

30/05/23

Tuesday

Sampling Techniques

- 1) Concept of population and sample, need for sampling, complete enumeration versus sampling, basic concepts in sampling, sampling and non-sampling error, methodologies in sample surveys (questionnaires, sampling design and methods followed in field investigation) by NSSO.
- 2) Subjective or purposive sampling, probability sampling or random sampling, simple random sampling with or without replacement, estimation of population mean, population proportions and their standard errors. Stratified random sampling, proportional and optimum allocation, comparison with simple random sampling for fixed sample size. Covariance and variance function.
- 3) Ratio, product and regression methods of estimation, estimation of population mean, evaluation of Bias and variance to the first order of approximation, comparison with simple random sampling.
- 4) Systematic sampling (when $N = nk$, k is an integer). Estimation of population mean and S.E. of this estimate comparison with SRS.
- 5) Equal size clusters sampling, estimators of population mean and total and their standard errors. Comparison of clusters sampling with SRS in terms of intra class correlation coefficient.
- 6) Concept of multi stage sampling and its application.

Page - 2

- Two stage sampling with equal no. of SSVs.
Estimation of population mean and total.
- Double sampling in ratio and regression method of estimation.

Varying Probability Sampling.

- 7) Sampling with probability proportional to size (with and without replacement method), Desh Raji and Das estimators for $n=2$, Horvitz and Thompson's estimator.
- 8) Interpenetrating sub-sampling - concept.

everywhere in this register
↳ $\text{pop}^n \xrightarrow{\text{implies}} \text{population}$ Page 3

Population - (or universe)

In a statistical investigation, there are individuals belonging to a group, the interest usually lies in the assessment of magnitude and study of variation with respect to one or more characteristics relating to individuals belonging to that group. This group of individuals under study is called population.

The population may be finite or infinite.

Need for sampling?

Case of Infinite pop^n

- If pop^n is infinite, complete enumeration is not possible in a statistical investigation.
- If the units are subject to destruction in the course of inspection, 100% ~~inspect~~ enumeration is not possible.
- Administrative and Financial implications, time constraint etc. (Time and Cost).

Sample

A finite subset of statistical individuals in a population.

(Sample size:- No. of individuals in a sample).

For determining the population characteristics instead of enumerating the entire pop^n , characteristics the individuals in the sample are studied or observed. Then, the sample characteristics are utilized to approximately estimate the population.

The error involved in such approximation is called sampling error. It is inherent in every sampling scheme.

Parameters and Statistics

Parameters :- Statistical constants of the population. e.g. popⁿ mean, popⁿ variance, etc.

Statistics :- Statistical measures computed from the sample observations.

A statistic is a function of the sample values. Hence, it varies from sample to sample as there can be multiplied choices of the sample.

$$t = t(x_1, x_2, \dots, x_n).$$

Sampling Distribution

No. of samples of size n from a popⁿ of size N can be ${}^N C_n$. For each of these sample a statistic 't' can be computed, which will vary from sample to sample.

The aggregate of the various values of the statistic 't' can be grouped into a frequency distribution which is known as the sampling distribution of the statistic.

Standard Error

The standard deviation of the sampling distribution of a statistic

eq.	Statistic	Standard Error
	\bar{x}	σ/\sqrt{n}
	p	$\sqrt{PQ/n}$

$n \rightarrow$ sample size, $\sigma^2 \rightarrow$ population variance.
 $P, Q \rightarrow$ popⁿ proportion, $Q = 1 - P$,
 $p \rightarrow$ sample proportion.

Utility of Standard Error

In large sample theory, S.E. forms the basis for testing the hypothesis.

If t is any statistic, then

$$Z = \frac{t - E(t)}{S.E.(t)} \sim N(0,1)$$

If $|t - E(t)| > 1.96 \times S.E.(t)$, the hypothesis is rejected at 5% level of significance.

It means that, if the difference between the observed and expected value of the statistic is greater than 1.96 times the S.E., the hypothesis is rejected at 5% l.o.s.

Principal Steps in a Sample Survey

(Steps involved in planning and execution).

- 1) Objective of the survey.
- 2) Defining the popⁿ to be surveyed / sampled.
- 3) Sampling frame and Sampling units.

- Sampling frame :- It is the source material from which the sample is drawn. It is the list of all those within a popⁿ who can be sampled.
- Sampling Unit :- Finite no. of distinct and identifiable units.
- 4) Data to be collected - Data should be collected keeping in view the objective of the survey.
(F4T Data collect kiya Jaye).
- 5) Questionnaire or Schedule.
- 6) Method of collecting Information.
 - (i) Interview Method :- enumerators visit home to home and interview the individual personally.
 - (ii) Mailed Questionnaire Method.
 - API - It is a face to face data collection method in which the interviewer uses a tablet or a mobile phone to record answers given during the interview.
- 7) Non ^{respondents} ~~respondants~~ :- The data can't be collected for all the sampled units. Respondents may not be available or refuse to answer. Reasons for non-response should be recorded by the investigator.
- 8) Selection of Proper sampling Design.
A no. of designs are available for the selection of a sample. Relative costs and time involved should be considered before making a selection.
- 9) Organisation of Field Work or (AP)
Training of the personal.

- 10) Summary and Analysis of the data.
Editing, Tabulation, Analysis, reporting.
- 11) Information gained for future surveys.

⇒ Principle of sample survey.

The theory of sampling is based upon following important principles

- 1) Principle of Statistical Regularity :- It lays down that a moderately large no. of units chosen at random from a large group are almost sure on the average to possess the characteristics of the large group.

(Imp. of selecting the sample at random).

Eg. Political Polling

- 2) Principle of validity :-

Obtain the valid test and estimates about the parameters of the population.

Prob. sampling satisfy this principle.

Eg. Health Surveys

- 3) Principle of Optimisation :-

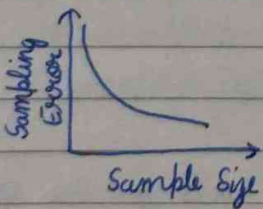
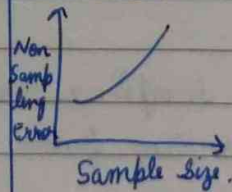
- 1) Minimising cost with given level of efficiency.
- 2) Maximising efficiency with given level of cost.

Eg. Government census

Sampling V/s Complete Census.

Basis	Sampling V/s. Complete Census.
1. Large Population	Complete enumeration of the population is not feasible in case of large population.
2. Testing in Destructive	In case, the quality of an article can be determined only by destroying the article, then complete enumeration is impracticable, e.g. crackers, life of bulb.
3. Time Constraint	Sampling results in reduction of cost in terms of money and man hours
4. Reduction Cost	Sampling results can be obtained more quickly and it saves time, since fewer data have to be collected and processed.
5. Greater Accuracy	a) In sampling, it is possible to determine the extent of sampling errors. b) In sampling, non-sampling errors can be controlled which is difficult to control in complete enumeration.
6. Greater Scope	As sample survey saves time and money, it has a greater scope.
7. Error	Census are subjected to only Non-Sampling Error.

Sampling Error v/s Non Sampling Error

Basis For Comparison	Sampling Error	Non-Sampling Error
1. Meaning	An error which occurs due to the sample selected doesn't perfectly represents the pop ⁿ of interest.	An error which occurs due to sources other than sampling, while conducting survey activities.
2. Cause	Deviation between sample mean and population mean.	Deficiency and inappropriate analysis of data. (Due to accumulation of errors bcoz we estimate pop ⁿ parameters from survey data.)
3. Type	Random	Random or Non-Random.
4. Occurs	When sample is selected. Sample survey. Absent in complete enumeration	Both in sample and census. Sample survey Census survey.
5. Sample Size	Possibility of error reduced with the increase in sample size. 	It has nothing to do with the sample size that generally increase. 
6. Determination	It is determined by the magnitude mean square error.	No General method. * Cause :- It arises at any stage of conducting a survey such as ^{processing} planning, data collection, tabulation etc.

Reasons :- only a part of popⁿ is used to estimate the popⁿ parameter

1) Sampling Errors (due to following reasons)

F → (a) Faculty selection of the sample.

S → (b) Substitution

O → (c) Family Demarcation of the sampling units

I → (d) Improper statistic to estimate the population parameter.

2) Non - Sampling Errors :-

a) Faculty Planning

b) Response Error

c) Compiling Error

d) Duplication Error

e) No Response Error.

Remarks :-

1) As the sample size increases, the sampling errors are subject to decrease while the non-sampling errors are subject to increase.

2) Quite often, non-sampling errors in census is greater than both the sampling and non-sampling errors taken together in a sample survey.

Sources of Non-Sampling Error

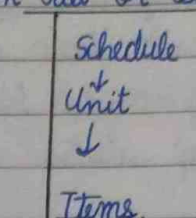
1) Inappropriate sampling frame - List of all the units of the target population

2) Response Errors - Investigators report untrue values respondent report ages multiple of five.

3) Error in Data processing :- while entering data in the computer.

4) Non-Response Error:- Failure to information on all or some of the items in the schedule.

Due to two reasons $\begin{cases} \rightarrow \text{Non-response} \\ \rightarrow \text{Non-coverage} \end{cases}$



Two types $\begin{cases} \rightarrow \text{Unit Non-Response} \\ \rightarrow \text{Item " "} \end{cases}$

Types of Sampling:-

The procedure of drawing a sample can be classified under :-

- Subjective (or purposive or judgement) sampling
- Probability sampling.
- Mixed sampling

Non-Probability sampling

1) Subjective Sampling eg. Quota Sampling

- In this sampling, the sample is drawn with definite purpose in view.
- The choice of sample depends on the discretion and judgement of the investigator.
- This type of sampling suffers from nepotism and thus, does not give the fair representation of the population.

2) Probability sampling

- In this sampling, the sample is drawn according to laws of chance in which each unit in the popⁿ has a pre-defined probability of getting selected in the sample.

The different types are:-

- i) each unit has a equal chance of being selected.
- ii) sampling units have different chance of being selected.
- iii) Prob. of selection of a unit is proportional to the sample size.

3) Mixed Sampling

In this sampling, samples are drawn partly according to the laws of chance and partly according to fixed sampling rule.

Sample Design Specifies.

- 1) Popⁿ to be surveyed.
- 2) Sample Selection Method.
- 3) Sampling Frame to use.
- 4) Size of the sample.
- 5) Stratification.
- 6) Allocation of size to different strata.
- 7) Statistic to be used.
- 8) Treatment of Non-Respondents.
- 9) Estimation of Sampling Error.

Sample Selection Procedure

- 1) Simple Random sampling.
- 2) Stratified sampling
- 3) Systematic sampling
- 4) Cluster sampling.
- 5) Multi-Stage sampling
- 6) Prob. proportional-to-size sampling.

Example:- Multi-Stage Sampling :- To obtain a sample of households from Rural areas of a state particularly FSUs may be districts, SSUs may be villages in the districts and TSUs may be the households.

Sampling Design in N.S. Surveys :-

- A stratified multi-stage sampling design is used in socio-economic surveys by NSO.
- FSU are the census villages in Rural sector. Urban Frame Survey Block in Urban sector.
- USU are households in both Rural & Urban sectors.