readily available and convenient.
- The researcher using such a sample cannot scientifically make generalizations about the total population from this sample because it would not be representative enough.

This is termed a convenience sample, it has obvious advantage in cost and logistics, and it is a good choice for many research questions .

A convenience sample can minimize volunteerism and other selection biases by consecutively selecting every accessible person who meets the entry criteria .

Convenience Sampling

-- Patients are selected, in part or in whole, at the convenience of the researcher;
no/limited attempt to ensure that sample is an accurate representation of Population

-- For example, standing at a shopping mall and selecting shoppers as they walk by to fill out a survey

QuestionPro

Criticisms

The most obvious criticism about convenience sampling is sampling bias and that the sample is not representative of the entire population.

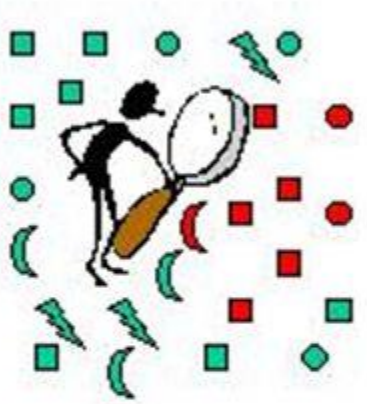
Another significant criticism about using a convenience sample is the limitation in generalization and inference making about the entire population.

Non-probability Sampling Methods

Convenience sampling method

the selecting on the basis of convenience

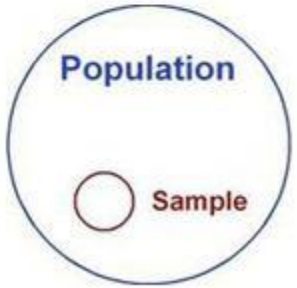
the selection at familiar locations and to choose respondents who are like themselves



often used during preliminary research efforts to get a gross estimate of the results

Judgment method

selecting samples that require a judgment or an "educated guess"



must be confident that the chosen sample is truly representative of the entire population.

Consecutive sampling is defined as a non-probability sampling technique where samples are picked at the ease of a researcher more like convenience sampling, only with a slight variation. Here, the researcher selects a sample or group of people, conducts research over a period, collects results, and then moves on to another sample.

A consecutive sample is especially desirable when it amounts to taking the entire accessible population over along enough period to include the seasonal variations or other changes over time considered important to the study question.

This sampling technique is also useful in documenting that a particular quality of a substance or phenomenon occurs within a given sample. Such studies are also very useful for detecting relationships among different phenomena.

For example, if the interviewer was to conduct a survey at a shopping center early in the morning on a given day, the people that he/she could interview would be limited to those given there at that given time, which would not represent the views of other members of society in such an area, if the survey was to be conducted at different times of day and several times per week.

This type of sampling is most useful for pilot testing.

Advantages of consecutive sampling

Here are the four advantages of consecutive sampling

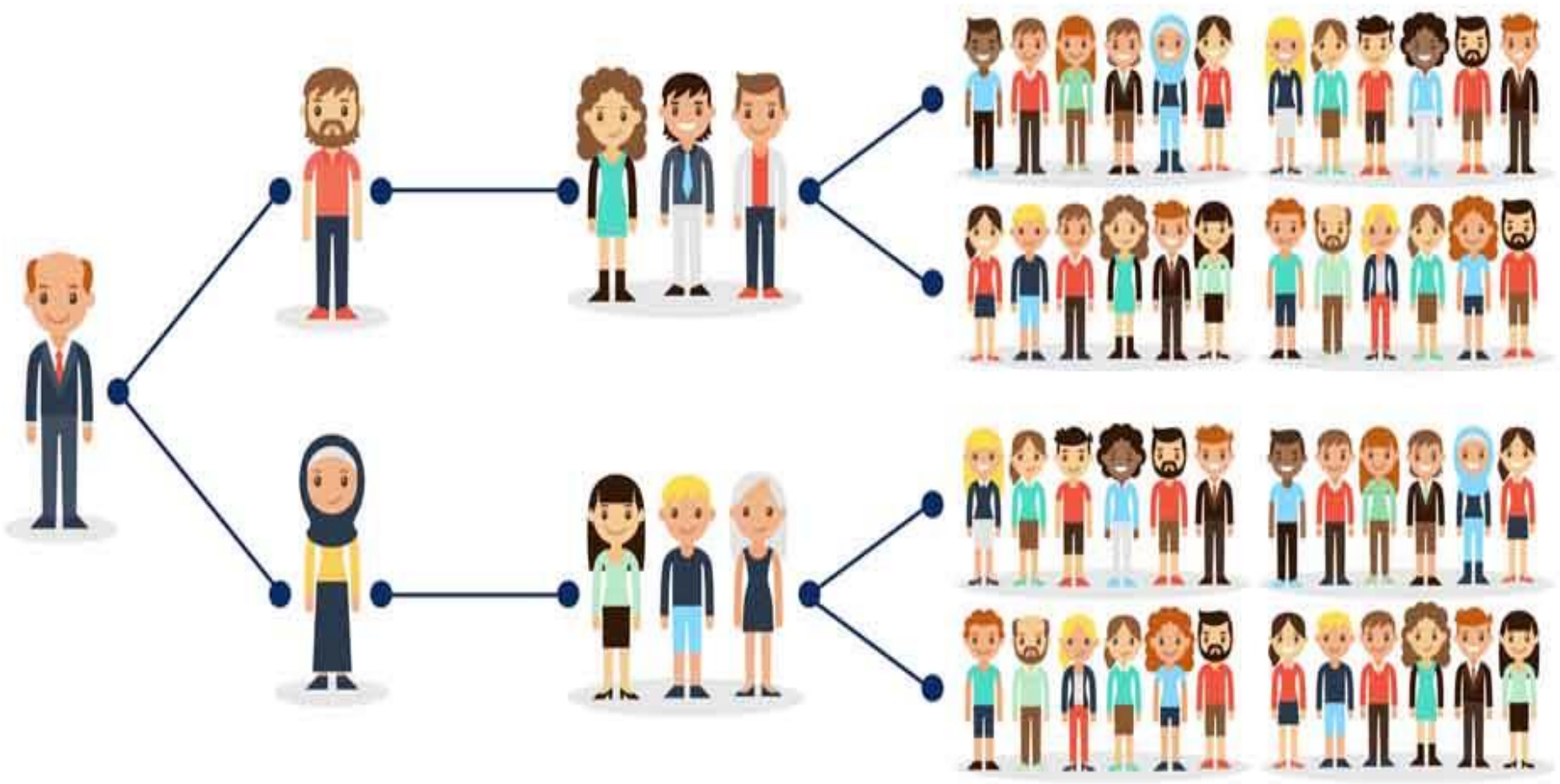
- In a consecutive sampling technique, the researcher has many options when it comes to sampling size and sampling schedule. The sample size can vary from a few to a few hundred, that the kind of range of sample size we are talking about here.
- In this sampling technique, sampling schedule is completely dependent on the nature of the research, a researcher is conducting. If a researcher is unable to obtain conclusive results with one sample, he/she can depend on the second sample and so on for drawing conclusive results.
- In consecutive sampling, a researcher can fine-tune his/her researcher. Due to its repetitive nature, minor changes and adjustments can be made right at the beginning of the research to avoid considering research bias.
- Very little effort is needed from the researcher's end to carry out the research. This technique is not time-consuming and doesn't require an extensive workforce.

Snowball sampling or chain-referral sampling is defined as a non-probability sampling technique in which the samples have traits that are rare to find. This is a sampling technique, in which existing subjects provide referrals to recruit samples required for a research study.

This sampling technique is often used in hidden population which are difficult for researchers to access.

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.

SNOWBALL SAMPLING



QuestionPro

- **Judgmental or purposive sampling:** are formed by the discretion of the researcher. Researchers purely consider the purpose of the study, along with the understanding of the target audience. For instance, when researchers want to understand the thought process of people interested in studying for their master's degree. The selection criteria will be: "Are you interested in doing your masters in …?" and those who respond with a "No" are excluded from the sample.

- - The researcher chooses the sample based on who they think would be appropriate for the study.

- This is used primarily when there is a limited number of people that have expertise in the area being researched

When selecting a sampling strategy it is necessary that it fits the purpose of the study, the resources available, the question being asked and the limitations being faced.

This holds true for sampling strategy as well as sample size.

A sample is expected to mirror the population from which it comes, however, there is no guarantee that any sample will be precisely representative of the population which it comes.

In practice, it is rarely known when a sample is unrepresentative and should be discarded.

Uses of non-probability sampling

Non-probability sampling is used for the following:

- **Create a hypothesis:** Researchers use the non-probability sampling method to create an assumption when limited to n prior information is available. This method helps with the immediate return of data and builds a base for further research.
- **Exploratory research**: Researchers use this sampling technique widely when 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 must 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.

Difference between Probability Sampling and Non-Probability Sampling Methods

Definition

Probability Sampling Methods	Non-Probability Sampling Methods
Probability Sampling is a sampling technique in which samples from a larger population are chosen using a method based on the theory of probability.	Non-probability sampling is a sampling technique in which the researcher selects samples based on the researcher's subjective judgment rather than random selection.

Population selection	The population is selected randomly.	The population is selected arbitrarily.

Sample	**Since there is a method for deciding the sample, the population demographics are conclusively represented.**	**Since the sampling method is arbitrary, the population demographics representation is almost always skewed.**
Time Taken	**Takes longer to conduct since the research design defines the selection parameters before the market research study begins.**	**This type of sampling method is quick since neither the sample or selection criteria of the sample are undefined.**
Results	**This type of sampling is entirely unbiased and hence the results are unbiased too and conclusive.**	**This type of sampling is entirely biased and hence the results are biased too, rendering the research speculative.**

ERRORS IN SAMPLING

'Sampling error' doesn't mean mistakes researchers have made when selecting or working with a sample.

What can make a sample unrepresentative of its population?

One of the most frequent causes is

sampling error

Sampling Errors

Sampling error is the error that arise in a data collection process as a result Of taking sample from a population rather than using a whole population.

Sampling error comprises the differences between the sample and the population that are due totally to the particular units that happen to have been selected.

Error is the difference between observed value and true value.

Sampling error

- The statistics of different samples from same population: different each other!
- The statistics: different from the parameter!

For example, suppose that a sample of 100 women are measured and are all found to be taller than178cm.

It is very clear even without any statistical prove that this would be a highly unrepresentative to invalid conclusions.

The sampling error exists in any sampling research.

It can not be avoided but may be estimated.

Three kinds of error:

(1) Systematic error (fixed)

(2) Measurement error (random)

(Observational error)

(3) Sampling error (random)

Errors in Sampling

There are two basic causes for sampling error:

1-Sampling bias is a tendency to favor the units that have particular characteristics.

Sampling bias is usually the result of a poor sampling plan.

BIAS IN SAMPLING-5 sources

- Any deviation from rules-
  self selection → volunteers
- Elimination of hard to identify people
  missing persistent absentees
- Replacement of previously selected individuals
  - Difficult to trace after being included in frame/uncooperative
- Large scale refusal
- List/sampling frame goes out of date

2-chance : That is the error that occurs just because of bad luck.

This may result in untypical choices.

Unusual units in a population do exist and there is always a possibility that an abnormally large number of them will be chosen.

The main protection against this kind of error is to use a large enough sample.

FACTORS THAT INFLUENCING THE SAMPLING ERROR ARE:

1- The size of the sample

2- The natural variability of the individual readings

As the size of the sample increases , sampling error will decrease.

As the individual reading vary widely from one another, we get more variability from one sample to another.

Errors of Observation

- Interview error- interaction between interviewer and person being surveyed
- Respondent error: respondents have difficult time answering the question
- Measurement error: inaccurate responses when person doesn't understand question or poorly worded question
- Errors in data collection

Non-Observation Errors

- Sampling error: naturally occurs
- Coverage error: people sampled do not match the population of interest
- Underrepresentation
- Non-response: won't or can't participate

Non sampling error (measurement error)

The other main cause of unrepresentative samples is non sampling error.

Like sampling error, non sampling error may either be produced by participants in the statistical study or be an innocent by product of the sampling procedures.

These are often more important than the sampling errors.

A non sampling error is an error that results from the manner in which the observations are made.

SAMPLING VS NON-SAMPLING ERROR

- **Sampling error:** This error arises when a sample is not representative of the population.
- **Non-sampling error:** This error arises not because a sample is not a representative of the population but because of other reasons. Some of these reasons are listed below:
    - Plain lying by the respondent.
- The error can arise while transferring the data from the questionnaire to the spreadsheet on the computer.
- There can be errors at the time of coding, tabulation and computation.
- Population of the study is not properly defined
- Respondent may refuse to be part of the study.
- There may be a sampling frame error.

Induced bias : Finally, it should be noted that the personal preconceptions of either the designer of the study or the data collector may tend to induce bias.

To protect against induced bias, advice of an individual trained in statistics should be required in the design and someone else aware of search pitfalls should serve in an auditing capacity.

What sampling method u recommend?

- Determining proportion of undernourished five year old in a village.
- Investigating nutritional status of preschool children.
- Selecting maternity records for the study of previous abortions or duration of postnatal stay.
- In estimation of immunization coverage in a province, data on seven children aged 12-23 months in 30 clusters are used to determine proportion of fully immunized children in the province.
- Give reasons why cluster sampling is used in this survey.

Practice

A large elementary school has 15 classrooms, with 24 children in each classroom. A sample of 30 children is chosen by the following procedure:

Each of the 15 teachers selects 2 children from his or her classroom to be in the sample by numbering the children from 1 to 24, using a random digit table to select two different random numbers between 01 and 24. The 2 children with those numbers are in the sample.

Did this procedure give a simple random sample of 30 children from the elementary school?

a) No, because the teachers were not selected randomly

b) No, because not all possible groups of 30 children had the same chance of being chosen

c) No, because not all children had the same chance of being chosen

d) Yes, because each child had the same chance of being chosen

e) Yes, because the numbers were assigned randomly to the children