Unit-4
SAMPLING

Q1. Define sampling with its merits and demerits.

Ans:- When one by one study of all units of a population is not ~~pop~~ possible due to some factors like time, cost, manpower, resources and destructive nature of study, we take a small representative part from the population for the study. This small representative part selected for the study from the population is called sample. The process of selecting a sample from a population is called sampling.

For example: A pathologist takes a syringe of blood as a sample to find out a disease.

Merits of sample survey

i) It is less expensive.
ii) It requires less time to get result.
iii) It needs smaller number of resources.
iv) It is the only method for large population.

Demerits of sample survey

i) It is not suitable method if information about all units of population is required.
ii) It is not suitable for rare events.
iii) It is not appropriate for historical data.

Q.2 Describe the types of errors in statistics.

Ans: There are two types of errors involved in collection, processing and ~~at~~ analysis of data in a survey ~~no~~ in statistics as: Sampling error and Non-sampling error.

i) Sampling error → Usually, different samples selected from the same population will give different results because they contain different elements. The result obtained from any one sample will generally be different from the result obtained from the corresponding population. The difference between the values of the sample statk obtained from a sample and the value of corresponding population parameter obtained from the population is called the sampling error.

Sampling error is the difference between the value of a sample statistic and the value of the corresponding population parameter. In case of mean, Sampling error = $\bar{X} - \mu$.

The sampling error occurs because of chance and is inherent and unavoidable in any and every sampling scheme. A sample with the smallest sampling error will always be considered a good representative of population. This error can be reduced by increasing the size of the sample. When the sample survey becomes a census survey, the sampling error becomes zero.

Increasing sample size, cross check, use of unbiased sampling, using appropriate sampling design and making clear questionnaire are required to reduce the sampling error.

Non-sampling error → The errors that occur in the collection, recording, and tabulation of data are called non-sampling errors. The non-sampling error is likely to increase with increase in sample size, while sampling errors decreases with increase in sample size. Non-sampling errors can be attributed to many factors such as inability to obtain information about all cases in the sample, definitional difficulties, inability or unwillingness on the part of the ~~respod~~ respondents to provide correct information, inability of recall information, errors made in collection of data, errors made in estimating value of parameter etc.

Non-sampling error is occured in research from other sources than the sample, which can be minimized by checking the process, preparing the questionnaire properly, performing pilot survey, fixing procedure, by using competent manpower and experts etc.

Q3 Parameter and statistic

There are various statistical measures in statistics such as mean, median, mode, standard deviation, proportion, correlation coefficient etc. These measures can be computed from both population data and sample data.

Any statistical measure computed from population data is called parameter. Let $X_1, X_2, \ldots, X_N$ be the values in the population of size N, then,

(a) Population Mean $(\mu) = \frac{X_1 + X_2 + \ldots + X_n}{N}$

(b) Population Variance $(\sigma^2) = \frac{\sum (X - \mu)^2}{N}$

Here, population mean (μ) and population variance (σ^2) are called population parameters.

Any statistical measure computed from sample data is statistic. Let $x_1, x_2, \ldots x_n$ be the values in the sample of size n, then

(a). Sample Mean $(\bar{X}) = \frac{x_1 + x_2 + \ldots + x_n}{n}$

(b). Sample variance $(s^2) = \frac{\sum (X - \bar{X})^2}{n-1}$

= sample mean square.

Here, sample mean $(\bar{X})$, sample variance (s^2) are called statistic. The value of every statistic varies from one sample to another whereas the value of a parameter is always constant for a population. There is a natural correspondence between statistic and parameter. For each population parameter there is a sample statistic computed from data that represents the best information about the unknown population parameter.

Q4 What are the advantages and limitations of sampling?

Ans: Advantages

i) Reduce cost → If data are secured from only a small fraction of the aggregate, expenditure are smaller if a complete census is attempted. With large population, results accurate enough to be useful can be obtained from samples that represent a small fraction of the population.

ii) Greater speed → for some reason, the data can be collected and summarized more quickly with a sample that with complete count. This is a vital consideration when information is urgently needed.

iii) Greater scope → In certain type of inquiry highly trained personnel or highly specialized equipment limited in availability, must be used to obtain the data. A complete census is impracticable. Thus surveys that rely on sampling have more scope and flexibility regarding the types of information that can be obtained.

iv) Greater Accuracy → The field work and processing of results become feasible when volume of work is reduced, a sample may produce more accurate result that the kind of complete enumeration that can be taken.

v) If population is too large, as for example, of trees in a jungle, we are left with no way but to resort to sampling.

vi) If the population is hypothetical in nature, for example rolling of dice, tossing of coin etc.

Limitations

i) Proper care should be taken in the planning and execution of the sample survey, otherwise the results obtained might be inaccurate and misleading.

ii) Sampling theory requires the services of trained and qualified personnel and sophisticated equipment for it's planning, execution and analysis. In absence of these, the result of the sample survey are not dependable.

iii) If the information is required about each and every units of the universe, there is no way but to use complete enumeration. The complete enumeration may be better than any sampling technique of time, money and labour.

Q5 Describe basic steps in sampling.

The basic steps involved in the planning and execution of a sample survey ~~etc~~ are as follows:-

i) Objective of the survey → The first step in sampling is to define the objective of the survey in clear & Lucid term. The sponsor of the survey should take care that these objective are commensurate with the available resources in terms of money, manpower and the time limit required for availability of the results of the survey.

ii) Defining population to be sampled → The rules ~~mut~~ must be setup to define regarding shape, size etc, keeping in mind that the boarder-line cases so as to enable the investigator to decide in the field.

iii) Data to be collected → It is well to verify that all the data are relevant to the purpose of the survey and that no essential data are omitted. There is frequently a tendency, particularly with human population to ask too many question, some of which are never subsequently analyzed.

iv) Degree of precision desired → The uncertainity can be reduced by taking larger samples and by using superior instruments of measurement.

v) Method of measurement → There may be choice of measuring instrument & of method of approach to the population. The approach may be by mail, by telephone, by personal visit or its combinations.

vi) Frame

vii) Selection of proper sampling design.

viii) Organization of field work.

ix) Summary and Analysis of data.

Q.6 Merits and Demerits of Simple Random Sampling.

Merits

i) Simple random sampling gives each unit an equal chance of being selected and personal bias is completely eliminated.

ii) The statistician can ascertain the efficiency of the estimate of the parameters by considering the sampling distribution of the statistics.

Demerits

i) The selection of simple random sampling requires an upto date frame.

ii) The sampling units which are widely spread geographically and in such case the cost of collecting the data may be much in terms of time and money.

iii) At times, a simple random sample might give most non-random looking results.

iv) For a given precision, simple random sampling usually requires larger sample size as compared to stratified random sampling.

⊛. Types of Sampling Methods :-

1) Random or Probability Sampling Technique:

It is defined as the method of sampling technique in which each unit of the population has fixed probability of selecting in the sample.

@. Simple random sampling → It is the most common and the simplest method of sampling in which each sample unit is selected from a population with equal probability.

(b) Stratified Random Sampling → When units in the population are not similar in nature, the population is divided into sub groups called strata before the sample is drawn. Then a simple random sample is drawn from each stratum in proportion to its size. The stratification of population should be done as follows:

→ The strata should be non-overlapping and should together comprise the whole population.

→ The strata should be as possible as homogenous within groups and heterogenous between the groups.

Purpose of stratification

i) To make more representative.

ii) For Greater accuracy.

iii) For administrative convenience.

(c) Systematic Random Sampling → A sampling technique in which only first unit is selected with the help of random numbers and the rest get selected automatically according to some pre-designed pattern is known as systematic random sampling.

(d) Cluster Random Sampling → In cluster random sampling, we divide the population into discrete groups prior to sampling. The groups are termed as clusters and then we select some clusters as sample from all clusters using simple random sampling. This's method of sampling is useful when the population consists of very large number of similar groups which are geographically distant.

(e) Multi-stage random sampling → Multi-stage random sampling, sometimes called multi-stage cluster sampling, is a development of cluster sampling. A sample of desirable size can be drawn from selected clusters which may increase the efficiency is called multi-stage random sampling.

2) Non-random or Non-probability Sampling Technique:

It is defined as the method of sampling technique in which each unit in a sample is selected on the basis of personal judgment. There are several non-random sampling methods for selecting samples from a population as:

(a) Judgment Sampling → A sampling process in which the sample units are selected according to researchers or investigators personal judgment is called the judgment sampling. Only the average units are considered and extreme units are omitted in judgment sampling. This method is suitable for a small number of sampling units in the population, solving everyday business problems and making public policy decisions.

(b) Convenience sampling → A sampling process in which the units are selected neither by probability nor by personal judgment but by convenience is called the convenience sampling. Sample obtained from easily available lists such as telephone directories, automobile registerations etc. is an example of convenient samples.

(c) Quota sampling → Quota sampling may be considered as a special of stratified sampling in which any investigator is told to examine or enumerate a fixed number of units called quotas from each stratum. The sampling quotas may be fixed according to some specified characteristics such as income, sex, occupation, religion etc.

(d) Snowball Sampling → Snowball sampling technique is a non-random sampling technique generally used in the case where it is difficult to reach into exact population. In this method survey subjects are selected on referral from other survey respondents.