UNIT 2: STATISTICAL SURVEY

UNIT STRUCTURE

2.1 Learning Objective

2.2 Introduction

2.3 Definition of Statistical Survey

2.4 Stages of Statistical Survey

2.5 Planning of a Statistical Survey

2.6 Execution of Statistical Survey

2.7 Basic terms used in Statistical Survey

2.8 Collection of data

2.9 Primary Data

2.10 Secondary Data

2.11 Pilot Survey

2.12 Scrutiny and Editing of Secondary data

2.13 Let us sum up

2.14 Further Readings

2.15 Answer to check your progress

2.16 Model Question

2.1 LEARNING OBJECTIVES

After going through this unit, you will be able to -

- understand the meaning of Statistical Survey
- know the various stages of Statistical Survey
- know the definition of basic terms used in Statistical Survey
- understand the meaning of Primary Data and Secondary data
- learn the various methods of collecting Primary data
- know about the precautions to be taken while using secondary data.

2.2 INTRODUCTION

Statistics are collected incidentally or intentionally. A considerable amount of statistics are available as a result of administrative functions of the government and are collected only incidertally, e.g. statistics of crime, accidents, taxation, imports-exports etc. These statistics are not collected primarily for research purposes, although they may be very useful for research. On the other hand, some statistics are collected for their own sake. These intentionally collected statistics are the results of statistical survey.

2.3 DEFINATION OF STATISTICAL SURVEY

Statistical Survey means some sort of investigation by an individual or agency wherein the relevant information is collected in quantitative terms rather than in words.

2.4 STAGES OF STATISTICAL SURVEY

Statistical Survey is a technical job which requires specialized knowledge and skill. The most imprtant task in any statistical survey is data collection. If the method of collecting data is not scientific, then the very purpose of the servey will not be fulfilled. A statistical survey passes through several stages. These stages can be summerised under two broad heads.

(a) Planning the survey

(b) Executing the survey

2.5 PLANNING OF A STATISTICAL SURVEY

Proper planning of a survey is of great importance because the quality of a survey depends mostly on planning. The matters which require careful consideration at the planning stage are :

Purpose of the survey : In any statistical survey, first essential thing is to understand the nature of the problem to be investigated and to define it in clear-cut language. This will help in identifying the relevant data.

Scope of the survey : The techniques to be adopted in collection and analysis of data depend on the scope of the survey. Scope of the survey refers to the geographical area to be coverd, the period of the study, the items to be covered, the type of information to be collected etc.

Sources of data : When the purpose and scope of a statistical survey have been stated and agreed upon, the next step is to determine the sources from which the data are to be collected. One must take decision whether the sources of data will be primary or secondary. (*A broad discussion on primary data and secondary data will be made afterwards*)

Types of survey : There are various factors that determine the type of survey to be conducted. One very important factor is object and scope and another is nature and cost factor. The cost factor determines the source of survey and its throughness.

The various type of statistical survey on which decision has to be taken are – (a) Census sample (b) Confidential or open survey (c) Direct or indirect survey (d) regular or adhoc survey (e) offical, semi-official or non-official (f) Extensive or Limited survey

Unit of data collection : To carry out a statistical survey, the statistical unit or units must be clearly defined. The unit in terms of which the investigation counts or measures the variables or attributes selected for enumeration, analysis and interpretation is known as *statistical unit*. For example, in a population census, the statistical unit is a person.

To be an ideal statistical unit, it must possess the following qualities:

(i) It should be clear and unambiguous

(ii) It should be specific.

(iii) It should be stable.

(iv) It should be uniform.

(v) It should be appropriate to the enquiry.

Statistical units can be broadly classified under two heads :

(a) UNITS OF COLLECTION

and

(b) UNITS OF ANALYSIS AND INTERPRETATION.

(a) **UNITS OF COLLECTION** : Data for enquiry are collected in terms of units of collection. The units of collection may be *simple*, *compound* or *composite*. A simple unit represents a single condition without qualifications. Examples of such units are – house, worker, hours etc. A simple unit with some qualifying words is called a compound unit. Example of such units are – skilled worker, man-hours etc.

(b) **UNITS OF ANALYSIS AND INTERPRETATION** : Statistical data are ultimately analysed and intimately analysed and interpreted with the units of analysis and interpretation. These units facilitate comparisions between different sets of data with respect to time and place. Genearlly the units of analysis are – rata, rations, percentages and co-efficients.

SOURCES OF DATA : For any statistical survey the sources of data may be either primary or secondary Data is termed primary when reference is to data collected for the first time by the investigator and termed secondary when the information are taken from available records on publised material or data already available. The census department, at at regular intervals, collected data on population of the country. These constitute primary data to the census department. But these data will be seconday to someone who is using these for some other investigation. The choice of the source of data depends largely on pupose and scope of the survey.

TECHNIQUE OF DATA COLLECTION : Mainly there are two techniques of data collection (i) census technique and (ii) sample technique. In census technique, information is obtained from each and every unit of the population which forms the subject matter of the study. Whereas, in the sample technique, information is obtained only from a representative part of the population and based on that, inference is drawn for the entire population. The census method is costlier and more time consuming. The choice of technique of data collection depends upon factors like – (a) the availability of resources (b) the time factor (c) degree of accuracy desired and (d) nature and scope of the enquiry.

CHOICE OF FRAME : The term frame or population frame refers to all the units in the population under study. The whole structure of enquiry is to a considerable extent detemined by he frame. Detailed planning of a survey can not be undertaken unless we know the nature and accuracy of the available frames. If there does not exist any frame, then the construction of a frame suitable for the purpose of the enquiry will constitute a major part of planning.

DEGREE OF ACCURACY DESIRED : The investigation must determine the degree of accuracy he wants to attain. In any statistical work, it is very difficult to attain 100% accuracy. The object of enquiry primarily determines the degree of accuracy. A very high degree of accuracy may mean a lot of time and cost. The necessary degree of accuracy in counting or measuring depends upon practical value of accuracy in relation to its cost.

2.6 EXECUTION OF STATISTICAL SURVEY

After planning the survey the next step is to execute the plan. The various phrases of work at execution stage are as follows :

(a) **Setting up and administrative Organisation** : An administrative organisation is needed for an investigation on which the complete control and administration of the enquiry falls. The size of the organisation depends upon the nature and scope of the enquiry.

(b) **Designing of forms** : The various forms specially the forms of questionnaires that is used in the course of enquiry should be designed with utmost care and should be drafted by skilled and experienced person.

(c) **Selection, traning and supervision of filed investigation** : The success of a survey depends upon the work of the enumerators. Therefore, these enumerators should be properly selected an throughly trained for the field work. Constant supervision is essential to achieve high quality work.

(d) **Control on the quality of field work** : Steps must be taken to ensure that the survey is under statistical control i.e. the errors if any, in the

survey are due to the presence of random variation and no assignable causes of variation are present. A system of field checks by the supervisors should also be introduced to maintain the standard.

(e) **Follow up of non-response** : Inspite of best efforts, there may be respondents who do not supply desired information. To deal with this problem, a list of non-respondents is made and then with the help of supervisory staff efforts can be made for getting response from such non responsed members.

(f) **Processing of data** : After collection of data, by the process of classification and tabulation, the collected data are summerised an arranged in a systematic order.

(g) **Preparation of report** : After the data have been collected and analysed, the results of the survey is drafted in the form of a report, which is the final step in the execution of any survey.



CHECK YOUR PROGRESS

Q 1: Name the various types of statistical survey on which decision has to be taken.

Q 2: Fill in the blank :

(a) The unit of collection may be _____ or _____

(b) Skilled worker is an example of _____ unit.

(c) Rates, ratios and percentages are examples of units of _____

(d) Results of population data constitute _____ data for the census department.

(e) A _____ is representative part of the population.

2.7 BASIC TERMS USED IN STATISTICAL SURVEY

Associated with statistical survey there are certain terms which are frequently used :

Characteristic : A quality possessed by an individual or object is called a *characteristic*. Height, weight, temperature, nationality, honesty etc. are examples of characteristic.

Attribute : A non measurable characteristic like sex, region, honesty, nationality etc. are examples of attribute.

Variable : A measurable characteristic like height, weight, temperature, rainfall etc. are examples of variable.

Variables are of two types – (i) Disrete (ii) Continuous.

(i) A variable which takes specific values only (usually integral values only) between two fixed limits is called a *discrete variable*. Size of a family (i.e. no. of persons in a family), Number of rooms in a college etc. are examples of discrete variable (as the value of such variable cannot be like 4/5, 7/8 etc.)

(ii) A variable which can take any value between two fixed limits is called a continous variable. Height of students, Temperature of a particular place etc. (as the values of such variable may be like 32.8, 160.5 etc)

Population : The collection of all items related to any survey is called *population* or *universe*. For example, if our co-survey relates to the average monthly income of like workers in a factory, then the group of all the workers in that factory will constitute our population. The number of elements in a population is called its size.

Sample : A part of selected items taken from a population for a survey is called a *sample*.

Sampling : The technique of selecting sample from a population is called *sampling*.

Parameter and Statistic : Any statistical measure like mean, median, standard deviation etc relating to a population is called *parameter*. If such statistical measures are related to a sample, these are called *statistic*.

Data : Observation expressed in numerical figures are called *data.* So, variables may be termed as data.

Investigator : Investigator is a person who plans and conducts any statistical survey independent by or with the help of others.

Enumerator : *Enumerator* is a person, who actually collects the desired statistical data. Often enumerators are trained personnel hired by the investigator for collecting data.

Respondent : *Respondent* is a person who answers/responds to the set of questions included in the questionnaire/schedule.

2.8 COLLECTION OF DATA

Collection of data is the first step in any statistical survey. Utmost care must be taken while collecting data because any statistical survey are based on these data.

For any statistical survey, two types of data may be collected. These are *Primary Data* and *Secondary Data*.

2.9 PRIMARY DATA

The data which are collected for the first time from the source of origin for a specific purpose are called *Primary Data*. Some examples of Primary data are –

(i) the census statistics collected by the office of teh Registrar General and Census Commissioner, Govt of India.

(ii) Reserve Bank of India Bulletine (monthly) issued by the Reserve Bank of India, Mumbai etc.

2.9 METHODS OF COLLECTING PRIMARY DATA

The methods commonly used for collecting Primary data are –

(a) Direct Personal Investigation

(b) Indirect Oral Investigation

(c) Mailed Questionnaire

(d) Schedule sent through enumerator

(e) By Local Agents

(a) Direct Personal Investigation : In this method, the investigator collects the information directly from the sources concerned. For instance, if one wants to study the living conditions of the people in a village, he has to go the village, contact the people and get the required information. Data collected by this method are original in character and very much reliable and accurate. But this method involves much cost and labour. So, this method can not be adopted when the enquiry is extensive or is to be done in large areas. This method does not give accurate result if the investigator is bias.

(b) Indirect Oral Investigation : In this method, information is collected indirectly by interviewing persons who are supposed to have knowledge about the problem under investigation. For example, if an accident takes place in a certain area and the cause of accident is not known, then the investigator may collect information from person who might have witnessed the accident or from those living in the neighbourhood to find the possible cause of the accident. Most of the commissions of enquiry appointed by Government collect primary data by this method. This method is less expensive and time saving.

(c) Mailed Questionnaire : In this method, a list of questions pertaining to the survey is prepared an is sent to various respondent by post. This list of questions is called a questionnaire and it provides space for answers. A mailed questionnaire should be accompanied by two things – (i) a polite covering letter – requesting the informats to fill up the questionnaire and (ii) a self addressed stamped envelope – so that the informates may post it back to the investigator.

This is the cheapest method by which data may be collected from a wide geographical area. This method can not be adopted if the respondents are illiterate. Non-response from the respondents is a major drawback of this method. It is difficult to verify the accuracy of the answers given by the

respondents as there is no contact between the investigator and the respondents.

(d) Schedule sent through enumerator : A schedule is also a list of questions pertaining to a survey. However, in this method, the enumerators contact informants personally, get replies to the questions contained in the schedule and fill them up in their own handwriting in the questionnaire form. In other words, a schedule is a list of questions where the answers to the questions are filled by the enumerators in the presence of the respondents.

Schedule can be used when the informants are illiterate. Non-response is minimised as the enumerators ersonally go to the informants. This is the most expensive method of collecting primary data as the enumerators are paid persons. The success of this method largely depends upon the training impaited to the enumerators.

(e) By Local Agents : Under this method, the investigator appoints local agents in different places to collect information. Generally this method is adopted by the newspaper agencies.

DIFFERENCE BETWEEN QUESTIONNARIRE AND SCHEDULE :

(i) A questionnaire is a list of questions which is sent to the respondents by post to be filled in by themselves	(i) A schedule is a list of questions which are carried by the enumerators personally to the respondents.
(ii) It may be used only for literate informants	(ii) It may be used for literate as well as illiterate informants.
(iii) It is the cheapest method of collecting primary data	(iii) It is the costliest method of collecting primary data.
(iv) Non-response from the respondents is a major drawback of this method.	(iv) Non response is minimised as enumerators personally go to the infomants.

FRAMING OF A QUESTIONNAIRE OR SCHEDULE :

Great care is to be taken in drafting a questionnaire or schedule as this is the medium through which information is collected. The success of a statistical investigation and the quality of its results depend largely upon how appropriately a questionnaire has been drafted and put to use.

The basic requirements of a good questionnaire or a schedule are as follows :

(i) The questions should be clear and brief.

(ii) The questions should be of objective type – either yes or no or multiple choice type.

(iii) The questions should be arranged in a logical manner.

(iv) The unit of measurment, if any, in which the respondent has to record the information should be clearly indicated e.g. kg, cm etc.

(v) The questionnaire should be attractive to look at.

(vi) The questions which involve numerical calculation should be avoided.

(vii) Number of questions should not be too many.

(viii) Personal questions or sensitive questions should not be included.

(ix) Necessary instructions for filling up the questionnaire should be given

(x) Some questions which cross-check the information on vital issues should be included.

2.10 SECONDARY DATA

The data which were previoulsy collected by some individual or agency as primary data for a certain purpose and are now used by someone else for another purpose are termed as *Secondary data*.

SOURCE OF SECONDARY DATA : The sources of secondary data may be classified under two heads –

(a) Published sources.

(b) Unpublished sources.

(a) **Published sources** :

(I) Reports and official publication of –

(i) International bodies like WHO, ILO, World Bank, statistical office of the United Nations etc.

(ii) Central and State Governments.

(iii) Ad-hoc committees and Commissions appointed by government.

(iv) Semi govt or local bodies like Municipal Corporation.

(v) Autonomous and private institutions like the Institution of Chartered Accountants, The Institute of Foreign Trade, Institute of Economic Growth etc.

(II) Financial and Economic Jornals like Indian Economic Revdew, Bulletin of Reserve Bank of India and various National and Regional magazines and newspapers.

(b) Unpublished Sources : Records maintained by various Government and Private institutions and offices, studies conducted by research scholars and research institutions etc.

PRECAUTIONS IN USE OF SECONDARY DATA : Utmost care should be taken in using secondary data. The degree of reliability of the secondary data must be assessed at the source. Prof A.L. Bowley pointed out that – "*Secondary data should not be accepted at their face value*." Therefore, before using secondary data, the investigators should consider the following factors :

(a) **Suitability of the data** : Before using secondary data the investigator must ensure whether the data are suitable for the purpose of the enquiry. The suitability of the data may be judged by comparing the nature and scope of the present enquiry with that of the original enquriry. For example, if the object of the prsent enquiry is to study the cost of living of the people, and if the data provide only the wholesale peices, then such data are not suitable as for calculating C.L.I, retail price is necessary.

(b) **Adequacy of the data** : The adequacy of the secondary data may be judged by studying the geographical area covered by the original enquiry and the expected coverage of the present enquiry. For example, if we want

to study the consumption pattern of the rice by the people of India, and if the available data relates only to the consumption pattern of rice by the people of Assam, then our data is not adequate. Another important factor to decide about the adequacy of the available data for the given enquiry is the time period for which the data are available.

(c) **Reliability of the data** : The reliability of the data can be judged by finding out whether the data collecting agency is free from personal bias. Moreover, it should also be verified that the data relates to a normal time period, that is free from economic boom or depression or natural calamities. If the enumeration is based on sample, it must be adequate and a representative sample. To be reliable, the data must be collected by trained and experienced investigator or enumerator.

2.11 PILOT SURVEY

It is the survey which is conducted on a small scale before conducting the main survey. Pilot survey is also known as pre-testing. Such a survey becomes useful in many ways, some of which are the following :

(i) Pilot survey becomes useful in estimating the time and cost involved in the main survey.

(ii) On the basis of pilot survey the investigator can decide whether requisite information can be collected by mailed questionnaire or by schedule sent through enumerator.

(iii) It provides insights into the problems faced by the interviewers.

(iv) With the help of pilot survey the investigator can adopt appropriate measures to avert certain problems that might involve in the execution of main survey.

On the whole, pilot survey is extremely helpful for planning and executing a survey.

2.12 SCRUTINY AND EDITING OF SECONDARY DATA

The secondary data must be used with caution. It is usually very difficult to verify such data and to edit them to find out inconsistencies,

problems errors and omission. Scrutiny of the secondary data is essential because the data might be inaccurate, unsuitable or inadequate.

Data editing is defined as the process involving the review and adjustment of collected survey data. The purpose is to control the quality of the collected data. Data editing may be performed manually, with the assistance of a computer or a combination of both.

Editing is concerned with removal of redundant data and filling of missing data.



CHECK YOUR PROGRESS

Q 3: Fill in the blanks :

(a) Data originally collected for an investigation is known as _____ data.

(b) Measurable charateristics are called _____.

(c) Population characteristics are called _____.

(d) Sample characteristics are called _____.

(e) There are _____ methods of collecting primary data.

(f) Primary data are _____ is nature.

(g) _____ survey is also known as pre-testing.



2.13 LET US SUM UP

In this unit, we have learnt about the following points related to a statistical survey or Statistical enquiry.

- Statistical survey means some sort of investigation by an individual or agency where in the relevant information is collected in quantitative terms.
- Two main stages of statistical survey are – Planning the survey and executing the survey.

- There are various types of statistial survey.
- The unit in terms of which the investigation counts or measures the variable selected for enumeration, analysis and interpretation is done known as statistical unit.
- Depending on the sources, the data may be classified as primary data and secondary data.
- There are five methods of collecting Primary data.
- Precautions must be taken while using secondary data.
- Measurable characteristics are called variable and non measurable characteristics are called attribute.
- The collection of all items related to any survey is called population or universe.
- A part of the selected items taken from a population is called a sample.
- Any statistical investigation where in data is collected from each unit of the population, is known as census survey.
- When in a statistical investigation, data is collected from a part of the population, it is known as sample survey.
- Investigator is a person who plans and conducts any statistical survey and enumerator is a person who actually collects the desired statistical data.
- The success of statistical survey depends largely upon drafting the questionnaire.
- Before using secondary data, the reliability, adequacy and accuracy of the data should be checked.
- Pilot survey is conducted on a small scale before conducting the main survey.
- The task of data editing is highly specialised and requires great care and attention. It is concerned with removal of redundant data and filling of missing data.



2.14 FURTHER READINGS

1) Bussiness Statistics. S.C. Gupta, Indira Gupta. Himalaya Publishing House.
2) An Introduction to commercial Mathematics and Statistics. Ajanta Mazumdar, Runjun Phookun, S. Chand and Company.
3) A text book of Bussiness Statistics : Dr Padmalochan Hazarika. S.Chand and Company.
4) Fundamentals of Statistics. S.C. Gupta. Himalya Publishing House.



2.15 ANSWER TO CHECK YOUR PROGRESS

Ans to Q No. 2: (a) Simple, compound. (b) Compound unit, (c) Analysis, (d) Primary, (e) Sample.

Ans to Q No. 3: (a) Primary, (b) Variable, (c) Parameter, (d) Statistic, (e) five, (f) orginial, (g) Pilot



2.16 MODEL QUESTIONS

Q 1: What is statistical survey? Mention the main stages of statistical survey.

Q 2: Describe the process of planning a statistical survey, with special reference to its scope and reference.

Q 3: Distinguish between (i) primary and secondary data (ii) sampling and census data.

Q 4: Define primary and secondary data. What are the precautions to be taken while using secondary data?

Q 5: Define secondary data. State their chief sources.

Q 6: Explain various methods of collecting primary data.

Q 7: Define : Primary data, sample, statistic, attribute, variable, parameter, investigator, respondent, enumerator, sample.

Q 8: What is the difference between Questionnaire and schedule.

Q 9: Write any five essential characteristics of an ideal questionnaire.

Q 10: What is the diffrence between population and sample.

Q 11: Explain the advantages of personal investigation as compared with other methods usually used in collecting primary data.

Q 12: Name the various types of statistical survey.

Q 13: What is meant by statistical survey. Name the various stages in such a survey.

Q 14: What are schedules and questionnaire? Discuss the important points to be noted while preparing a schedues and questionnaire.

Q 15: Distinguish between census and sample survey. Give reasons why sample survey is preferred.

Q 16: What is a statistical unit? What do you mean by units of collection and units of analysis. Discuss their relative uses.

Q 17: What is a statistical unit and explain the essential requirements of a good statistical unit.

Q 18: Describe the various stages of execution of a statistical survey.

Q 19: Write short notes on (i) Mailed questionnaire (ii) Schedule sent through enumerator.

Q 20: (i) Write a brief note on scrutiny and editing of secondary data.

(ii) Write a brief note on execution of statistical survey.

(iii) Write a brief note on the usefulness of Pilot survey.

Q 21: Fill in the blanks :

(i) Data originally collected for an investigation is known as _____

(ii) Population characteristics are known as _____.

(iii) _____ is a suitable method of collecting data in cases where the informants are literate and spread over a vast area.

(iv) Variable is a _____ characteristic.

(v) A non-measurable characteristic is called an _____

(vi) The _____ data should be used after careful scrutiny.

(vii) There are _____ methods of collecting primary data.

Q22: (i) Name the two things which should be sent along with a mailed questionnaire.

(ii) Which method of collecting primary data should be used by enquiry commisions?

(iii) Which method of collecting data should be used when field of enquiry is small?

(iv) Which method of collecting data should be used when the informats are illiterate and spread over a vast area?

Q23: Indicate whether the following statements are true or false :

(i) Mailed questionnaire should be accompanied by a covering letter.

(ii) Secondary data do not need much scrutiny and should be accepted at their face value.

(iii) The difference between primary and secondary data is only that degree.

(iv) A sample survey is less expensive than census survey.

(v) The task of editing data is highly specialised and requres great care and attention.

(vi) The results obtained by census study are always more reliable than those obtained in a sample study.

(vii) An attribute can be measured.

(viii) A discrete variable can take any fractional value.

*** ***** ***