BUSINESS RESEARCH MODULE- I (PART A)

By Asst. Prof. (Eco) Mrs. Tapaswini Nayak.

- What is research?

✓ It is a systematic search for truth.

✓ Through research, new and original information, ideas about the world we live in, are obtained.

✓ Research is search for knowledge.

✓ Research means finding answers to the questions.

✓ Research is defined as a scientific and systematic search for information on a specific topic.

- **MEANING OF RESEARCH:**
- Research refers to a search for knowledge.
- One can also define research as a scientific and systematic search for pertinent information on a specific topic. In fact, research is an art of scientific investigation.
- According to **Clifford Woody** research comprises defining and redefining problems, formulating hypothesis or suggested solutions; collecting, organizing and evaluating data; making deductions and reaching conclusions; and at last carefully testing the conclusions to determine whether they fit the formulating hypothesis.

- **OBJECTIVES OF RESEARCH**
- The objective of research is to find answers to the questions by applying scientific procedures. In other words, the main aim of research is to find out the truth which is hidden and has not yet been discovered.
- To gain familiarity with new insights into a phenomenon (i.e., exploratory or formulative research studies);
- To accurately portray the characteristics of a particular individual, group, or a situation (i.e., descriptive research studies or '**ex post facto research**');
- To analyze the frequency with which something occurs (i.e., diagnostic research studies); and
- To examine the hypothesis of a causal relationship between two variables (i.e., hypothesis-testing research studies).

- **RESEARCH METHODS VERSUS METHODOLOGY**

- Research methods include all those techniques/methods that are adopted for conducting research. On the other hand, research methodology is the way in which research problems are solved systematically.

- **SIGNIFICANCE OF RESEARCH**

- Research encourages scientific and inductive thinking. The role of research in applied economics in the context of an economy or business is greatly increasing in modern times.
- Research assumes significant role in the formulation of economic policy for both, the government and business. Government budget formulation, for example, depends particularly on the analysis of needs and desires of people, and the availability of revenues, which requires research.
- Research also helps in the proper allocation of a country’s scarce resources.

- Research is also necessary for collecting information on the social and economic structure of an economy to understand the process of change occurring in the country.

- Research also assumes significance in solving various operational and planning problems associated with business and industry. In several ways, operations research, market research and motivational research are vital and their results assist in taking business decisions.

- Market research refers to the investigation of the structure and development of a market for the formulation of efficient policies relating to purchases, production and sales.

- Operational research relates to the application of logical, mathematical, and analytical techniques to find solution to business problems, such as cost minimization or profit maximization, or the optimization problems.

- Motivational research helps to determine why people behave in the manner they do with respect to market characteristics.

- Research is equally important to social scientists for analyzing the social relationships and seeking explanations to various social problems.

- **TYPES OF RESEARCH:**

- There are different types of research.

1. **DESCRIPTIVE VERSUS ANALYTICAL:**

- Descriptive research consists of surveys and fact-finding enquiries of different types. The main objective of descriptive research is describing the state of affairs as it prevails at the time of study. The term '**ex post facto research**' is quite often used for descriptive research studies in social sciences and business research. The researcher has no control over the variables here. He/she has to only report what is happening or what has happened. Majority of the ex post facto research projects are used for descriptive studies in which the researcher attempts to examine phenomena, such as the consumers' preferences, frequency of purchases, shopping, etc.

- Meanwhile in the Analytical research, the researcher has to use the already available facts or information, and analyze them to make a critical evaluation of the subject.

2. APPLIED VERSUS FUNDAMENTAL:

- Research can also be applied or fundamental in nature. An attempt to find a solution to an immediate problem encountered by a firm, an industry, a business organization, or the society is known as applied research.

- On the other hand, fundamental research mainly concerns generalizations and formulation of a theory. In other words, “Gathering knowledge for knowledge’s sake is termed ‘pure’ or ‘basic’ research” Researches relating to pure mathematics or concerning some natural phenomenon are instances of Fundamental Research. Likewise, studies focusing on human behaviour also fall under the category of fundamental research. Thus, while the principal objective of applied research is to find a solution to some pressing practical problem, the objective of basic research is to find information with a broad base of application and add to the already existing organized body of scientific knowledge.

3. QUANTITATIVE VERSUS QUALITATIVE

- Quantitative research relates to aspects that can be quantified or can be expressed in terms of quantity. It involves the measurement of quantity or amount.

- On the other hand, Qualitative research is concerned with qualitative phenomena, or more specifically, the aspects related to or involving quality or kind. For example, an important type of qualitative research is ‘Motivation Research’, which investigates into the reasons for certain human behaviour.

4. CONCEPTUAL VERSUS EMPIRICAL

- The research related to some abstract idea or theory is known as Conceptual Research. Generally, philosophers and thinkers use it for developing new concepts or for reinterpreting the existing ones.

- Empirical Research, on the other hand, exclusively relies on the observation or experience with hardly any regard for theory and system. Such research is data based, which often comes up with conclusions that can be verified through experiments or observation. Empirical research is also known as experimental type of research, in which it is important to first collect the facts and their sources, and actively take steps to stimulate the production of desired information. In this type of research, the researcher first formulates a working hypothesis, and then gathers sufficient facts to prove or disprove the stated hypothesis.

5. <u>One Time Research or Longitudinal Research</u>

- In the formal case the research is confined to a single time period, whereas the later case the research is carried on over several time periods.

6. <u>Laboratory Research and Field setting Research</u>

- This classification is based on the environment in which research is carried out.

7. <u>Historical Research</u>

- Historical Research is that which utilities historical sources like documents remains etc to study events ideas of the past including the philosophy of persons and groups at any remote point of time.

- **RESEARCH PROCESS**
- Research process consists of a series of steps or actions required for effectively conducting research.
- The following are the steps that provide useful procedural guidelines regarding the conduct of research:

- Formulating the research problem
- Extensive literature survey
- Developing hypothesis
- Preparing the research design
- Determining sample design
- Collecting data
- Execution of the project
- Analysis of data
- Hypothesis testing
- Generalization and interpretation
- Preparation of the report or presentation of the results.

- **FORMULATING THE RESEARCH PROBLEM**
- There are two types of research problems, viz., those which relate to states of nature and those which relate to relationships between variables. At the very outset the researcher must find out the problem he wants to study, i.e., he must decide the general area of interest or aspect of a subject-matter that he would like to inquire into.

- The best way of understanding the problem is to discuss it with one's own colleagues or with those having some expertise in the matter. In an academic institution the researcher can seek the help from a guide who is usually an experienced man and has several research problems in mind.

- **EXTENSIVE LITERATURE SURVEY**
- For this purpose, the abstracting and indexing journals and published or unpublished bibliographies are the first place to go to. Academic journals, conference proceedings, government reports, books etc., must be tapped depending on the nature of the problem. A good library will be a great help to the researcher at this stage.

- **DEVELOPMENT OF WORKING HYPOTHESES**
- After extensive literature survey, researcher should state in clear terms the working hypothesis or hypotheses.
- Working hypothesis is tentative assumption made in order to draw out and test its logical or empirical consequences.
- The role of the hypothesis is to guide the researcher by delimiting the area of research and to keep him on the right track.

- **PREPARING THE RESEARCH DESIGN**
- It refers to the conceptual structure within which research would be conducted.
- The preparation of such a design facilitates research to be as efficient as possible yielding maximal information.
- The preparation of the research design, appropriate for a particular research problem, involves usually the consideration of the following:
  - The means of obtaining the information.
  - The availability and skills of the researcher and his staff (if any).
  - Explanation of the way in which selected means of obtaining information will be organised and the reasoning leading to the selection.
  - The time available for research.
  - The cost factor relating to research, i.e., the finance available for the purpose.

- **<u>DETERMINING SAMPLE DESIGN</u>**
- All the items under consideration in any field of inquiry constitute a 'universe' or 'population'. A complete enumeration of all the items in the 'population' is known as a census inquiry.

- Census inquiry is not possible in practice under many circumstances. For instance, blood testing is done only on sample basis. Hence, quite often we select only a few items from the universe for our study purposes. The items so selected constitute what is technically called a sample.

- The researcher must decide the way of selecting a sample or what is popularly known as the sample design

- Samples can be either probability samples or non-probability samples.

- With probability samples each element has a known probability of being included in the sample but the non-probability samples do not allow the researcher to determine this probability.
- Probability samples are those based on simple random sampling, systematic sampling, stratified sampling, cluster/area sampling whereas non-probability samples are those based on convenience sampling, judgment sampling and quota sampling techniques. A brief mention of the important sample designs is as follows:

I. *Deliberate sampling:*

- Deliberate sampling is also known as purposive or non-probability sampling. This sampling method involves purposive or deliberate selection of particular units of the universe for constituting a sample which represents the universe.

- When population elements are selected for inclusion in the sample based on the ease of access, it can be called *convenience sampling*.

- On the other hand, in *judgement sampling* the researcher’s judgement issued for selecting items which he considers as representative of the population. For example, a judgement sample of college students might be taken to secure reactions to a new method of teaching. Judgement sampling is used quite frequently in qualitative research where the desire happens to be to develop hypotheses rather than to generalize to larger populations.

II. Simple Random Sampling:

- This type of sampling is also known as chance sampling or probability sampling where each and every item in the population has an equal chance of inclusion in the sample.

- For example, if we have to select a sample of 300 items from a universe of 15,000 items, then we can put the names or numbers of all the 15,000 items on slips of paper and conduct a lottery.

III. Systematic sampling

- In some instances the most practical way of sampling is to select every 15th name on a list, every 10th house on one side of a street and so on. Sampling of this type is known as systematic sampling.

IV. Stratified sampling:

- If the population from which a sample is to be drawn does not constitute a homogeneous group, then stratified sampling technique is applied.

- In this technique, the population is stratified into a number of non- overlapping subpopulations or strata and sample items are selected from each stratum.

V. Quota sampling:

- In stratified sampling the cost of taking random samples from individual strata is often so expensive that interviewers are simply given quota to be filled from different strata, the actual selection of items for sample being left to the interviewer's judgement. This is called quota sampling.

- The size of the quota for each stratum is generally proportionate to the size of that stratum in the population. Quota sampling is thus an important form of non-probability sampling.

VI. *Cluster sampling and area sampling:*

- Cluster sampling involves grouping the population and then selecting the groups or the clusters rather than individual elements for inclusion in the sample.

- Suppose some departmental store wishes to sample its credit card holders. It has issued its cards to 15,000 customers. The sample size is to be kept say 450. For cluster sampling this list of 15,000 card holders could be formed into 100 clusters of 150 card holders each. Three clusters might then be selected for the sample randomly.

- *Area sampling* is quite close to cluster sampling and is often talked about when the total geographical area of interest happens to be big one.

- Under area sampling we first divide the total area into a number of smaller non-overlapping areas, generally called geographical clusters, then a number of these smaller areas are randomly selected, and all units in these small areas are included in the sample.

VII. Multi-stage sampling:

- This is a further development of the idea of cluster sampling. This technique is meant for big inquiries extending to a considerably large geographical area like an entire country.

- Under multi-stage sampling the first stage may be to select large primary sampling units such as states, then districts, then towns and finally certain families within towns.

VIII. Sequential sampling:

- This is somewhat a complex sample design where the ultimate size of the sample is not fixed in advance but is determined according to mathematical decisions on the basis of information yielded as survey progresses.

- This design is usually adopted under acceptance sampling plan in the context of statistical quality control.

- **COLLECTING THE DATA:**

- In dealing with any real life problem it is often found that data at hand are inadequate, and hence, it becomes necessary to collect data that are appropriate.

- There are several ways of collecting the appropriate data which differ considerably in context of money costs, time and other resources at the disposal of the researcher.

- Primary data can be collected either through experiment or through survey.

- in the case of a survey, data can be collected by any one or more of the following ways:

I. ***By observation:*** This method implies the collection of information by way of investigator's own observation, without interviewing the respondents.
II. ***Through personal interview:*** The investigator follows a rigid procedure and seeks answers to a set of pre-conceived questions through personal interviews.

III. Through telephone interviews:

- This method of collecting information involves contacting the respondents on telephone itself.

IV. By mailing of questionnaires:

- The researcher and the respondents do come in contact with each other if this method of survey is adopted.
- Questionnaires are mailed to the respondents with a request to return after completing the same. It is the most extensively used method in various economic and business surveys.
- Before applying this method, usually a Pilot Study for testing the questionnaire is conduced which reveals the weaknesses, if any, of the questionnaire.

- ***V. Through schedules:***
- Under this method the enumerators are appointed and given training. They are provided with schedules containing relevant questions.
- These enumerators go to respondents with these schedules. Data are collected by filling up the schedules by enumerators on the basis of replies given by respondents

- **EXECUTION OF THE PROJECT:**

➢ Execution of the project is a very important step in the research process.

➢ If the execution of the project proceeds on correct lines, the data to be collected would be adequate and dependable.

➢ The researcher should see that the project is executed in a systematic manner and in time.

- **ANALYSIS OF DATA:**
- After the data have been collected, the researcher turns to the task of analyzing them.
- The analysis of data requires a number of closely related operations such as establishment of categories, the application of these categories to raw data through coding, tabulation and then drawing statistical inferences.

- **HYPOTHESIS-TESTING:**
- After analyzing the data as stated above, the researcher is in a position to test the hypotheses, if any, he had formulated earlier.
- Various tests, such as Chi square test, *t*-test, *F*-test, have been developed by statisticians for the purpose.
- The hypotheses may be tested through the use of one or more of such tests, depending upon the nature and object of research inquiry.
- Hypothesis-testing will result in either accepting the hypothesis or in rejecting it.

- **<u>GENERALISATIONS AND INTERPRETATION:</u>**
- If a hypothesis is tested and upheld several times, it may be possible for the researcher to arrive at generalisation, i.e., to build a theory.
- As a matter of fact, the real value of research lies in its ability to arrive at certain generalisations. If the researcher had no hypothesis to start with, he might seek to explain his findings on the basis of some theory. It is known as interpretation.

- **PREPARATION OF THE REPORT OR THE THESIS:**
- Finally, the researcher has to prepare the report of what has been done by him. Writing of report must be done with great care keeping in view the following:
- The layout of the report should be as follows:
  - ***The Preliminary Pages***
  - ***The Main Text***
  - ***The End Matter***
- ***The Preliminary Pages***
- *In its preliminary pages* the report should carry title and date followed by acknowledgements and foreword. Then there should be a table of contents followed by a list of tables and list of graphs and charts, if any, given in the report.

- ***<u>The Main Text</u>***
- *The main text of the report* should have the following parts:
- ***(a) Introduction:***
- It should contain a clear statement of the objective of the research and an explanation of the methodology adopted in accomplishing the research. The scope of the study along with various limitations should as well be stated in this part.
- ***(b) Summary of findings:***
- After introduction there would appear a statement of findings and recommendations in non-technical language. If the findings are extensive, they should be summarised.

- ***(c) Main report:***
- The main body of the report should be presented in logical sequence and broken-down into readily identifiable sections.
- ***(d) Conclusion:***
- Towards the end of the main text, researcher should again put down the results of his research clearly and precisely. In fact, it is the final summing up.

- ***The End Matter***
- *At the end of the report*, appendices should be enlisted in respect of all technical data. Bibliography, i.e., list of books, journals, reports, etc., consulted, should also be given in the end. Index should also be given specially in a published research report.

- **RESEARCH PROBLEM**

- A research problem, in general, refers to some difficulty which a researcher experiences in the context of either a theoretical or practical situation and wants to obtain a solution for the same.
- **It is the first step taken in research process as a defining the research problem. Research problem is used in two field i.e (i) Academic (ii) Business Field**
- **In academic field of research study, while formulating the research problem, some times the ideas taken from research guide by the researcher.**
- **In Business Field the researcher had taken the helps of management or administrative personnel while formulating the research problem.**

- While doing research, defining the problem is very important because "problem clearly stated is half-solved". This shows how important it is to **"define the problem correctly".** While defining the problem, it should be noted that definition should be unambiguous. If the problem defining is ambiguous, then the researcher will not know "what data is to be collected" or "what technique is to be used" etc.
- ***Example of an ambiguous definition:*** "Find out by how much sales has declined recently". Let us suppose that the research problem is defined in a broad and general way as follows:

- "Why is the productivity in Korea much higher than that in India"? In this type of question, a number of ambiguities are there, such as:
- 1. What sort of productivity is to be specified; is it men, machine, materials?
- 2. To which type of industry is the productivity related to?
- 3. In which time-frame are we analysing the productivity?

❖ ***Example of an unambiguous definition:*** On the contrary, a problem will be as follows:

- "What are the factors responsible for increased labour productivity in Korean textile manufacturing industries during 1996-07 relative to Indian textile industries?“

▪ Selecting or identifying the Research Problem

- The research problem undertaken for study must be carefully selected. The task is a difficult one, although it may not appear to be so. Help may be taken from a research guide in this connection.
- **Selection Criteria**

➢ 1. Your genuine enthusiasm for the problem.

➢ 2. Controversial subject should not become the choice of an average researcher.

➢ 3. The degree to which research on this problem benefits the profession and society.

➢ 4. The degree to which research on this problem will assist your professional goals and career objectives.

➢ 5. Too narrow or too vague problems should be avoided.

➢ 6. The degree to which this research will interest superiors and other leaders in the field.

➢ 7. The degree to which the research builds on your experience and knowledge.

➢ 8. Ease of access to the population to be studied and the likelihood that they will be cooperative Affordability.

- 9. Likelihood of publication.
- 10. Relationship to theories or accepted generalizations in the field.
- 11. Degree to which ethical problems are involved.
- 12. Degree to which research is unique or fills a notable gap in the literature.
- 13. Degree to which the research builds on and extends existing knowledge before the final selection of a problem is done, a researcher must ask himself the following questions:
- (a) Whether he is well equipped in terms of his background to carry out the research?
- (b) Whether the study falls within the budget he can afford?
- (c) Whether the necessary cooperation can be obtained from those who must participate in research as subjects?

- **FORMULATION OF HYPOTHESIS:**
- **Introduction**

➢ When a researcher observes known facts and takes up a problem for analysis, he first has to start somewhere and this point of starting is **Hypothesis**. In other words, one has to proceed to formulate tentative solution. This purposed solutions constitute the Hypothesis.

➢ The collection of facts (data) will be fruitful if they are either for or against this proposed solution. The tentative explanation or solutions are the very basis for research process.

➢ A research hypothesis is quite often a predictive statement, which is capable of being tested using scientific methods that involve an independent and some dependent variables. For instance, the following statements may be considered:

- **i. “Students who take tuitions perform better than the others who do not receive tuitions” or,**
- **ii. “The female students perform as well as the male students”.**

- These two statements are hypotheses that can be objectively verified and tested. Thus, they indicate that a hypothesis states what one is looking for. Besides, it is a proposition that can be put to test in order to examine its validity.

Meaning of Hypothesis

- Hypothesis are tentative, intelligent guesses as to the solution of the problem.
- Hypothesis is a specific statement of prediction. It describes in concrete terms what you expect to happen in the study.
- Hypothesis is an assumption about the population of the study.
- It delimits the area of research and keeps the researcher on the right track.

- **Hypothesis –Necessary or Not**

➢ Is the formulation of useful hypotheses always necessary and possible? It is true that hypotheses are useful and they guide the research process in the proper direction. But can hypotheses may not arise. Similarly, in exploratory studies, initially it may not be possible to set up any worthwhile hypotheses. In fact, the very purpose of such exploratory studies may be to formulate meaningful hypotheses for further formal studies. But strictly speaking, the mere fact-finding and the exploratory studies cannot be considered to be typical research studies. In all **analytical and experimental studies**, hypotheses should be set up in order to give a proper direction to them.

- **PROBLEM (VS) HYPOTHESIS**

- Hypothesis is an assumption, that can be tested and can be proved to be right or wrong.

- A problem is a broad question which cannot be directly tested. A problem can be scientifically investigated after converting it into a form of hypothesis.

- **CHARACTERISTICS OF HYPOTHESIS**

1. Conceptual Clarity - It should be clear and precise.

2. Specificity - It should be specific and limited in scope.

3. Consistency - It should be consistent with the objectives of research.

4. Testability - It should be capable of being tested.

5. Expectancy - It should state the expected relationships between variables.

6. Simplicity - It should be stated as far as possible in simple terms.

7. Objectivity - It should not include value judgments, relative terms or any moral preaching.

8. Theoretical Relevance - It should be consistent with a substantial body of established or known facts or existing theory.

9. Availability of Techniques – Statistical methods should be available for testing the proposed hypothesis.

TYPES OF HYPOTHESIS

Null Hypothesis

When a hypothesis is stated negatively, it is called null hypothesis. It is a 'no difference', 'no relationship' hypothesis. i.e., It states that, no difference exists between the parameter and statistic being compared to or no relationship exists between the variables being compared.

It is usually represented as H_O or H_0 .

Example:

- H_0: There is no relationship between a family's income and expenditure on recreation.

Alternate Hypothesis

It is the hypothesis that describes the researcher's prediction that, there exist a relationship between two variables or it is the opposite of null hypothesis. It is represented as H_A or H_1.

Example:

H_A: **There is a definite relationship between family's income and expenditure on recreation.**

- **Formulate the Hypothesis**
- The normal approach is to set two hypotheses instead of one, in such a way, that if one hypothesis is true, the other is false. Alternatively, if one hypothesis is false or rejected, then the other is true or accepted. These two hypotheses are:
- 1. Null hypothesis
- 2. Alternate hypothesis
- Let us assume that the mean of the population is m_0 and the mean of the sample is x. Since we have assumed that the population has a mean of m_0, this is our null hypothesis. We write this as
- $H_0 m = m_0$, where H_0 is the null hypothesis. Alternate hypothesis is $H_A = m$. The rejection of null hypothesis will show that the mean of the population is not m_0. This implies that alternate hypothesis is accepted.

- **FUNCTIONS OR ROLE OF HYPOTHESIS**

- It gives a definite point to the investigation and provides direction to the study.
- It determines the data needs.
- It specifies the sources of data.
- It suggests which type of research is likely to be more appropriate.
- It determines the most appropriate technique of analysis.
- It contributes to the development of theory.

- ***<u>RESEARCH DESIGN:</u>***
- The most important step after defining the research problem is preparing the design of the research project, which is popularly known as the ‘research design’.
- A research design helps to decide upon issues like what, when, where, how much, by what means etc. with regard to an enquiry or a research study. In fact, research design is the conceptual structure within which research is conducted.

Specifically, the research design highlights decisions which include:

- ✓ The nature of the study
- ✓ The purpose of the study
- ✓ The location where the study would be conducted
- ✓ The nature of data required
- ✓ From where the required data can be collected
- ✓ What time period the study would cover
- ✓ The type of sample design that would be used
- ✓ The techniques of data collection that would be used
- ✓ The methods of data analysis that would be adopted and
- ✓ The manner in which the report would be prepared In view of the stated research design decisions, the overall research design may be divided into the following (Kothari 1988):

a. The sampling design that deals with the method of selecting items to be observed for the selected study.

b. The observational design that relates to the conditions under which the observations are to be made.

c. The statistical design that concerns with the question of how many items are to be observed, and how the information and data gathered are to be analyzed.

d. The operational design that deals with the techniques by which the procedures specified in the sampling, statistical and observational designs can be carried out.

- **FEATURES OF RESEARCH DESIGN:**
- The important features of Research Design may be outlined as follows:
  - It constitutes a plan that identifies the types and sources of information required for the research problem
  - It constitutes a strategy that specifies the methods of data collection and analysis which would be adopted.
  - It also specifies the time period of research and monetary budget involved in conducting the study, which comprise the two major constraints of undertaking any research.

- **CONCEPTS RELATING TO RESEARCH DESIGN:**
- Some of the important concepts relating to Research Design are discussed below:
- **Dependent And Independent Variables:**
- A magnitude that varies is known as a variable. The concept may assume different quantitative values like height, weight, income etc. Qualitative variables are not quantifiable in the strictest sense of the term. However, the qualitative phenomena may also be quantified in terms of the presence or absence of the attribute(s) considered. The phenomena that assume different values quantitatively even in decimal points are known as 'continuous variables'. But all variables need not be continuous. Values that can be expressed only in integer values are called 'non-continuous variables'. In statistical terms, they are also known as 'discrete variables'.

- For example, age is a continuous variable, whereas the number of children is a non-continuous variable. When changes in one variable depend upon the changes in other variable or variables, it is known as a dependent or endogenous variable, and the variables that cause the changes in the dependent variable are known as the independent or explanatory or exogenous variables.
- For example, if demand depends upon price, then demand is a dependent variable, while price is the independent variable. And, if more variables determine demand, like income and price of the substitute commodity, then demand also depends upon them in addition to the price of original commodity. In other words, demand is a dependent variable which is determined by the independent variables like price of the original commodity, income and price of substitutes.

- **Extraneous Variables:**

- The independent variables which are not directly related to the purpose of the study but affect the dependent variables, are known as extraneous variables.
- For instance, assume that a researcher wants to test the hypothesis that there is a relationship between children's school performance and their self- confidence, in which case the latter is an independent variable and the former, a dependent variable. In this context, intelligence may also influence the school performance. However, since it is not directly related to the purpose of the study undertaken by the researcher, it would be known as an extraneous variable.
- The influence caused by the extraneous variable(s) on the dependent variable is technically called the 'experimental error'

- **Control**:
- One of the most important features of a good research design is to minimize the effect of extraneous variable(s) which is known as control.
- **Confounded Relationship:**
- The relationship between the dependent and independent variables is said to be confounded by an extraneous variable, when the dependent variable is not free from its effects.
- **Research Hypothesis:**
- When a prediction or a hypothesized relationship is tested by adopting scientific methods, it is known as research hypothesis. Generally, a research hypothesis must consist of at least one dependent variable and one independent variable.
- Whereas, the relationships that are assumed but not to be tested are predictive statements that are not to be objectively verified, thus are not classified as research hypotheses.

- **<u>Experimental and Non-experimental Hypothesis Testing Research</u>**:
- When the objective of a research is to test a research hypothesis, it is known as hypothesis-testing research.

- Such research may be in the nature of experimental design or non- experimental design.

- The research in which the independent variable is manipulated is known as ‘experimental hypothesis-testing research’, whereas the research in which the independent variable is not manipulated is termed as ‘non-experimental hypothesis testing research’.

- For example, assume that a researcher wants to examine whether family income influences the school attendance of a group of students, by calculating the coefficient of correlation between the two variables. Such an example is known as a non-experimental hypothesis testing research, because the independent variable - family income is not manipulated here.
- Again assume that the researcher randomly selects 150 students from a group of students who pay their school fees regularly and then classifies them into two sub- groups by randomly including 75 in Group A, whose parents have regular earning, and 75 in Group B, whose parents do not have regular earning. Assume that at the end of the study, the researcher conducts a test on each group in order to examine the effects of regular earnings of the parents on the school attendance of the student. Such a study is an example of experimental hypothesis-testing research, because in this particular study the independent variable regular earnings of the parents have been manipulated.

- **Experimental And Control Groups:**
- When a group is exposed to usual conditions in an experimental hypothesis-testing research, it is known as ‘control group’. On the other hand, when the group is exposed to certain new or special condition, it is known as an ‘experimental group’. In the afore-mentioned example, Group A can be called as control group and Group B as experimental group. If both the groups, A and B are exposed to some special feature, then both the groups may be called as ‘experimental groups’. A research design may include only the experimental group or both the experimental and control groups together.

- **<u>Treatments</u>**:
- Treatments refer to the different conditions to which the experimental and control groups are subject to. In the example considered, the two treatments are the parents with regular earnings and those with no regular earnings.
- Likewise, if a research study attempts to examine through an experiment the comparative effect of three different types of fertilizers on the yield of rice crop, then the three types of fertilizers would be treated as the three treatments.

- **<u>Experiment:</u>**
- Experiment refers to the process of verifying the truth of a statistical hypothesis relating to a given research problem.
- For instance, an experiment may be conducted to examine the yield of a certain new variety of rice crop developed.

- **<u>Experimental Unit(s):</u>**
- Experimental units refer to the pre-determined plots, characteristics or the blocks, to which different treatments are applied. It is worth mentioning here that such experimental units must be selected with great caution.

- **TYPES OF RESEARCH DESIGN**:
- There are different types of research designs. They may be broadly categorized as:
    - Exploratory Research Design;
    - Descriptive and Diagnostic Research Design
    - Hypothesis-Testing Research Design.

- **<u>Exploratory Research Design:</u>**
- The Exploratory Research Design is known as formulative research design. The main objective of using such a research design is to formulate a research problem for an in-depth or more precise investigation, or for developing a working hypothesis from an operational aspect.
- The major purpose of such studies is the discovery of ideas and insights.
- Usually, the following three methods are considered in the context of a research design for such studies. They are

a) A survey of related literature

b) Experience survey

c) Analysis of ‘insight-stimulating’ instances.

- **<u>Descriptive And Diagnostic Research Design:</u>**
- A Descriptive Research Design is concerned with describing the characteristics of a particular individual or a group.

- Meanwhile, a diagnostic research design determines the frequency with which a variable occurs or its relationship with another variable. In other words, the study analyzing whether a certain variable is associated with another comprises a diagnostic research study.

- On the other hand, a study that is concerned with specific predictions or with the narration of facts and characteristics related to an individual, group or situation, are instances of descriptive research studies. Generally, most of the social research design falls under this category.

- As a research design, both the descriptive and diagnostic studies share common requirements, hence they are grouped together. However, the procedure to be used and the research design need to planned carefully.
- The research design must also make appropriate provision for protection against bias and thus maximize reliability, with due regard to the completion of the research study in an economical manner.

- The research design in such studies should be rigid and not flexible. Besides, it must also focus attention on the following:
  - Formulation of the objectives of the study,
  - Proper designing of the methods of data collection,
  - Sample selection,
  - Data collection,
  - Processing and analysis of the collected data, and
  - Reporting the findings.

- **<u>Hypothesis-Testing or experimental Research Design</u>**:
- Hypothesis-Testing Research Designs are those in which the researcher tests the hypothesis of causal relationship between two or more variables.

- These studies require procedures that would not only decrease bias and enhance reliability, but also facilitate deriving inferences about the causality. Generally, experiments satisfy such requirements. Hence, when research design is discussed in such studies, it often refers to the design of experiments.

- **BASIC PRINCIPLES OF EXPERIMENTAL DESIGNS**

- Professor Fisher has enumerated three principles of experimental designs:

1) The Principle of Replication
2) The Principle of Randomization
3) Principle of Local Control.

- **THE PRINCIPLE OF REPLICATION** :

❖ According to the Principle of Replication, the experiment should be repeated more than once. Thus, each treatment is applied in many experimental units instead of one. By doing so the statistical accuracy of the experiments is increased.

❖ For example: suppose we are to examine the effect of two varieties of rice. For this purpose we may divide the field into two parts and grow one variety in one part and the other variety in the other part. We can then compare the yield of the two parts and draw conclusion on that basis. But if we are to apply the principle of replication to this experiment, then we first divide the field into several parts, grow one variety in half of these parts and the other variety in the remaining parts. We can then collect the data of yield of the two varieties and draw conclusion by comparing the same. The result so obtained will be more reliable.

- **The Principle of Randomization:**
  - ❖ It provides protection, when we conduct an experiment, against the effect of extraneous factors by randomization. In other words, this principle indicates that we should design or plan the experiment in such a way that the variations caused by extraneous factors can all be combined under the general heading of "chance."
  - ❖ For instance, if we grow one variety of rice, say, in the first half of the parts of a field and the other variety is grown in the other half, then it is just possible that the soil fertility may be different in the first half in comparison to the other half. If this is so, our results would not be realistic. In such a situation, we may assign the variety of rice to be grown in different parts of the field on the basis of some random sampling technique i.e., we may apply randomization principle and protect ourselves against the effects of the extraneous factors (soil fertility differences in the given case). As such, through the application of the principle of randomization, we can have a better estimate of the experimental error.

- **The Principle of Local Control** :
  - ❖ It is another important principle of experimental designs. Under it the extraneous factor, the known source of variability, is made to vary deliberately over as wide a range as necessary and this needs to be done in such a way that the variability it causes can be measured and hence eliminated from the experimental error.
  - ❖ In other words, according to the principle of local control, we first divide the field into several homogeneous parts, known as blocks, and then each such block is divided into parts equal to the number of treatments. Then the treatments are randomly assigned to these parts of a block. Dividing the field into several homogenous parts is known as ‘blocking’.

- **IMPORTANCE OF RESEARCH DESIGN:**
  - ❖The need for a research design arises out of the fact that it facilitates the smooth conduct of the various stages of research.
  - ❖It contributes to making research as efficient as possible, thus yielding the maximum information with minimum effort, time and expenditure.
  - ❖A research design helps to plan in advance, the methods to be employed for collecting the relevant data and the techniques to be adopted for their analysis.

- ❖ This would help in pursuing the objectives of the research in the best possible manner, provided the available staff, time and money are given.

- ❖ Hence, the research design should be prepared with utmost care, so as to avoid any error that may disturb the entire project.

- ❖ Thus, research design plays a crucial role in attaining the reliability of the results obtained, which forms the strong foundation of the entire process of the research work.

- **CHARACTERISTICS OF A GOOD RESEARCH DESIGN:**

  - ❖A good research design often possesses the qualities of being flexible, suitable, efficient, economical and so on.

  - ❖Generally, a research design which minimizes bias and maximizes the reliability of the data collected and analyzed is considered a good design.

Thank you.

BUSINESS RESEARCH MODULE – I (PART B)

By Asst. Prof. (Eco) Mrs. Tapaswini Nayak.

Methods Of Data Collection

- A researcher uses 2 types of data
- Primary Data
- Secondary Data

- **PRIMARY DATA**
- Primary data are the fresh data obtained for first time.
- It is expensive and time consuming.
- These are original in character.

- **SECONDARY DATA**
- Secondary data are second hand data which are already been collected.

- **COLLECTION OF PRIMARY DATA**
- **Questionnaire Method**
  - It is popular in big inquiries.
  - This method is adopted by private individual, research workers, private and public organization and by govt.
  - In this method a questionnaire is sent to the person with a request to answer the questions and return it back.
  - A questionnaire contains a number of questions printed in a logical order.
  - The questionnaire is mailed to the respondents who are expected to read and understand questions and write down replay in space given,
  - Respondent has to answer the questions as per his own knowledge.

- **Advantages**
  - It is less costly and covers large area.
  - Free from bias.
  - Who are not easily approachable they can be approached through questionnaire.
  - A large sample size can be covered.
- **Disadvantages**
  - Low rate of return (no guaranty of reply).
  - This method can only be used when respondents are educated and cooperative.
  - No control over the questionnaire once it is sent.
  - Not flexible.
  - May not get correct answer.

- **Main aspects of a questionnaire**
  - **General form**
- It can be either structured or unstructured questionnaire

- **Structured Questionnaire**

- Structured questionnaire are definite, concrete and predetermined questions.
- Questions are presented with exact wordings to each respondents.
- Question may be closed (yes or no type) or open (inviting free response i.e., long and descriptive.
- The questions may have alternative answers.
- By using the above structure use of own word by the respondent is minimized.

- **Unstructured Questionnaire**
- Unstructured questionnaire uses open questions.
- Questionnaire may contain different questions for different respondent.
- **Question Sequence**
  - A proper sequence of questions reduces chances of misunderstanding.
  - The question sequence must be clear and logical.

- **Example :**
- ✓ What is your name?
- ✓ What is your qualification?

- ❖ The first question should be placed first before second question.
- ❖ If we interchange the sequence, it may not be logical.
- ❖ The following questions should be avoided during the beginning.

- ▪ Questions that put great stress on memory.
- ▪ Questions of personal character.
- ▪ Questions related to personal wealth.
- ▪ Question sequence can be set by use of pilot survey.

- **Question Formulation And Wording**
  - Questions must follow the following standard.
    - ➢ It should be easily understand.
    - ➢ It should be simple (should have one meaning).
    - ➢ It should be concrete.

- **Example :**

❖ How many razor blades do you use annually instead of this question, the question may be.

❖ How many razor blades did you use last week.

❖ Abused words should not be used.

- **Essentials of A Good Questionnaire**
  - Must be short and simple.
  - Question should process in a logical order.
  - Question should be easily understood.
  - Question should move from easy to difficult.
  - Personal questions should be avoided.
  - Highly technical questions should be avoided.
  - Adequate space must be given for answer.
  - Alternatives like – don’t know, no preference should be there.
  - Quality of paper and colour must be attractive.

- **Interview Method**
  - **Personal Interview**
    - This interview is based on direct or face to face contact with respondent.
    - At a point of time questions may be asked to one or more persons.
    - It is a method of direct personal investigation.
    - These are conducted usually in a very structured way.
    - All questions are pre-planned and cross questions may arise.
    - Unstructured interview don not follow any planned format.
    - In unstructured interview question asked arbitrarily.
    - All question and answers are recorded for further analysis.

- **Focused interview**
  - Interviewer focuses on the experience of interviewee.
  - Interviewer decides the manner and sequence of questions.

- **Example**

❖ Interview of a political leader, sports person etc.

- **Clinical Interview**
  - This interview deals with knowing experience and feelings.
  - It is a mixture of observation and interview.

- **Example**

❖ Doctor’s treatment

- **Non Directive Interview**
  - Here the interviewer encourage the respondent to talk on a topic with bare minimum questions.
  - Few questions are asked and more explanations are desired.

- **Advantages of Personal Interview**
  - More information, greater depth.
  - Interviewer uses own skill to get answer.
  - Greater flexibility.
  - Verbal answers can be recorded (Audio).
  - Personal information can be obtained.
  - More control over questions.

- **Disadvantages of Personal Interview**
  - Expensive.
  - Does not cover wide geographical area.
  - Possibility of bias of interviewer and respondent.
  - Some respondent may not give time.
  - Less time for response.
  - Some answers can’t be spoken but written.
  - Time consuming.

- **Pre-Requisites of a Good Interview**
  - Interviewer should be carefully selected, trained and educated.
  - Technical competency and experience
  - Interviewers should be enthusiastic, hardworking honest and polite.
  - Interviewer must be interesting
  - Questions must not be irritating.
  - Respondent should fell that his answers are real, valuable and listen carefully.
  - Doubtful or cross questions asked by respondent must be handled carefully and politely.

- **Telephonic Interview**
  - Respondents should be contacted through telephone.
  - It is mainly used in industrial sector like telecom, financial corporation like LIC, SAHARA etc.

- **Advantages of Telephonic Interview**
  - More flexible
  - Wide coverage
  - Less cost than personal interview
  - High response
  - No fields staff
  - Time saving

- **Disadvantages of Telephonic Interview**
  - Limited time for response
  - Respondent may not have telephone
  - Chances of wrong information is higher.
  - Large geographical area not touched because of more cost and possibility of bias.

- **Other Methods for Primary Data Collection**
  - **Warranty Cards**
    - It is in postcard sized.
    - It is sent to the customers to fill aspects of product
    - Customers fill-up required field and send it back.
  - **Distributors or Store Audit**
    - Distributors timely send auditors to retailers or stockist.
    - They ask for product information and response of customer.
  - **Free Samples**
    - Some free samples are distributed among the customers, distributors, retailers, doctors etc.
    - After use of the product their feedback is taken.
    - As per information improvement measures are taken.

- **Feedback**
  - After use of a particular services or good feedback is taken from consumers.
  - Customers have to fill the required filed and send it back to the retailer.
  - Mechanical device: it is a popular method of data collection.
  - This methods uses camera, audio recording devices, mobile phones to record data,

Example

- ❖ Journalist
- ▪ **Depth Interview**
  - It is used to discover underlying truth, attribute tendency.
  - A large variety of direct and indirect questions are asked.
  - It needs high degree of skill and presence of mind.

- **SCHEDULE**
  - Schedule method is very much similar to questionnaire method.
  - Schedule means proforma containing a set of questions.
  - It is filled by enumerators who are specially appointed.
  - Enumerators along with schedule goes to the respondent ask questions and record the answers.
  - In some situations schedule is handed over to the respondent and enumerator guides to fill it.
  - Enumerator must be carefully selected and trained.
  - He should verify the answers on the spot.
  - He should be honest, sincere and hardworking.
  - It is useful in extensive enquiries and its very costly.
  - It is conducted by government and big industries.
  - Population census all over the world made by this method.

- **Difference Between Questionnaire and Schedule**

Questionnaire	Schedule
❖ Questionnaireissentthroughmail ❖ It is cheap ❖ Nonresponse chance is higher ❖ Questionnairemethodisveryslow ❖ Personal contact maynot possible ❖ Not helpful if respondent is illiterate ❖ Helpfulinlargegeographicalarea.	❖ Enumerators directly goes tothe respondent ❖ It is expensive ❖ Certain response ❖ Schedule method is fast ❖ Personal contact is essential ❖ Literacy nevermatters. ❖ Not possible in schedule.

- **SECONDARY DATA**
- When the researcher used data which are previously collected by some other researcher, institution or agencies for their own purpose are called secondary data.
- The researcher can collect secondary data either form internal source of an organization i.e., unpublished source or from the published sources likes reports, journals, magazine, internet etc.,

- # Difference Between Primary and Secondary Data

Primary Data	Secondary Data
❖ The data is collected by the researcher himself for finding the solution of a particular problem is know as primary data	❖ When a data which was earlier collected by some researcher or organizations for their own purpose is used in current research for similar purpose is called secondary data.
❖ More costly	❖ Less costly
❖ It is directly collected from the respondents	❖ It is collected from some published or unpublished sources (magazine, journal, internet etc.,)

❖ Methods used for primary data collection are interview, mailing questionnaire method, observation and survey.	❖ Methods for secondary data collection, include study of journals, reports, census and different database.
❖ It is more original, reliable data.	❖ It is lessreliable
❖ Form of primary data is raw in nature and needs to be processed.	❖ Secondary data is already processed. It only needs to be analysed
❖ More accuracy of data.	❖ Less accuracy of data.

- **SAMPLING DESIGN**
- **Meaning**
- Research objectives are generally translated into research questions that enable the researchers to identify the information needs. Once the information needs are specified, the sources of collecting the information are sought. Some of the information may be collected through secondary sources (published material). Whereas the rest may be obtained through primary sources.
- The primary methods of collecting information could be the observation method, personal interview with questionnaire, and their analysis plays a vital role in finding answers to research questions. Survey respondents should be selected using the appropriate procedures, otherwise the researchers may not be able to get the right information to solve the problem under investigation. The process of selecting the right individuals, objects or events for the study is known as sampling. Sampling involves the study of a small number of individuals, objects chosen from a large group.

- **Sampling Concepts**
- Before we get into the details of various issues pertaining to sampling, it would be appropriate to discuss some of the sampling concepts.

➢ **Population :** Population refers to any group of people or object that form the subject of study in a particular survey and are similar in one or more ways. For example, the number of fulltime MBA students, the population size would be 200. We may be interested in understanding their perceptions about business education. If there are 200 class IV employees would form the population of interest. If a TV manufacturing company produces 150 TVs per week and we are interested in estimation the proportion of defective TVs produced per week, all the 150 TVs would form our population. If, in an organization there are 1000 engineers, out of which 350 are mechanical engineers and we are interested in examining the proportion of mechanical engineers who intend to leave the organization within six months, all the 350 how the patients in a hospital are looked after, then all the patients of the hospital would fall under the category of population.

- **Element:** An element comprises a single member of the population. Out of the 350 mechanical engineers mentioned above, each mechanical engineer would form an element of the population. In the example of MBA students whose perception about the management education is of interest to us, each of the 200 MBA students will be an element of the population. This means that there will be 200 elements of the population.
- **Sampling frame:** Sampling frame comprises all the elements of a population with proper identification that is available to us for selection at any stage of sampling. For example, the list of registered voters in a constituency could form a sampling frame; the telephone directory; the number of students registered with university; the attendance sheet of a particular class and the payroll of an organization are examples of sampling frames. When the population size is very large, it becomes virtually impossible to form a sampling frame. We know that there is a large number of consumers of soft drinks and, therefore, it becomes very difficult to for the sampling frame for the same.

- **Sample:** It is a subset of the population. It comprises only some element of the population. If out of the 350 mechanical engineers employed in an organization, 30 members would constitute the sample.
- **Sampling unit:** A sampling unit is a single member of the sample. If a sample of 50 students is taken from a population of 200 MBA students in a business school, then each of the 50 students is a sampling unit. Another example could be that if a sample of 50 patients is taken from a hospital to understand their perception about the services of the hospital, each of 50 patients is a sampling unit.

- **Sampling:** It is process of selection an adequate number of elements from the population so that the study of the sample will not only help in understanding the characteristics of t population but will also enable us to generalize the results. We will see later that there are two types of sampling designs-probability sampling design and non-probability sampling design.
- **Census (or complete enumeration):** An examination of each and every element of the population is called census or complete enumeration. Census is an alternative to sampling.

- **STEPS IN SAMPLING DESIGN**
- The sampling design process includes five steps that shown sequentially in figure 2.5 these steps are closely interrelated and relevant to all aspects of the management research project, from problem definition to the presentation of the results. Therefore, sample design decision should be integrated with all other decisions in a research project.

- Define the target population
- ↓
- Determine the sampling frame
- ↓
- Select a sampling technique(s)
- ↓
- Determine the sample size
- ↓
- Execute the sampling process

- **Define the target population**
- Sampling design begins by specifying the target population. The target population is the collection of elements or objects that possess the information sought by the researcher and about which inferences are to be made. The target population must be defined precisely. Imprecise definition of the target population will result in research that is ineffective at best and misleading at worst. Defining the target population involves translating the problem definition into a precise statement of who should and should not be included in the sample. The target population should be defined in terms of elements. Sampling units, extent and tie. An element is the object about which of from which the information is desired.
- Suppose that Revlon wanted to assess consumer response to a new line of lipsticks and wanted to sample females over 18 years of age. It may be possible to sample females over 18 directly, in which case a sampling unit would be the same as an element. Alternatively, the sampling unit might be households. In the latter case, households would be sampled and all females over 18 in each selected household would be interviewed. Here, the sampling unit and the population element are different.

- **Determine the Sampling Frame**
- **A sampling frame** is a representation of the elements of the target population. It consists of a list or set of directions for identifying the target population. Examples of a sampling frame include the telephone book, an association directory listing the firms in an industry, a mailing list purchased from a commercial organization, a city directory, or a map.
- If a list cannot be compiled, then at least some directions for identifying the target population should be specified, such as random digit dialing procedures in telephone surveys. In the opening duck stamp example, the sampling frame consisted of a computer program for randomly and efficiently generating telephone numbers, excluding nonworking and non household numbers.

- **<u>SAMPLE DESIGN</u>**
  - A sample design is a definite plan for obtaining a sample from a given population.
  - It refers to the technique or the procedure which the researcher would adopt for selecting the items for the sample.
  - Before data collection sample design is done.
- **Steps of sample design :**
  - **1. Objective**
    - Objective of the survey must be clear of concrete.
    - Researcher should confirm that the objectives are according to money, man power and time limit available for the survey.
  - **2. Population**
    - Population should be clearly defined according to the objective of the survey.
    - **3. Sampling Units And Frame**
  - Before selecting sample sampling units should be fixed.
  - Sampling unit may be geographical one such as state, district, village etc., or a construction unit such as a house, flat etc. it may be a social unit such as family, club, school etc. or it may be an individual.
  - The researchers have to decide one or more of such unit that he has to select for his study.
  - The list of sampling unit is called as sampling frame.

- **4. Size of Sample**
  - It refers to the number of items to be selected from the universe to constitute a sample.
  - The size of sample should not be excessively large or too small.
  - The sample size should be optimum.
- **5. Parameters of Interest**
  - Statistical constant of the population are called as parameters.

Example

- ❖ Population mean, population standard deviation.
- ❖ For determining sample design one should specify the population parameter such as average or proportion.

- **6. Data Collection**
  - Only essential information about the sample should be collected.
-
  - **7. Non-Respondents**
    - Due to some practical difficulties data may not be collected for all the sample units.
    - Such cases should be handled with care.
  - **8. Selection of Proper Sampling Design**
    - The researcher must decide the technique to the used in selecting the items for the sample.
- **9. Organizing Field Work**
  - There should be efficient supervisory staff and trained personnels for the success of field work.

- **10. Pilot Survey**
  - It is always helpful to try the sample design on small scale before going to the field to know the practical problems and troubles which is called pilot survey.
- **11. Budgetary Constraint**
  - Costs have a major impact on decision relating to not only the size of the sample but also the type of sample.
  - For less cost non-probability sampling is used.
- **FUNDAMENTALS OF SAMPLING**
  - Sampling error
  - Non sampling error

- **Sampling Error**
  - This error arises in drawing or selecting sample.
  - It arises because we don’t take total elements of population into our study. We choose subpart from population which represents the total population.
  - This may be due to enumerator bias or by chance.
  - Similarly chance of error rises, when sample size reduces.

Example 1

- Suppose your study is fuel efficiency of cars in Bhubaneswar.
- Population is 50,000/- and sample is 300.
- By chance in drawing random sample all the Maruti cars are selected, then sampling error arises.

Example 2

- If your study is average height of trees in chandaka forest.
- By chance you selected an area in which all trees are of short height.
- It may lead to wrong result.
  - **Non Sampling Error**
    - Error may arise in processing of data, interpretation of result, data collection, hypothesis testing etc.
    - When error arises due to factors other than drawing sample, is called non sampling error.

- **Difference Between Sampling and Complete Enumeration**

Sampling	Complete Enumeration
❖ It a part of the population selected for researcher.	❖ It is complete study of all elements of population for research.
❖ Less costly	❖ More costly
❖ Less time consuming	❖ More time consuming
❖ It can be applied for most of the studies like economy, medicines, market, consumers etc., (scope is brand)	❖ It’sscopeisveryless.Itisusedonly ofcensuspopulationcalculation.
❖ Helpful for infinite population	❖ Nothelpfulforinfinitepopulation.
❖ Theoretically less accuracy	❖ Practically lessaccuracy
❖ Sampling error is possible	❖ Sampling error is absent

- **Types of Sample Design**
  - Simple random sampling
  - Deliberate sampling
  - Judgment sampling
  - Systematic sampling
  - Stratified sampling
  - Quota sampling
  - Cluster sampling
  - Area sampling / geographical sampling
  - Multi-stage sampling
  - Sequential sampling

- **Simple Random Sampling**
  - It is a chance or probability sampling
  - Each item of the population has equal chance to be selected.
  - It is based on lottery system and hence bias free.

  Example

  - ❖ All students have equal chance to be selected randomly.

- **Deliberate Sampling**
  - It is also called purposive or non-probability sampling.

  Example

  - ❖ Suppose researcher enters into some class and select some students as per his wish.
  - ❖ He may or may not enter into some classes.

- **Judgment Sampling**
  - Researcher chose elements on the basis of his judgment.

  Example

  - ❖ He decides which students will be taken for an appropriate research.

- **Systematic Sampling**
  - Entire population is numbered and out of that some specific numbers are selected.

Example

❖ all students are numbered from 1 to 15,00 like

0, 1, 2, 3, 4, 5,……………..,1500.

❖ Every even number student will be taken for research.

- **Stratified Sampling**
  - If the population is not in a homogenous group the stratified sampling is taken.

Example

❖ Students are stratified into different groups according to their course like MBA, MCA, ECE, EEE, MECH etc., then some samples are taken on random basis.

MBA	MCA	ECE	EEE	MECH
5	3	7	10	11

- **Quota Sampling**

  It is just like stratified sampling but here some quota is assigned.

Example

- ❖ 1 student is taken as sample out of each 30 students.

MBA	MCA	ECE	EEE	MECH
30	60	120	90	120
1	2	4	3	4

- **Cluster Sampling**
  - Here the entire population is grouped into some clusters.
- Some clusters are selected for observation. In the above example sample size = 300 We made 15 cluster of 100 students each

- Out of 15 clusters 3 clusters are selected.

- **Area Sampling**
  - Here the entire population is divided into some geographical areas.
  - Then randomly some geographical areas are selected for sample.
- **Multistage Sampling**
  - It is meant for a large geographical area like country.
  - Here some samples are collected from country, then state, then district then village etc.
- **Sequential Sampling**
  - In this technique the size of sample is not prefixed.
  - As per the requirement the size is determined.
  - Used for statistical quality control.

- **COLLECTING DATA**
  - Collection of data is needed for conducting any type of research.
- Data may be of 2 types
  - Primary
  - Secondary
- **Primary**
  - Primary data can be collected through experiment or through survey.
- **Secondary**
  - Secondary data can be obtained from annual reports, news papers, journals, magazines, internets etc.

The methods for collecting primary data are :

- **By Observation :**
  - By own observation of the researcher data are collected

Example

❖ Doctor's treatment

- **Personal Interview**
  - Questions are asked to the respondents and answers are recorded.
  - Success of this method id dependent on ability of the researcher.

  Example

  - News reporters.

- **Telephonic Interview**
- In this method interview is taken by telephone.
- It is suitable where the where the respondent are from wide geographical areas.
- **By Mailing Questionnaire**
- Here there is no direct contact between researcher and respondent.
- A questionnaire is sent by post or e-mail with a request to give answer.
- By filling the relevant answers the respondent sent it back to researcher.
- **Schedule**
- Here enumerators are appointed and trained.
- A schedule is prepared with relevant questions.
- A schedule contains questions I n table format.
- Enumerator ask questions to respondent and record answer in schedule.

Example

- Census survey

- **Execution of Project :**
  - Researcher has to supervise the research work.
  - Suitable questionnaire is prepared with probable answers.
  - Answers are recorded properly and stored safely.
- **Data Analysis**
  - The collected data are analyzed properly.
  - The data are grouped, tabulated and coded.
  - Bulk data are converted into meaningful groups.
  - Unwanted data are removed from data base.
  - Data analysis involves calculation of correlation, standard deviation, mean etc.
- **Hypothesis Testing**
  - Hypothesis can be tested by chi-square test, t-test, f-test, z-test etc.
  - Hypothesis testing is to know whether hypothesis will be accepted or rejected.

- **Measurement and Scaling Technique**
  - We use some yardstick to determine weight, height etc.,
  - We also judge how well we like a song a painting or personality.
  - We measures both physical objects and abstract concepts (ego, attitude, personality etc.)
  - But measuring abstract concepts are more complex and complicated.
  - Properties like weight, height etc., can be measured directly with some standard units of measurement but it is not easy to measure attitude, personality, motivation, feelings.
  - Measurement is made by scale.

Example

- Rice, wheat, sugar is measured in kg, height, length is measured in meter.

- **Measurement is Scales**
  - **Nominal Scale**
    - In nominal scale simply numbers symbols are assigned.

Example

- Assigning number to Football players in order to identify them.
- Such numbers are just for convenient labels.
- If signifies nothing.
- Roll number and registration number are also nominal scales.
- There is no arithmetic origin of this scale.

  - **Ordinal Scale**
    - It places events in order.

Example

- OJEE rank
- We can say that rank 5 is better than rank 10.
- Ordinal scale means ranking from lowest to highest.
- A particular rank signifies greater than or less than a rank.
- Difference between rank 1 and 2 may not be the exact difference between 5 and 6.

- **Internal Scale**
  - In interval scale zero is in interval. It is having an arbitrary zero.
  - Both the ends are open.
  - Better example is Fahrenheit scale or Celsius scale.
  - Here zero is not absolute. Means zero is not the end temperature goes in positive side or negative side.
  - In both sides there is no limit.

-4 -3 -2 -1 0 1 2 3 4

- **Ratio scale**
  - Ratio scale have an absolute zero.
  - It is having one end closed and one open.

Example

- C.M. Scale.
- ‘O” indicates absence of height or length.
- It represents actual amount of variables.

- **Example**
  - 10. Cm. is exactly double of 5 Cm.
  - It measures physical dimensions like weight, height distance etc.

0 1 2 3 4 5 6 8

- **Errors in Measurement**
  - Measurement should be precise, unambiguous in an ideal research.
  - After lots of care error may occur in measurement.
  - Researchers must be aware of the possible errors in measurement.

Following may be the sources of errors.

- **Response**
  - Respondent may hide information in the field in which researchers has less knowledge.
- **Situation**
  - In the presence of another person he may not say truth.
- **Attitude of enumerator**
  - His behaviours, style, way of talking, looks encourage of discourage respondent in providing information.
- **Instrument**
- Errors may arise because of the defective measurement tools.

- **Steps of Development of Measurement Tools**
  - **Concepts Development**
    - At first concept should be developed.

Example

❖ Jeans buying behaviours in Bhubaneswar.

❖ There are so many brands, outlet and variety available.

-
  - **Dimension of the Concept**
    - Different dimensions may be price, variety, availability, brands etc.
-
  - **Selection of Indicator**
    - For each dimension an indicator is fixed.

Example

❖ Price by Rupees or Dollar, Brands by Rank of Company etc.

- **Scaling**
  - In research more often we face measurement problem.
  - Scaling means describing the procedure of assigning numbers to various degrees of opinion and other concepts.
- **Example**
  - We are going to measure teaching quality of a teacher.

The dimensions may be teaching skill, time management, working (communication), body language, handling students.

Then we assign scale to each dimension.

- Highly good
- Good
- Average
- Below average
- Poor

- Or we may assign some points varying from 0 to 10.
- **Scaling Technique**
  - Rating scale
  - Itemized scale : a. Likert scale b. Semantic differential scale c. Stapled scale
  - Arbitrary scales
  - Differential scale
  - Summated or likert type scale
  - Cumulative scale
  - Factor scale
  - Multidimensional scaling
- **Rating scale**
  - This scaling technique involve description of a limited number of aspects of a thing or trait of a person.

- **Rating scale**
  - This scaling technique involve description of a limited number of aspects of a thing or trait of a person.

- The rating may be in the following forms.
- ➢ Like very much ->Excellent
- ➢ Like some what ->Good
- ➢ Neutral ->Average
- ➢ Dislike some what ->Below average
- ➢ Dislike very much ->Poor

- There is no specific rule to use 2 point scale, 3 point scale, 5 point scale or more than that.
- **The Graphic Rating Scale**
  - In this method the various points are usually put along a line.

Like Like Neutral Dislike Dislike
Very much some what some what very much

- **The itemized rating scale**
  - It presents a series of statements from which respondent select one.
- **Example**
    - How well a worker mixes with his co-worker.
    - He always makes quarrel with others.
    - He rarely make quarrel with others.
    - He never make quarrel with others.
    - Others are quarrelling with him.
- **Arbitrary scales**
  - It means random.
  - This scale is developed on ad hock basis.
  - Researcher arbitrarily select some statements.
  - Some statements are selected instantly.
  - It is easy and quick.

- **Differential scale (Thurston type scale)**
  – In this method selection of items are made by a panel of judges.
  – Judges evaluate whether items are related to topic or not.
-
- **Procedure**
  – Researcher gather a large number of statements regarding a particular thought or group ideas usually 20.
- **Example**
    - Opinion regarding war.
    - War creates destruction.
    - War raises power.
    - War should be abolished.
    - Nuclear war is harmful.
    - War creates division.
    - War is loss both parties.
    - War should be made by UNO approval.

- ✓ Such statements are given to a panel of judges.
- ✓ A copy is given to each jury.
- ✓ They eliminate some statements which are not relevant.
- ✓ They retain N statements.
- ✓ The N statement are ranked 1^{st} to Nth.
- ✓ 1^{st} is most unfavourable.
- ✓ 2^{nd} is next unfavourable.
- ✓ Nth is most favourable.
- ✓ When ambiguity arise regarding rank of a statement that is eliminated.
- ✓ Finally 5 to 6 statements are selected for final statements for scale.
- ✓ This is a filtration process.

- **Summated or Liker Type Scale**
  - It consist of a number of statements which express favourableness or unfavourableness of attitude.
  - Respondent indicates his agreement or disagreement in particular subject matter.
  - Each response is given a numerical score.
  - The overall score represents idea about a particular subject matter.
- **Example**
  - ❖Job satisfaction in a particular company

Strongly Agree(5) Agree(4) Undecided(3)
Disagree(2) Strongly Agree(1)

- ✓ Each point carry some score.
- ✓ Suppose we made a research on 30 people.

  All give (1)   30 x 1 = 30

  All give (5)   30 x 5 = 150

- ✓ Score varies from 30 to 150
- ✓ A less score represent high job satisfaction.
- ✓ A high score represents less job satisfaction.

- **Procedure for Likert Scale :**
  - Researcher collects large number of statements which are relevant to the attitude being studied. Each statements expresses definite favourableness and unfavourableness to a particular point of view or the attitude.
  - After the statement have been gathered a trial test should be administrated to a number of subject.
  - The response to various statements are scored in such a way that a response indicative of most favourable attitude is given the highest score of 5 and most unfavourable attitude is given the lowest score i.e., 1.
  - Then total score of each respondent is obtained by adding his scores that he received for separate statements.
  - The next step is to array these total scores and find out those statements which have a high discriminatory power.
  - The researcher may select some part of the highest and lowest total score.
  - Only those statement that correlate with the total test should be retained in the final instrument and all others must be discarded from it.

- **Advantages**
  - It is easy to construct the likert type scale because panel judges are not required.
  - It is more reliable because under it respondents answer each statement included in the instrument.
  - It takes very less time.
- **Limitations**
  - In likert type scale we can examine whether response are more or less favourable to a topic but we can not measure how much more or less.
- **Cumulative Scale**
  - It is also called luis Gutlaman's scalogeram.
  - It contains a series of statements to which respondent expresses his agreement or disagreement.
  - Statements are in cumulative form.
  - In this scale who says item (3), then he is agree with 1,2,3 points.
  - Who says item 4, he is agree with 1,2,3,4 points.

- **Example**
- Facilities in an outlet.
    - It is having fans.
    - It is having fans and ACs.
    - It is having fans, ACs and lift.
    - It is having Fan, ACs, Lift and cafeteria.
    - Who says point (iv), he is automatically agree with point (i), (ii) and (iii).
- **Semantic Differential Scale**
    – It is developed to measure the psychological meanings of an object to an individual.
    – This scale is based on the presumption that an object can have different dimensions of lornotative meanings which can be located in multi- dimensional property space.

- **Leadership Position of Candidate**
- **Example :**

3	2	1	0	-1	-2	-3
Successful						Unsuccessful
Strict						Lenient
Progressive						Regressive
Strong						Weak
Active						Passive
Fast						Slow
True						False
Sociable						Unsociable

- We have to score the candidates from +3 to -3. On the basis of above state scales.
- Numeric values are not written in actual scale

Bipolar adjectives

- **Multidimensional Scaling**
  - When more than one scale is associated with a particular aspect then multidimensional scale is appropriate.
- **Example**
  - ❖ Liquid can be measure in both liter and K.G.
- Both physical and abstract concept may have more than one scale associated.
- It is also used when all the variables in a study are to be analyzed simultaneously.
- Through MDS technique one can represent geometrically the locations and interrelationships among the set of points.
- It is not used widely because of complexity of procedure.

- Dimensions
  - Undimensional
  - Multidimensional scaling
- In undimensional scale, only one attribute of the object is measured.
- While in multidimensional scale, the object is described with several dimensions.
- **Example**
  - ❖ Popularity of a restaurant can be measured by a single measure Food taste. It can be also defined on multiple dimension like food taste, service, cleanliness, ambience etc.

- **Staple Scale**
  - This scale is developed by John Staple.
  - In this scale 10 categories are numbered from -5 to +5.
  - This scale does not have zero or the neutral point.
  - Respondents rate how each term describes the objects by selecting appropriate number.
  - In this scale positive number means the term describes the object accurately.
  - Which negative number implies that the term describes the objects inaccurately.
  - +5 means highest degree of accuracy while -5 means highest degree of inaccuracy.

- **Example**
- Let 3 phrases

I. Tasty Food

II. Fast Service

III. Good ambience for a restaurant

- A respondent is asked to rate how accurately these terms or phrases describe a specific restaurant.
- +5 +5 +5
- +4 +4 +4
- +3 +3 +3
- +2 +2 +2
- <u>**+1**</u> +1 +1

- Tasty Food Fast Service Good Ambience

Tasty Food	Fast Service	Good Ambience
-1	-1	-1
-2	<u>**-2**</u>	-2
-3	-3	-3
-4	-4	-4
-5	-5	-5

- Item Number
- 4 3 2 1 Respondent Score
- X X X X 4
- - X X X 3
- - - X X 2
- - - - X 1
- - - - - 0

- **Factor Scales**
  - It is like a score card given to each respondent.
  - It contains a series of favourable and unfavourable variables.
  - One side favourable and other side unfavourable.

Lenient							Strict
Silent Not helping Passive Proactive							Talkative Helpful Active Reactive
Fast							Slow
	3	2	1	-1	-2	-3	

- **Single/ Multiple Category Scales**
  - Here we have 2 or more mutually exclusive responses.
- **Example**
    - Yes or No
    - True of false

- The respondent has to choose any one out of the given categories
- Do you play cricket? Yes No

- What is you Marital Status? Unmarried Married Divorce Widower

- It should be ensured that the provided choice on, should cover almost all possible answers of the asked question.

- **Verbal Frequency Scale**
  - This scale is used when the respondent is unable or unwilling to give the exact numbers in the answer.
- **Example**

❖ How often do you eat ext side in a day?

1. Frequently 2. Sometimes 3.Rarely 4. Never

- **Comparative Scales**
  - In comparative scaling, the respondent is asked to compare one object with another.

- Common comparative scaling techniques are
  - Paired comparisons
  - Rank order
  - Constant sum scale

- **Paired Comparison Scale**
  - This is a comparative scaling technique in which a respondent is presented with 2 objects at a time and asked to select any one object according to some criteria.
  - The data obtained are ordinal in nature.

- **Example**
  - There are 4 types of cold drinks
  - Coke, Pepsi, sprite and Limca
  - The respondents can prefer Pepsi to coke or coke to sprite etc.

Brand	Coke	Pepsi	Sprite	Limca
Coke				
Pepsi				
Sprite				
Limca				
Number of times preferred				

- Number of judgements required in paired comparisons is given by formula N =
- $\frac{N(n-1)}{2}$
- Where, N = Number of judgements
- n = number of objects or items to be judged.
- $\frac{10\ (10-1)}{2} = \frac{10 \times 9}{2} = \frac{90}{2}$

- **Rank Order**

o Under this method of comparative scaling the respondents are asked to rank their choices.

- **Example**

o To judge 10 times it takes 45 pair comparison to complete the task.

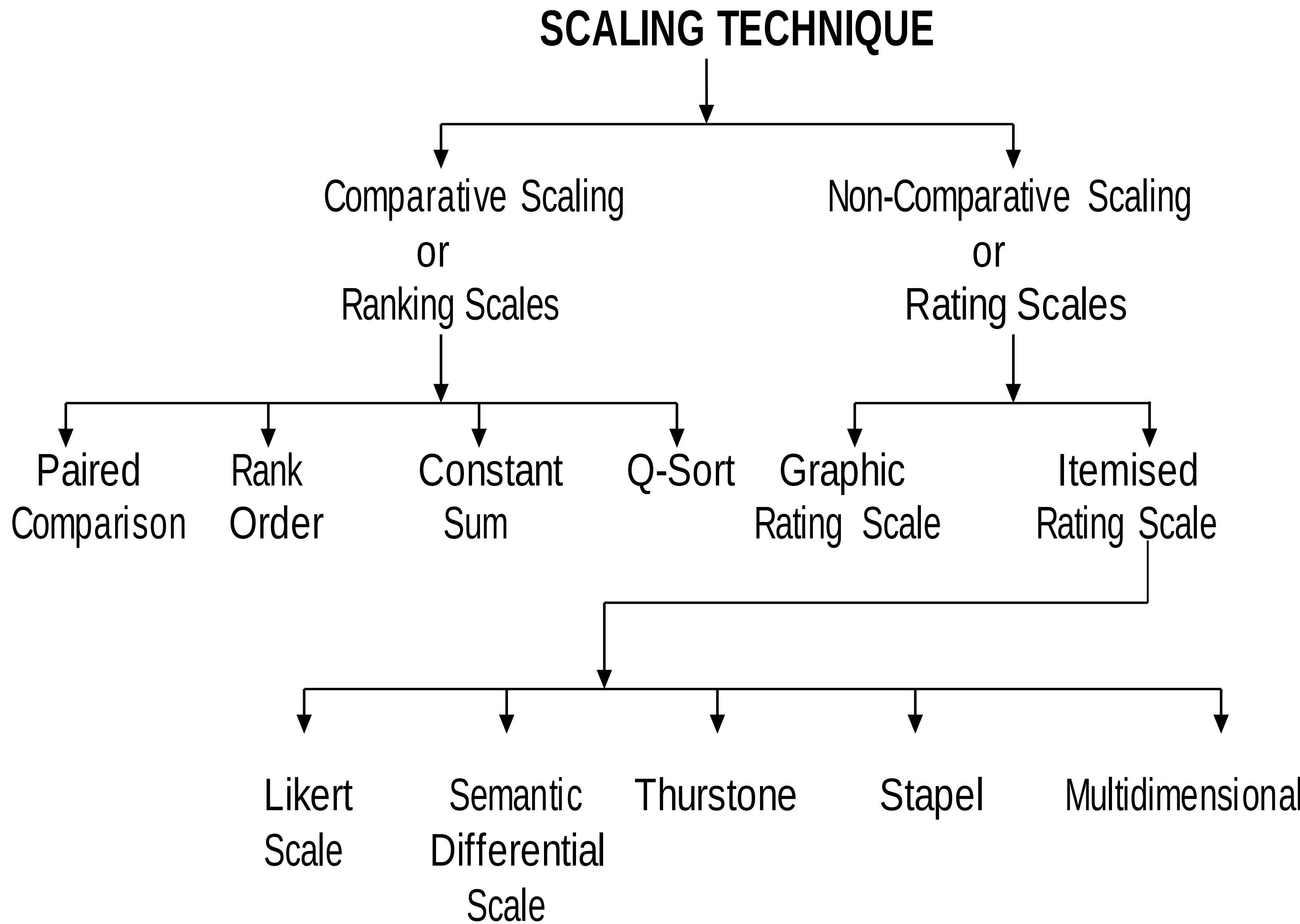
o Whereas the rank order method simple requires the ranking of 10 times only.

- **Constant Sum**
  - The respondent gives certain points to each object out of a fixed sum of points.
  - This fixed sum is usually taken as 100.
- **Example**
  - o a family planning for a holiday, fixes the budget of Rs. 50,000/-
  - o They wish to plan the expenditure on transportation, accommodation, food, drinks and others.

- The constant sum here is 50,000/- which could be divided as

- ✓ Transportation – 10,000.00
- ✓ Accommodation – 20,000.00
- ✓ Food – 12,000.00
- ✓ Drink – 0.00
- ✓ Others – 8,000.00

- **Q-Sort Scaling**
  - In this method the participants are asked to sort the objects or statements in several categories.



SCALING TECHNIQUE
Comparative Scaling or Ranking Scales
Non-Comparative Scaling or Rating Scales
Paired Comparison
Rank Order
Constant Sum
Q-Sort
Graphic Rating Scale
Itemised Rating Scale
Likert Scale
Semantic Differential Scale
Thurstone
Stapel
Multidimensional

- **Data Processing**
  - After data collection, the data are to be analysed.
  - Data processing involves
  - Questionnaire checking
  - Editing
  - Coding
  - Classification
  - Tabulation
  - Graphical representation
  - Data cleaning
  - Data adjusting
- **Questionnaire Checking**
  - This involves the examination of all the questionnaire for their completeness and interviewing quality.

- **Editing**
  - It is a process of examining the collection raw data to detect errors and correct these when possible.
    - Editing is of 2 types
- Field editing
- Central editing
- Field editing
  - It refers to review of the collected data on the spot.
  - This sort of editing should be done as soon as possible after the interview on the same day or the very next day.
- Central editing
  - It is done after collection of all the data.
  - This form of editing is done by a single person in case of small study and by a group of editors in case of a large study.
  - Editors must keep a view some points during editing.
  - They should be familiar with instructions given to the interviewers, coders as well as the editing instructions given to them.
  - While deleting an appropriate data sufficient reason must be there.
  - Put signature after editing.

- **Coding**
  - It refers to the process of assigning number and symbols to the edited data to make the data different from other data.
  - Coding is essential in case of large number of heterogeneous data.
- **Example**
  - If we will make a study on students of BIITM.
  - The registration number starting with OT coded as “A”
  - The registration number starting with 10 coded as “B”
  - The registration number starting with 11 coded as “C”
  - The registration number starting with 12 coded as “D”
- **Classification**
  - Classification means dividing homogenous data into meaningful group.
  - Data having a common characteristic are placed in class one class and in the way the entire data is classified into number of groups.

- Classification according to attributes
  - It means classification of data as per common characteristics.
  - Common characteristics may be descriptive (such as literacy, sex, honesty etc.) or numerical such as weight, height, income, etc.) i.e., qualification.
- Classification according to class interval
  - When classification is done by quantitative characteristics it is called class interval.

- **Example**
  - People having income from 4000 – 8000 per month falls under one class.
  - People having income 8000 – 12000 per month belong to another class.

- Each group in class interval are having an upper limit and a lower limit which is called class limit.

- Entire data can be divided into a number of groups or classes called class interval.

- Difference between 2 class limit is called class magnitude.

- We may have classes with equal class magnitude or unequal classes magnitude.
- The number of items which fall in a given class is called frequency of class.

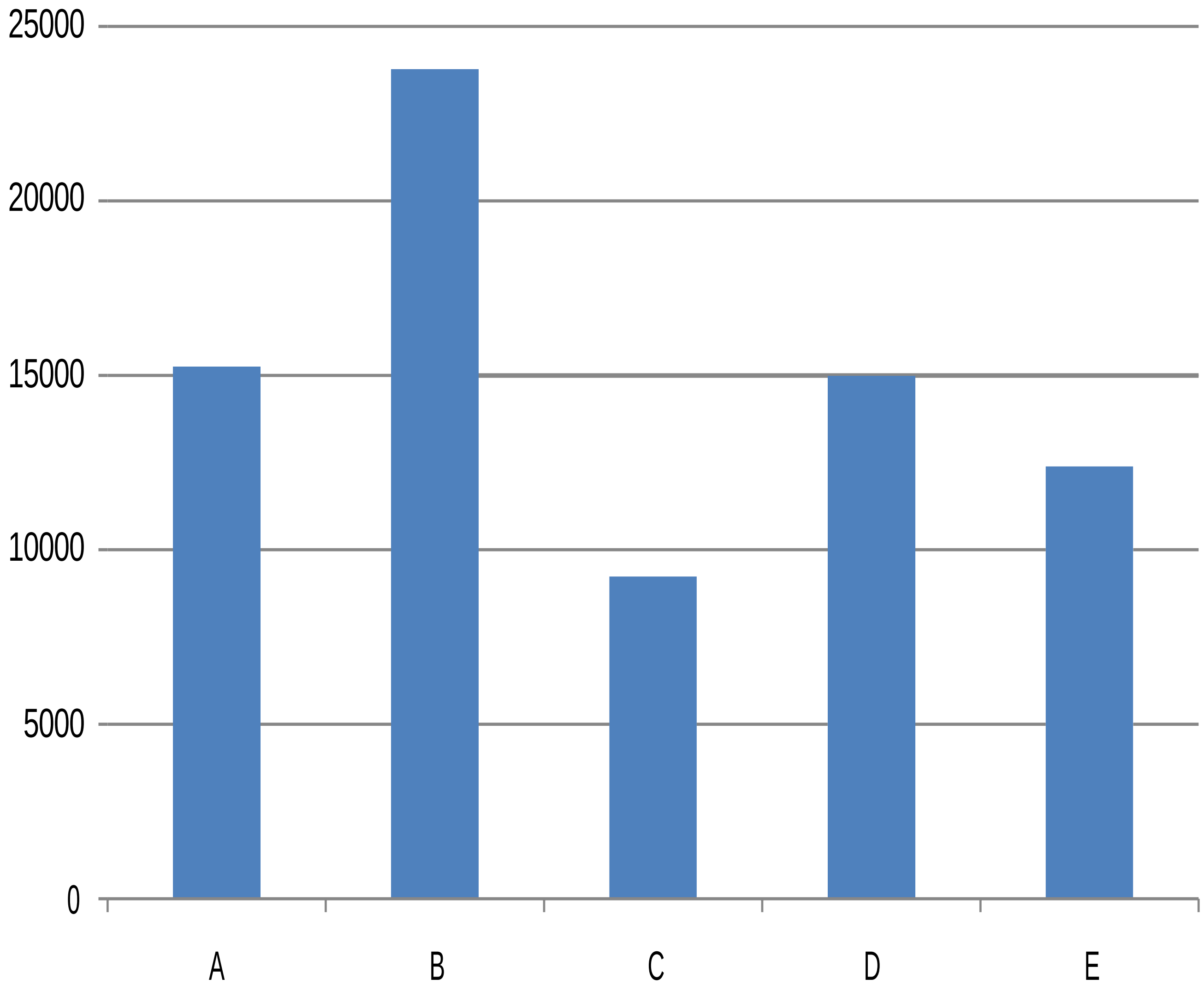
Income Group	Number of families (Frequency)
Below 4000	13
4000 – 8000	15
8001 – 12000	07
12001 – 16000	14
16000 above	08
Total	57

- **Tabulation**

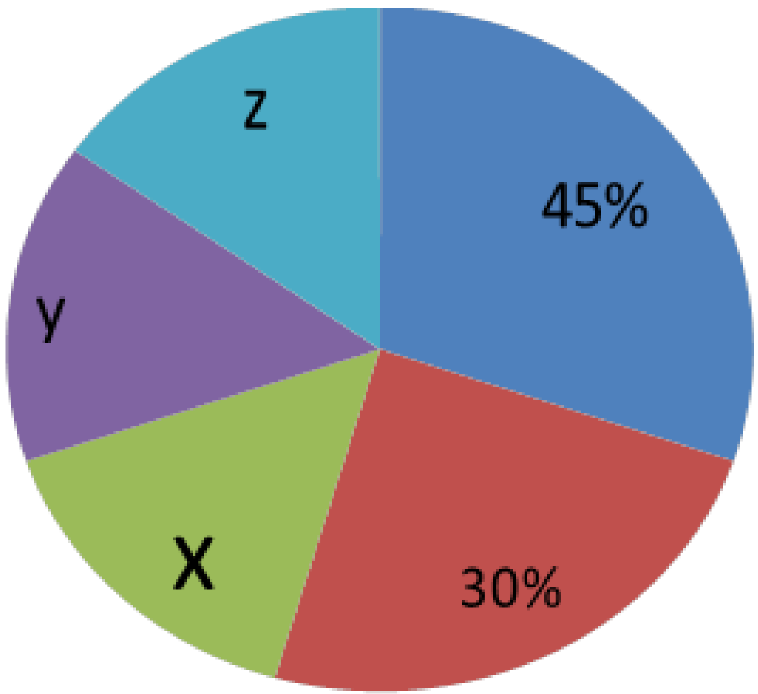
- When a mass of data has been assembly it become necessary for the researcher to arrange in some kind of concise and logical order which is called tabulation.
- Tabulation is the process of summarising raw data, displaying in compact form (i.e., for of statistical table)
- Tabulation is an orderly arrangement of data in column and rows.

- Tabulation is essential because :
  - It saves space, reduce description and explanation.
  - It facilitates comparison.
  - Detection of error.
  - It provides a basis for various statistical computations.
  - Tabulation can be done by hand or computer.
  - Hand tabulation is preferred in small inquiries.
- Principles of Tabulation :
  - Every table should have a clear and concise title.
  - Each table should have a table number.
  - Column heading row heading should be brief.
  - Unit of measurement should be given.
  - Emplanarotry foot notes should be placed.
  - Source must be there in case of secondary data.
  - Headings should be written in bold letter.
  - Thick lines should be used to differentiate data.
  - Columns may be numbered.
  - Abbreviation should be avoided.
  - Table should be logical, clear and accurate.
  - Arrangement of items may be chronological, alphabetical, geographical etc.

- **Graphical representation:**
  - Graphs helps to understand the data easily.
  - Most common graphs are bar chart and pie.
- **Bar chart**
  - It consists of a series of rectangles or bars.
  - The height of each rectangle is determined by frequency of that category.
- **Example**
- o Sales of Pepsi in year 2015-16 is 5 areas of Bhubaneswar i.e., A, B, C, D and E are 15245, 23762, 9231, 14980, 12387. So bar chart will be


25000
20000
15000
10000
5000
0
A
B
C
D
E

- **Pie chart**
  - It is a circular chart divided into sector which shows relative frequency.



- **Data Cleaning**
  - It means checking the data for consistency and treatment for missing value.
  - Missing value refers to the value which are unknown or not answered by the respondent.
  - In place of such missing values some neutral value or mean value may be used.
- **Data Adjusting**
  - To improve the quality of data analysis it is needed some times.
- **Weight Assigning**
  - Each respondent is assigned weight to know his importance in comparison to other respondents.
- **Example**
  - In response of educated people will be given higher weight and uneducated people will be given lower weight.

- **Variable Re-Specification**
  - It means creating new variables or modifying existing variables.
- **Example**
  - ❖ If usefulness of a certain product is measured on 10 point scale then it may be reduced to 4 point or 5 point.
- **Scale Transformation**
  - It is done to ensure the comparability with other scales.
  - Or to make the data suitable for analysis.
  - The variables which are measured on different scales can not be compared.
- **Laddering**
  - Developed by T.J. Reynolds and J. Gutman.
  - It is an interview technique where simple response to a question is used by the interviewer to find subconscious motives.
  - It begins with a simple question and then another question is asked about that response.

- **Example**
- The interviewer may ask the question

- How come you skipped the class? Response : I went out with my friends.

- The next question will be

- Why did you go out with you friends?
- Response : I went for watching movie, or for Pizza. The format is
- Q : Why x ?
- Ans : Because Z.

- Q : Why Z?
- Ans : Because b.

- Q : Why b?

- **Use in Marketing**
  - Laddering technique is used in marketing to judge what values inspire the consumer to consume a particular product.
  - The laddering technique allows business to know their customers better by asking them simple direct questions.
- Laddering framework can be applied to market research by asking following questions.
- Why did you choose this product/services?
- Why is it good/bad?
- Why is this important to you/your business?
- **Conclusion**
- It is helpful in business decision making.

Thank you.

BUSINESS RESEARCH

By Asst. Prof. (Eco) Mrs. Tapaswini Nayak.

Module II
Hypothesis Testing:
(Fundamental Concepts)

- **Population / Universe :**
- Population in research is a complete set of elements that possess a standard parameter between them.
- The ‘population’ in research doesn’t necessarily have to be human. It can be any parameter of data that possesses a common trait.
- **Sample :**
- A sample is a smaller part of the whole, i.e., a subset of the entire population. It is representative of the population in a study. When conducting surveys, the sample is the members of the population who are invited to participate in the survey.
- Hence said, a sample is a subgroup or subset within the population.

- **Example :** Research in XYZ university. 10 colleges under university.

- 1 college = 100 students. Total = 10,000 students.

- I have to study out of 10,000 students how many males and how many female.

- Out of 10 college we can take 3 colleges and 300 students from each college we study total students we take for study = 900

- So, instead of 10,000 students we are studying 900 . So, 10,000 = universe/ population
- 900 = sample

- **Types of Population:** (2 types)
- i) Finite(Countable) ii)Infinite(Uncountable)
- **Small & large sample:**
- When sample size(n) is less than 30,then it is called small sample and when sample size(n)is more than 30,then it is called large sample.
- **Parameter and statistic :**
- Statistical value calculated related to population is called parameter. Statistical value calculated related to sample is called statistic.
- Example: in the above example our population = 10,000 suppose they are giving exam.
- When we are calculating average mark of 10,000 students then we are calculating statistical value (mean) of population so, it will be called parameter.

- On the other hand when we are taking sample = 900
- And we are calculating average mark of 900 students, then we are calculating statistical value of sample, which is called statistic.
- Parameter is not generally calculated, calculation of statistic is done generally in research.

- **Parametric test and non parametric test:** when we have some assumption (idea) about the population then parametric test is used.

- When we don't have any idea (we can't assume anything about the population) then non parametric test is used.

- **Hypothesis** : it is an assumption yet to be tested.
- We assume something before doing research, we have to test whether it is true of not.

- Example: before doing research when we are assuming that out of 10,000 students 50% may be girls, 50% may be boys. Which is called Hypothesis.
- By doing research we will try to the hypothesis whether its true or not.
- **Null Hypothesis**
- The null hypothesis (H_0) is a statement of no effect, relationship, or difference between two or more groups or factors. In research studies, a researcher is usually interested in disproving the null hypothesis.
- **Alternative Hypothesis**
- The alternative hypothesis (H_1) is **the statement that there is an effect or difference.** This is usually the hypothesis the researcher is interested in proving. The alternative hypothesis can be **one-sided**(only provides one direction, e.g., lower) **or two-sided.**

- **Level of confidence and level of significance.**
- Level of confidence refers to how much confident (%) you are about your result (correct result)
- Error is possible every where. Level of confidence can be 99%. (1% error is possible)
- 1% = error level of significance.

- **Precision limit**
- Limit between which the research results are expected to lie.
- Example: suppose we have expected that our research result will be lie between 40% to 60%.
- So, the upper limit and lower limit is called precision limit.

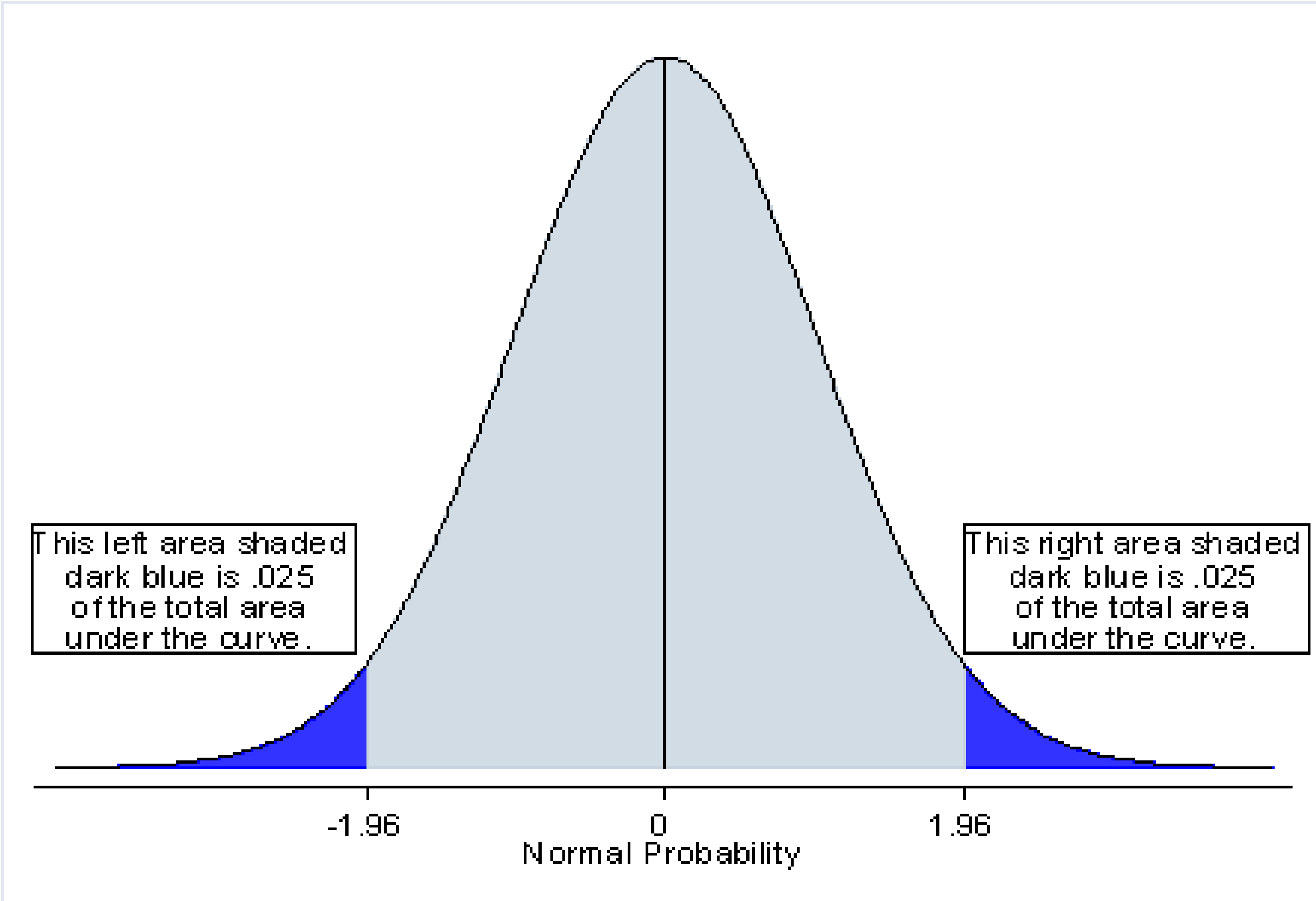
- **Finite Population Correction factor**
- The Finite Population Correction Factor, sometimes just called the FPC factor, is used when the sample size is large relative to the population size.
- We need to use the FPC when the ratio of the sample size n to the population size N is greater than 5%. For example, if the population size is 300 and the sample size is 30, we have a ratio of 10% and thus need to use the FPCF.
- The most common formula for calculating the FPC is

$$\sqrt{\frac{N-n}{N-1}}$$

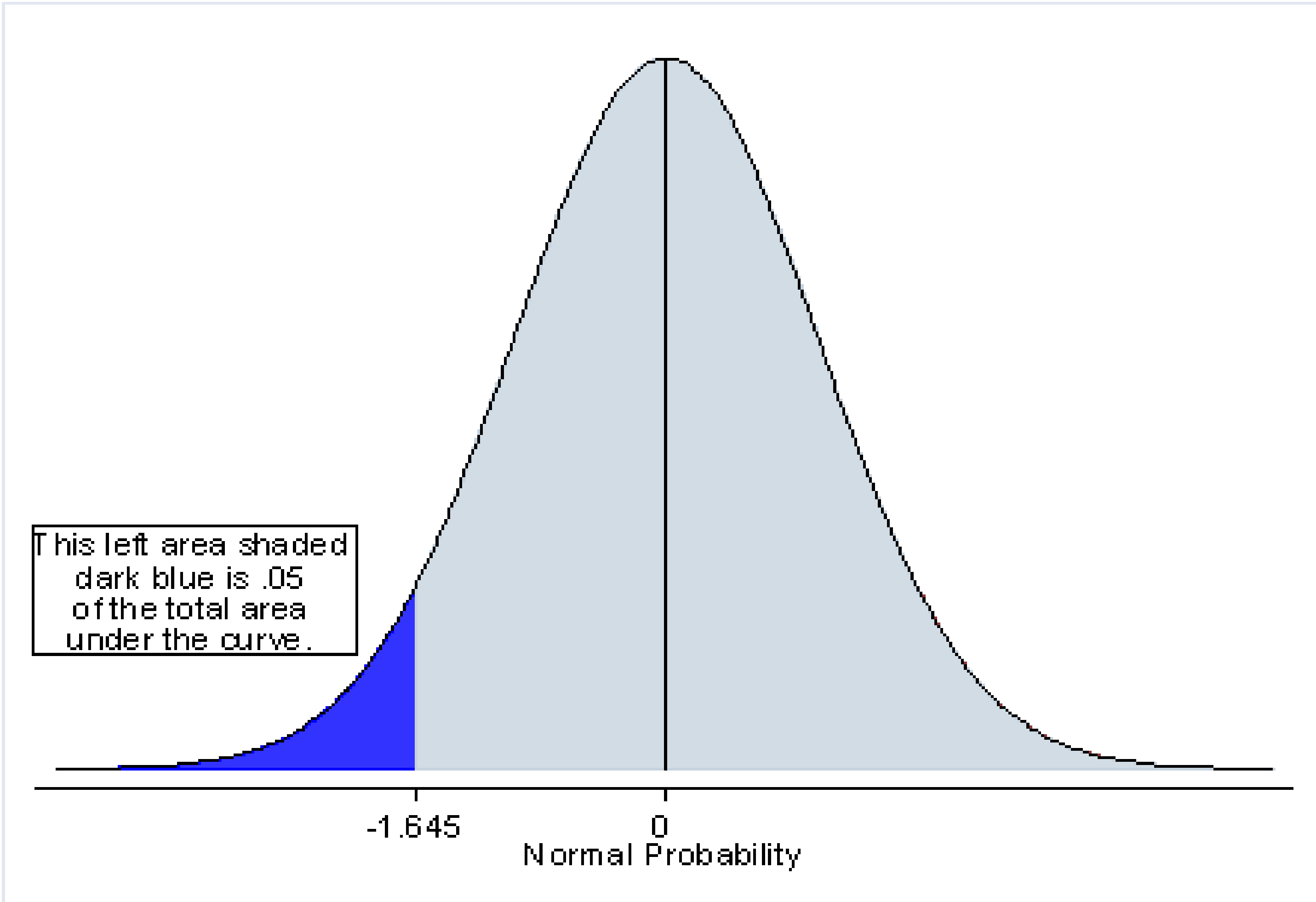
- **Test of Hypothesis or Decision Rule:**
- Suppose the given hypothesis is H_0 and the alternative hypothesis H_1, then the researcher has to make a rule known as the decision rule. According to the decision rule, the researcher accepts or rejects H_0.
- For example, if the H_0 is that certain students are good against the H_1 that all the students are good, then the researcher should decide the number of items to be tested and the criteria on the basis of which to accept or reject the hypothesis.

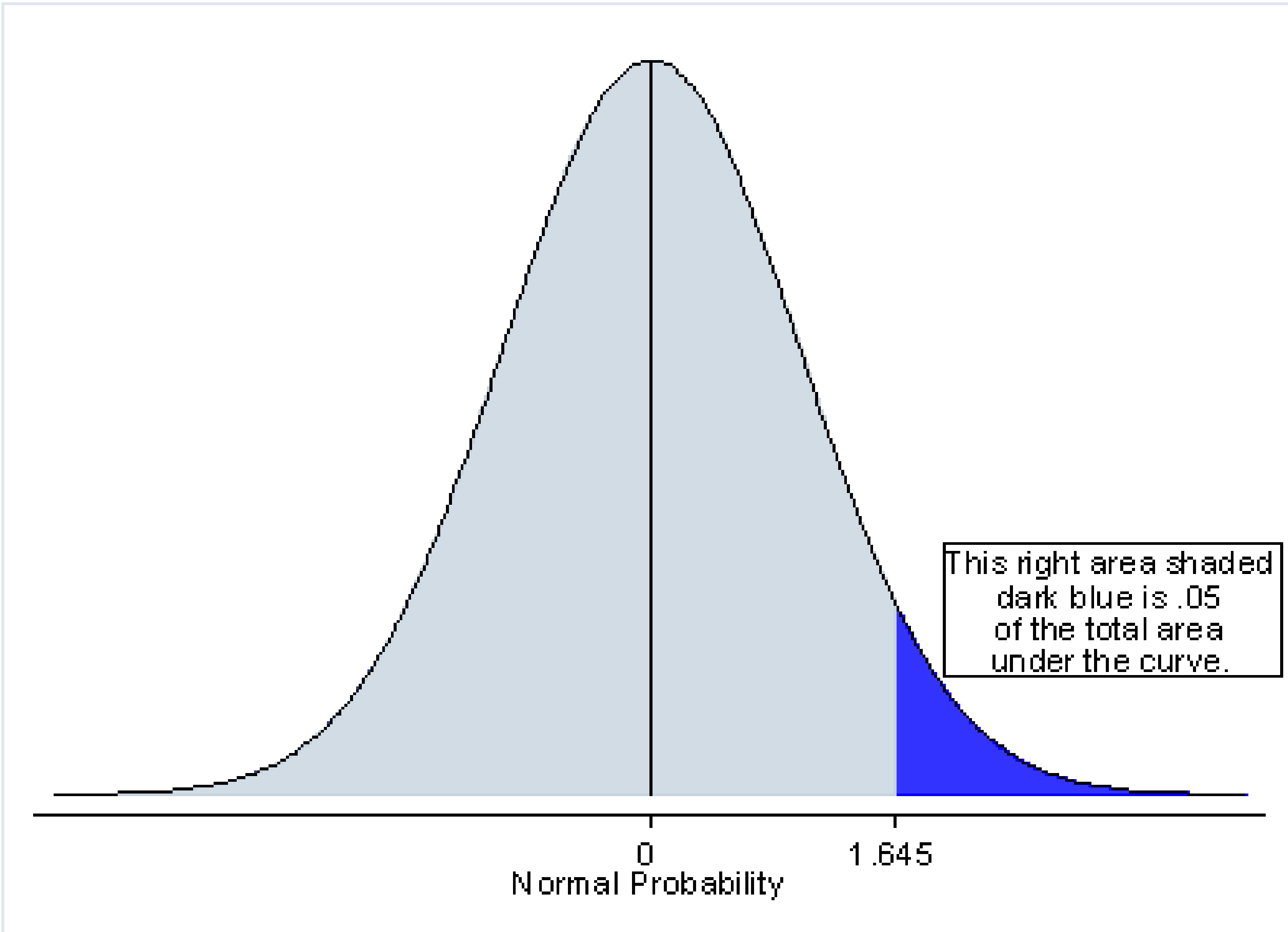
- **Type I and Type II Errors:**
- As regards the testing of hypotheses, a researcher can make basically two types of errors. He/she may reject H0 when it is true, or accept H0 when it is not true. The former is called as **Type I error** and the latter is known as **Type II error**.
- In other words, Type I error implies the rejection of a hypothesis when it must have been accepted, while Type II error implies the acceptance of a hypothesis which must have been rejected.
- Type I error is denoted by α (alpha) and is known as α error, while Type II error is usually denoted by β (beta) and is known as β error.

- **One-Tailed and Two-Tailed Tests:**
- **•What is a two-tailed test?**
- First let's start with the meaning of a two-tailed test. If you are using a significance level of 0.05, a two-tailed test allots half of your alpha to testing the statistical significance in one direction and half of your alpha to testing statistical significance in the other direction. This means that .025 is in each tail of the distribution of your test statistic. When using a two-tailed test, regardless of the direction of the relationship you hypothesize, you are testing for the possibility of the relationship in both directions.
- For example, we may wish to compare the mean of a sample to a given value x using a t-test. Our null hypothesis is that the mean is equal to x. A two-tailed test will test both if the mean is significantly greater than x and if the mean significantly less than x. The mean is considered significantly different from x if the test statistic is in the top 2.5% or bottom 2.5% of its probability distribution, resulting in a p-value less than 0.05.


This left area shaded
dark blue is .025
of the total area
under the curve.
This right area shaded
dark blue is .025
of the total area
under the curve.
-1.96
0
1.96
Normal Probability

- **What is a one-tailed test?**
- Next, let's discuss the meaning of a one-tailed test. If you are using a significance level of .05, a one-tailed test allots all of your alpha to testing the statistical significance in the one direction of interest. This means that .05 is in one tail of the distribution of your test statistic. When using a one-tailed test, you are testing for the possibility of the relationship in one direction and completely disregarding the possibility of a relationship in the other direction.
- Let's return to our example comparing the mean of a sample to a given value x using a t-test. Our null hypothesis is that the mean is equal to x. A one-tailed test will test either if the mean is significantly greater than x or if the mean is significantly less than x, but not both. Then, depending on the chosen tail, the mean is significantly greater than or less than x if the test statistic is in the top 5% of its probability distribution or bottom 5% of its probability distribution, resulting in a p-value less than 0.05. The one-tailed test provides more power to detect an effect in one direction by not testing the effect in the other direction. A discussion of when this is an appropriate option follows.


This left area shaded dark blue is .05 of the total area under the curve.
-1.645
0
Normal Probability


This right area shaded
dark blue is .05
of the total area
under the curve.
0
1.645
Normal Probability

- **Critical or Rejection Region** – the range of values for the test value that indicate a significant difference and that the null hypothesis should be rejected.
- **Non-critical or Non-rejection Region** – the range of values for the test value that indicates that the difference was probably due to chance and that the null hypothesis should not be rejected.
- **Critical Value (CV)** – separates the critical region from the non-critical region, i.e., when we should reject H0 from when we should not reject H0.
- The location of the critical value depends on the inequality sign of the alternative hypothesis.
- Depending on the distribution of the test value, you will use different tables to find the critical value.

- **Procedure of Hypothesis Testing:**
- Testing a hypothesis refers to verifying whether the hypothesis is valid or not. Hypothesis testing attempts to check whether to accept or not to reject the null hypothesis. The procedure of hypothesis testing includes all the steps that a researcher undertakes for making a choice between the two alternative actions of rejecting or accepting a null hypothesis. The various steps involved in hypothesis testing are as follows:
- **1) Making a Formal Statement:** This step involves making a formal statement of the null hypothesis (H0) and the alternative hypothesis (Ha). This implies that the hypotheses should be clearly stated within the purview of the research problem. For example, suppose a school teacher wants to test the understanding capacity of the students which must be rated more than 90 per cent in terms of marks, the hypotheses may be stated as follows:
- Null Hypothesis H0 : = 100 Alternative Hypothesis H1 : > 100

- ***2) Selecting a Significance Level*:**
- •The hypotheses should be tested on a pre-determined level of significance, which should be specified. Usually, either 5% level or 1% level is considered for the purpose. The factors that determine the levels of significance are: (a) the magnitude of difference between the sample means; (b) the sample size: (c) the variability of measurements within samples; and (d) whether the hypothesis is directional or non-directional. In sum, the level of significance should be sufficient in the context of the nature and purpose of enquiry.
- **•3) Deciding the Distribution to Use:**
- •After making decision on the level of significance for hypothesis testing, the researcher has to next determine the appropriate sampling distribution. The choice to be made generally relates to normal distribution and the t-distribution. The rules governing the selection of the correct distribution are similar to the ones already discussed with respect to estimation.

- **4) Selection of a Random Sample and Computing An Appropriate Value:**
- •Another step involved in hypothesis testing is the selection of a random sample and then computing a suitable value from the sample data relating to test statistic by using the appropriate distribution. In other words, it involves drawing a sample for furnishing empirical data.
- **•5) Calculation of the Probability:**
- •The next step for the researcher is to calculate the probability that the sample result would diverge as far as it can from expectations, under the situation when the null hypothesis is actually true.

- **6) Comparing the Probability:**
- •Another step involved consists of making a comparison of the probability calculated with the specified value of α, i.e. The significance level. If the calculated probability works out to be equal to or smaller than the α value in case of one-tailed test, then the null hypothesis is to be rejected. On the other hand, if the calculated probability is greater, then the null hypothesis is to be accepted.
- •In case the null hypothesis H0 is rejected, the researcher runs the risk of committing the Type I error. But, if the null hypothesis H0 is accepted, then it involves some risk (which cannot be specified in size as long as H0 is vague and not specific) of committing the Type II error.

- **Discuss the Z -test.**
- Ans. The Z-test compares sample and population means to determine if there is a significant difference. It requires a simple random sample from a population with a Normal distribution and where the mean is known. The z measure is calculated as:
- **Suppose** that in a particular geographic region, the mean and standard deviation of scores on a reading test are 100 points, and 12 points, respectively. Our interest is in the scores of 55 students in a particular school who received a mean score of 96. We can ask whether this mean score is significantly lower than the regional mean — that is, are the students in this school comparable to a simple random sample of 55 students from the region as a whole, or are their scores surprisingly low?

- We begin by calculating the standard error of the mean:

$$\mathrm{SE} = \frac{\sigma}{\sqrt{n}} = \frac{12}{\sqrt{55}} = \frac{12}{7.42} = 1.62$$

- Next we calculate the z-score, which is the distance from the sample mean to the population mean in units of the standard error:

$$z = \frac{M - \mu}{\mathrm{SE}} = \frac{96 - 100}{1.62} = -2.47$$

- In this example, we treat the population mean and variance as known, which would be appropriate either if all students in the region were tested, or if a large random sample were used to estimate the population mean and variance with minimal estimation error.
- The classroom mean score is 96, which is −2.47 standard error units from the population mean of 100. Looking up the *z*-score in a table of the standard normal distribution, we find that the probability of observing a standard normal value below -2.47 is approximately 0.5 - 0.4932 = 0.0068. This is the one-sided p-value for the null hypothesis that the 55 students are comparable to a simple random sample from the population of all test-takers. The two-sided p-value is approximately 0.014 (twice the one-sided p-value).

- Another way of stating things is that with probability $1 - 0.014 = 0.986$, a simple random sample of 55 students would have a mean test score within 4 units of the population mean. We could also say that with 98% confidence we reject the null hypothesis that the 55 test takers are comparable to a simple random sample from the population of test-takers.

- **TEST OF SIGNIFICANCE FOR PROPORTIONS USING NORMAL DISTRIBUTION**
- **Test for a Single Proportion**
- Suppose a large sample of size n its taken from a normal population to test the significance difference between a sample population Q and the population proportion P we use statistic.

$$z = \frac{\hat{p} - p}{\sqrt{\frac{pq}{n}}}$$

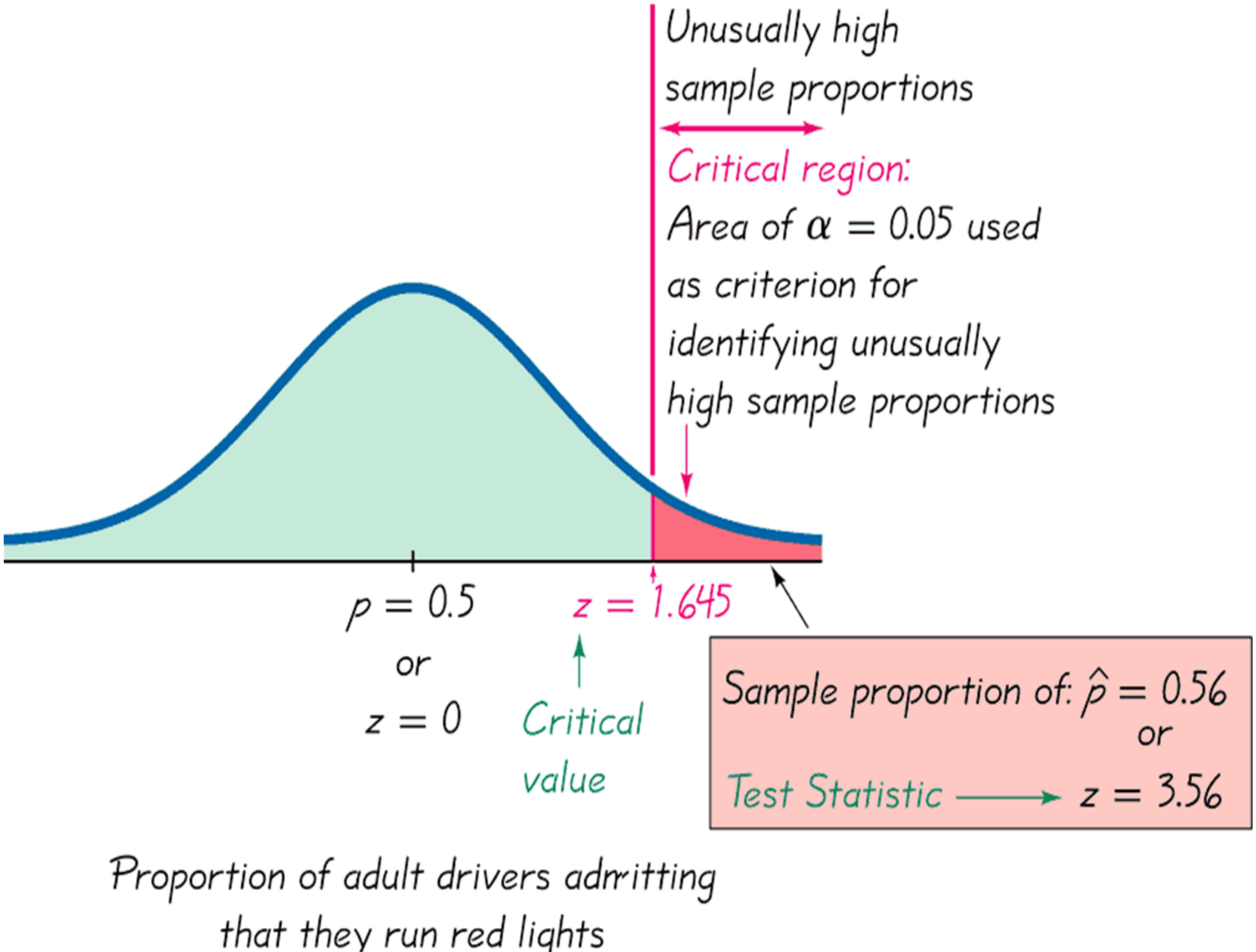
- Example: A survey of n = 880 randomly selected adult drivers showed that 56% (or p = 0.56) of those respondents admitted to running red lights. Find the value of the test statistic for the claim that the majority of all adult drivers admit to running red lights. (In Section 8-3 we will see that there are assumptions that must be verified. For this example, assume that the required assumptions are satisfied and focus on finding the indicated test statistic.)

- Solution: The preceding example showed that the given claim results in the following null and alternative hypotheses: H_0: p = 0.5 and H_1: p > 0.5. Because we work under the assumption that the null hypothesis is true with p = 0.5, we get the following test statistic:

$$z = \frac{\hat{p} - p}{\sqrt{\frac{pq}{n}}} = \frac{0.56 - 0.5}{\sqrt{\frac{(0.5)(0.5)}{880}}} = 3.56$$

- Interpretation: We know from previous chapters that a *z* score of 3.56 is exceptionally large. It appears that in addition to being “more than half,” the sample result of 56% is significantly more than 50%.

  See figure following.


Unusually high
sample proportions
Critical region:
Area of $\alpha = 0.05$ used
as criterion for
identifying unusually
high sample proportions
$p = 0.5$
or
$z = 0$
$z = 1.645$
Critical
value
Sample proportion of: $\hat{p} = 0.56$
or
Test Statistic → $z = 3.56$
Proportion of adult drivers admitting
that they run red lights

- The critical region (or rejection region) is the set of all values of the test statistic that cause us to reject the null hypothesis. For example, see the red-shaded region in the previous figure.

- **Z TEST FOR MEAN**
- **Z Test Statistics Formula – Example #1**
- **Principal at school claims that students in his school are above average intelligence and a random sample of 30 students IQ scores have a mean score of 112.5 and mean population IQ is 100 with a standard deviation of 15. Is there sufficient evidence to support the principal claim? Test at 5 per cent level of significance.**

	A	B
1	Sample Mean ($\bar{x}$)	112.5
2	Population Mean (μ)	100
3	Standard Deviation of Population (σ)	15
4	Number of Observation (n)	30
5		

- **Solution:**
- Z Test Statistics is calculated using the formula given below
- **Z Test = ($\bar{x}$ – μ)** / **(σ / √n)**
- Z Test = (112.5 – 100) / (15 / √30)
- Z Test **= 4.56**
- Compare the z test results with z test standard table and you can come to the conclusion in this example null hypothesis is rejected and the principal claim is right.

- **Z Test Statistics Formula – Example #2**
- Suppose an investor looking to analyze the average daily return of the stock of one the company is greater than 1% or not? So investors picked up a random sample of 50 and return is calculated and has a mean of 0.02 and investors considered standard deviation of mean is 0.025.
- So, in this case, the null hypothesis is when the mean is 3% and the alternative hypothesis is that of mean return is higher than 3%. Investors assume alpha of 0.05% is selected as a two-tailed test and 0.025% of the sample in each tail and alpha critical value is either 1.96 or -1.96. So if the result of the Z test is less or greater than 1.96 null hypothesis will be rejected.

	A	B
1	Sample Mean ($\bar{x}$)	0.02
2	Population Mean (μ)	1%
3	Standard Deviation of population (σ)	0.025
4	Number of Observation (n)	50
5		

- **Solution:**
- Z Test Statistics is calculated using the formula given below
- **Z Test = $(\bar{x} – \mu)$** / **$(\sigma / \sqrt{n})$**
- Z Test = (0.02 – 1%) / (0.025/ √50)
- Z Test **= 2.83**
- So from the above calculation investors will come to conclusion and he will reject the null hypothesis because the result of z is greater than 1.96 and come to an analysis that the average daily return of the stock is more than 1%.

- **Z Test Statistics Formula – Example #3**
- An insurance company is currently reviewing its current policy rates when originally settings the rate they believe that the average claim amount will be a maximum of Rs 180000. The company is concern about that true mean actually higher than this. The company randomly select 40 sample claim and calculate sample mean of Rs 195000 assuming a standard deviation of Claim is Rs 50000 and set alpha as 0.05. So z test to be performed to see insurance company should be concerned or not.

	A	B
1	Sample Mean ($\bar{x}$)	195000
2	Population Mean (μ)	180000
3	Standard Deviation of population (σ)	50000
4	Number of Observation (n)	40
5		

- **Solution:**
- Z Test Statistics is calculated using the formula given below
- **Z Test = $(\bar{x} - \mu) / (\sigma / \sqrt{n})$**
- Z Test = $(195000 - 180000) / (50000 / \sqrt{40})$
- Z Test **= 1.897**

- **Step – 1** Set the Null hypothesis
- **Step – 2** calculate the test statistics
- So if you put all available figures in z test formula it will give us z test results as**<u>1.897</u>**
- **Step – 3** Set Rejection region
- Considering alpha as 0.05 let's say rejection region is 1.65
- **Step – 4** Conclude
- As per z test results, we can see that 1.897 is greater than the rejection region of 1.65 so the company fails to accept the null hypothesis and the insurance company should be concerned about their current policies.

- **‘T’ TEST**
- A t-test is any statistical hypothesis test in which the test statistic follows a Student's t distribution under the null hypothesis. It can be used to determine if two sets of data are significantly different from each other.
- A t-test is most commonly applied when the test statistic would follow a normal distribution if the value of a scaling term in the test statistic were known. When the scaling term is unknown and is replaced by an estimate based on the data, the test statistics (under certain conditions) follow a Student's t distribution.
- **Types of t-tests**
- A t-test is a hypothesis test of the mean of one or two normally distributed
- populations. Several types of t-tests exist for different situations, but they all use a test
- statistic that follows a t-distribution under the null hypothesis:

Test	Purpose	Example
1sample t-test	Tests whether the mean of a single population is equal to a target value	Is the mean height of female college students greater than 5.5 feet?
2 sample t-test	Tests whether the difference between the means of two independent populations is equal to a target value	Does the mean height of female college students significantly differ from the mean height of male college students?
paired t-test	Tests whether the mean of the differences between dependent or paired observations is equal to a target value	If you measure the weight of male college students before and after each subject takes a weight-loss pill, is the mean weight loss significant enough to conclude that the pill works?

- **Discuss the uses of t-test.**
- **Independent one sample *t*-test**
- In testing the null hypothesis that the population means is equal to a specified value $\mu 0$, one uses the statistic

$$t = \frac{\bar{x} - \mu}{\frac{s}{\sqrt{n}}}$$

- where s is the sample standard deviation of the sample and n is the sample size. The degrees of freedom used in this test is $n - 1$.
- ***Unequal sample sizes, unequal variance***
- This test is used only when the two population variances are assumed to be different (the two sample sizes may or may not be equal) and hence must be

$$t = \frac{\overline{x_1} - \overline{x_2}}{\sqrt{\frac{S_1{}^2}{N_1} + \frac{S_2{}^2}{N_2}}}$$

- where
  $\bar{x}_1$ is the mean of first data set
  $\bar{x}_2$ is the mean of first data set
  S_1^2 is the standard deviation of first data set
  S_2^2 is the standard deviation of first data set
  N_1 is the number of elements in the first data set
  N_2 is the number of elements in the first data set

- **One Sample T Test Example**
- **Sample question**: your company wants to improve sales. Past sales data indicate that the average sale was Rs100 per transaction. After training your sales force, recent sales data (taken from a sample of 25 salesmen) indicates an average sale of Rs130, with a standard deviation of Rs15. Did the training work? Test your hypothesis at a 5% level of significance.
- **Step 1**: Write your null hypothesis statement . The accepted hypothesis is that there is no difference in sales, so:
  H_0: $\mu = 100$.
- **Step 2**: Write your alternative hypothesis statement. This is the one you're testing. You think that there *is* a difference (that the mean sales increased), so:
  H_1: $\mu > 100$.
- **Step 3**: Identify the following pieces of information you'll need to calculate the test statistic. The question should give you these items:

- **The sample mean**($\bar{x}$). This is given in the question as 130.
- **The population mean**(μ). Given as 100 (from past data).
- **The sample standard deviation**(s) = 15.
- **Number of observations**(n) = 25.
- **Step 4**: Insert the items from above into the t score formula

$$t = \frac{\bar{x} - \mu}{\frac{s}{\sqrt{n}}}$$

- $t = (130 – 100) / ((15 / \sqrt{(25)})$
  $t = (30 / 3) = 10$
  This is your **calculated t-value**.
- **Step 5:** Find the t table value. You need two values to find this:
- The alpha level: given as 5% in the question.
- The degree of freedom, which is the number of items in the sample (n) minus 1: 25 – 1 = 24.
- Look up 24 degrees of freedom in the left column and 0.05 in the top row. The intersection is 1.711. This is your one-tailed critical t-value.
- What this critical value means is that we would expect most values to fall under 1.711. If our calculated t-value (from Step 4) falls within this range, the null hypothesis is likely true.
- Step 5: Compare Step 4 to Step 5. The value from Step 4 **does not** fall into the range calculated in Step 5, so we can reject the null hypothesis. The value of 10 falls into the rejection region (the left tail).
- In other words, it's highly likely that the mean sale is greater. The sales training was probably a success.

- Independent Samples t-Test (or 2-Sample t-Test)
- **Example:**
  Calculate the T test value whose inputs are 10, 20,30,40,50 and 1, 29,46,78,99
  First calculate standard deviation & mean of the given set,

  For the data set **10, 20,30,40,50**
  Total Inputs(N) =(10,20,30,40,50)
  Total Inputs(N)=5
  Mean(x_m)= (x_1+x_2+x_3...x_n)/N
  Mean(x_m)= 150/5
  Means(x_m)= 30
  SD=sqrt(1/(N-1)*((x_1-x_m)^2+(x_2-x_m)^2+..+(x_n-x_m)^2))
  =sqrt(1/(5-1)((10-30)^2+(20-30) ^{2}+(30-30) ^{2}+(40-30) ^{2}+(50-30) ^{2}))

- $=sqrt(1/4((-49.6)^2+(-21.6)^2+(-4.6)^2+(27.4)^2+(48.4)^2))$
  $=sqrt(1/4((2460.16)+(466.56)+(21.16)+(750.76)+(2342.56)))$
  $=sqrt(1510.3)$
  $=38.8626$
  $Variance=SD^2$
  $Variance=38.8626^2$
  $Variance=1510.3$

- **To find T Test:**

  From above we know that,

  $x_1 = 30$

  $x_2 = 50.6$

  $S_1^2 = 250$

  $S_2^2 = 1510.3$

  $N_1 = 5$

  $N_2 = 5$
- Substitute these values in the above formula,

$$t = \frac{\overline{x_1} - \overline{x_2}}{\sqrt{\frac{S_1^{\,2}}{N_1} + \frac{S_2^{\,2}}{N_2}}}$$

- T = (30 - 50.6)/√((250/5) + (1510.3/5)) = -1.0979

  When you try such calculations on your own, this t-test calculator can be used to verify your results of calculations.

- **Chi-Square Distribution**
- Test Statistic

$$\chi^2 = \frac{(n-1)\,s^2}{\sigma^2}$$

n = sample size

s^2 = sample variance

σ^2 = population variance
(given in null hypothesis)

- ***P*-Values and Critical Values for Chi-Square Distribution**

❖ **Use Table .**

❖ **The degrees of freedom = *n* –1**

- Properties of Chi-Square Distribution

❖ **All values of χ^2 are nonnegative, and the distribution is not symmetric (see Figure 8-13, following).**

❖ **There is a different distribution for each number of degrees of freedom (see Figure 8-14, following).**

❖ **The critical values are found in Table using n – 1 degrees of freedom.**

- **Properties of Chi-Square Distribution - cont.**

Properties of the Chi-Square Distribution

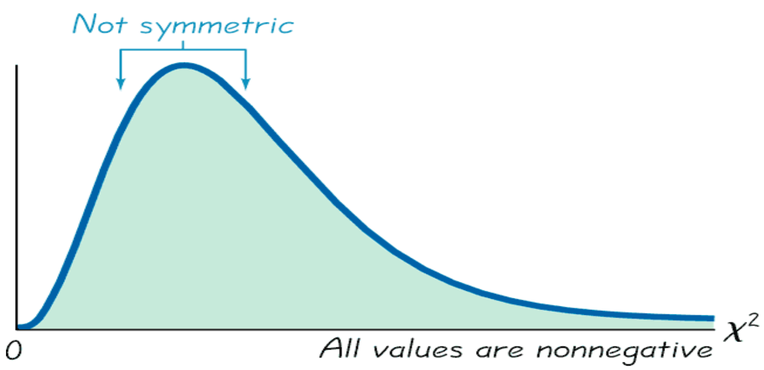
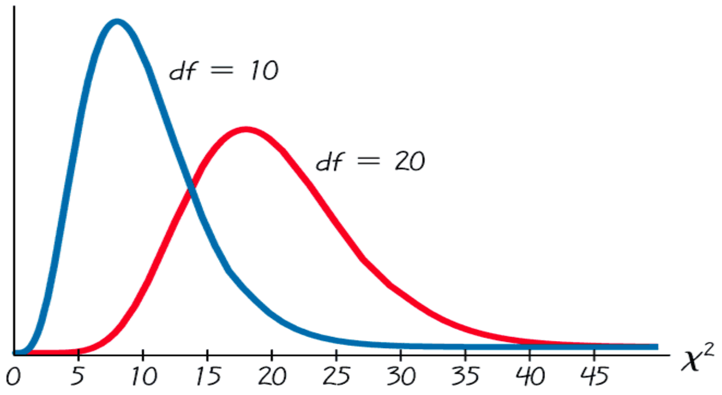


Figure 8-13

Chi-Square Distribution for 10 and 20 Degrees of Freedom



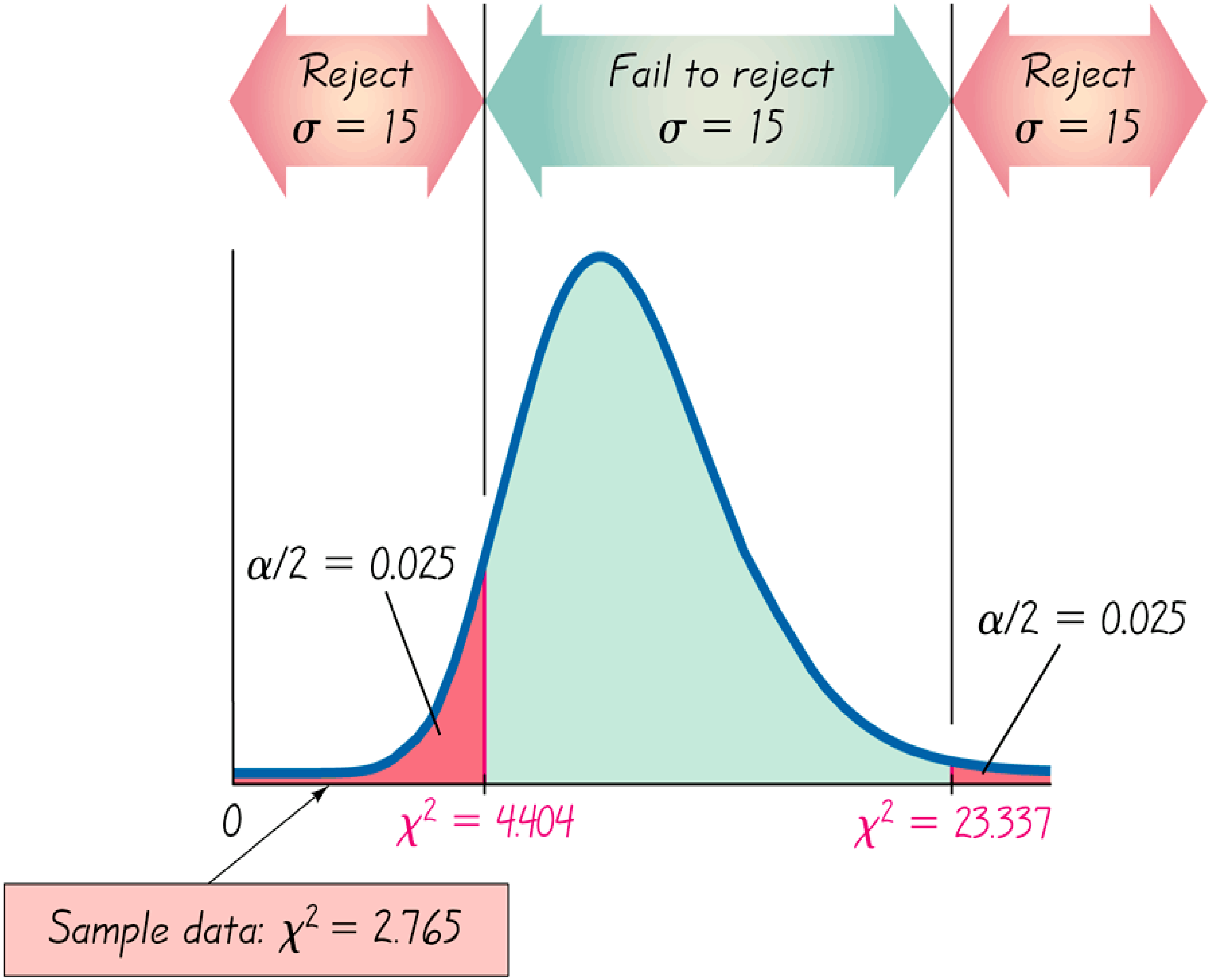
There is a different distribution for each number of degrees of freedom.

Figure 8-14

Example: For a simple random sample of adults, IQ scores are normally distributed with a mean of 100 and a standard deviation of 15. A simple random sample of 13 statistics professors yields a standard deviation of s = 7.2. Assume that IQ scores of statistics professors are normally distributed and use a 0.05 significance level to test the claim that $\sigma = 15$.

H_0: $\sigma = 15$
H_1: $\sigma \neq 15$
$\alpha = 0.05$
$n = 13$
$s = 7.2$

$$\chi^2 = \frac{(n-1)s^2}{\sigma^2} = \frac{(13-1)(7.2)^2}{15^2} = 2.765$$


Reject
σ = 15
Fail to reject
σ = 15
Reject
σ = 15
α/2 = 0.025
α/2 = 0.025
0
χ² = 4.404
χ² = 23.337
Sample data: χ² = 2.765

The critical values of 4.404 and 23.337 are found in Table , in the 12th row (degrees of freedom = *n* – 1) in the column corresponding to 0.975 and 0.025.

Because the test statistic is in the critical region, we reject the null hypothesis. There is sufficient evidence to warrant rejection of the claim that the standard deviation is equal to 15.

- **Uses**
- $\chi 2$ test of goodness of fit. Through the test we can find out the deviations between the observed values and expected values. Here we are not concerned with the parameters but concerned with the form of distribution. Karl Pearson has developed a method to test the difference between the theoretical value (hypothesis) and the observed value. The test is done by comparing the computed value with the table value of $\chi 2$ for the desired degree of freedom. A Greek letter $\chi 2$ is used to describe the magnitude of difference between the fact and theory.
- The $\chi 2$ may be defined as

$$\chi^2 = \sum \frac{(O-E)^2}{E}$$

Where:

χ^2 = Chi Square obtained
$\sum$ = the sum of
O = observed score
E = expected score

Steps

- A hypothesis is established along with the significance level.
- Compute deviations between observed value and expected value (O-E).
- Square the deviations calculated (O-E)2.
- Divide the (O-E)2 by its expected frequency.
- Add all the values obtained in step 4.
- Find the value of χ2, from χ2 table at certain level of significance, usually 5% level.

If the calculated value of χ2 is greater than the tabled value of χ2, at certain level of significance, we reject the hypothesis. If the computed value of χ2 value is zero then, the observed value and expected values completely coincide. If the computed value of χ2 is less than the table value, at a certain degree of level of significance, it is said to be non significant. This implies that the discrepancy, between the observed and expected frequencies in simple sampling.

- **χ2 as a test of independence**
- χ2 test can be used to find out whether one or more attributes are associated or not. For example, coaching class and successful candidate, marriage and failure, etc.; we can find out whether they are related or independent we take a hypothesis that the attributes are independent. If the calculated value of x2 is less than the tabled value at a certain level of significance, the hypothesis is correct and vice versa.

$$\chi_c^2 = \sum \frac{(O_i - E_i)^2}{E_i}$$

- Expected value = $\dfrac{\text{Row Total (R.T) x Column Total (C.T)}}{\text{Grand Total (G.T)}}$

- **Example 1:** A certain drug was administered to 500 people out of a total of 800 included in the sample to test its efficiency against typhoid. The result are given below.

	Typhoid	No. Typhoid	Total
Drug	200	300	500
No Drug	280	20	300
Total	480	320	800

- On the basis of these data, can it be concluded that the drug is effective in preventing typhoid (Given for V = 1, χ^2 0.05 = 3.84).

- **Solution :**
- Let the hypothesis be "The drug is not effective in preventing typhoid". "The table of expected frequencies are:

480 X 500 -------------- =300 800	320 X 500 ------------- = 200 800	500
480 X 300 -------------- = 180 800	320 X 300 ------------- = 120 800	300
480	320	800

- N.B Alternatively, after finding out the first value, the remaining can be obtained in the following manner.

O	E	(O-E)	$(O-E)^2$	$(O-E)^2/E$
200	300	-100	10000	33.33
280	180	+100	10000	55.56
300	200	+100	10000	50.00
20	120	-100	10000	83.33
800	800	$\chi 2 =$ 222.22		

- V = (r-l) (C-l .)
- = (2-1) (2-1)
- = 1 x 1
- = 1.
- d.f= l , $\chi2$ 0.05 = 3.84
- Calculated value of $\chi2$ = 222.22, Table value of $\chi2$ = 3.84 (from table) The calculated value of $\chi2$ is much greater than the table value. Therefore the hypothesis the drug is not effective is rejected. Hence we conclude that the drug is effective in preventing typhoid:

- **ANOVA**
- ANOVA is a statistical technique. It is used to test the equality of three or more sample means.
- Based on the means, inference is drawn whether samples belongs to same population or not.
- ***Conditions for using ANOVA***
- 1. Data should be quantitative in nature.
- 2. Data normally distributed.
- 3. Samples drawn from a population follow random variation.

- ***ANOVA can be discussed in two parts:***
- 1. One-way classification
- 2. Two and three-way classification.

- **One-way ANOVA**
- Following are the steps followed in ANOVA:
- 1. Calculate the variance between samples.
- 2. Calculate the variance within samples.
- 3. Calculate F ratio using the formula.
- F = Variance between the samples/Variance within the sample
- 4. Compare the value of F obtained above in (3) with the critical value of F such as 5% level of significance for the applicable degree of freedom.
- 5. When the calculated value of F is less than the table value of F, the difference in sample means is not significant and a null hypothesis is accepted. On the other hand, when the calculated value of F is more than the critical value of F, the difference in sample means is considered as significant and the null hypothesis is rejected.

- *Example:* ANOVA is useful.
- 1. To compare the mileage achieved by different brands of automotive fuel.
- 2. Compare the first year earnings of graduates of half a dozen top business schools.
- **Application in Market Research**
- Consider the following pricing experiment. Three prices are considered for a new toffee box introduced by Nutrine company. Price of three varieties of toffee boxes are 39, 44 and 49. The idea is to determine the influence of price levels on sales. Five supermarkets are selected to exhibit these toffee boxes. The sales are as follows:

Price ( )	1	2	3	4	5	Total	Sample mean $\bar{x}$	
39	8	12	10	9	11	50	10	
44	7	10	6	8	9	40	8	
49	4	8	7	9	7	35	7	

- What the manufacturer wants to know is: (1) Whether the difference among the means is significant? If the difference is not significant, then the sale must be due to chance. (2) Do the means differ? (3) Can we conclude that the three samples are drawn from the same population or not?

- *Example:* In a company there are four shop floors. Productivity rate for three methods of incentives and gain sharing in each shop floor is presented in the following table. Analyze whether various methods of incentives and gain sharing differ significantly at 5% and 1% F-limits.

Shop Floor	Productivity rate data for three methods of incentives and gain sharing		
	X1	x2	x3
1	5	4	4
2	6	4	3
3	2	2	2
4	7	6	3

- ***Solution:***
- ***Step 1:*** Calculate mean of each of the three samples (i.e., x1, x2 and x3, i.e. different methods of incentive gain sharing).
- $\bar{\bar{X}} = \bar{x}1 + \bar{x}2 + \bar{x}3$
- k

Thank you.

BR MODULE III

By Asst. Prof. (Eco) Mrs. Tapaswini Nayak.

- **Factor Analysis**
- The main purpose of Factor Analysis is to group large set of variable factors into fewer factors. Each factor will account for one or more component. Each factor is a combination of many variables.
- There are two most commonly employed factor analysis procedures or methods. They are:
- **1. Principle component analysis**
- **2. Common factor analysis**.
- When the objective is to summarize information from a large set of variables into fewer factors, **principle component factor analysis** is used. On the other hand, if the researcher wants to analyze the components of the main factor, **common factor analysis** is used.

- ***Example:*** **Common factor** – Inconvenience inside a car. The components may be:
- 1. Leg room
- 2. Seat arrangement
- 3. Entering the rear seat
- 4. Inadequate dickey space
- 5. Door locking mechanism.

- **Principle Component Factor Analysis**
- ***Purposes:*** Customer feedback about a two-wheeler manufactured by a company.
- ***Method:*** The MR manager prepares a questionnaire to study the customer feedback. The researcher has identified six variables or factors for this purpose. They are as follows:
- 1. Fuel efficiency (A)
- 2. Durability (Life) (B)
- 3. Comfort (C)
- 4. Spare parts availability (D)
- 5. Breakdown frequency (E)
- 6. Price (F)

- The questionnaire may be administered to 5,000 respondents. The opinion of the customer is gathered. Let us allot points 1 to 10 for the variables factors A to F. 1 is the lowest and 10 is the highest. Let us assume that application of factor analysis has led to grouping the variables as follows:

- A, B, D, E into factor-1
- F into Factor -2
- C into Factor - 3
- Factor - 1 can be termed as Technical factor;
- Factor - 2 can be termed as Price factor;
- Factor - 3 can be termed as Personal factor.

- For future analysis, while conducting a study to obtain customers' opinion, three factors mentioned above would be sufficient. One basic purpose of using factor analysis is to reduce the number of independent variables in the study.

- **Types**
- Factor analysis may be **R-type factor analysis** or it may be **Q-type factor analysis.**
- **In R-type factor analysis**, high correlations occur when respondents who score high on variable 1 also score high on variable 2 and respondents who score low on variable 1 also score low on variable 2. Factors emerge when there are high correlations within groups of variables.
- **In Q-type factor analysis**, the correlations are computed between pairs of respondents instead of pairs of variables. High correlations occur when respondent 1's pattern of responses on all the variables is much like respondent 2's pattern of responses. Factors emerge when there are high correlations within groups of people.
- **Q-type analysis** is useful when the object is to sort out people into groups based on their simultaneous responses to all the variables.
- **Factor analysis** has been mainly used in developing psychological tests (such as IQ tests, personality tests, and the like) in the realm of psychology. In marketing, this technique has been used to look at media readership profiles of people.

- ***Merits:*** The main merits of factor analysis can be stated thus:
- 1. The technique of factor analysis is quite useful when we want to condense and simplify the multivariate data.
- 2. The technique is helpful in pointing out important and interesting, relationships among observed data that were there all the time, but not easy to see from the data alone.
- 3. The technique can reveal the latent factors (i.e., underlying factors not directly observed) that determine relationships among several variables concerning a research study. For example, if people are asked to rate different cold drinks (say, Limca, Nova-cola, Gold Spot and so on) according to preference, a factor analysis may reveal some salient characteristics of cold drinks that underlie the relative preferences.
- 4. The technique may be used in the context of empirical clustering of products, media or people i.e., for providing a classification scheme when data scored on various rating scales have to be grouped together.

- ***Limitations:*** One should also be aware of several limitations of factor analysis. Important ones are as follows:
- 1. Factor analysis, like all multivariate techniques, involves laborious computations involving heavy cost burden. With computer facility available these days, there is no doubt that factor analysis has become relatively faster and easier, but the cost factor continues to be the same i.e., large factor analyses are still bound to be quite expensive.
- 2. The results of a single factor analysis are considered generally less reliable and dependable for very often a factor analysis starts with a set of imperfect data. "The factors are nothing but blurred averages, difficult to be identified." To overcome this difficulty, it has been realised that analysis should at least be done twice. If we get more or less similar results from all rounds of analyses, our confidence concerning such results increases.
- 3. Factor-analysis is a complicated decision tool that can be used only when one has thorough knowledge and enough experience of handling this tool. Even then, at times it may not work well and may even disappoint the user.

- **To conclude**, we can state that in spite of all the said limitations “when it works well, factor analysis helps the investigator make sense of large bodies of intertwined data. When it works unusually well, it also points out some interesting relationships that might not have been obvious from examination of the input data alone”.

- **Application Of Factor Analysis:**
- **1. Model Building For New Product Development:**
- As pointed out earlier, a real life situation is highly complex and it consists of several variables. A model for the real life situation can be built by incorporating as many features of the situation as possible. But then, with a multitude of features, it is very difficult to build such a highly idealistic model. A practical way is to identify the important variables and incorporate them in the model. Factor analysis seeks to identify those variables which are highly correlated among themselves and find a common factor which can be taken as a representative of those variables. Based on the factor loading, some of variables can be merged together to give a common factor and then a model can be built by incorporating such factors. Identification of the most common features of a product preferred by the consumers will be helpful in the development of new products.

- **2. Model Building For Consumers:**
- Another application of factor analysis is to carry out a similar exercise for the respondents instead of the variables themselves. Using the factor loading, the respondents in a research survey can be sorted out into various groups in such a way that the respondents in a group have more or less homogeneous opinions on the topics of the survey. Thus a model can be constructed on the groups of consumers. The results emanating from such an exercise will guide the management in evolving appropriate strategies towards market segmentation.

• Multiple regression analysis:

- **INTRODUCTION**
- Multiple regression analysis is a powerful technique used for predicting the unknown value of a variable from the known value of two or more variables.
- It also called as predictors.
- Method used for studying the relationship between a dependent variable and two or more independent variables.
- **Purposes:**

❖ Prediction
❖ Explanation
❖ Theory building

- You use **correlation analysis** to find out if there is a statistically significant relationship between TWO variables.
- You use **linear regression analysis** to make predictions based on the relationship that exists between two variables. The main limitation that you have with correlation and linear regression as you have just learned how to do it is that it only works when you have TWO variables. The problem is that most things are way too complicated to "model" them with just two variables.
- **Multiple Regression** (R) A statistical tool that allows you to examine how multiple independent variables are related to a dependent variable. Once you have identified how these multiple variables relate to your dependent variable, you can take information about all of the independent variables and use it to make much more powerful and accurate predictions about why things are the way they are. This latter process is called **"Multiple Regression".**

- The variable whose value is to be predicted is known as the **dependent variable.**
- The ones whose known values are used for prediction are known **Independent (exploratory) variables**.
- **Regression analysis** is a statistical tool used for the investigation of relationships between variables. Usually, the investigator seeks to ascertain the causal effect of one variable upon another — the effect of a price increase upon demand, for example, or the effect of changes in the money supply upon the inflation rate.
- **Regression analysis** is used to estimate the strength and the direction of the relationship between two linearly related variables: X and Y. X is the "independent" variable and Y is the "dependent" variable.

- The two basic types of regression analysis are:
- **Simple regression analysis:** Used to estimate the relationship between a dependent variable and a single independent variable; for example, the relationship between crop yields and rainfall.
- **Multiple regression analysis:** Used to estimate the relationship between a dependent variable and two or more independent variables; for example, the relationship between the salaries of employees and their experience and education.
- Multiple regression analysis introduces several additional complexities but may produce more realistic results than simple regression analysis.

- Multiple regression models thus describe how a single response variable Y depends linearly on a number of predictor variables.
- **Examples:**

❖ The selling price of a house can depend on the desirability of the location, the number of bedrooms, the number of bathrooms, the year the house was built, the square footage of the lot and a number of other factors.

❖ The height of a child can depend on the height of the mother, the height of the father, nutrition, and environmental factors.

- Regression analysis is based on several strong assumptions about the variables that are being estimated. Several key tests are used to ensure that the results are valid, including hypothesis tests. These tests are used to ensure that the regression results are not simply due to random chance but indicate an actual relationship between two or more variables.
- An estimated regression equation may be used for a wide variety of business applications, such as:
  - Measuring the impact on a corporation’s profits of an increase in profits
  - Understanding how sensitive a corporation’s sales are to changes in advertising expenditures
  - Seeing how a stock price is affected by changes in interest rates

- Regression analysis may also be used for forecasting purposes; for example, a regression equation may be used to forecast the future demand for a company’s products.

- There are 3 major uses for multiple linear regression analysis.
- **First,** it might be used to identify the strength of the effect that the independent variables have on a dependent variable.
- **Second**, it can be used to forecast effects or impacts of changes. That is, multiple linear regression analysis helps us to understand how much will the dependent variable change when we change the independent variables. For instance, a multiple linear regression can tell you how much GPA is expected to increase (or decrease) for every one point increase (or decrease) in IQ.
- **Third,** multiple linear regression analysis predicts trends and future values. The multiple linear regression analysis can be used to get point estimates. An example question may be "what will the price of gold be 6 month from now?"

- **Discriminant Analysis**
- In this analysis, two or more groups are compared. In the final analysis, we need to find out whether the groups differ one from another.
- ***Example:*** Where discriminant analysis is used
- 1. Those who buy our brand and those who buy competitors’ brand.
- 2. Good salesman, poor salesman, medium salesman
- 3. Those who go to Food World to buy and those who buy in a Kirana shop.
- 4. Heavy user, medium user and light user of the product.

- **Suppose** there is a comparison between the groups mentioned as above along with demographic and socio-economic factors, then discriminant analysis can be used. One way of doing this is to proceed and calculate the income, age, educational level, so that the profile of each group could be determined. Comparing the two groups based on one variable alone would be informative but it would not indicate the relative importance of each variable in distinguishing the groups. This is because several variables within the group will have some correlation which means that one variable is not independent of the other.

- If we are interested in segmenting the market using income and education, we would be interested in the total effect of two variables in combinations, and not their effects separately. Further, we would be interested in determining which of the variables are more important or

- had a greater impact. To summarize, we can say, that **Discriminant Analysis** can be used when we want to consider the variables simultaneously to take into account their interrelationship.
- Like regression, the value of dependent variable is calculated by using the data of independent variable.

- $Z = b_1x_1 + b_2x_2 + b_3x_3 + \ldots\ldots\ldots\ldots.$
- Z = Discriminant score
- b_1 = Discriminant weight for variable
- x = Independent variable
- As can be seen in the above, each independent variable is multiplied by its corresponding weightage.
- This results in a single composite discriminant score for each individual. By taking the average of discriminant score of the individuals within a certain group, we create a group mean. This is known as centroid. If the analysis involves two groups, there are two centroids. This is very similar to multiple regression, except that different types of variables are involved.

- **Application**
- A company manufacturing **FMCG** products introduces a sales contest among its marketing executives to find out "How many distributors can be roped in to handle the company's product". Assume that this contest runs for three months. Each marketing executive is given target regarding number of new distributors and sales they can generate during the period. This target is fixed and based on the past sales achieved by them about which, the data is available in the company. It is also announced that marketing executives who add 15 or more distributors will be given a Maruti Swift as prize. Those who generate between 5 and 10 distributors will be given a two-wheeler as the prize. Those who generate less than 5 distributors will get nothing. Now assume that 5 marketing executives won a Maruti Swift and 4 won a two-wheeler.

- The company now wants to find out, “Which activities of the marketing executive made the difference in terms of winning a prize and not winning the prize”. One can proceed in a number of ways. The company could compare those who won the Maruti Swift against the others. Alternatively, the company might compare those who won, one of the two prizes against those who won nothing. It might compare each group against each of the other two.
- Discriminant analysis will highlight the difference in activities performed by each group members to get the prize. The activity might include:

- 1. More number of calls made to the distributors.
- 2. More personal visits to the distributors with advance appointments.
- 3. Use of better convincing skills.
- **Discriminant analysis** answers the following questions:
- 1. What variable discriminates various groups as above; the number of groups could be two or more? Dealing with more than two groups is called Multiple Discriminant Analysis (M.D.A.).
- 2. Can discriminating variables be chosen to forecast the group to which the brand/person/place belong to?
- 3. Is it possible to estimate the size of different groups?

• Research Report

- **Introduction**
- A report is a very formal document that is written for a variety of purposes, generally in the sciences, social sciences, engineering and business disciplines. Generally, findings pertaining to a given or specific task are written up into a report. It should be noted that reports are considered to be legal documents in the workplace and, thus, they need to be precise, accurate and difficult to misinterpret.
- There are three features that, together, characterize report writing at a very basic level: a predefined structure, independent sections, and reaching unbiased conclusions.
- ***Predefined structure:*** Broadly, these headings may indicate sections within a report, such as an introduction, discussion, and conclusion.
- ***Independent sections:*** Each section in a report is typically written as a stand-alone piece, so the reader can selectively identify the report sections they are interested in, rather than reading the whole report through in one go from start to finish.
- ***Unbiased conclusions:*** A third element of report writing is that it is an unbiased and objective form of writing.

- **Types of Report Notes**
- There are two types of reports (1) Oral report (2) Written report.
- **Oral Report**
- This type of reporting is required, when the researchers are asked to make an oral presentation. Making an oral presentation is somewhat difficult compared to the written report. This is because the reporter has to interact directly with the audience. Any faltering during an oral presentation can leave a negative impression on the audience. This may also lower the self confidence of the presenter. In an oral presentation, communication plays a big role. A lot of planning and thinking is required to decide 'What to say', 'How to say', 'How much to say'. Also, the presenter may have to face a barrage of questions from the audience. A lot of preparation is required; the broad classification of an oral presentation is as follows.

- **Nature of an Oral Presentation**
- ***Opening:*** A brief statement can be made on the nature of discussion that will follow. The opening statement should explain the nature of the project, how it came about and what was attempted.
- ***Finding/Conclusion:*** Each conclusion may be stated backed up by findings.
- ***Recommendation:*** Each recommendation must have the support of conclusion. At the end of the presentation, question-answer session should follow from the audience.
- ***Method of presentation:*** Visuals, if need to be exhibited, can be made use of. The use of tabular form for statistical information would help the audience.

- (a) What type of presentation is a root question? Is it read from a manuscript or memorized or delivered ex-tempo. Memorization is not recommended, since there could be a slip during presentation. Secondly, it produces speaker-centric approach. Even reading from the manuscript is not recommended, because it becomes monotonous, dull and lifeless. The best way to deliver in ex-tempo, is to make main points notes, so that the same can be expanded. Logical sequences should be followed.

- **Written Report**
- Following are the Various Types of Written Reports:
- (A) ***Reports can be classified based on the time-interval such as:***
- (1) Daily
- (2) Weekly
- (3) Monthly
- (4) Quarterly
- (5) Yearly
- (B) ***Type of reports:***
- (1) Short report
- (2) Long report
- (3) Formal report
- (4) Informal report
- (5) Government report

- **1. *Short report:*** Short reports are produced when the problem is very well defined and if the scope is limited. For example, Monthly sales report. It will run into about five pages. It consists of report about the progress made with respect to a particular product in a clearly specified geographical locations.
- **2. *Long report:*** This could be both a technical report as well as non-technical report. This will present the outcome of the research in detail.
- (a) ***Technical report:*** This will include the sources of data, research procedure, sample design, tools used for gathering data, data analysis methods used, appendix, conclusion and detailed recommendations with respect to specific findings. If any journal, paper or periodical is referred, such references must be given for the benefit of reader.
- (b) ***Non-technical report:*** This report is meant for those who are not technically qualified. E.g. Chief of the finance department. He may be interested in financial implications only, such as margins, volumes, etc. He may not be interested in the methodology.

- 3. ***Formal report:***
- *Example:* The report prepared by the marketing manager to be submitted to the Vice- President (marketing) on quarterly performance, reports on test marketing.
- 4. ***Informal report:*** The report prepared by the supervisor by way of filling the shift log book, to be used by his colleagues.
- 5. ***Government report:*** These may be prepared by state governments or the central government on a given issue.
- *Example:* Programme announced for rural employment strategy as a part of five-year plan.

- **Distinguish between Oral and Written Report**

Oral report	Written report
No rigid standard format.	Standard format can be adopted.
Remembering all that is said is difficult if not impossible. This is because the presenter cannot be interrupted frequently for clarification.	This can be read a number of times and clarification can be sought whenever the reader chooses.
Tone, voice modulation, comprehensibility and several other communication factors play an important role.	Free from presentation problems.
Correcting mistakes if any, is difficult.	Mistakes, if any, can be pinpointed and corrected.
The audience has no control over the speed of presentation.	Not applicable.
The audience does not have the choice of picking and choosing from the presentation.	The reader can pick and choose what he thinks is relevant to him. For instance, the need for information is different for technical and nontechnical persons.

- **Differentiate between Technical Report & Popular Report :**
- **A *technical report*** is used whenever a full written report of the study is required whether for recordkeeping or for public dissemination.
- **A *popular report*** is used if the research results have policy implications. We give below a few details about the said two types of reports:

- ***Technical Report***

In the technical report the main emphasis is on

i. the methods employed,
ii. assumptions made in the course of the study,
iii. the detailed presentation of the findings including their limitations and supporting data.

- A general outline of a **technical report** can be as follows:
- ***1. Summary of results:*** A brief review of the main findings just in two or three pages.
- ***2. Nature of the study:*** Description of the general objectives of study, formulation of the problem in operational terms, the working hypothesis, the type of analysis and data required, etc.
- ***3. Methods employed:*** Specific methods used in the study and their limitations. For instance, in sampling studies we should give details of sample design viz., sample size, sample selection, etc.
- ***4. Data:*** Discussion of data collected, their sources, characteristics and limitations. If secondary data are used, their suitability to the problem at hand be fully assessed. In case of a survey, the manner in which data were collected should be fully described.

- ***5. Analysis of data and presentation of findings**:* The analysis of data and presentation of the findings of the study with supporting data in the form of tables and charts be fully narrated. This, in fact, happens to be the main body of the report usually extending over several chapters.
- ***6. Conclusions:*** A detailed summary of the findings and the policy implications drawn from the results be explained.
- ***7. Bibliography:*** Bibliography of various sources consulted be prepared and attached.
- ***8. Technical appendices:*** Appendices be given for all technical matters relating to questionnaire, mathematical derivations, elaboration on particular technique of analysis and the like ones.
- ***9. Index:*** Index must be prepared and be given invariably in the report at the end.

- The order presented above only gives a general idea of the nature of a technical report; the order of presentation may not necessarily be the same in all the technical reports. This, in other words, means that the presentation may vary in different reports; even the different sections outlined above will not always be the same, nor will all these sections appear in any particular report. It should, however, be remembered that even in a technical report, simple presentation and ready availability of the findings remain an important consideration and as such the liberal use of charts and diagrams is considered desirable.

- ***Popular Report***
- The popular report is one which gives emphasis on simplicity and attractiveness. The simplification should be sought through clear writing, minimization of technical, particularly mathematical, details and liberal use of charts and diagrams. Attractive layout along with large print, many subheadings, even an occasional cartoon now and then is another characteristic feature of the popular report. Besides, in such a report emphasis is given on practical aspects and policy implications. We give below a general outline of a **popular report.**

- ***1. The findings and their implications:*** Emphasis in the report is given on the findings of most practical interest and on the implications of these findings.
- ***2. Recommendations for action:*** Recommendations for action on the basis of the findings of the study is made in this section of the report.
- ***3. Objective of the study:*** A general review of how the problem arise is presented along with the specific objectives of the project under study.
- ***4. Methods employed:*** A brief and non-technical description of the methods and techniques used, including a short review of the data on which the study is based, is given in this part of the report.

- ***5. Results:*** This section constitutes the main body of the report wherein the results of the study are presented in clear and non-technical terms with liberal use of all sorts of illustrations such as charts, diagrams and the like ones.
- ***6. Technical appendices:*** More detailed information on methods used, forms, etc. is presented in the form of appendices. But the appendices are often not detailed if the report is entirely meant for general public.

- There can be several variations of the form in which a popular report can be prepared. The only important thing about such a report is that it gives emphasis on simplicity and policy implications from the operational point of view, avoiding the technical details of all sorts to the extent possible.

- **Significance of Report Writing**
- Preparation and presentation of a research report is the most important part of the research process. No matter how brilliant the hypothesis and how well designed is the research study, they are of little value unless communicated effectively to others in the form of a research report. Moreover, if the report is confusing or poorly written, the time and effort spent on gathering and analysing data would be wasted. It is therefore, essential to summarise and communicate the result to the management in the form of an understandable and logical research report.

- **Research report** is regarded as a major component of the research study for the research task remains unfinished till the report has been presented and/or written. As a matter of fact even the most brilliant hypothesis, very well designed and conducted research study, and the most striking generalizations and findings are of modest value unless they are effectively communicated to others. The rationale of research is not well served unless the findings are made known to others. Research results must customarily enter the general store of knowledge. All this explains the importance of writing research report. There are people who do not consider writing of report as an essential part of the research process. But the general opinion is in favour of treating the presentation of research results or the writing of report as division and parcel of the research project. Writing of report is the final step in a research study and requires a set of skills somewhat different from those called for in respect of the former stages of research. This task should be accomplished by the researcher with extreme care; he may seek the assistance and guidance of experts for the reason.

- **STRUCTURE OF A RESEARCH REPORT**
- The Research report layout must necessarily be conveyed enough about the study so that he can place it in its general scientific context, judge the adequacy of its methods and thus form an opinion of how seriously the findings are to be taken.
- For this purpose there is the need of proper layout of the report. The layout of the report means as to what the research report should contain. A comprehensive layout of the research report should comprise preliminary pages, the main text and the end matter. Let us deal with them separately.

- ***A. Preliminary Pages***
- In its preliminary pages the report should carry a *title and date,* followed by acknowledgements in the form of 'Preface' or 'Foreword'. Then there should be a *table of contents* followed by *list of tables and illustrations* so that the decision-maker or anybody interested in reading the report can easily locate the required information in the report.
- ***B. Main Text***
- The main text provides the complete outline of the research report along with all details. Title of the research study is repeated at the top of the first page of the main text and then follows the other details on pages numbered consecutively, beginning with the second page. Each main section of the report should begin on a new page. The main text of the report should have the following sections:

1. Introduction
2. Statement of findings and recommendations
3. The results
4. The implications drawn from the results; and
5. The summary.

- ***1. Introduction:*** The purpose of introduction is to introduce the research project to the readers. It should contain a clear statement of the objectives of research i.e., enough background should be given to make clear to the reader why the problem was considered worth investigating. A brief summary of other relevant research may also be stated so that the present study can be seen in that context. The hypotheses of study, if any, and the definitions of the major concepts employed in the study should be explicitly stated in the introduction of the report.

- The methodology adopted in conducting the study must be fully explained. The scientific reader would like to know in detail about such thing: How was the study carried out? What was its basic design? If the study was an experimental one, then what were the experimental manipulations? If the data were collected by means of questionnaires or interviews, then exactly what questions were asked (The questionnaire or interview schedule is usually given in an appendix)? If measurements were based on observation, then what instructions were given to the observers?
- ***2. Statement of findings and recommendations:*** After introduction, the research report must contain a statement of findings and recommendations in non-technical language so that it can be easily understood by all concerned. If the findings happen to be extensive, at this point they should be put in the summarised form.

- ***3. Results:*** A detailed presentation of the findings of the study, with supporting data in the form of tables and charts together with a validation of results, is the next step in writing the main text of the report. This generally comprises the main body of the report, extending over several chapters. The result section of the report should contain statistical summaries and reductions of the data rather than the raw data.
- All the results should be presented in logical sequence and splitted into readily identifiable sections. All relevant results must find a place in the report. But how one is to decide about what is relevant is the basic question. Quite often guidance comes primarily from the research problem and from the hypotheses, if any, with which the study was concerned.
- But ultimately the researcher must rely on his own judgement in deciding the outline of his report. “Nevertheless, it is still necessary that he states clearly the problem with which he was concerned, the procedure by which he worked on the problem, the conclusions at which he arrived, and the bases for his conclusions.

- ***4. Implications of the results:*** Toward the end of the main text, the researcher should again put down the results of his research clearly and precisely. He should, state the implications that flow from the results of the study, for the general reader is interested in the implications for understanding the human behaviour. Such implications may have three aspects as stated below:

- A statement of the inferences drawn from the present study which may be expected to apply in similar circumstances.
- The conditions of the present study which may limit the extent of legitimate generalizations of the inferences drawn from the study.
- The relevant questions that still remain unanswered or new questions raised by the study along with suggestions for the kind of research that would provide answers for them. It is considered a good practice to finish the report with a short conclusion which summarises and recapitulates the main points of the study. The conclusion drawn from the study should be clearly related to the hypotheses that were stated in the introductory section. At the same time, a forecast of the probable future of the subject and an indication of the kind of research which needs to be done in that particular field is useful and desirable.

- ***5. Summary:*** It has become customary to conclude the research report with a very brief summary, resting in brief the research problem, the methodology, the major findings and the major conclusions drawn from the research results.

- ***C. End Matter***
- At the end of the report, appendices should be enlisted in respect of all technical data such as questionnaires, sample information, mathematical derivations and the like ones. Bibliography of sources consulted should also be given. Index (an alphabetical listing of names, places and topics along with the numbers of the pages in a book or report on which they are mentioned or discussed) should invariably be given at the end of the report. The value of index lies in the fact that it works as a guide to the reader for the contents in the report.

- **ORAL PRESENTATION OF RESEARCH REPORT**

- Oral presentation of the results of the study is considered effective, particularly in cases where policy recommendations are indicated by project results.
- The merit of this approach lies in the fact that it provides an opportunity for give-and-take decisions which generally lead to a better understanding of the findings and their implications.
- But the main demerit of this sort of presentation is the lack of any permanent record concerning the research details and it may be just possible that the findings may fade away from people's memory even before an action is taken.
- In order to overcome this difficulty, a written report may be circulated before the oral presentation and referred to frequently during the discussion.
- Oral presentation is effective when supplemented by various visual devices. Use of slides, wall charts and blackboards is quite helpful in contributing to clarity and in reducing the boredom, if any. Distributing a board outline, with a few important tables and charts concerning the research results, makes the listeners attentive who have a ready outline on which to focus their thinking.

- This very often happens in academic institutions where the researcher discusses his research findings and policy implications with others either in a seminar or in a group discussion.

Thus, research results can be reported in more than one ways, but the usual practice adopted, in academic institutions particularly, is that of writing the Technical Report and then preparing several research papers to be discussed at various forums in one form or the other. But in practical field and with problems having policy implications, the technique followed is that of writing a popular report. Researches done on governmental account or on behalf of some major public or private organisations are usually presented in the form of technical reports.

Thank you.