

---

## UNIT 4 HYPOTHESIS FORMULATION AND SAMPLING

---

### Structure

- 4.0 Introduction
- 4.1 Objectives
- 4.2 Meaning and Characteristics of Hypothesis
- 4.3 Formulation of Hypothesis
- 4.4 Possible Difficulties in Formulation of a Good Hypothesis
- 4.5 Types of Hypotheses
  - 4.5.1 Null Hypothesis
  - 4.5.2 Alternative Hypothesis
- 4.6 Errors in Testing a Hypothesis
- 4.7 Importance of Hypothesis Formulation
- 4.8 Sampling
  - 4.8.1 Definition of Sampling
  - 4.8.2 Sampling Terminology
  - 4.8.3 Purpose of Sampling
- 4.9 Sampling Methods
  - 4.9.1 Non Probability Sampling
  - 4.9.2 Probability Sampling
- 4.10 Importance of Sampling
- 4.11 Let Us Sum Up
- 4.12 Unit and Questions
- 4.13 Glossary
- 4.14 Suggested Readings and References

---

### 4.0 INTRODUCTION

---

Scientific process or all empirical sciences are recognised by two inter-related concepts, namely; (a) context of discovery (getting an idea) and (b) context of justification (testing and results). Hypotheses are the mechanism and container of knowledge moving from the unknown to known. These elements form techniques and testing ground for scientific discovery. Hypotheses are tentative explanation and potential answer to a problem. Hypothesis gives the direction and helps the researcher interpret data. In this unit, you will be familiarised with the term hypothesis and its characteristics. It is, then, followed by the hypothesis formulation and types of hypothesis. Errors in hypothesis testing are also highlighted.

Further, In order to test the hypothesis, researcher rarely collects data on entire population owing to high cost and dynamic nature of the individual in population. Therefore, they collect data from a subset of individual – a sample - and make the inferences about entire population. This leads us to what we should know about the population and sample. So, researcher plans sample design and uses

various method of sampling. This unit will acquaint you with the meaning of sampling and basic terminology which is used in sampling design.

Now, it will move to purpose of sampling. And finally, various probability and non-probability sampling methods along with advantages and disadvantages are described.

---

## **4.1 OBJECTIVES**

---

After reading this unit, you will be able to:

- Define and describe hypothesis and its characteristics;
- explain formulation of hypothesis;
- Enumerate the possible difficulties in formulating hypothesis;
- Explain types of hypotheses;
- Identify in hypotheses testing;
- Define sampling;
- Explain the purpose of sampling; and
- Analyse various probability and non-probability sampling methods.

---

## **4.2 MEANING AND CHARACTERISTICS OF HYPOTHESIS**

---

In conducting research, the important consideration after the formulation of a research problem is the construction of hypothesis. As you know, any scientific inquiry starts with the statement of a solvable problem, when the problem has been stated, a tentative solution in the form of testable proposition is offered by the researcher.

Hypothesis is often considered a tentative and testable statement of the possible relationship between two or more events / variables under investigation.

According to Mcguigan (1990), ‘a testable statement of a potential relationship between two or more variables, i.e. advance as potential solution to the problem’. Kerlinger (1973) defined ‘a hypothesis is a conjectural statement of the relation between two or more variables’. In order to be useful in any study, the hypothesis needs to be stated in such a way that it might be subjected to empirical testing. The researcher is responsible to suggest or find some way to check how the hypothesis stands against empirical data. When a hypothesis is formulated, the investigator must determine usefulness of the formulated hypothesis. There are several criteria or characteristics of a good research hypothesis. A good hypothesis is one which meets such criteria to a large extent. Some of these characteristics are enumerated below:

- 1) Hypothesis should be conceptually clear;
- 2) Hypothesis must be testable;
- 3) Hypothesis should be related to the existing body or theory and impact;
- 4) Hypothesis should have logical unity and comprehensiveness;
- 5) Hypothesis should be capable of verification; and
- 6) Hypothesis should be operationisable.

---

## 4.3 FORMULATION OF HYPOTHESIS

---

Science proceeds with observation, hypothesis formulation and hypothesis testing. After testing the hypothesis, through various statistical tests, researcher can accept or reject the hypothesis. If the hypothesis is accepted then researcher can replicate the results, if hypothesis is rejected then researcher can refined or modify the results.

By stating a specific hypothesis, the researcher narrows the focus of the data collection effort and is able to design a data collection procedure which is aimed at testing the plausibility of the hypothesis as a possible statement of the relationship between the terms of the research problem.

It is, therefore, always useful to have a clear idea and vision about the hypothesis. It is essential for the research question as the researcher intends to verify, as it will direct and greatly help to interpretation of the results.

---

## 4.4 POSSIBLE DIFFICULTIES IN FORMULATION OF A GOOD HYPOTHESIS

---

There are three major possible difficulties; a researcher could face during formulation of hypothesis. First, the absence of knowledge of a theoretical framework is a major difficulty in formulating a good research hypothesis. Second, if detailed theoretical evidences are not available or if the investigator is not aware of the availability of those theoretical evidences, a research hypothesis cannot be formulated. Third, when the investigator is not aware of the scientific research techniques, she/he will not be able to frame a good research hypothesis.

Despite these difficulties, the investigator attempts in her/his research to formulate a hypothesis. Usually the hypothesis is derived from the problem statement. The hypothesis should be formulated in a positive and substantive form before data are collected. In some cases additional hypothesis may be formulated after collection of data, but they should be tested on a new set of data and not on the old set which has suggested it. The formulation of a hypothesis is a creative task and involves a lot of thinking, imagination and innovation. Reichenbach (1938) has made a distinction between the two processes found commonly in any hypothesis formulation task. One is the context of discovery and another is the context of justification. The manner or the process through which a scientist arrives at a hypothesis illustrates the context of justification. A scientist is concerned more with a context of justification in the development of a hypothesis. He never puts his ideas or thoughts as they nakedly occur in the formulation of a hypothesis. Rather, he logically reconstructs his ideas or thoughts and draws some justifiable inferences from those ideas and thoughts. He never cares to relate how he actually arrived at a hypothesis. He does not say, for example, that while he was shaving, this particular hypothesis occurred to him. He usually arrives at a hypothesis by the rational reconstruction of thoughts. When a scientist reconstructs his thoughts and communicates them in the form of a hypothesis to others, he uses the context of justification. When he arrives at a hypothesis, he extensively as well as intensively surveys a mass of data, abstracts them, tries to find out similarities among the abstracted data and finally makes a generalisation or deduces a preposition in the form of a hypothesis.

Here is an important distinction to be made between formulating a hypotheses and choosing one. Although a researcher often becomes interested in a question about human behaviour for personal reasons, the ultimate value of research study depends on the researcher bringing methodological criteria to bear on the selection of the hypothesis to be tested. In other words, Good hypothesis are made, not born.

Hypothesis plays a key role in formulating and guiding any study. The hypotheses are generally derived from earlier research findings, existing theories and personal observations and experience. For instance, you are interested in knowing the effect of reward on learning. You have analysed the past research and found that two variables are positively related. You need to convert this idea in terms of a testable statement. At this point you may develop the following hypothesis.

Those who are rewarded shall require lesser number of trials to learn the lesson than those who are not rewarded.

A researcher should consider certain points while formulating a hypothesis:

- i) Expected relationship or differences between the variables.
- ii) Operational definition of variable.
- iii) Hypotheses are formulated following the review of literature

The literature leads a researcher to expect a certain relationship.

Hypotheses are the statement that is assumed to be true for the purpose of testing its validity.

As suggested by Russell and Reichenback (1947), the hypotheses should be stated in the logical form on the general implications. A hypothesis can be put in the form of an if ..... then statement; if A is true then B should follow. For example, verbal development theory of amnesia states that childhood amnesia caused by the development of language. To test this theory, researcher can make a hypothesis like this – if the lack of verbal ability is responsible for childhood amnesia, then the children should not be able to verbally recall events usually words that they did not know at the time of events.

### **Self Assessment Questions**

#### **Fill in the blanks**

- 1) Hypothesis is considered as ..... and ..... statement of the possible relationship between two or more variables.
- 2) Hypothesis can be put in the form of an if ..... statement.
- 3) Hypothesis is formulated..... to review of literature.
- 4) ..... of knowledge of a theoretical framework is a major difficulty in formulating a hypothesis.
- 5) Formulation of a hypothesis enhances ..... in the study.

**Answers:** (1) tentative, testable, (2) then, (3) Prior, (4) Absence, (5) Objectivity.

## 4.5 TYPES OF HYPOTHESES

As explained earlier, any assumption that you seek to validate through investigation is called hypotheses. Hence theoretically, there should be one type of hypotheses on the basis of the investigation that is, research hypothesis. However, because of the conventions in scientific enquiries and wording used in the constructions of the hypothesis, Hypotheses can be classified into several types, like; universal hypotheses, existential hypotheses, conceptual hypotheses etc. Broadly, there are two categories of the hypothesis:

- i) Null hypothesis
- ii) Alternative hypothesis

### 4.5.1 Null Hypothesis

Null hypothesis is symbolised as  $H_0$ . Null hypothesis is useful tool in testing the significance of difference. In its simplest form, this hypothesis asserts that there is no true difference between two population means, and the difference found between sample means is, accidental and unimportant, that is arising out of fluctuation of sampling and by chance. Traditionally null hypothesis stated that there is zero relationship between terms of the hypothesis. For example, (a) schizophrenics and normal do not differ with respect to digit span memory (b) There is no relationship between intelligence and height.

The null hypothesis is an important component of the decision making methods of inferential statistics. If the difference between the samples of means is found significant the researcher can reject the null hypothesis. It indicates that the differences have statistically significant and acceptance of null hypothesis indicates that the differences are due to chance. Null hypothesis should always be specific hypothesis i.e. it should not state about or approximately a certain value.

The null hypothesis is often stated in the following way:

$$H_0: \mu_{HV} \leq \mu_{LV}$$

Thus, the null hypothesis is that mean of the population of those children who have the high vocabulary (group1) is less than or equal to mean of those who lack the vocabulary (Group 2 ).

### 4.5.2 Alternative Hypothesis

Alternative hypothesis is symbolised as  $H_1$  or  $H_a$ , is the hypothesis that specifies those values that are researcher believes to hold true, and the researcher hopes that sample data will lead to acceptance of this hypothesis as true. Alternative hypothesis represents all other possibilities and it indicates the nature of relationship.

The alternative hypothesis is stated as follows:

$$H_1: \mu_{HV} > \mu_{LV}$$

The alternative hypothesis is that the mean of population of those who have the vocabulary is greater than the mean of those to lack the vocabulary. In this example the alternative hypothesis is that the experimental population had higher mean than the controls. This is called directional hypothesis because researcher predicted

that the high vocabulary children would differ in one particular direction from the low vocabulary children. Sometimes researcher predicts only that the two groups will differ from each other but the researcher doesn't know which group will be higher. This is non directional hypothesis.

The null and alternative hypothesis in this case would be stated as follows:

$$H_0: \mu_1 = \mu_2$$

$$H_1: \mu_1 \neq \mu_2$$

Thus, the null hypothesis is that mean of group 1 equals the mean of group 2, and the alternative hypothesis is that the mean of group 1 does not equal the mean of group 2.

## 4.6 ERRORS IN TESTING A HYPOTHESIS

You have already learned that hypotheses are assumptions that may be prove to be either correct or incorrect. It is possible to arrive at a incorrect conclusion about a hypothesis for the various reasons if –

- Sampling procedure adopted faulty
- Data collection method inaccurate
- Study design selected is faulty
- Inappropriate statistical methods used
- Conclusions drawn are incorrect

Two common errors exist when testing a hypothesis.

Type I error – Rejection of a null hypothesis when it is true.

Type II error - Acceptance of a null hypothesis when it is false.

### Self Assessment Questions (State True and False)

- |   |       |
|---|-------|
| 1) Null hypothesis is denoted by $H_1$ .  | T/F   |
| 2) If the hypothesis is accepted then researcher can replicate the results.       | T/F   |
| 3) Rejection of a null hypothesis when it is true is called type II error.        | T / F |
| 4) Hypothesis can be stated directional and non directional.                      | T / F |
| 5) Alternative hypothesis specifies values that researcher believes to hold true. | T / F |

**Answers :** (1) F, (2) T, (3) F, (4) T, (5) T

## 4.7 IMPORTANCE OF HYPOTHESIS FORMULATION

Hypothesis is the basic function of the scientific research. If simple, brief and clear scientific hypothesis has been formulated, there shall be no problem for the investigator to proceed in the research field. Its utility or importance for and research may be studied as under.

Accordingly to Goode and Hatt ('without' hypothesis formulation the research is unfocussed, a random empirical wandering. The results can not be studied as facts with clear meaning. Formulation of hypothesis links between theory and investigation which lead to discovery of addition to knowledge.

### Self Assessment Questions

Study the following research questions and state the possible hypothesis with specify their types specify their types.

1) Is physical attractiveness related to friendship?

.....

.....

.....

.....

.....

.....

2) Does meaningful of material affect the rate of learning?

.....

.....

.....

.....

.....

.....

3) Does reinforcement improve the learning for solving simple discrimination task?

.....

.....

.....

.....

.....

4) Does onset of fatigue reduce the efficiency of the worker?

.....

.....

.....

.....

.....

.....



## 4.8 SAMPLING

Researchers usually can not make direct observation of every individual in the population under study. Instead, they collect data from a subset of individuals- a sample – and use those observations to make inferences about the entire population.

Ideally, the sample corresponds to the larger population on the characteristics of (s) of interest. In that case, the researcher's conclusions from the samples are properly applicable to the entire population.

Sampling is the process of selection of units (e.g. people, organisation) from a population of interest so that by studying the sample may fairly generate results back to the population from which they were chosen.

### 4.8.1 Definition of Sampling

According to Young (1992) “A statistical sample is miniature picture of cross selection of the entire group or aggregate from which the sample is taken”.

According to Goode and Hatt(1981) “A sample, as the name implies, is a smaller representative of a large whole”.

According to Blalock (1960) “It is a small piece of the population obtained by a probability process that mirrors with known precision, the various patterns and sub-classes of population”.

### 4.8.2 Sampling Terminology

Before we explain the purpose, uses and method of sampling, it will be better to describe those fundamental terms which are concerned to sampling concepts and principles.

*Population:* Population is a well defined set up of all elements pertaining to a given characteristic. It refers to the whole that include all observations or measurements of a given characteristic. Population is also called universe or population. It may be defined as any identifiable and well specified group of individual for example. All primary teachers, nonnumber of all college teachers and all university students are the example of population. A population may be finite or infinite.

A finite population is one where all the members can be easily counted. An infinite population is one whose size is unlimited, and can not count easily. Population of college teachers is an example of finite population and production of wheat, and fishes in river are the example of infinite population. A measure based upon the entire population is called a parameter.

*Sample:* A sample is any number of persons selected to represent the population according to some rule of plan. Thus, a sample is a smaller representation of the population. A measure based upon a sample is known as a statistic.

*Sample size:* No. of selected individual for example, no. of students, families from whom you obtain the require information is called the sample size and usually denoted by the letter (n).



*Sampling design or strategy:* The way researcher selects the sample or students or families etc. is called the sampling design and strategy. It refers to the techniques or procedures the researcher would adopt in selecting some sampling units from which inferences about the population are drawn.

*Sampling unit:* Each individual or case that becomes the basis for selecting a sample is called sampling unit or sampling elements.

*Sampling frame:* The list of people from which the sample is taken. It should be comprehensive, complete and up-to-date. Examples of sampling frame: Electoral Register; Postcode Address File; telephone book.

#### **Self Assessment Questions (Fill in the blanks)**

- 1) Any identifiable and well specified group of individual is known as .....
- 2) List of all the units of the population is called .....
- 3) Purposes of sampling is to derive the desired information about the population at the minimum ..... and maximum .....
- 4) The way the researcher selects the sample is known as .....
- 5) ..... is the miniature picture of entire group.

**Answers:** (1) population, (2) sampling frame, (3) cost, reliability, (4) sampling design, (5) sample.

### **4.8.3 Purpose of Sampling**

The objective of sampling is to derive the desired information about the population at the minimum cost or with the maximum reliability. Further, the aims in selecting a sample are to achieve maximum precision in estimates within a given sample size and to avoid bias in the selection of sample. Bias in the selection of sample can take place if: (a) the researcher selects the sample by non random method and influenced by human choice. (b) The researcher does not cover the sampling population accurately and completely (c) A section of a sample population is impossible to find or refuses to cooperate.

---

## **4.9 SAMPLING METHODS**

---

Blalock (1960) indicated that most sampling methods could be classified into two categories:

- i) Non probability sampling methods
- ii) Probability sampling methods

### **4.9.1 Non Probability Sampling Methods**

Non probability sampling is one in which there is no way of assessing the probability of the element or group of elements, of population being included in the sample. In other words, non-probability sampling methods are those that provide no basis for estimating how closely the characteristics of sample approximate the parameters of population from which the sample had been obtained. This is because non probability sample do not use the techniques of random sampling. Important techniques of non probability sampling methods are:

**i) Haphazard, Accidental, or Convenience Sampling**

Haphazard sampling can produce ineffective, highly unrepresentative samples and is not recommended. When a researcher haphazardly selects cases that are convenient, he or she can easily get a sample that seriously misrepresents the population. Such samples are cheap and quick; however, the systematic errors that easily occur make them worse than no sample at all. The person-on-the-street interview conducted by television programs is an example of a haphazard sample. Likewise, television interviewers often select people who look “normal” to them and avoid people who are unattractive, poor, very old, or inarticulate. Such haphazard samples may have entertainment value, but they can give a distorted view and seriously misrepresent the population.

For example, an investigator may take student of class X into research plan because the class teacher of the class happens to be his / her friend. This illustrates accidental or convenience sampling.

**ii) Quota Sampling**

Quota Sampling is an improvement over haphazard sampling. In quota sampling, a researcher first identifies relevant categories of people (e.g., male and female; or under age 30, ages 30 to 60, over age 60, etc.), then decides how many to get in each category. Thus, the number of people in various categories of the sample is fixed. For example, a researcher decides to select 5 males and 5 females under age 30, 10 males and 10 females aged 30 to 60, and 5 males and 5 females over age 60 for a 40-person sample. It is difficult to represent all population characteristics accurately.

Quota sampling ensures that some differences are in the sample. In haphazard sampling, all those interviewed might be of the same age, sex, or background. But, once the quota sampler fixes the categories and number of cases in each category, he or she uses haphazard or convenience sampling. Nothing prevents the researcher from selecting people who act friendly or who want to be interviewed. Quota sampling methods are not appropriate when the interviewers choose who they like (within above criteria) and may therefore select those who are easiest to interview, so, sampling bias can take place. Because not using the random method, it is impossible to estimate the accuracy. Despite these limitations, quota sampling is a popular method among non-probability methods of sampling, because it enables the researcher to introduce a few controls into his research plan and this methods of sampling are more convenient and less costly than many other methods of sampling.

**iii) Purposive sampling**

Purposive sampling is a valuable kind of sampling for special situations. It is used in exploratory research or in field research. It uses the judgment of an expert in selecting cases or it selects cases with a specific purpose in mind. With purposive sampling, the researcher never knows whether the cases selected represent the population. Purposive sampling is appropriate to select unique cases that are especially informative.

For example, a researcher wants to study the temperamental attributes of certain problem behaviour children. It is very difficult to list all certain problem behaviour children and sample randomly from the list. The researcher uses many different

methods to identify these cases and approach them to obtain the relevant information. The primary consideration in purposes sampling is the judgment of researcher as to who can provide the best information to achieve the objectives of the study. The researcher only goes to those people who in his / her opinion are likely to have the required information and be willing to share it.

For studying attitude toward any national issue, a sample of journalists, teacher and legislators may be taken as an example of purposive sampling because they can more reasonably be expected to represent the correct attitude than other class of people residing in country.

Purposes sampling is somewhat less costly, more readily accessible, more convenient and select only those individual that are relevant to research design.

Despite these advantages of purposes sampling, there is no way to ensure that the sample is truly represent of the population and more emphasis is placed on the ability of researcher to assess the elements of population.

#### iv) **Snowball sampling**

Snowball sampling is also known as network, chain referral or reputation sampling method. Snowball sampling which is a non probability sampling method is basically sociometric. It begins by the collection of data on one or more contacts usually known to the person collecting the data. At the end of the data collection process (e.g., questionnaire, survey, or interview), the data collector asks the respondent to provide contact information for other potential respondents. These potential respondents are contacted and provide more contacts. Snowball sampling is most useful when there are very few methods to secure a list of the population or when the population is unknowable.

Snowball sampling has some advantages— 1) Snowball sampling, which is primarily a sociometric sampling technique, has proved very important and is helpful in studying small informal social group and its impact upon formal organisational structure, 2) Snowball sampling reveals communication pattern in community organisation concepts like community power; and decision-making can also be studied with the help of such sampling technique.

Snowball sampling has some limitations also— 1) Snowball sampling becomes cumbersome and difficult when is large or say it exceeds 100, 2) This method of sampling does not allow the researcher to use probability statistical methods. In fact, the elements included in sample are not randomly drawn and they are dependent on the subjective choices of the originally selected respondents. This introduces some bias in the sampling.

#### v) **Systematic sampling**

Systematic sampling is another method of non-probability sampling plan, though the label 'systematic' is somewhat misleading in the sense that all probability sampling methods are also systematic sampling methods. Due to this, it often sounds that systematic sampling should be included under one category of probability sampling, but in reality this is not the case.

Systematic sampling may be defined as drawing or selecting every ninth person from a predetermined list of elements or individuals. Selecting every 5<sup>th</sup> roll

number in a class of 60 students will constitute systematic sampling. Likewise, drawing every 8<sup>th</sup> name from a telephone directory is an example of systematic sampling. If we pay attention to systematic sampling plan, it become obvious that such a plan possesses certain characteristics of randomness (first element selected is a random one) and at the same time, possesses some non-probability traits such as excluding all persons between every ninth element chosen.

Systematic sampling is relatively quick method of obtaining a sample of elements and it is very easy to check whether every ninth number or name has been selected. Further Systematic sampling is easy to used.

Despite these advantages, systematic sampling ignores all persons between every ninth element chosen. Then it is not a probability sampling plan. In Systematic sampling there is a chance to happen the sampling error if the list is arranged in a particular order.

**Activity**

Make a list of some research studies where some of the non probability methods could be used. Also justify the choice of particular sampling method you have selected for the study.

## 4.9.2 Probability Sampling

Probability sampling methods are those that clearly specify the probability or likelihood of inclusion of each element or individual in the sample. Probability sampling is free of bias in selecting sample units. They help in estimation of sampling errors and evaluate sample results in terms of their precision, accuracy and efficiency and hence, the conclusions reached from such samples are worth generalisation and comparable to similar population to which they belong. Major probability sampling methods are:

### i) Simple random sampling

A simple random sample is a probability sample. A simple random sample requires (a) a complete listing of all the elements (b) an equal chance for each elements to be selected (c) a selection process whereby the selection of one element has no effect on the chance of selecting another element. For example, if we are to select a sample of 10 students from the seventh grade consisting of 40 students, we can write the names (or roll number) of each of the 40 students on separate slips of paper – all equal in size and colour – and fold them in a similar way. Subsequently, they may be placed in a box and reshuffled thoroughly.

A blindfolded person, then, may be asked to pick up one slip. Here, the probability of each slip being selected is  $1/40$ . Suppose that after selecting the slip and noting the name written on the slip, he again returns it to the box. In this case, the probability of the second slip being selected is again  $1/40$ . But if he does not return the first slip to the box, the probability of the second slip becomes  $1/39$ . When an element of the population is returned to the population after being selected, it is called sampling with replacement and when it is not returned, it is called sampling without replacement.

Thus random sampling may be defined as one in which all possible combinations of samples of fixed size have an equal probability of being selected.

Advantages of simple random sampling are:

- 1) Each person has equal chance as any other of being selected in the sample.
- 2) Simple random sampling serves as a foundation against which other methods are sometimes evaluated.
- 3) It is most suitable where population is relatively small and where sampling frame is complete and up-to-date.
- 4) As the sample size increases, it becomes more representative of universe.
- 5) This method is least costly and easily assessable of accuracy.

Despite these advantages, some of the disadvantages are:

- 1) Complete and up-to-date catalogued universe is necessary.
- 2) Large sample size is required to establish the reliability.
- 3) When the geographical dispersion is so wider therefore study of sample item has larger cost and greater time.
- 4) Unskilled and untrained investigator may cause wrong results.

#### **Activity**

In a class of 140 students, select a simple random sample of size 20 students with replacement technique. Also mention the probability of each one of 140 students being included in the sample.

#### **ii) Stratified random sampling**

In stratified random sampling the population is divided into two or more strata, which may be based upon a single criterion such as sex, yielding two strata-male and female, or upon a combination of two or more criteria such as sex and graduation, yielding four strata, namely, male undergraduates, male graduates, female undergraduates and female graduates. These divided populations are called subpopulations, which are non-overlapping and together constitute the whole population.

Having divided the population into two or more strata, which are considered to be homogeneous internally, a simple random sample for the desired number is taken from each population stratum. Thus, in stratified random sampling the stratification of population is the first requirement.

There can be many reasons for stratification in a population.

Two of them are:

- 1) Stratification tends to increase the precision in estimating the attributes of the whole population.
- 2) Stratification gives some convenience in sampling. When the population is divided into several units, a person or group of person may be deputed to supervise the sampling survey in each unit.

Advantages of stratified Random Sampling are:

- 1) Stratified sampling is more representative of the population because formation of stratum and random selection of item from each stratum make it hard to exclude in strata of the universe and increases the sample's representation to the population or universe.

- 2) It is more precise and avoids the bias to great extent.
- 3) It saves time and cost of data collection since the sample size can be less in the method.

Despite these advantages, some of the disadvantages of stratified sampling are:

- 1) Improper stratification may cause wrong results.
- 2) Greater geographical concentration may result in heavy cost and more time.
- 3) Trained investigators are required for stratification.

### iii) **Cluster sampling**

A type of random sample that uses multiple stages and is often used to cover wide geographic areas in which aggregated units are randomly selected and then sample are drawn from the sampled aggregated units or cluster

For example, if the investigator wanted to survey some aspect of 3<sup>rd</sup> grade elementary school going children. First, a random sample of number of states from the country would be selected. Next, within each selected state, a random selection of certain number of districts would be made. Then within district a random selection of certain number of elementary schools would be made. Finally within each elementary school, a certain number of children would be randomly selected. Because each level is randomly sampled, the final sample becomes random. However, selection of samples is done to different stages. This is also called multi stage sampling.

This sampling method is more flexible than the other methods. Sub-divisions at the second stage unit needs be carried out only those unit selected in the first stage. Despite these merits, this sampling method is less accurate than a sample, containing the same number of the units in single stage samples.

#### **Self Assessment Questions**

- 1) Non probability sampling is one which there is way of assessing the probability of the element or group of element of population, being included in the sample. T/F
- 2) Simple random sampling is the core technique and attaches equal probability to each unit of the population to be selected. T/F
- 3) Cluster sampling method sometimes known as multi stage sampling method. T/F
- 4) Snowball technique is a probability sampling method. T/F
- 5) Stratified sampling is more representative for the population than other methods. T/F

**Answer:** (1) F, (2) T, (3) T, (4) F, (5) T.

## **4.10 IMPORTANCE OF SAMPLING**

In research, sampling method has obtained great importance. Sampling studies are becoming more and more popular in all type of studies. The vastness of the population, the difficulties of contacting people, high refusal rate, difficulties of ascertaining the universe make sampling the best alternative in case of social studies. The census method is rarely, if ever tried in matters of social research.



Recent developments in sampling technique have made this method more reliable and valid. The results of sampling have attained a sufficiently high standard of accuracy.

The three main advantage of sampling are that cost is lowest, data collection is faster, and since the data set is smaller, it is possible to ensure homogeneity and to improve the accuracy and quality of data (Ader, Mellenbergh & Hard (2008)

---

## **4.11 LET US SUM UP**

---

In this unit you have learnt about hypothesis formulation and sampling. A hypothesis is a speculative statement that is subjected to verification through a research study. In formulating a hypothesis it is important to ensure that it is simple, specific and conceptually clear; is able to be verified; is rooted in an existing body of knowledge; and able to be operational zed. There are two broad types of hypothesis: a null hypothesis and an alternate hypothesis.

Sampling is the act, process, or technique of selection a suitable sample, or a representative part of a population for the purpose of determining parameters or characteristics of the whole population.

Researchers rarely survey the entire population for two reasons: The cost is too high, and the population is dynamic in that the individual making up the population may change over time. Sampling methods are of two types i.e. Non probability and probability sampling methods. Probability sampling methods are those in which some probability to each unit of the population to be included in the sample and this is more representative. Three different probability sampling method are discussed as simple random sampling, stratified random sampling and cluster / multi stage sampling. The other non probability sampling methods discussed are convenience sampling, Quota sampling, Purposive sampling, Snowball sampling and systematic sampling. These methods are also used but lack the representative character of samples.

---

## **4.12 UNIT END QUESTIONS**

---

- 1) Define hypothesis and explain its characteristics.
- 2) Write short notes on:
  - a) Formulation of hypothesis
  - b) Null hypothesis
  - c) Alternative hypothesis
- 3) Write a hypothesis which incorporates each pair of concepts listed below:
  - a) academic achievement and teaching methods
  - b) education and social prestige
  - c) frustration and need for achievement
- 4) What is sampling? Discuss its importance.
- 5) What is simple random sampling. Discuss its advantages and disadvantages.
- 6) What do you mean by probability sampling method. Discuss any two types of probability sampling methods.



7) Define the following:

- 1) Sampling unit
- 2) Population
- 3) Sampling frame

---

## 4.13 GLOSSARY

---

<b>Hypothesis</b>	:	A tentative and testable statement of a potential relationship between two or more variables.
<b>Null hypothesis</b>	:	The hypothesis that is of no scientific interest; sometimes the hypothesis of no difference.
<b>Alternative hypothesis</b>	:	Statistical term for research hypothesis that specifies values that researcher believes to hold true.
<b>Population</b>	:	It is the aggregate from which a sample is drawn. In statistics, it refers to any specified collection of objects, people, organisation etc.
<b>Population size</b>	:	It is the total number of units present in the population.
<b>Sampling units</b>	:	They are members of the population.
<b>Sampling frame</b>	:	It is the list of all the units of population.
<b>Sampling design</b>	:	It is a definite plan for obtaining a sample from a given population.
<b>Sample size</b>	:	It is the total number of units in the sample.
<b>Simple random sample</b>	:	It is a sample in which each unit of the population has an equal chance of being selected in the sample.

---

## 4.14 SUGGESTED READINGS AND REFERENCES

---

Ader, H.J., Mellenbergh, G.J., & Hand, D.J. (2008). *Advising on Research Methods: A Consultant's Companion*. Huizen, The Netherlands: Johannes van kessel Publishing.

Blalock, HM (1960), *Social Statistics*, NY: McGraw-Hill.

Goode, WJ & Hatt, PK (1981). *Methods in Social Research*. Tokyo: McGraw hill Book Company.

### References

Kerlinger, FN (1973). *Foundations of Behavioural Research*. New York: Rinehart and Winston.

McGuigan, FJ (1990). *Experimental Psychology: A Methodological Approach*, Englewood Cliffs, New Jersey: Prentice-hall.

Reichenbach, H (1938). *Experience and Prediction*. Chicago: University of Chicago press.

Reichenbach, H (1947). *Elements of Symbolic Logic*, New York: Macmillan.

Young, PV (1992). *Scientific Social Survey and Research*, New Delhi: Prentice-hall of India.