

18/12

Unit-1

population: A group of objects (or) aggregates. The size of the population is denoted with ' N '.

sample: A part (or) subset of population and size of the sample is denoted with ' n '.

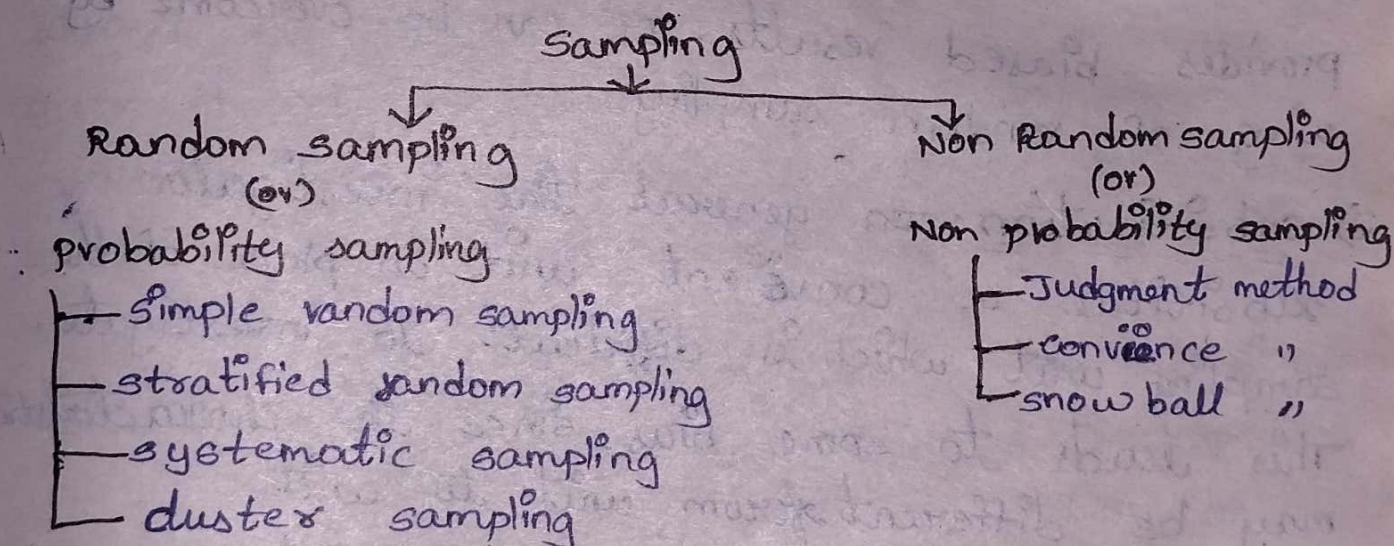
ex: testing the taste of curry of small portion (or) part.
selecting the ^{random} students for quiz.

parameter: The characteristic of population.

statistic: The characteristic of sample.

sampling: It is a process of taking sample from population.

ex: A sample of 10 students is taken from a class of average marks 40, within (parameter) The average marks of the sample is 38. (statistic).



sampling distribution: The distribution of statistic is called sampling distribution.

census: The process of collecting information of each and every object in a population.

In India census are conducted for

every 10 years

19/12
Scope of Errors The error can be classified into 2 types 1) Sampling error
2) Non sampling error

sampling errors:- The sampling errors arise due to the fact that only a portion of the population has been used to estimate the parameters. so these errors are present in sample survey and absent in census survey. sampling errors are due to the following reasons:

- 1) Faulty selection of sample: some of the bias is ~~introduced~~ introduced by the use of improper sampling techniques for the selection of sample. In most of the situations judgement sampling provides biased results. It can be overcome by using of random sampling.
- 2) substitution: In general the investigator substitute a convenient unit in place of the sampling unit which is difficult to investigate. This leads to some bias since the characteristic may be different from unit to unit.
- 3) Faulty demarcation of sampling unit: There will be some bias due to defective faulty demarcation of sampling units. It will be happen in the most area survey's dealing with border lines.

Improper choice of statistic

A constant error due to improper selection of statistic for estimating the population parameters. For example, We know that sampling variance s^2 is biased, for estimating population variance σ^2 . Where $\frac{ns^2}{n-1}$ is unbiased.

Scope of Errors (Sampling and Non sampling Errors)

The errors involved in the collection, processing and analysis of a data can be classified into two heads.

i) sampling errors ii) Non sampling errors

sampling errors: The sampling errors arise due to fact that only a portion of the population has been used to estimate the parameters and draw inferences about the population.

Thus these errors are present in sample survey and absent in census survey.

sampling errors are due to the following reasons:-

i) Faulty selection of sample: some of the bias is introduced by the use of improper sampling technique for the selection of sample. In most of situations, expect few, judgement sampling provides biased results. It can be overcome by using a random sampling.

2) substitution: In general the investigator substitute a convenient unit in place the sampling unit which is difficult to investigate. This leads to some bias since the characteristic may be different from unit to unit.

3) Faulty demarcation of sampling units:

There will be some bias due to defective demarcation of sampling units. It will be happen in most area surveys dealing with border lines.

4) Improper choice of statistics: A constant error due to improper selection of statistic for estimating population parameter.

For example, we know that sampling variance (s^2) is biased for estimating population variance σ^2 where as $ns^2/(n-1)$ is unbiased.

* the sampling errors can be reduced by increasing the sample size, since the sampling errors (S.E.) is inversely proportional to the square root of sample size.

Non-sampling Errors :-

The non-sampling errors arise at the stages present in both the census and sample surveys. Non-sampling errors can occur at every stage of planning and execution of census and statistic survey. These arise from the following factors

1. Faulty planning or definitions: The planning of the survey consists of definitions various objectives. Hence the non-sampling errors can be due to:
 - a) Errors due to location of units, errors in recording measurements, errors due to ill designed questionnaire etc.
 - b) Data specification being inadequate and unreasonable.
 - c) Lack of trained and qualified personnel.

Response errors:- These errors resulting from responses may be due to the following reasons.

- * May be accidental
- * Due to the prestige of informant (it will upgrade or downgrade the response).
- * Self interest of informant.
- * Bias due to interviewer.

3) Non-Response Bias: - It occurs if full information is not obtained on all the selected sampling units.

4) compiling Errors: The various operations of data such as editing, coding, tabulation, etc are potential sources of error. These can be control through verification, consistency check, etc.

5) publication Errors: - The publication errors (errors committed during presentation and printing of results) due to the two sources - mechanics of publication and failure of survey organisation to point out the limitations of statistics.

* The non-sampling errors can be reduced by assigning trained, experienced and skilled personnel.

The data obtained in a complete census, although free from sampling errors, would still be subject to non-sampling errors where as data obtained in a sampling survey would be subject to both sampling and non-sampling errors.

Advantages of sampling over census:-

1) Less time:- There is considerable saving in time and labour since only a portion of population has to be examined at the same time results can be obtained rapidly and analyzed much faster.

2) Reduction in cost: Sampling usually results in reduction in cost in terms of money and hours since in most of cases our resources are limited in terms of money and the time. sampling is more advantageous than census.

3) Greater Accuracy of results:

The results of a sample survey are usually much more reliable than those obtained from a complete census due to the following reasons.

→ It is possible to determine the extent of the sampling errors.

→ scope of non-sampling errors is less in sampling compared with census.

4) Greater scope:

The complete census is impracticable if the survey requires a highly trained personnel and more sophisticated equipment for collection and analysis of data. It is possible to have a thorough and intensive enquiry because a more detailed information can be obtained from a small group of respondents.

- 5) If population is too large, if testing is destructive if population is hypothetical, sampling only the way.

Principal steps in a sample survey
The following are main steps involved in planning and execution of a sample survey.

1. Objectives of the survey.

The first and foremost step is to define the objectives of the survey in clear and concrete terms. The sponsoring agency of survey, should take care that these objectives are corresponding with the available resources in terms of money, man power and time limit.

2. Defining the sampled population.

The population from which sample is chosen should be defined in clear and unambiguous terms. Some practical difficulties in handling certain segments of population (special border cases) may point to their elimination from the scope of survey. This for reasons of practicability or convenience the population to be sampled may be different than the population for which results are wanted (target population).

3. Sampling units and the frame:

The units of sampled population are called sampling units. These must cover the entire sampled population and they must be distinct,

non-overlapping and unambiguous. so that every element of the population belongs to only one sample unit.

ex: In socio-economic survey for selecting people in a city, the sampling unit might be an individual person, a family or a block in a locality.

The list of sampling units is called frame. It may be in form of a map or any other acceptable material. The frame serves as a guide to the population to be covered. This frame should be accurate and dynamic.

4. selection of proper sampling design:-

A No. of designs (probability sampling methods and non-probability sampling methods) for selection of sample are available. A proper design should be selected considering efficiency cost and time.

5. selection of method of collecting information:-

An appropriate method should be selected for collecting information keeping accuracy, cost & time. A proper information care should be taken in case of non-respondents.

Some important methods are interview method and questionnaire method.

6. data to be collected.

The data should be collected keeping view the objectives of the survey. A practical method is to check out an outline of the tables that the survey should produce. This would help in elimination the collection of irrelevant and too many information and ensure that no essential data.

7. organization of field work.

It is absolutely essential that the personnel should be well trained in locating sample units, recording the measurements and the methods of collection of required data before starting field work. The success of a survey to a great extent depends upon the reliable field work. It is necessary to inspect the field work after completion by supervisory staff.

8. the pre-test:

From practical point of view a small pre-test should be conducted. Here pre-test means trying out the questionnaire and field methods on a small scale. It always helps to decide upon effective method of asking questions and disclose certain problems, otherwise it will be quite serious on a large scale sample survey.

9. summary and analysis of the data
the analysis of the data may be classified as.

i) scrutiny and editing of data.

An initial quality check should be carried out by the supervisory staff while the investigators are in the field. the scrutiny and editing of completed schedules and questionnaires will help in eliminating the erroneous and inconsistent data.

ii) Tabulation of data:

Before carrying out the tabulation of data. we must decide the procedure of tabulation of the data which are incomplete due to non-response to certain items in questionnaire. For a large scale sample survey advanced software tools will help to tabulation the data and these need coding the qualitative variables.

iii) Statistical Analysis:

After the data has been properly scrutinized edited and tabulate every careful statistical analysis is to be made. Appropriate formulae should be used to provide unbiased and accurate results.

10. Reporting and conclusions:

Finally a report incorporating detailed statement of the different stages of the survey should be prepared. In the presentation of the results, it is good practice to report the technical aspect of the design i.e, types of the estimators used

along with expected error.

II. Information gained for future survey:

The information gained from any complicated survey and sample in the form of the data regarding means, SD's, costs, time etc. serve as a potential guide for future surveys - It serves as a lesson to the organiser for future surveys in recognising and rectifying the mistakes committed in the execution of the survey.