

RESEARCH METHODOLOGY

M.A., (Journalism & Mass Communication)

Semester – III, Paper- I

Lesson Writers

Dr. Padma Ponugoti

Assistant Professor
Dept. of Journalism and
Mass Communication
Dr.B.RAmbedkar University
Srikakulam, AP.

Dr.Tarakeswara Rao Ippili

Assistant Professor
Dept. of Journalism and
Mass Communication
Central Tribal University of A.P.
Vizayanagaram, AP.

Dr. K. John Babu

Assistant Professor
Dept. of Convergent Journalism
School of Media Studies
Central University of Kashmir
Tumula Campus.

Lesson Writer & Editor

Dr. J. Madhu Babu

Assistant Professor,
Dept. of Journalism and Mass Communication
Acharya Nagarjuna University
Nagarjuna Nagar, Guntur, AP.

Director

Dr. NAGARAJU BATTU

MBA., MHRM., LLM., M.Sc. (Psy).,MA (Soc)., M.Ed., M.Phil., Ph.D

CENTRE FOR DISTANCE EDUCATION

ACHARAYANAGARJUNAUNIVERSITY

NAGARJUNANAGAR – 522510

Ph:0863-2293299,2293214,

0863-2346259(Study Material)

Website: www.anucde.info

e-mail:anucdedirector@gmail.com

M.A. (JMC) – Research Methodology

First Edition : 2023

No. of Copies :

©Acharya Nagarjuna University

This book is exclusively prepared for the use of students of M.A.(Journalism & mass Communication) Centre for Distance Education, Acharya Nagarjuna University and this book is meant for limited Circulation only.

Published by:

Dr. NAGARAJU BATTU,

Director

**Centre for Distance Education,
Acharya Nagarjuna University**

Printed at:

FOREWORD

Since its establishment in 1976, Acharya Nagarjuna University has been forging a head in the path of progress and dynamism, offering a variety of courses and research contributions. I am extremely happy that by gaining 'A' grade from the NAAC in the year 2016, Acharya Nagarjuna University is offering educational opportunities at the UG, PG levels apart from research degrees to students from over 443 affiliated colleges spread over the two districts of Guntur and Prakasam.

The University has also started the Centre for Distance Education in 2003-04 with the aim of taking higher education to the door step of all the sectors of the society. The centre will be a great help to those who cannot join in colleges, those who cannot afford the exorbitant fees as regular students, and even to housewives desirous of pursuing higher studies. Acharya Nagarjuna University has started offering B.A., and B.Com courses at the Degree level and M.A., M.Com., M.Sc., M.B.A., and L.L.M., courses at the PG level from the academic year 2003-2004onwards.

To facilitate easier understanding by students studying through the distance mode, these self-instruction materials have been prepared by eminent and experienced teachers. The lessons have been drafted with great care and expertise in the stipulated time by these teachers. Constructive ideas and scholarly suggestions are welcome from students and teachers involved respectively. Such ideas will be incorporated for the greater efficacy of this distance mode of education. For clarification of doubts and feedback, weekly classes and contact classes will be arranged at the UG and PG levels respectively.

It is my aim that students getting higher education through the Centre for Distance Education should improve their qualification, have better employment opportunities and in turn be part of country's progress. It is my fond desire that in the years to come, the Centre for Distance Education will go from strength to strength in the form of new courses and by catering to larger number of people. My congratulations to all the Directors, Academic Coordinators, Editors and Lesson-writers of the Centre who have helped in these endeavors.

Prof. P. Raja Sekhar
Vice-Chancellor
Acharya Nagarjuna University

III SEMESTER
PAPER - I
301JM21 : RESEARCH METHODOLOGY
Syllabus

Unit 1: Introduction to Mass Communication research: Meaning of research, scientific method – Characteristics – Steps in research identification and formulation of research problem in communication research.

Unit 2: Basic elements of research – concepts, definitions, variables, hypothesis and causation. Hypothesis – type of hypothesis – hypothesis testing - Research designs in Mass Communication - Survey research. Content analysis and historical method.

Unit 3: Sampling in communications Research: Types, their applications and limitations. Methods of data Collection: Interview, Questionnaire, Observation and Case study – Applications and limitations of different methods.

Unit 4: Use of statistics in communication research: Basic statistical tools, Measures of central tendency (mean, mode and median), measure of dispersion (standard deviation, correlation and chi-square).

Unit 5: Data processing, analysis, presentation and interpretation of data. Use of graphs in data presentation – Writing a research proposal; writing research report – Its components and style.

Reference Books:

1. Research Methods in the Behavioural Science -Festinger L.and Katz.D
2. Foundation of Behavioural Research -Kerlinger.F.N
3. Content Analysis. An Introduction to Methodology – Krippendorff.K
4. Research Methods in Mass Communication – Westley Bruce.N and Guido.H
5. Introduction to Mass Communication Research-Nafziger, Ralph, David and White.
6. Using Statistics for Psychological Research –Walker.J.T.
7. Methods in Social Research: W.J Goode&P.K. Hatt.
8. Methodology and Techniques of Social Science research: Wilkison and Bhandarkar
9. Research Methodology, Methods and Techniques: C.R.Kotari
10. Mass Media Research, an introduction: Winner & Dominic

CONTENTS

Sl.No	Lesson	Page No.
1.	Concepts, Nature, and Scope of Research	1.1 – 1.9
2.	Scientific Method	2.1 – 2.7
3.	Research Problem	3.1 – 3.11
4.	Elements of Research	4.1 – 4.9
5.	Hypothesis and Hypothesis Testing	5.1 – 5.7
6.	Research Design	6.1 – 6.10
7.	Research Approaches and Methods	7.1 – 7.12
8.	Sampling	8.1 – 8.8
9.	Methods of Data Collection	9.1 – 9.8
10.	Observation and Case Study Methods	10.1 – 10.9
11.	Statistical Methods in Media Research	11.1 – 11.11
12.	Standard Deviation and Variance	12.1 – 12.14
13.	Chi- Square Test	13.1 – 13.12
14.	Anova and Multiple Regressions	14.1 – 14.8
15.	Data Processing and Analysis	15.1 -15.13
16.	Writing A Research Proposal	16.1 – 16.11

LESSON - 1

CONCEPTS, NATURE, AND SCOPE OF RESEARCH

OBJECTIVES :

After reading this lesson, the reader should be able to gain an overview of Research Concept, Nature, and Scope.

- To define and understand Research Concept, Nature, and Scope.
- To analyse the scope of the Research
- To get an idea on the overview of Research

STRUCTURE :

- 1.1 Introduction
- 1.2 Definitions
 - 1.2.1 Nature and characteristics
 - 1.2.2 Purpose of research
- 1.3 Criteria of good research
- 1.4 Summary
- 1.5 Self-assessment Questions
- 1.6 Reference and Suggested readings

1.2 INTRODUCTION :

- ❖ Research is a scientific investigation. Investigation means a search for new facts and ideas in any branch of knowledge. Thus, we can say that research is a search for knowledge. Research may be considered as a movement, a movement from the unknown to the known. It is actually a voyage of discovery.
- ❖ Research is carried out for two purposes; one is the discovery of new facts and the second, verification of the old ones. The object of every business organization, of course, is the discovery of new facts, new relationship, and new laws governing the business phenomena. But constant verification of the old concepts is also needed especially in dynamic business environment.
- ❖ Common sense knowledge, based on the accumulated experiences, prejudices and beliefs of the people is often contradictory and inconsistent. On the other hand, scientific observations are based on verifiable evidence or systematic body of proof that can be cited. For example, some common-sense statements are: man is more intelligent than woman; married men remain happier than single people; rural people are more hardworking than urban people etc. Contrary to this, the scientific research or scientific inquiry finds that woman is as intelligent as man; there is no association in happiness and marriage; hard work is not related to environment alone. Thus, a statement based on common sense is just a guess or prejudice or mistaken interpretation, though at times it may be true, wise and a useful bit of knowledge. But it is not based on any scientific evidence. A scientific statement is based on accumulated systematic knowledge through research.

1.2 DEFINITIONS OF RESEARCH :

The following are the important definitions of research :

- ❖ **According to Francis Rommel** : “Research is an endeavour /attempt to discover, develop and verify knowledge. It is an intellectual process that has developed over hundreds of years ever-changing in purpose and form and always researching to the truth.”
- ❖ **According to P.M. Cook** : “Research is an honest, exhaustive, intelligent search for facts and their meanings or implications with reference to a given problem. The product or findings of a given piece of research should be an authentic, verifiable contribution to knowledge in the field studied.”
- ❖ **According to W.S. Monroe’s** : “Research may be defined as a method of studying problems whose solutions are to be derived partly or wholly from facts.”
- ❖ **John W. Best** : “ Research is considered to be the more formal, systematic intensive process of carrying on the scientific method of analysis. It involves a more systematic structure of investigation, usually resulting in some sort of formal record of procedures and a report of results or conclusion.”
- ❖ **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 careful testing the conclusions to determine whether they fit the formulated hypothesis.”
- ❖ **Redman& Mori** : “Research is a systematic effort to gain new knowledge.”
- ❖ **P.V. Young** : “Social research may be defined as a scientific undertaking which by means of logical and systematized techniques aims to discover new facts or verify and test old facts, analyse their sequences, interrelationships and casual explanation which were derived within an appropriate theoretical frame of reference, develop new scientific tools, concepts and theories which would facilitate reliable and valid study of human behaviour.”

Meaning of Research :

Research is undertaken within most professions. More than a set of skills, it is a way of thinking: examining critically the various aspects of professional work. It is a habit of questioning what you do, and a systematic examination of the observed information to find answers with a view to instituting appropriate changes for a more effective professional service.

When you say that you are undertaking a research study, to find answers to a question, you are implying that the process:

1. Is being undertaken within a framework of a set of philosophies (approaches);
2. Uses procedures, methods and techniques that have been tested for their validity and reliability;
3. Is designed to be unbiased and objective.
4. The word research is composed of two syllables, *re* and *search*. ‘**re**’ is a prefix meaning again, anew or over again ‘**search**’ is a verb meaning to examine closely and carefully, to test and try, or to probe.
5. Together they form a noun describing a careful, systematic, patient study and investigation in some field of knowledge, undertaken to establish facts or principles.
6. Research is a structured enquiry that utilizes acceptable scientific methodology to solve problems and create new knowledge that is generally applicable. Scientific methods consist of systematic observation, classification and interpretation of data.

- ❖ Research is, thus, an original contribution to the existing stock of knowledge making for its advancement. It is the pursuit of truth with the help of study, observation, comparison and experiment. In short, the search for knowledge through objective and systematic method of finding solution to a problem is research.
- ❖ Research is the systematic process of collecting and analysing information to increase our understanding of the phenomenon under study. It is the function of the researcher to contribute to the understanding of the phenomenon and to communicate that understanding to others.
- ❖ Research may be defined as a systematic approach/method consisting of enunciating the problem, formulating a hypothesis, collecting the facts or data, analysing the facts and reaching certain conclusions, either in the form of solution towards the concerned problem or in certain generalizations for some theoretical formulation. Research may also be defined as a scientific study, which by means of logical and systematized techniques, aims to :
 1. Discover new facts or verify and test old facts,
 2. Analyse their sequences, inter-relationships and explanations which are derived within an appropriate theoretical frame of reference,
 3. Develop new scientific tools, concepts, and theories which would facilitate reliable and valid study of human behaviour in decision making.

1.2.1 Nature or Characteristics of Research :

Research is a process of collecting, analysing and interpreting information to answer questions. But to qualify as research, the process must have certain characteristics: it must, as far as possible, be controlled, rigorous, systematic, valid and verifiable, empirical and critical.

1. **Controlled** - in real life there are many factors that affect an outcome. The concept of control implies that, in exploring causality in relation to two variables (factors), you set up your study in a way that minimizes the effects of other factors affecting the relationship. This can be achieved to a large extent in the physical sciences (cookery, bakery), as most of the research is done in a laboratory. However, in the social sciences (Hospitality and Tourism) it is extremely difficult as research is carried out on issues related to human beings living in society, where such controls are not possible. Therefore, in Hospitality and Tourism, as you cannot control external factors, you attempt to quantify their impact.
2. **Valid and verifiable** - this concept implies that whatever you conclude on the basis of your findings is correct and can be verified by you and others.
3. **Empirical** - this means that any conclusions drawn are based upon hard evidence gathered from information collected from real life experiences or observations.
4. **Critical** - critical scrutiny of the procedures used and the methods employed is crucial to a research enquiry. The process of investigation must be fool proof and free from drawbacks. The process adopted and the procedures used must be able to withstand critical scrutiny.
5. **Rigorous** - you must be scrupulous in ensuring that the procedures followed to find answers to questions are *relevant, appropriate and justified*. Again, the degree of rigor varies markedly between the physical and social sciences and within the social sciences.
6. **Systematic** - this implies that the procedure adopted to undertake an investigation follow a certain logical sequence. The different steps cannot be taken in a haphazard way. Some procedures must follow others.

Aim of Research :

The purpose of research is to discover answers to questions through the application of scientific procedures. The main aim of research is to find out the truth which is hidden and which has not been discovered as yet. Though each research study has its own specific purpose, we may think of research objectives as falling into a number of following broad groupings:

1. To gain familiarity with a phenomenon or to achieve new insights into it (Studies with this object in view are termed as exploratory or formulate research studies);
2. To portray accurately the characteristics of a particular individual, situation or a group (Studies with this object in view are known as descriptive research studies);
3. To determine the frequency with which something occurs or with which it is associated with something else (Studies with this object in view are known as diagnostic research studies);
4. To test a hypothesis of a causal relationship between variables (Such studies are known as hypothesis-testing research studies)

Types of Research :

The basic types of research are as follows:

1.2.1 Descriptive vs. Analytical :

Descriptive research includes surveys and fact-finding enquiries of different kinds. The major purpose of descriptive research is description of the state of affairs as it exists at present. In social science and business research we quite often use the term *ex post facto* research for descriptive research studies. The main characteristic of this method is that the researcher has no control over the variables; he can only report what has happened or what is happening. Most *ex post facto* research projects are used for descriptive studies in which the researcher seeks to measure such items as, for example, frequency of shopping, preferences of people, or similar data. *Ex post facto* studies also include attempts by researchers to discover causes even when they cannot control the variables. The methods of research utilized in descriptive research are survey methods of all kinds, including comparative and correlational methods. In analytical research, on the other hand, the researcher has to use facts or information already available, and analyse these to make a critical evaluation of the material.

Applied vs. Fundamental :

Research can either be applied (or action) research or fundamental (to basic or pure) research. Applied research aims at finding a solution for an immediate problem facing a society or an industrial/business organisation, whereas fundamental research is mainly concerned with generalisations and with the formulation of a theory. "Gathering knowledge for knowledge's sake is termed 'pure' or 'basic' research." Research concerning some natural phenomenon or relating to pure mathematics are examples of fundamental research. Similarly, research studies, concerning human behaviour carried on with view to make generalizations about human behaviour, are also examples of fundamental research, but research aimed at certain conclusions facing a concrete social or business problem is an example of applied research. Research to identify social, economic or political trends that may affect a particular institution or copy research or the marketing research are examples of applied research. Thus, the central aim of applied research is to discover a solution for some pressing practical problems. Whereas basic research is directed towards finding information

that has a broad base of applications and thus, adds to the already existing organized body of scientific knowledge.

Quantitative vs. Qualitative :

Quantitative research is based on the measurement of quantity or amount. It is applicable to phenomena that can be expressed in terms of quantity. Qualitative research, on the other hand, is concerned with qualitative phenomenon, i.e., phenomena relating to or involving quality or kind. For instance, when we are interested in investigating the reasons for human behaviour, we quite often talk of 'Motivation Research', an important type of qualitative research. This type of research aims at discovering the underlying motives and desires, using in-depth interviews for the purpose. Other techniques of such research are word association tests, sentence completion tests, story completion tests and similar other projective techniques. Attitude or opinion research, i.e., research designed to find out how people feel or what they think about a particular subject or institution is also qualitative research.

Qualitative research is especially important in the behavioural sciences where the aim is to discover the underlying motives of human behaviour. Through such research we can analyse the various factors which motivate people to behave in a particular manner or which make people like or dislike a particular thing. It may be stated, however, that to apply qualitative research in practice is relatively a difficult job and therefore, while doing such research, one should seek guidance from experimental psychologists.

Conceptual vs. Empirical :

Conceptual research is related to some abstract idea(s) or theory. It is generally used by philosophers and thinkers to develop new concepts or to reinterpret existing ones. On the other hand, empirical research relies on an experience or observation alone, often without due regard for system and theory. It is data based research, coming up with conclusions which are capable of being verified by observation or experiment. We can also call it as experimental type of research, in such research it is necessary to get at facts first-hand, at their source, and actively to go about doing certain things to stimulate the production of desired information. In such *research*, the *researcher* must first provide himself with a working hypothesis or guess as to the probable results. He then works to get enough facts (data) to prove or disprove his hypothesis. He then sets up experimental designs which he thinks will manipulate the persons or the materials concerned so as to bring forth the desired information.

Such research is thus characterised by the experimenter's control over the variables under study and his deliberate manipulation of one of them to study its effects. Empirical research is appropriate when proof is sought that certain variables affect other variables in some way. Evidence gathered through experiments or empirical studies is today considered to be the most powerful support possible for a given hypothesis

Some Other Types of Research: All other types of research are variations of one or more of the above stated approaches, based on either the purpose of research, or the time required to accomplish research, on the environment in which research is done, or on the basis of some other similar factor.

1. **One Time Research:** From the point of view of time, we can think of research either as one-time research or longitudinal research. In the former case the research is confined to a single time-period, whereas in the latter case the research is carried on over several time-periods.
2. **Laboratory Research:** Research can be *field-setting research* or *laboratory research* or *simulation research*, depending upon the environment in which it is to be carried

out. Research can as well be understood as *clinical or diagnostic research*. Such research follows case-study methods or in-depth approaches to reach the basic causal relations. Such studies usually go deep into the causes of things or events that interest us, using very small samples and very deep probing data gathering devices.

3. **Exploratory Research:** The research may be *exploratory* or it may be formalized. The objective of exploratory research is the development of hypotheses rather than their testing, whereas formalized research studies are those with substantial structure and with specific hypotheses to be tested.
4. **Historical Research:** *Historical research* is that which utilizes historical sources like documents, remains, etc., to study events or ideas of the past, including the philosophy of persons and groups at any remote point of time.
5. **Conclusion-oriented Research:** Research can also be classified as *conclusion-oriented* and decision-oriented. While doing conclusion-oriented research, a researcher is free to pick up a problem, redesign the enquiry as he proceeds and is prepared to conceptualize as he wishes. Decision-oriented research is always for the need of a decision maker and the researcher in this case is not free to embark upon research. According to his own inclination. Operations research is an example of decision oriented research since it is a scientific method of providing executive departments with a quantitative basis for decisions regarding operations under their control.

1.2.2 Purpose of Research :

The purpose of research is to discover answers to questions through the application of scientific procedures. The main aim of the research is to find out the truth which is hidden and which has not been discovered as yet. Though each research study has its own specific purpose, we may think of research objectives as falling into a number of following broad groupings:

1. To gain familiarity with a phenomenon or to achieve new insights into it (studies with this object in view are termed exploratory or formulate research studies);
2. To portray accurately the characteristics of a particular individual, situation, or group (studies with this object in view are known as descriptive research studies);
3. To determine the frequency with which something occurs or with which it is associated with something else (studies with this object in view are known as diagnostic research studies);
4. To test a hypothesis of a causal relationship between variables (such studies are known as hypothesis-testing research studies).

Motivation in Research :

What makes people undertake research? This is a question of fundamental importance. The possible motives for doing research may be either one or more of the following:

1. Desire to get a research degree along with its consequential benefits;
2. Desire to face the challenge in solving unsolved problems, i.e., concern over practical problems initiates' research;
3. Desire to get intellectual joy of doing some creative work;
4. Desire to be of service to society;
5. Desire to get respectability. However, this is not an exhaustive list of factors motivating people to undertake research studies.

Many more factors such as directives of government, employment conditions, curiosity about new things, desire to understand causal relationships, social thinking and awakening, and the like may as well motivate (or at times compel) people to perform research operations

1.3 CRITERIA OF GOOD RESEARCH :

Whatever may be the types of research works and studies, one thing that is important is that they all meet on the common ground of scientific method employed by them. One expects scientific research to satisfy the following criteria:

- ❖ The purpose of the research should be clearly defined and common concepts be used.
- ❖ The research procedure used should be described in sufficient detail to permit another researcher to repeat the research for further advancement, keeping the continuity of what has already been attained.
- ❖ The procedural design of the research should be carefully planned to yield results that are as objective as possible.
- ❖ The researcher should report with complete frankness, flaws in procedural design and estimate their effects upon the findings.
- ❖ The analysis of data should be sufficiently adequate to reveal its significance and the methods of analysis used should be appropriate. The validity and reliability of the data should be checked carefully.
- ❖ Conclusions should be confined to those justified by the data of the research and limited to those for which the data provide an adequate basis.
- ❖ Greater confidence in research is warranted if the researcher is experienced, has a good reputation in research and is a person of integrity.

In Other Words, We Can State the Qualities of A Good Research¹² as Under:

- i. **Good research is systematic** : It means that research is structured with specified steps to be taken in a specified sequence in accordance with the well-defined set of rules. Systematic characteristic of the research does not rule out creative thinking but it certainly does reject the use of guessing and intuition in arriving at conclusions.
- ii. **Good research is logical** : This implies that research is guided by the rules of logical reasoning and the logical process of induction and deduction are of great value in carrying out research. Induction is the process of reasoning from a part to the whole whereas deduction is the process of reasoning from some premise to a conclusion which follows from that very premise. In fact, logical reasoning makes research more meaningful in the context of decision making.
- iii. **Good research is empirical** : It implies that research is related basically to one or more aspects of a real situation and deals with concrete data that provides a basis for external validity to research results.
- iv. **Good research is replicable** : This characteristic allows research results to be verified by replicating the study and thereby building a sound basis for decisions.

Problems Encountered By Researchers in India :

Researchers in India, particularly those engaged in empirical research, are facing several problems. Some of the important problems are as follows:

The lack of a scientific training in the methodology of research is a great impediment for researchers in our country. There is paucity of competent researchers. Many researchers

take a leap in the dark without knowing research methods. Most of the work, which goes in the name of research, is not methodologically sound. Research to many researchers and even to their guides, is mostly a scissor and paste job without any insight shed on the collated materials. The consequence is obvious, viz., the research results, quite often, do not reflect the reality or realities. Thus, a systematic study of research methodology is an urgent necessity. Before undertaking research projects, researchers should be well equipped with all the methodological aspects. As such, efforts should be made to provide short duration intensive courses for meeting this requirement.

- ❖ There is insufficient interaction between the university research departments on one side and business establishments, government departments and research institutions on the other side. A great deal of primary data of non-confidential nature remains untouched/untreated by the researchers for want of proper contacts. Efforts should be made to develop satisfactory liaison among all concerned for better and realistic researches. There is need for developing some mechanisms of a university/industry interaction programme so that academics can get ideas from practitioners on what needs to be researched and practitioners can apply the research done by the academics.
- ❖ Most of the business units in our country do not have the confidence that the material supplied by them to researchers will not be misused and as such they are often reluctant in supplying the needed information to researchers. The concept of secrecy seems to be sacrosanct to business organisations in the country so much so that it proves an impermeable barrier to researchers. Thus, there is the need for generating the confidence that the information/data obtained from a business unit will not be misused.
- ❖ Research studies overlapping one another are undertaken quite often for want of adequate information. This results in duplication and fritters away resources. This problem can be solved by proper compilation and revision, at regular intervals, of a list of subjects on which and the places where the research is going on. Due attention should be given toward identification of research problems in various disciplines of applied science which are of immediate concern to the industries.
- ❖ There does not exist a code of conduct for researchers and inter-university and interdepartmental rivalries are also quite common. Hence, there is need for developing a code of conduct for researchers which, if adhered sincerely, can win over this problem.
- ❖ Many researchers in our country also face the difficulty of adequate and timely secretarial assistance, including computerial assistance. This causes unnecessary delays in the completion of research studies. All possible efforts are made in this direction so that efficient secretarial assistance is made available to researchers and that too well in time. University Grants Commission must play a dynamic role in solving this difficulty.
- ❖ Library management and functioning is not satisfactory at many places and much of the time and energy of researchers are spent in tracing out the books, journals, reports, etc., rather than in tracing out relevant material from them.
- ❖ There is also the problem that many of our libraries are not able to get copies of old and new Acts/Rules, reports and other government publications in time. This problem is felt more in libraries which are away in places from Delhi and/or the state capitals. Thus, efforts should be made for the regular and speedy supply of all governmental publications to reach our libraries.

- ❖ There is also the difficulty of timely availability of published data from various government and other agencies doing this job in our country. Researcher also faces the problem on account of the fact that the published data vary quite significantly because of differences in coverage by the concerning agencies.

There may, at times, take place the problem of conceptualization and also problems relating to the process of data collection and related things.

Research Methodology & Research Methods :

Research Methods : The Research methods may be defined as all those methods/techniques that are used for conducting the research. Research methods can be put into the following three groups

- ❖ In the first group, we include those methods which are concerned with the collection of data; these methods will be used where the data already available are not sufficient to arrive at the required solution.
- ❖ The second group consists of those Statistical Techniques which are used for establishing relationships between variables.
- ❖ The third group consists of those methods which are used to evaluate the accuracy of the results obtained.

Research Methodology : Research methodology may be defined as a way to systematically solve the research problem. Research methodology constitutes of research methods, selection criterion of research methods, used in context of research study and explanation of using of a particular method or technique and why other techniques are not used so that research results are capable of being evaluated either by researcher himself or by others.

1.4 SUMMARY :

Research originates in a decision process. Usually, research is said to begin with a question or a problem. In research process, management problem is converted into a research problem which is the major objective of the study. Research question is further subdivided, covering various facets of the problem that need to be solved. The role and scope of research has greatly increased in the field of business and economy as a whole. The study of research methods provides you with knowledge and skills you need to solve the problems and meet the challenges of today is modern pace of development.

1.5 SELF-ASSESSMENT QUESTIONS :

1. Define Research?
2. What are the objectives of Research?
3. State the significance of Research

1.6 REFERENCES /SUGGESTED READINGS :

1. Allen, T. Harrell, New Methods in Social Science Research, New York: Praeger Publishers, 1978
2. Bailey, Kenneth D., "Methods of Social Research," New York, 1978
3. Best, John W., and Kahn, James V., "Research in Education," 5th Ed., New Delhi: Prentice-Hall of India Pvt. Ltd., 1986
4. Burgess, Ernest W., "Research Methods in Sociology" in Georges Gurvitch and W.E. Moore (Ed.), Twentieth Century Sociology, New York: New York Philosophical Library, 1949.
5. C R Kotari, Research Methodology, VishwaPrakashan.

LESSON - 2

SCIENTIFIC METHOD

OBJECTIVES :

After reading this lesson, the reader should be able to gain an overview of understand, scope, and importance of the Scientific Method

- To define and understand Scientific Method
- To analyze the scope of the Scientific Method
- To get an idea and importance of the Scientific Method

STRUCTURE :

- 2.1 Introduction
- 2.2 Characteristics of Scientific Method
- 2.3 Induction and Deduction
- 2.4 Scope of the Scientific Method
- 2.5 Summary
- 2.6 Self-Assessment Questions
- 2.7 Reference and Suggested Readings

2.1 INTRODUCTION TO SCIENTIFIC METHOD :

To be termed scientific, a method of inquiry must be based on empirical and measurable evidence subject to specific principles of reasoning. The “scientific method” attempts to minimize the influence of the researchers’ bias on the outcome of an experiment. The researcher may have a preference for one outcome or another, and it is important that this preference not bias the results or their interpretation. Sometimes “common sense” and “logic” tempt us into believing that no test is needed. Another common mistake is to ignore or rule out data which do not support the hypothesis.

The scientific method is the process by which scientists, collectively and over time, endeavor to construct an accurate (that is, reliable, consistent and non-arbitrary) representation of the world.

Recognizing that personal and cultural beliefs influence both our perceptions and our interpretations of natural phenomena, we aim through the use of standard procedures and criteria to minimize those influences when developing a theory. As a famous scientist once said, “Smart people (like smart lawyers) can come up with very good explanations for mistaken points of view.” In summary, the scientific method attempts to minimize the influence of bias or prejudice in the experimenter when testing a hypothesis or a theory.

Clover and Basely define “scientific method is a systematic step by step procedure following the logical process of reasoning.”

Difference between Scientific and non-scientific Method :

The scientific approach to knowledge is empirical. The empirical approach emphasizes direct observation and experimentation as a way of answering questions. Intuition can play a role in idea formation, but eventually the scientist is guided by what direct observation and experimentation reveal to be true. Their findings are often counterintuitive.

When observing phenomena, a scientist likes to exert a specific level of control. When utilizing control, scientists investigate the effects of various factors one by one. A key goal for the scientist is to gain a clearer picture of those factors that actually produce a phenomenon. It has been suggested that tight control is the key feature of science. Non-scientific approaches to knowledge are often made unsystematically and with little care. The non-scientific approach does not attempt to control many factors that could affect the events they are observing (don't hold conditions constant). This lack of control makes it difficult to determine cause-and-effect relationships (too many confounds, unintended independent variable).

How can two people witness the same event but see different things? This often occurs due to personal biases and subjective impressions. These characteristics are common traits among nonscientists. Their reports often go beyond what has just been observed and involve speculation. Scientific reporting attempts to be objective and unbiased. One way to lessen the chance of biased reporting is checking to see if other independent observers report the same findings. Even when using this checkpoint the possibility of bias is still present. Following strict guidelines to prevent bias reporting decreases the chances of it occurring. Although I would say 100% unbiased reports rarely, if ever, occur.

It is not unusual for people in everyday conversation to discuss concepts they really don't understand. Many subjects are discussed on a routine basis even though neither party knows exactly what the subject means. They may have an idea of what they are discussing (even though their ideas may be totally opposite). Although they cannot precisely define the concepts they are talking about. The scientist attaches an operational definition (a definition based on the set of operations that produced the thing defined) to concepts. An example of an operational definition follows: hunger a physiological need for food; the consequence of food deprivation. Once an operational definition has been established communication can move forward.

2.2 CHARACTERISTICS OF SCIENTIFIC METHOD :

- ❖ **Empirical** : Scientific method is concerned with the realities that are observable through “sensory experiences.” It generates knowledge which is verifiable by experience or observation. Some of the realities could be directly observed, like the number of students present in the class and how many of them are male and how many female. The same students have attitudes, values, motivations, aspirations, and commitments. These are also realities which cannot be observed directly, but the researchers have designed ways to observe these indirectly. Any reality that cannot be put to “sensory experience” directly or indirectly (existence of heaven, the Day of Judgment, life hereafter, God's rewards for good deeds) does not fall within the domain of scientific method.
- ❖ **Verifiable** : Observations made through scientific method are to be verified again by using the senses to confirm or refute the previous findings. Such confirmations may have to be made by the same researcher or others. We will place more faith and credence in those findings and conclusions if similar findings emerge on the basis of data collected by other researchers using the same methods. To the extent that it does happen (i.e. the results are replicated or repeated) we will gain confidence in the scientific nature of our research. Replicability, in this way, is an important characteristic of scientific method. Hence revelations and intuitions are out of the domain of scientific method.

- ❖ **Cumulative** : Prior to the start of any study the researchers try to scan through the literature and see that their study is not a repetition in ignorance. Instead of reinventing the wheel the researchers take stock of the existing body of knowledge and try to build on it. Also, the researchers do not leave their research findings into scattered bits and pieces. Facts and figures are to be provided with language and thereby inferences drawn. The results are to be organized and systematized. Nevertheless, we don't want to leave our studies as standalone. A linkage between the present and the previous body of knowledge has to be established, and that is how the knowledge accumulates. Every new crop of babies does not have to start from a scratch; the existing body of knowledge provides a huge foundation on which these researchers build on and hence the knowledge keeps on growing.
- ❖ **Deterministic** : Science is based on the assumption that all events have antecedent causes that are subject to identification and logical understanding. For the scientist, nothing "just happens" – it happens for a reason. The scientific researchers try to explain the emerging phenomenon by identifying its causes. Of the identified causes which ones can be the most important? For example, in the 2006 BA/BSC examination of the Mumbai University 67 per cent of the students failed. What could be the determinants of such a mass failure of students? The researcher may try to explain this phenomenon and come up with variety of reasons which may pertain to students, teachers, administration, curriculum, books, examination system, and so on. Looking into such a large number of reasons may be highly cumbersome model for problem solution. It might be appropriate to tell, of all these factors which one is the most important, the second most important, the third most important, which two in combination are the most important. The researcher tries to narrow down the number of reasons in such a way that some action could be taken. Therefore, the achievement of a meaningful, rather than an elaborate and cumbersome, model for problem solution becomes a critical issue in research. That is parsimony which implies the explanation with the minimum number of variables that are responsible for an undesirable situation.
- ❖ **Ethical and Ideological Neutrality** : The conclusions drawn through interpretation of the results to data analysis should be objective; that is, they should be based on the facts of the findings derived from actual data, and not on our own subjective or emotional values. For instance, if we had a hypothesis that stated that greater participation in decision making will increase organizational commitment, and this was not supported by the results, it makes no sense if the researcher continues to argue that increased opportunities for employee participation would still help. Such an argument would be based, not on the factual, data based research findings, but on the subjective opinion of the researcher. If this was the conviction of the researcher all along, then there was no need to do the research in the first place. Researchers are human beings, having individual ideologies, religious affiliations, cultural differences which can influence the research findings. Any interference of their personal likings and dis-likings in their research can contaminate the purity of the data, which ultimately can affect the predictions made by the researcher. Therefore, one of the important characteristics of scientific method is to follow the principle of objectivity, uphold neutrality, and present the results in an unbiased manner.

2.3 INDUCTION AND DEDUCTION :

From the Oxford English Dictionary (OED); to induce (in relation to science and logic) means "to derive by reasoning, to lead to something as a conclusion, or inference, to suggest or imply," and induction "as the process of inferring a general law or principle from observation of particular instances." Another version is the "adducing (pulling together) of a number of separate facts, particulars, etc. especially for the purpose of proving a general

statement.” E. Mayr in his *Growth of Biologic Thought* offers this definition: “inductivism claims that (we) can arrive at objective unbiased conclusions only by...recording, measuring, and describing what we encounter without any root hypothesis...”

Meaning : It is one of the scientific methods. It follows the logical reasoning process. It is a process of reasoning whereby the researcher arrives at universal generalizations from particular facts. In other words, this method involves studying several individual cases and drawing a generalization. Thus, it involves two elements, i.e., observation and generalization. Conclusions drawn from induction and tentative inferences and they are subject to further confirmation based on more evidence.

Essential Conditions :

According to Clover and Basely, four conditions are essential for valid induction. These are:

1. **Accuracy:** Observations must be correctly performed and recorded, and data collected should be accurate. Errors in observations, experiments or interviews and faulty recording of the information can affect the conclusions drawn.
2. **Representative:** Observations must cover representative cases drawn from a specific universe. For instance, to conduct a survey on brand loyalty among young females for a particular product category, then there should be proper representation of young females comprising of college going females, young working females, graduates, non-graduates, and other representations. Appropriate sampling technique can be used for this purpose, such as stratified random sampling.
3. **Appropriate Sample Size:** Observations must cover an adequate number of cases. The sample size must be large enough to make it representative so as to obtain reliable results.
4. **Proper Conclusions:** Conclusions must be confined to inferences drawn from the findings. The conclusions must be drawn only after proper analysis of the collected data relating to a particular study. They should not be generalized to apply to cases not covered in the sample. For instance, if the study is conducted to know the brand loyalty of young females for a particular product, then the conclusions drawn from such a study cannot be applied to young males, or elderly ladies, and so on, as they are not covered in the sample.

Merits :

- Induction provides the universal premise and is helpful in finding out the material truth.
- It points out the relativity of generalizations. It indicates that a particular generalization is valid in certain situation.
- It is scientific in nature. Some of the important laws or principles of physical and social sciences have been developed through the use of inductive approach.
- Inductive generalizations are precise and accurate as they are based on observed facts and realistic foundations.

Demerits :

- The collection of data for induction is a complex job. This method requires high degree of competence on the part of the researcher.
- Induction is not useful without deduction. Without deduction, induction produces only a mass of unrelated facts.
- It is a time consuming and expensive process; as a lot of time, effort and money is required to collect and analyze a large volume of data to arrive at generalization.

Deduction :

- ❖ The OED definition of to deduce is “to show or hold a thing to be derived from etc...” or “to draw as a conclusion from something known or assumed, to infer”; deduction thus is “inference by reasoning from generals to particulars,” or “the process of deducing from something known or assumed...”
- ❖ Deductive reasoning is a basic form of valid reasoning. Deductive reasoning, or deduction, starts out with a general statement, or hypothesis, and examines the possibilities to reach a specific, logical conclusion. The scientific method uses deduction to test hypotheses and theories.
- ❖ In deductive reasoning, if something is true of a class of things in general, it is also true for all members of that class. For example, “All men are mortal. Harold is a man. Therefore, Harold is mortal.” For deductive reasoning to be sound, the hypothesis must be correct. It is assumed that the premises, “All men are mortal” and “Harold is a man” are true. Therefore, the conclusion is logical and true.
- ❖ It’s possible to come to a logical conclusion even if the generalization is not true. If the generalization is wrong, the conclusion may be logical, but it may also be untrue. For example, the argument, “All bald men are grandfathers. Harold is bald. Therefore, Harold is a grandfather,” is valid logically but it is untrue because the original statement is false.

Essential conditions :

The following are the essential conditions for valid deduction:

- The general premise or assumption must be correct. IF the general premise is correct, then the conclusion drawn can be right.
- The general premise must be applied only to the cases that properly come under it.

Merits :

Deduction is relatively simple and less time consuming as compared to inductive method.

- It is precise and accurate in generalization as it makes use of logic and mathematical tools of analysis.
- Its social sciences, where there is limited scope for experimentation, this method becomes the only method for the development of generalizations.

Demerits :

- It may result in inaccurate generalizations, if the premise are incorrect or partially correct.
- This method is abstract. Therefore, if a large amount of abstraction were used in generalizations, then such generalizations would be useless.

2.4 SCOPE OF SCIENTIFIC METHODS :

Social science research has a vast scope in respect of areas of application. The social science research can be useful in a number of areas such as:

Economic Planning : Social science research can be of immense use in economic planning in a given society. Economy planning requires basic data on the various aspects of our society and economy, resource endowment and the needs, hopes and problems of the people, etc. Economic planning is undertaken to achieve certain objectives such as:

1. To bring about regional development.
2. To make optimum use of available resources.
3. To bring out self-reliance.

4. To generate employment, etc.

Systematic research provides the required data for planning and developing various schemes or programs such as employment generation programs, rural development programs, etc.

- ❖ **Control over Social Phenomena** : Through social science research, first-hand information can be obtained in respect of the working of institutions and organization, which in turn provides greater power of control over the social phenomena. The social science research has practical implications for formal and informal styles of managing, organization structures, and introduction of changes in the organization.
- ❖ **Social Welfare** : Social research can be used to collect the required data on different aspects of social life in a given society, so as to develop social welfare programs. For instance, in a developing country like India, there are various social welfare problems such as low literacy, law and order problems due to caste, religion, and other conflicts, social evils like child marriages, abuse of women, and so on. Therefore, to overcome social problems the Government and other organizations can collect relevant data through systematic research, and accordingly develop various social welfare programs, such as family welfare campaigns, literacy programs, women and children welfare programs, etc.
- ❖ **Helps to Solve Problems** : Research can be undertaken to find solutions to solve specific problems. For instance, an organization may initiate research to find solution to the problem of declining sales of their products in the market. An educational institution can undertake research to find out the causes of low attendance or poor results. A government organization may undertake research to solve the problem or to ascertain the impact of slums on the quality of life in a particular city, and such other research activities. The research enables to find appropriate solutions to specific problems which in turn helps to improve the quality of performance in various organizations or institutions.
- ❖ **Verifies and Tests Existing Laws** : Research may be undertaken to verify and test existing laws or theories. Such verification and testing of existing theories help to improve the knowledge and ability to handle situations and events. This is true when the existing theories may not be sufficient or relevant to handle certain situations and events, and therefore, through research, improvements or modifications can be made in the existing laws or theories.
- ❖ **Develops New Tools and Theories** : Research helps to develop new tools, concepts and theories for a better study of an unknown phenomenon. For this purpose, exploratory research is undertaken to achieve new insights into such phenomenon.
- ❖ **Helps to Predict Events**: Research may be undertaken to predict future course of events. For instance, research may be undertaken to find out the impact of growing unemployment of educated youth on the social life of the society in future. The findings of such research would not only indicate the possible impact, but would also make the concerned authorities to take appropriate measures to reduce unemployment, to reduce the growth of population and to overcome the negative consequences, as and when they take place.
- ❖ **Extends Knowledge** : Researchers undertake research to extend the existing knowledge in physical sciences (such as physics, chemistry, mathematics, etc.).As well as in social sciences (like sociology, management, psychology) etc. The knowledge can be enhanced by undertaking research in general and by fundamental research in particular.

2.5 SUMMARY :

The scientific method, as applied to social sciences, includes a variety of research approaches, tools, and techniques, for collecting and analysing qualitative or quantitative data. These methods include laboratory experiments, field surveys, case research, ethnographic research, action research, and so forth. Much of this book is devoted to learning about these different methods. However, recognize that the scientific method operates primarily at the empirical level of research, i.e., how to make observations and analyse these observations. Very little of this method is directly pertinent to the theoretical level, which is really the more challenging part of scientific research.

2.6 SELF-ASSESSMENT QUESTIONS :

1. What are the characteristics of the Scientific method?
2. Explain the need of the Scientific Method?
3. Describe the scope of the Scientific Method?
4. What do your meaning by Scientific Method? Distinguish between induction and deduction?
5. What are the merits and demerits of Scientific Method?

2.7 REFERENCES /SUGGESTED READINGS :

1. Allen, T. Harrell, *New Methods in Social Science Research*, New York: Praeger Publishers, 1978
2. Bailey, Kenneth D., "Methods of Social Research," New York, 1978
3. Best, John W., and Kahn, James V., "Research in Education," 5th Ed., New Delhi: Prentice-Hall of India Pvt. Ltd., 1986
4. Burgess, Ernest W., "Research Methods in Sociology" in Georges Gurvitch and W.E. Moore (Ed.), *Twentieth Century Sociology*, New York: New York Philosophical Library, 1949.
5. C R Kotari, *Research Methodology*, VishwaPrakashan.

Dr.PadmaPonugoti

LESSON - 3

RESEARCH PROBLEM

OBJECTIVES :

- ❖ The importance of formulating a research problem
- ❖ Sources of research problems
- ❖ Considerations in selecting a research problem
- ❖ Specific issues to consider when formulating a research problem in qualitative research
- ❖ Steps in formulating a research problem
- ❖ How to formulate research objectives

STRUCTURE :

- 3.1 Introduction
- 3.2 Definitions of the Research Problem
 - 3.2.1 The importance of formulating a research problem
 - 3.2.2 Sources of research problems
 - 3.2.3 Research problem defined
 - 3.2.4 Characteristics of a Problem
- 3.3 The necessity of defining a research problem
 - 3.3.1 Statement of the problem
 - 3.3.2 Identification of a Problem
 - 3.3.3 Criteria for selection of the problem
 - 3.3.4 Considerations in selecting a research problem
- 3.4 Summary
- 3.5 Self-assessment Questions
- 3.6 Reference and Suggested readings

3.1 INTRODUCTION :

The central aim of this chapter is to detail the process of formulating a research problem, even though the specific process that you are likely to adopt depends upon:

- your expertise in research methodology;
- Your knowledge of the subject area;
- Your understanding of the issues to be examined;
- The extent to which the focus of your study is predetermined.
- If you are not very familiar with the research process and/or do not have a very specific idea about what is to be researched, you need to follow every step detailed in this chapter.

However, more experienced researchers can take a number of shortcuts. The process outlined here assumes that you have neither the required knowledge of the process of formulating a research problem nor a specific idea about what is to be researched. If you have a specific idea for the basis of your inquiry, you do not need to go through this chapter. However, you should make sure that your idea is researchable as not all problems lend themselves to research methodologies.

The research problem broadly speaking, any question that you want to answer and any assumption or assertion that you want to challenge or investigate can become a research problem or a research topic for your study. However, it is important to remember that not all questions can be transformed into research problems and some may prove to be extremely difficult to study. According to Powers, Meaghan and Twoomey (1985: 38), 'Potential research questions may occur to us on a regular basis, but the process of formulating them in a meaningful way is not at all an easy task.' As a newcomer, it might seem easy to formulate a problem but it requires considerable knowledge of both the subject area and research methodology. Once you examine a question more closely you will soon realise the complexity of formulating an idea into a problem which is researchable. 'First identifying and then specifying a research problem might seem like research tasks that ought to be easy and quickly accomplished. However, such is often not the case (Yegidis&Weinback 1991: 35). It is essential for the problem you formulate to be able to withstand scrutiny in terms of the procedures required to be undertaken. Hence you should spend considerable time thinking it through.

3.2 DEFINITIONS OF THE RESEARCH PROBLEM :

The obstacles which hinder our path are regarded as problem. Different definitions of the problem are given below;

- ❖ "Problem is the obstacle in the path of satisfying our needs." **John Geoffery**
- ❖ "Problem is a question which is to be solved." **John. G. Tornsand**
- ❖ "To define a problem means to put a fence around it, to separate it by careful distinctions from like questions found in related situations of need." **Whitney**
- ❖ "A problem is a question proposed for a solution generally speaking a problem exists when there is no available answer to the same question." **J.C. Townsend**
- ❖ "A problem is an interrogative sentence or statement that asks: What relation exists between two or more variables?" **F.N. Kerlinger**
- ❖ "To define a problem means to specify it in detail and with precision, each question and subordinate question to be answered is to be specified, and the limits of the investigation must be determined. Frequently, it is necessary to review previous studies in order to determine just what is to be done. Sometimes it is necessary to formulate the point of view or educational theory on which the investigation is to be based. If certain assumptions are made, they must be explicitly noted." **Monero and Engelhart**

3.2.1 The importance of formulating a research problem :

The formulation of a research problem is the first and most important step of the research process. It is like the identification of a destination before undertaking a journey. In the absence of a destination, it is impossible to identify the shortest – or indeed any – route. Similarly, in the absence of a clear research problem, a clear and economical plan is impossible. To use another analogy, a research problem is like the foundation of a building. The type and design of the building are dependent upon the foundation. If the foundation is well designed and strong you can expect the building to be also. The research problem serves as the foundation of a research study: if it is well formulated, you can expect a good study to follow. According to Kerlinger:

If one wants to solve a problem, one must generally know what the problem is. It can be said that a large part of the problem lies in knowing what one is trying to do.

You must have a clear idea with regard to what it is that you want to find out about and not what you think you must find.

A research problem may take a number of forms, from the very simple to the very complex. The way you formulate a problem determines almost every step that follows: the type of study design that can be used; the type of sampling strategy that can be employed; the research instrument that can be used or developed; and the type of analysis that can be undertaken. Suppose your broad area of interest is depression. Further, suppose you want to conduct a research study regarding services available to patients with depression living in a community. If your focus is to find out the types of services available to patients with depression, the study will dominantly be descriptive and qualitative in nature.

These types of studies fall in the category of qualitative research and are carried out using qualitative research methodologies. On the other hand, if you want to find out the extent of use of these services, that is the number of people using them, it will dominantly use quantitative methodologies even though it is descriptive in nature describing the number of people using a service. If your focus is to determine the extent of use in relation to the personal attributes of the patients, the study will be classified as correlational (and quantitative). The methodology used will be different from the one used in the case of a descriptive study. Similarly, if your aim is to find out the effectiveness of these services, the study will again be classified as correlational and the study design used, methods of collecting data and its analysis will be a part of the quantitative methodology. Hence, it is important for you to understand that the way you formulate a research problem determines all the subsequent steps that you have to follow during your research journey.

The formulation of a problem is like the ‘input’ to a study, and the ‘output’ – the quality of the contents of the research report and the validity of the associations or causation established – is entirely dependent upon it. Hence the famous saying about computers, ‘garbage in, garbage out’, is equally applicable to a research problem.

Initially, you may become more confused but this is normal and a sign of progression. Remember: confusion is often but a first step towards clarity. Take time over formulating your problem, for the clearer you are about your research problem/question, the easier it will be for you later on. Remember, this is the most crucial step.

3.2.2 Sources of research problems :

This section is of particular relevance if you have not yet selected a research topic and do not know where to start. If you have already selected your topic or question, go to the next section.

Most research in the humanities revolves around four Ps:

1. People;
2. Problems;
3. Programmes;
4. Phenomena.

In fact, a closer look at any academic or occupational field will show that most research revolves around these four Ps. The emphasis on a particular ‘p’ may vary from study to study but generally, in practice, most research studies are based upon at least a combination of two ps. You may select a group of individuals (a group of individuals – or a community as such – ‘people’), to examine the existence of certain issues or problems

relating to their lives, to ascertain their attitude towards an issue ('problem'), to establish the existence of a regularity ('phenomenon') or to evaluate the effectiveness of an intervention ('program'). Your focus may be the study of an issue, an association, or a phenomenon per se; for example, the relationship between unemployment and street crime, smoking and cancer, or fertility and mortality, which is done on the basis of information collected from individuals, groups, communities or organizations. The emphasis in these studies is on exploring, discovering, or establishing associations or causation. Similarly, you can study different aspects of a programme: its effectiveness, its structure, the need for it, consumers' satisfaction with it, and so on. In order to ascertain these, you collect information from people.

Every research study has two aspects: the people provide you with the 'study population', whereas the other three ps furnish the 'subject areas. Your study population – individuals, groups, and communities – is the people from whom the information is collected. Your subject area is a problem, programme or phenomenon about which the information is collected.

Examine your own academic discipline or professional field in the context of the four ps in order to identify anything that looks interesting.

The concept of the 4Ps is applicable to both quantitative and qualitative research though the main difference at this stage is the extent of their specificity, dissection, precision and focus. In qualitative research, these attributes are deliberately kept very loose so that you can explore more as you go along, in case you find something of relevance. You do not bind yourself with constraints that would put limits on your ability to explore. There is a separate section on 'Formulating a research problem in qualitative research' later in the chapter, which provides further guidance on the process.

3.2.3 RESEARCH PROBLEM DEFINED :

A research problem is some difficulty either of a theoretical or practical nature that an individual or organization faces and wishes to obtain a solution for the same. A research problem must contain the following

1. An individual or an organization that has the problem.
2. They must occupy some environment/condition to which the difficulty pertains.
3. Some objective/goal to be attained.
4. Some alternative courses of action through which these objectives can be attained.
5. Researcher must have some doubts regarding the selection of possible alternatives.

A research problem is a situation that causes the researcher to feel apprehensive, confused, and will at ease. It is the demarcation of a problem area within a certain context involving the 'who or 'what ', the 'where', the 'when', and the 'why ' of the problem situation.

There are many problem situations that may give rise to research. Three sources usually contribute to problem identification. Own experience or the experience of others may be a source of problem supply. A second source could be scientific literature. You may read about certain findings and notice that a certain field was not covered. This could lead to a research problem. Theories could be a third source. Shortcomings in theories could be researched.

Research can thus be aimed at clarifying or substantiating an existing theory, clarifying contradictory findings, correcting a faulty methodology, correcting the inadequate or

unsuitable use of statistical techniques, reconciling conflicting opinions, or solving existing practical problems.

3.2.4 Characteristics of a Problem :

Although selecting a research problem is one of the most difficult steps for a student in a research process, it is unfortunately one for which the least guidance can be given.

A problem statement must have the following characteristics:

1. It should ask about a relationship between two or more variables.
2. It should be stated clearly and unambiguously, usually in question form.
3. It should be possible to collect data or answer the questions asked.
4. It should not represent a moral or ethical position.

Relationship between Variables :

In this kind of problem, the researcher manipulates a minimum of one variable to determine its effects on other variables, as opposed to a purely descriptive study in which the researcher observes, counts, or in some way measures the frequency of appearance of a particular variable in a particular setting. For example, how many students in school have I.Q.s in excess of 120?

Since no attempts need be made to deal with a relationship between variables, this problem requires only a “book-keeping” procedure, if however, the problem was worded; Are boys more likely than girls to have I.Q.’s in excess of 120 then it would involve the relationship between variables?

The Problem is stated in Question Form :

- ❖ The problem should be in question form as:
- ❖ What is the relationship between I.Q. and achievement?
- ❖ Do students learn more from a directive teacher or a nondirective teacher?
- ❖ Is there a relationship between racial background and dropout rate?
- ❖ Do more students continue in training programmes offering stipends or in programmes not offering stipends?
- ❖ What is the relationship between role learning ability and socio-economic status?

Empirical Testability :

A problem should be testable by empirical methods, that is, through the collection of data. Moreover, for a student’s purposes, it should lend itself to study by a single researcher, on a limited budget, within a year. The nature of the variables included in the problem is a good clue to its testability. An example of the kind of problem that is wise to avoid it: Does an extended experience in communal living improve a person’s outlook on life? In addition to the magnitude and probable duration of studying such a problem, the variable themselves would be difficult to manipulate or measure.

Avoidance of Moral or Ethical Judgements :

Questions about ideals or values are often more difficult to study than questions about aptitudes or performance, as, those that would be difficult to test are: Should men define their feelings? Ethical consideration should also be taken into consideration in defining or stating a problem.

3.3 THE NECESSITY OF DEFINING A RESEARCH PROBLEM :

It is important to formulate a research problem properly. In fact, the problem formulation is even more essential than its prospective solution. A carefully defined research problem does not let a researcher stray from the research path that should be followed. It is, therefore, concluded that only upon a detailed definition of the research problem, the researcher can progress with the design of the research methodology. This also leads to smoother progress on all the subsequent steps that are involved in completing a research project.

3.3.1 Statement of the problem :

The statement of the problem involves the demarcation and formulation of the problem, i.e. the 'who/what, 'where ', 'when ', and 'what '. It usually includes the statement of the hypothesis.

A problem statement is a concise description of the issues that need to be addressed by a problem-solving team and should be presented to them (or created by them) before they try to solve the problem. When bringing together a team to achieve a particular purpose provide them with a problem statement. A good problem statement should answer these questions:

- What is the problem? This should explain why the team is needed.
- Who has the problem or who is the client/customer? This should explain who needs the solution and who will decide whether the problem has been solved.
- What form can the resolution be? What are the scope and limitations (in time, money, resources, and technologies) that can be used to solve the problem? Does the client want white paper? A web-tool? A new feature for a product? Brainstorming on a topic? The primary purpose of a problem statement is to focus the attention of the problem-solving team.
- However, if the focus of the problem is too narrow or the scope of the solution too limited the creativity and innovation of the solution can be stifling. In project management, the problem statement is part of the project charter. It lists what's essential about the project and enables the project manager to identify the project scope as well as the project stakeholders.

3.3.2 Identification of a Problem :

The identification and analysing a research problem is the first and most crucial step of the research process. A problem cannot be solved effectively unless a researcher possesses the intellect and insight to isolate and understand the specific factors giving rise to the difficulty.

The present research scholars understand that identification of a problem means selecting a topic of research or statement of the problem. It is wrong to think so. A topic or statement of the problem and research problem are not synonymous but they are inclusive. The problem concerns the functioning of the broader area of the field studied whereas a topic or title or statement of the problem is the verbal statement of the problem. The topic is the definition of the problem that delimits or pinpoints the task of a researcher.

It is the usual practice of researchers they select the topic of the study from different sources, especially from research abstracts. They do not identify the problem, but a problem is made on the basis of the topic. It results that the researcher has no involvement in his research activities. Whatever they do, do mechanically.

Since identifying the exact nature and dimensions of a problem is of major importance in research work, it is very essential that an investigator should learn how to recognize and define a problem. He should proceed step by step in locating the research problem. The following steps are to be followed in identifying a research problem:

Step 1: Determining the field of research in which a researcher is keen to do the research work.

Step 2: The researcher should develop mastery in the area or it should be the field of his specialization.

Step 3: He should review the research conducted in the area to know the recent trend and studies in the area.

Step 4: On the basis of the review, he should consider the priority field of the study.

Step 5: He should draw an analogy and insight in identifying a problem or employ his personal experience of the field in locating the problem. He may take the help of a supervisor or expertise of the field.

Step 6: He should pinpoint specific aspects of the problem which is to be investigated.

For example, a researcher wants to work in the field of teacher education which is the field of his interest. He has deep insight and mastery over the area. On the basis of review and his personal experience, the researcher perceives a problem in the field of teacher-education programmes that training institutions and colleges of education are not able to produce effective teachers although a large number of such institutions have been opened after independence. This problem has several dimensions but these can be studied simultaneously. The researcher further visualizes that the potential candidates are not admitted to this program.

A question arises : Are the potential candidates admitted to our teacher education problem? Thus, the procedure of identification of a problem can be shown with the help of a paradigm.

3.3.3 Criteria for selection of the problem :

The factors to be considered in the selection of a research problem are both the criteria external and personal. External criteria have to do with such matters as novelty and importance for the field availability of data and method, and institutional or administrative cooperation. Personal criteria involved such considerations as interest, training, cost and time, etc.

Thus, the criteria for the selection of the problem suggested by Good and Scates are as follows:

- Novelty and avoidance of unnecessary duplications.
- Importance for the field represented and implementation
- Interest, intellectual curiosity, and drive.
- Training and personal qualifications.
- Availability of data and method.
- Special equipment and working conditions.
- Approachability of the sample.
- Sponsorship and administrative cooperation.
- Hazards, penalties, and handicaps.

- Cost and returns.
- Time factor.

- ❖ **Novelty and avoidance of unnecessary duplication** : The question of novelty or newness is not merely one of duplication of earlier investigations. It involves the regency of the data summarized especially in the case of survey studies made during a period of great Economic, Educational, and Social Change.
- ❖ **Importance for the field represented and implementation** : This criterion of importance in the choice of a problem involves such matters as significance for the field involved, timelines, and practical value in terms of application and implementation of the results. Scientific research in Education, psychology, and social sciences, in general, has an especially urgent obligation to play a social role in rendering service to society and humanity.
- ❖ **Interest, intellectual curiosity, and drive** : One of the personal motives of research most frequently mentioned by scientists themselves is pure curiosity, accompanied by genuine interest and a driven satisfaction and enjoyment.
- ❖ **Availability of data and method** : The data under consideration must meet certain standards of accuracy, objectivity, and verifiability.
- ❖ **Special equipment and working conditions** : The major purpose of the equipment is to define the process of observation to provide control of conditions and accuracy or permanence of recording.
- ❖ **Sponsorship and administrative cooperation** : It is a common practice for the thesis to be sponsored by a faculty adviser in whose area of specialization the problem lies.
- ❖ **Costs and returns** : The candidate must consider carefully his own financial resources in the light of such facilities and assistance as can be provided by the institution.
- ❖ **Time factor** : As a general rule, the minimum amount of graduate work for the Master's degree is one year and for the Doctor's degree three years

The grouping of the research problem :

- ❖ The need to communicate what will be studied in clear, concise, and unambiguous terms
- ❖ One or more sentences indicating the goal, purpose, or overall direction of the study
- ❖ General characteristics
 - Implies the possibility of empirical investigation
 - Identifies a need for the research
 - Provides focus
 - Provides a concise overview of the research
- ❖ Two ways of stating the problem
 - Research problems: typically, a rather general overview of the problem with just enough information about the scope and purpose of the study to provide an initial understanding of the research
 - Research statements and/or questions: more specific, focused statements and questions that communicate in greater detail the nature of the study
- ❖ Researchable and non-researchable problems
 - Researchable problems imply the possibility of empirical investigation
 - non-researchable problems include explanations of how to do something, vague propositions, and value-based concerns
- ❖ Comparing quantitative and qualitative research problems

- Quantitative problems
- Qualitative problems
- ❖ Sources of research problems
 - Casual observation
 - Deductions from Theory
 - Related literature
 - Current social and political issues
 - Practical situations
 - Personal interests and experience
 - Replication of previous studies
 - Clarification of contradictory research results

Quantitative research problems :

Three types of questions

- ❖ Descriptive
- ❖ Relational
- ❖ Causal
- ❖ Identifies specifically the type of research, the variables and relationships between them, and the subjects
- ❖ Variables
 - ❖ A variable is a label or name that represents a concept or characteristic that varies (e.g., gender, weight, achievement, attitudes toward inclusion, etc.)
 - Conceptual and operational definitions of variables
 - Conceptual (i.e., constitutive) definition uses words or concepts to define a variable
 - Operational definition is an indication of the meaning of a variable through the specification of the manner by which it is measured, categorized, or controlled
 - ❖ Types of variables
 - Three variable labels defined by the context within which the variable is discussed
 - ❖ Independent and dependent variables
 - ❖ Extraneous and confounding variables
 - ❖ Continuous and categorical variables
 - Independent and dependent (i.e., cause and effect)
 - ❖ Independent variables act as the "cause" in that they precede, influence, and predict the dependent variable
 - ❖ Dependent variables act as the effect in that they change as a result of being influenced by an independent variable
 - ❖ Examples
 - ❖ Some situations do not lend themselves to the use of the terms independent or dependent because it is difficult to discuss them in causal terms

Extraneous and confounding variables

- Extraneous variables are those that affect the dependent variable but are not controlled adequately by the researcher
- Confounding variables are those that vary systematically with the independent variable and exert influence on the dependent variable

Continuous and categorical

- Continuous variables are measured on a scale that theoretically can take on an infinite number of values

- Categorical variables are measured and assigned to groups on the basis of specific characteristics
- Continuous variables can be converted to categorical variables, but categorical variables cannot be converted to continuous variables.

3.3.4 Considerations in selecting a research problem :

When selecting a research problem/topic there are a number of considerations to keep in mind which will help to ensure that your study will be manageable and that you remain motivated. These considerations are:

- ❖ **Interest** – Interest should be the most important consideration in selecting a research problem. A research endeavour is usually time-consuming and involves hard work and possibly unforeseen problems. If you select a topic that does not greatly interest you, it could become extremely difficult to sustain the required motivation and put in enough time and energy to complete it.
- ❖ **Magnitude** – You should have sufficient knowledge about the research process to be able to visualize the work involved in completing the proposed study. Narrow the topic down to something manageable, specific, and clear. It is extremely important to select a topic that you can manage within the time and with the resources at your disposal. Even if you are undertaking a descriptive study, you need to consider its magnitude carefully.
- ❖ **Measurement of concepts** – If you are using a concept in your study (in quantitative studies), make sure you are clear about its indicators and their measurement. For example, if you plan to measure the effectiveness of a health promotion program, you must be clear as to what determines the effectiveness and how it will be measured. Do not use concepts in your research problem that you are not sure how to measure. This does not mean you cannot develop a measurement procedure as the study progresses. While most of the developmental work will be done during your study, it is imperative that you are reasonably clear about the measurement of these concepts at this stage.
- ❖ **Level of expertise** – Make sure you have an adequate level of expertise for the task you are proposing. Allow for the fact that you will learn during the study and may receive help from your research supervisor and others, but remember that you need to do most of the work yourself.
- ❖ **Relevance** – Select a topic that is of relevance to you as a professional. Ensure that your study adds to the existing body of knowledge, bridges current gaps or is useful in policy formulation. This will help you to sustain interest in the study.
- ❖ **Availability of data** – If your topic entails the collection of information from secondary sources (office records, client records, census or other already-published reports, etc.) make sure that this data is available and in the format, you want before finalizing your topic.
- ❖ **Ethical issues** – Another important consideration in formulating a research problem is the ethical issues involved. In the course of conducting a research study, the study population may be adversely affected by some of the questions (directly or indirectly); deprived of an intervention; expected to share sensitive and private information; or expected to be simply experimental ‘guinea pigs. How ethical issues can affect the study population and how ethical problems can be overcome should be thoroughly examined at the problem-formulation stage.

3.4 SUMMARY :

To select a research problem we may depend on various sources: our personal experiences, the existing theories in the field concerned, etc. A research problem should contribute to the existing body of knowledge and lead to further research. It should be testable and feasible

3.5 SELF-ASSESSMENT QUESTIONS :

A. Write short notes on the following

- i. Meaning of research problem
- ii. Problem of defining a research problem
- iii. Formulation of a research problem
- iv. Potential to be a research problem

B. Short answer type questions

- i. Define the term research problem and identify three characteristics of good research problems?
- ii. Differentiate research problems from research problem statements and/or questions?
- iii. Identify four common sources of research problems?
- iv. Describe the characteristics of quantitative research problems?
- v. State the criteria for evaluating quantitative research problems and evaluate specific problems using these criteria?
- vi. Describe the characteristics of a qualitative research problem?

3.6 REFERENCE AND SUGGESTED READINGS :

1. C R Kothari, Research Methodology, VishwaPrakashan.
2. Cooper and Schindler, Business Research Methods, TMH.
3. David Luck and Ronald Rubin, Marketing Research, PHI.
4. Naresh Amphora, Marketing Research, Pearson Education.
5. S. N. Murthy and U. Bhojanna, Business Research Methods, 3rd Edition, Excel Books.

Dr. Padma Ponugoti

LESSON - 4

ELEMENTS OF RESEARCH

OBJECTIVES :

After reading this lesson, the reader should be able to understand an overview of the basic elements of research. Key objectives of this lesson are:

- To define the scientific research and its application in media and social science research;
- To know the characteristics of scientific research
- To understand the significance of Digital Media Research
- To learn the basic elements of research such as concept, construct, variables and measurement;
- To know various levels and types of measurement and its usage in the research;
- To comprehend the importance of causation in research and *a cause-and-effect* relationship between variables;

STRUCTURE :

- 4.1 Introduction
- 4.2. What does research mean?
 - 4.2.1. Definitions of scientific research
- 4.3. Significance of Digital Media Research
- 4.4. Characteristics of scientific research
- 4.5. Basic elements of research
 - 4.5.1. Concept
 - 4.5.2. Construct
 - 4.5.3. Variables
 - 4.5.4. Measurement
- 4.7 Levels and type of measurement
- 4.8 Summary
- 4.9 Glossary
- 4.10 Self-Assessment Questions
- 4.11 Suggested Reading

4.1. INTRODUCTION :

When you see the world around you with the spirit of enquiry, you may observe certain facts which are known to you and some are hidden. We can say that research is the quest for the unknown, and also about the known. It often starts with observation and is done to reveal the secrets of this universe. It may be an observational study of natural phenomena or a rational study of the relations between various things and concepts in which these phenomena are expressed. The public always try to know what is happening and how was happened and what are the facts. For them, research is an honest, intelligent and exhaustive quest for facts. It is used to help raise awareness of issues that are happening day to day in our life. It gives you a solid foundation on which you can build your ideas and opinions. Research entails going beyond the raw data and sharing real-life stories that have a more personal impact on people.

4.2. WHAT DOES RESEARCH MEAN? :

- Research is an attempt to discover something.
- It literally means looking back.
- To some means an attitude of enquiry, but to others, research is an aspect of scientific quest for knowledge.
- It is a method to find the truth, and also a method to think critically.
- Research is a quest for ordered or systematic knowledge.
- Research is the process of finding an acceptable knowledge, which is gathered, organized and systematized.
- It is tested and validated with the help of research tools.
- It is the process of arriving at solutions of problems through planned collection, analysis and interpretation of data.
- It is a systematic way of collecting, classifying and analyzing information, either quantitative or qualitative.
- The process of research opens up new opportunities for learning and growth.
- Research refers to a systematic investigation carried out to discover new knowledge, expand existing knowledge, solve practical problems, and develop new things.
- Scholars from the social science defined research as a systematized effort to gain new knowledge.
- Research is anything that is original contribution to the existing stock of knowledge.

4.2.1. Definition of Scientific Research :

Research is a careful investigation and critical in finding the facts and the principles of an investigation to establish something very clever, (Webster's New International). For Clifford Woody, research comprises defining and redefining problems, formulating hypothesis or suggested solutions; collecting, organising and evaluating data; making deductions and reaching conclusions; and at last carefully testing the conclusions to determine whether they fit the formulating hypothesis, (Kothari, C.R., 2004: p1). D. Slesinger and M. Stephenson (1930) defined research as "The manipulation of things, concepts or symbols for the purpose of generalizing to extend, correct or verify knowledge, whether that knowledge aids in construction of theory or in the practice of an art." The authors have a different view of research as they suggest that it can be taken up by modifying, challenging and changing; available knowledge either to prove a process being appropriate or to develop it in entirety. Hillway (1969) defined research as "a method of study through the careful and exhaustive investigation of all ascertainable evidence bearing upon a definable problem, we reach a solution to that problem". Kerlinger (1986) defined scientific research as "a systematic, controlled, empirical and critical investigation of hypothetical propositions about presumed relations among observed phenomenon".

4.3. SIGNIFICANCE OF DIGITAL MEDIA RESEARCH :

Digital media research is a methodological approach that incorporates internet-based data which includes other communicative and social media platforms such as Facebook, Instagram, Vine, Twitter, jiffy, Periscope etc. Digital media methods also extend to database research, data generated by sensors, drones and autonomous automobiles. Contemporary research engaging digital media methods is built upon the 'computational turn' where 'computational approaches is increasingly reflected across a number of disciplines, including the arts, humanities and social sciences, which use technologies to shift the critical ground of their concepts and theories'. Digital media research is increasingly being mediated through

digital technology affecting both the epistemologies and ontologies that underlie a research program. According to Jonathon Hutchinson (2016), the significant points of departure for digital media methods in the media and communication discipline are: 1.The increasing need for typologies and ontologies in social media research; 2.The significance of mapping public issues; and 3.The difficulties researchers face as text-based communication shifts to visually oriented platforms.

4.4 CHARACTERISTICS OF SCIENTIFIC RESEARCH :

When we talk about the science, there are three primary characters we could find: exactness, regularity in behavior and verifiability through experiments. The same are applicable in any scientific research which inculcates:

- Scientific inductive thinking
- Promotes logic of thought
- Solves problems and aid to economic policy
- Aid to all government organization policies
- Study social relationship
- Seek answers to social problems
- For new ideas and insights
- Generalization new themes

Following are the main characteristics of scientific research:

- ❖ **Public:**Scientific research always for the cause of development and advancement of the public. Scientific advancement depends on freely available information. It must be verifiable
- ❖ **Objective:** It should be depended on facts rather than interpretation. Independent observers have to use rules for classifying behavior.
- ❖ **Empirical:**Researcher should be able to perceive and classify what they study. They should reject metaphysical and nonsensical explanations of events.
- ❖ **Systematic and cumulative:**By following the review of literature, scientists follow a set of rules. A theory is a set of related propositions that presents a systematic view of phenomena by specifying relationships among concepts.
- ❖ **Predictive:**Scientific research is concerned with relating the present to the future.

4.5 ELEMENTS OF RESEARCH :

In this section, we discuss four basic elements of research: concepts, constructs, measurement and variables.

4.5.1. Concepts :

A concept is a term that expresses an abstract idea formed by generalizing from particulars and summarizing related observations. Concepts are mental representations and are typically based on experience. Concepts can be of real phenomena (dogs, clouds, pain). Concepts can be of agree-upon phenomena (truth, beauty, justice, prejudice, value, etc.).

Example : Researcher might observe that a public speaker becomes restless, starts to perspire, and fidgets with a pencil just before giving a speech. The researcher might summarize these observed patterns of behavior and label them “speech anxiety.” Concepts in mass media research include terms such as advertising effectiveness, message length, media usage, and readability.

Why concept is importance? : Concept is important for two reasons.

- ❖ Simplify the research process by combining particular characteristics, objects, or people into general categories. *Example:* Families that own computers, modems, MP3 players, cell phones, DVD or Blue Ray machines. To make it easier to describe these families, the researcher calls them taffies and categorizes them under the concept of technologically advanced families.
- ❖ Concepts simplify communication among those who have a shared understanding of them. Researcher uses concepts to organize their observations into meaningful summaries and to transmit this information to others. *Example.* Agenda setting to describe a complicated set of audience and media activities find that their colleagues understand what is being discussed.

4.5.2. Construct :

A construct is a concept that has three distinct characteristics:

- ❖ It is an abstract idea that is usually broken down into dimensions represented by lower- level concepts.
- ❖ Because of its abstraction, a construct usually can't be observed directly.
- ❖ A construct is usually designed for a specific research purpose so that its exact meaning relates only to the context in which it is found.

Example 1: The construct “involvement” has been used in many advertising studies such as the Internet for “advertisement involvement”. It is difficult to see advertisement involvement directly. Advertisement involvement includes concepts of attention, interest, and arousal.

Example 2 : Authoritarianism: it represents a construct defined to describe a certain type of personality; it involves nine different concepts, including conventionalism, submission, superstition, and cynicism. Authoritarianism itself cannot be seen, so some type of questionnaire or standardized test is used to determine its presence. Both concepts and constructs are valuable tools in theoretical research.

4.5.3. Variables :

The purpose of all research is to describe and explain variance in the world. Variance is simply the difference. Variation may occur naturally in the world or change that we create as a result of a manipulation. Within the context of a research investigation, concepts are generally referred to as variables. A variable is, as the name applies, something that varies. Variables can represent any type of data, such as names, sounds, texts, numbers, arrays, images, or any item or class of objects supported by the computer language.

For example, variables are names that are given to the variance we wish to explain. “Any concept, which is made measurable by a definition, is a variable. A variable explains the difference between two concepts.” A variable is either a result of some force or is itself the force that causes a change in another variable. Distinction between types of variables depends on the purpose of the research.

But certain abstract concepts cannot be measured. For example: Beauty, Love, and Ego. That means, something that can assume different values. In this case, it is a concept to which numbers are assigned and that changes in value. For example: TV viewing is a variable with values that range from 0- 24 hours each day. When we measure a construct in a way that gives varying values, we call the measure variable. Simple we can say that an empirical counterpart of a concept or construct is calling a variable.

Variables are important to understand because they are the basic units of the information studied and interpreted in research studies. Researchers carefully analyze and interpret the value(s) of each variable to make sense of how things relate to each other in a descriptive study or what has happened in an experiment. Variables are classified in terms of their relationship with one another.

Causation and correlation : Causation means that a change in one variable causes a change in another variable. There is a cause-and-effect relationship between variables. The two variables are correlated with each other and there is also a causal link between them. Whereas, correlation describes an association between types of variables.

When one variable changes, so does the other. A correlation is a statistical indicator of the relationship between variables. These variables change together. But this covariation (correlated variation) isn't necessarily due to a direct or indirect causal link. A correlation doesn't imply causation, but causation always implies correlation.

Types of variables :

There are many types of variables that we can see in research such as: Independent and Dependent Variable, Qualitative and Quantitative Variables, Discrete and Continuous Variable, Background and Moderating Variable, Extraneous and intervening Variables and Suppressor Variable etc. We will discuss some important variable in this chapter.

Dependent and Independent Variable : Independent Variables are systematically varied by the researcher. Dependent Variable is observed and their values are presumed to depend on the effects (influence) of the independent variables. Dependent Variable is what the researcher wishes to explain.

Example 1 : A researcher is interested to determine how the angle of a camera shot affects an audience's perception of the credibility of a newscaster. Three versions of a newscast are recorded; one shot from a very low angle, another from a high angle and a third from eye level.

Groups of subjects are randomly assigned to view one of the three versions. They are asked to complete the questionnaire to measure the newscaster's credibility. Camera angle is independent variable. Dependent variable is perceived credibility of the newscaster as measured by the questionnaire. Here, credibility will vary according to the camera angle.

Example 2 : When a researcher shows a video of violence to one group of people and a video of non-violence, to another group of people, the *independent variable* is the showing video. Each person's response to the violent video and non-violent video is called the *dependent variable*. This could be many things depending upon what the video is for, such as to test the attitudes, emotions and feelings. Therefore, in media experiments, a researcher manipulates an independent variable to determine if it causes a change in the dependent variable.

In descriptive studies, all the variables are dependent variables because they are studied in relation to all the other variables that exist in the setting where the research is taking place. However, in descriptive studies, variables are not discussed using the terms "independent" or "dependent." Instead, the names of the variables are used when discussing the study. Example: there is more diabetes in people of Native Indian heritage than people who come from other countries. In a descriptive study, the researcher would examine how diabetes (a variable) is related to a person's genetic heritage (another variable).

Qualitative and Quantitative Variables : Quantitative Variables refer to as "numeric" variables, these are variables that represent a measurable quantity. Examples: number of TV

channels in a state, number of hours of radio listening etc. Similarly, **Qualitative Variables** refers to as “categorical” variables, these are variables that take on names or labels and can fit into categories. Example: Gender in media houses (Male and Female).

Discrete and Continuous Variable : Discrete variables are countable in a finite amount of time. For example, you can count the change in your pocket and money in your bank account. You could also count the amount of money in *everyone's* bank accounts. It might take you a long time to count that last item, but the point is—it's still countable. **Continuous Variables** would take forever to count. In fact, you would get to “forever” and never finish counting them. For example, age. You can't count “age”. Because it would literally take forever. You could be: 25 years, 10 months, 2 days by today. It will change after few days/months/years.

4.5.4 Measurement :

Measurement is the process observing and recording the observations that are collected as part of a research effort. Measurement is a procedure for assigning symbols, letters, or numbers to empirical properties of variables according to rules. Numerals are labels that have no inherent meaning, for example, in drivers' license numbers, zip codes, or social security numbers.

Numbers are numerals that have quantitative meaning and are amenable to statistical analysis, for example, age, height, or weight. Rules for assigning labels to properties of variables are the most important component of measurement, because poor rules can make the outcome meaningless. It is difficult to measure concepts directly, e.g., "work effort," so what are usually measured are indicators of concepts, such as speed, repetitiveness, etc.

We generally consider four broad categories of measurements. Survey research includes the design and implementation of interviews and questionnaires. Scaling involves consideration of the major methods of developing and implementing a scale. Qualitative research provides an overview of the broad range of non-numerical measurement approaches. And unobtrusive measures present a variety of measurement methods that don't intrude on or interfere with the context of the research.

4.6 LEVELS AND TYPES OF MEASUREMENT :

There are different levels of measurement. These levels differ as to how closely they approach the structure of the number system we use. It is important to understand the level of measurement of variables in research, because it determines the type of statistical analysis that can be conducted, and, therefore, the type of conclusions that can be drawn from the research.

We will discuss various levels of measurements :

- Nominal,
- Ordinal,
- Interval,
- Ratio Level.

Nominal Level : This is the most basic level of measurement. it is essentially *labeling*. It can only establish whether two observations are alike or different. For example: sorting a into two piles - red cards and black cards. A nominal level of measurement uses symbols to classify observations into categories that must be both mutually exclusive and exhaustive. Exhaustive means that there must be enough categories that all the observations will fall into some

category. Mutually exclusive means that the categories must be distinct enough that no observations will fall into more than one category.

For example : the sex of the journalist is determined by a question on the questionnaire. Observations were sorted into two mutually exclusive and exhaustive categories, male and female. Observations could be labeled with the letters M and F, or the numerals 0 and 1. The variable of marital status may be measured by two categories, married and unmarried.

But these must each be defined so that all possible observations will fit into one category but no more than one: legally married, common-law marriage, religious marriage, civil marriage, living together, never married, divorced, informally separated, legally separated, widowed, abandoned, annulled, etc. In nominal measurement, all observations in one category are alike on some property, and they differ from the objects in the other category (or categories) on that property (e.g. sex-male, female). There is no ordering of categories (no category is better or worse, or more or less than another).

Ordinal Level : An ordinal level of measurement uses *symbols* to classify observations into categories that are not only mutually exclusive and exhaustive; in addition, the categories have some explicit relationship among them. For example, observations may be classified into categories such as taller and shorter, greater and lesser, faster and slower, harder and easier etc. However, each observation must still fall into one of the categories (the categories are exhaustive) but no more than one (the categories are mutually exclusive). The military uses ranks as symbols to distinguish categories of soldiers.

Most of the commonly used questions which ask about job satisfaction use the ordinal level of measurement. For example, asking whether one is very satisfied, satisfied, neutral, dissatisfied, or very dissatisfied with one's job is using an ordinal scale of measurement.

Interval Level : An interval level of measurement classifies observations into categories that are not only mutually exclusive and exhaustive, and have some explicit relationship among them, but the relationship between the categories is known and exact. This is the first quantitative application of numbers. In the interval level, a common and constant unit of measurement has been established between the categories. For example, the commonly used measures of temperature are interval level scales. We know that a temperature of 75 degrees is one degree warmer than a temperature of 74 degrees, just as a temperature of 42 degrees is one degree warmer than a temperature of 41 degrees.

Numbers may be assigned to the observations because the relationship between the categories is assumed to be the same as the relationship between numbers in the number system. For example, $74+1=75$ and $41+1=42$. The intervals between categories are equal, but they originate from some arbitrary origin. That is, there is no meaningful zero point on an interval scale.

Ratio Level : The ratio level of measurement is the same as the interval level, with the addition of a meaningful zero point. There is a meaningful and non-arbitrary zero point from which the equal intervals between categories originate. For example, weight, area, speed, and velocity are measured on a ratio level scale.

In public policy and administration, budgets and the number of program participants are measured on ratio scales. In many cases, interval and ratio scales are treated alike in terms of the statistical tests that are applied. Variables measured at a higher level can always be converted to a lower level, but not vice versa. For example, observations of actual age (ratio

scale) can be converted to categories of older and younger (ordinal scale), but age measured as simply older or younger cannot be converted to measures of actual age.

Measurement Problems : Commonly encountered problems include a misplaced belief in precision. It is not usually necessary, for example, to measure annual income in dollars and cents. Another problem is measures that go against social conventions. It is often easier to ask people to check of categories than to supply specific information, for example, with regard to age, income, education, etc. It is a trade-off between gathering higher-level (interval or ratio) data and having a higher questionnaire completion rate (less missing data).

A third problem is when the operational definition does not correspond to the conceptual definition. It may be easier to measure the number of students suspended from school than to measure the concept of school violence. A fourth problem is when the researcher becomes addicted to certain statistics, and gathers only data measured at the level appropriate for those statistical formulas.

4.7 SUMMARY :

Scientific research is an academic activity and scientific methods used in the research attempt to achieve the ideal goal of finding the facts on specific topics by experimentation, observation, logical arguments from accepted postulates and a combination of these three in varying proportions.

As a researcher you should understand the meaning, conceptual definitions and the basic elements of research. Media research connects you to the society either directly or indirectly in adding new knowledge to the existing knowledge and acquiring new knowledge. As a researcher, you must pay all possible attention in order to understand the basic elements of research.

4.8. GLOSSARY :

- ❖ **Research:** Research is the process of finding an acceptable knowledge, which is gathered, organized and systematized.
- ❖ **Concept:** Research is the process of finding an acceptable knowledge, which is gathered, organized and systematized.
- ❖ **Construct:** A construct is the combination of various concepts. It has three distinct characteristics; an abstract idea, can't be observed directly and designed for a specific research purpose.
- ❖ **Variable:** Any concept, which is made measurable by a definition is a variable. A variable explains the difference between two concepts
- ❖ **Measurement:** Measurement is the process observing and recording the observations that are collected as part of a research effort. It is a procedure for assigning symbols, letters, or numbers to empirical properties of variables according to rules.
- ❖ **Causation:** A change in one variable causes a change in another variable. There is a cause-and-effect relationship between variables.
- ❖ **Correlation:** A correlation is a statistical indicator of the relationship between variables
- ❖ **Digital media research:** It is a methodological approach that incorporates internet-based data which includes social media platforms, websites, web-pages etc.

4.9 SELF-ASSESSMENT QUESTIONS :

- 1) What does research mean?
- 2) What are the characteristics of scientific research?
- 3) What are the elements of research?
- 4) What do you know about the digital media research?
- 5) What are the levels and types of measurement?

4.10. REFERENCE AND SUGGESTED READINGS :

1. Anders Hansen et al. (1998). Mass communication research methods. Macmillan
2. Anselm L. Strauss & Juliet M. Corbin.(1998). Basics of qualitative research; techniques and procedures for developing grounded theory, (2nd Revised Edition). Sage
3. Arthur Asa Berger. (2000). Media and communication research methods: An introduction to qualitative and quantitative approaches. Sage
4. Dennis McQuail. (1997). Audience analysis. Sage
5. Denzin, N.K., & Lincoln, Y.S. (Eds.).(2000). Handbook of qualitative research. Sage
6. Hillway (1969). *Handbook of educational research: A guide to methods and materials*. Mifflin Co., Houghton, Boston.
7. Jonathon Hutchinson (2016). An introduction to digital media research methods: how to research and the implications of new media data. Communication Research and Practice. Issue 1 (2), Retrieved from <https://www.tandfonline.com/doi/full>
8. Klaus Bruhn Jensen. (2002). A handbook of media and communication research. Routledge
9. Kothari, C.R. (2004). *Research Methodology: Methods and Techniques*. New Delhi: New age International Publishers.
10. Roger D. Wimmer & Joseph R. Dominick. (2005). Mass media research: An introduction: Wadsworth Pub Co.
11. Slesinger, D. and Stephenson, M. (1930) The Encyclopaedia of Social Sciences. Vol. IX, MacMillan Publications.

Dr. K. John Babu

LESSON - 5

HYPOTHESIS AND HYPOTHESIS TESTING

OBJECTIVES :

After reading this lesson, the reader should be able to understand an overview of the hypothesis and testing hypothesis. Key objectives of this lesson are:

- To define the hypothesis and its application in social science research;
- To comprehend the importance of hypothesis and Statistical Significance in social science/communication research;
- To know the types of hypothesis errors;
- To learn the characteristics of good hypothesis;
- To distinguish various types of hypothesis and hypothesis testing;
- To understand how to formulate and test the hypothesis;

STRUCTURE :

- 5.1 Introduction to research hypothesis
- 5.2 Definitions and characteristics of Hypothesis
- 5.3 Advantages of Hypotheses
- 5.4 Types of Hypothesis
 - 5.4.1 Simple hypothesis
 - 5.4.2 Complex hypothesis
 - 5.4.3 Null hypothesis
 - 5.4.4 Alternative hypothesis
- 5.5 Hypothesis Testing
- 5.6 Hypothesis errors
- 5.7 Importance of Statistical Significance
- 5.8 Types of Hypothesis Testing
 - 5.8.1 Normality
 - 5.8.2 T-test
 - 5.8.3 Chi-Square Test
 - 5.8.4 ANOVA
 - 5.8.5 Z-test
- 5.9 Formulation of Hypothesis
- 5.10 Summary
- 5.11 Glossary
- 5.12 Self-Assessment Questions
- 5.13 Suggested Reading

5.1 INTRODUCTION TO RESEARCH HYPOTHESIS :

Mass media researchers, most of the time start their studies based on existing theory and are able to make predictions about the outcome of their work. They begin with some tentative generalization regarding a relationship between two or more variables. These generalizations may take two forms: research questions and statistical hypotheses. The two are identical except for the aspect of prediction—hypotheses predict an outcome, whereas

research questions do not. Tukey (1986) said that hypotheses ask, Do we have firm evidence that is happening/has happened?

Hypothesis is a tentative solution to the topic that enables a researcher to kickstart the research process. Hypotheses suggest relationships that can be empirically verified (Singh 2006). The rationale generally comes from past theoretical explanation presented in the literature. Hypothesis plays an important role in the development of theory. It focuses the scope of the study and also suggests the empirical observations to be undertaken. Feldman (2006) was used in preparing some of the statements of hypothesis.

1. Watching too many soap operas on television creates antisocial behavior.
2. People who read newspapers wash their hands more frequently than those who do not read newspapers.
3. Einstein's theory of relativity would not have been developed if he had access to television.

5.2 DEFINITIONS AND CHARACTERISTICS OF HYPOTHESIS :

Hypothesis is a tentative generalization about the relationship between two or more variables that predicts an outcome. According to Bailey (1978) "A hypothesis is a proposition in testable form and predicts a particular relationship between two or more variables". Whereas, Grinnell (1988) points out, a hypothesis is written in such a way that it can be proven or disproven by valid and reliable data.

Following characteristics construct hypothesis:

- It should be logically consistent
- it should be stated concisely
- It should be testable.
- It should be of verification
- It should be derived from theory
- It should be precise, simple and specific
- It should be compatible with current knowledge in the area

5.3 ADVANTAGES OF HYPOTHESES :

There are a variety of benefits to the researcher with hypotheses like direct the researcher, saves time, quantification of variables and help rule out intervening and confounding variables.

- ❖ **Provide direction for a study** : Hypothesis development is usually the conclusion of a rigorous literature review and emerges as a natural step in the research process. Without hypotheses, research lacks focus and clarity.
- ❖ **Eliminate wastage of time** : Hypothesis helps the research to narrow down to the topic without any confusion. It eliminates the wasteful of time and energy.
- ❖ **Allow for quantification of variables** : All terms used in hypotheses must have an operational definition. Therefore, researcher has to give adequate operational definition to every term or concept or phenomenon in research. The well-defined concept or phenomenon can be quantified easily.

5.4 TYPES OF HYPOTHESIS :

Chris Drew (2022 October 31) noted 13 different types of hypothesis. These include simple, complex, null, alternative, composite, directional, non-directional, logical, empirical, statistical, associative, exact, and inexact. Few important types are:

5.4.1 Simple hypothesis :

A simple hypothesis is a hypothesis that predicts a correlation between two test variables: an independent and a dependent variable. Example: on average, people over 50 like Television more than people fewer than 50. Here, expected correlation between the dependent variable and the independent variable is clear.

5.4.2 Complex hypothesis :

A complex hypothesis is a hypothesis that contains multiple variables, making the hypothesis more specific but also harder to prove. You can have multiple independent and dependent variables in this hypothesis. Example: 1) Age and (2) weight increase your chances of getting (3) diabetes and (4) heart disease.

5.4.3 Null hypothesis :

The null hypothesis is the claim that no relationship exists between two sets of data or variables being analyzed. The null hypothesis is that any experimentally observed difference is due to chance alone, and an underlying causative relationship does not exist, hence the term "null". A null hypothesis predicts that there will be no significant relationship between the two test variables. For example, you can say that "The study will show that there is no correlation between marriage and happiness." Example: There is no significant relationship between watching TV serials and domestic violence (H_0). There is no statistically significant difference in political views between carpenters and plumbers.

5.4.4 Alternative hypothesis :

The hypothesis of difference is called alternate hypotheses. It is opposite of null hypothesis and is designated by H_1 or H_A . Hypothesis is may be directional which connote an expected direction the relationship or difference between the variables. Example: For the same work, male earns higher wages than female (H_1). There is an impact of radio listening on old age people.

5.5 HYPOTHESIS TESTING :

Sir Ronald A. Fisher developed the concept of significance testing in 1925 and formulated tables based on the areas under the normal curve defined by .01 and .05. Fisher called his statistical hypothesis testing as "tests of significance", in his publication *Statistical Methods for Research Workers*. Modern significance testing is largely the product of Karl Pearson (p-value, Pearson's chi-squared test), William Sealy Gosset (Student's t-distribution), and Ronald Fisher (null hypothesis, analysis of variance, significance test). Hypothesis testing was developed by Jerzy Neyman and Egon Pearson, the son of Karl, (Raymond Hubbard and Bavaria, 2013).

Hypothesis testing in statistics refers to analyzing an assumption about a population parameter and allows us to make probabilistic statements. It is a form of statistical inference that uses data from a sample to draw conclusions about a population parameter or a population probability distribution. With the use of sample data, hypothesis testing makes an assumption about how true the assumption is for the entire population from where the sample is being taken. Statistical hypothesis test is a method of statistical inference used to decide whether the data at hand sufficiently support a particular hypothesis and to determine a possible conclusion from two different, and likely conflicting, hypotheses.

It can be understood as quantitative results for a qualitative problem. It is used to make an educated guess about an assumption using statistics. It offers researcher a statistical

approach about the theoretical assumptions a researcher made. Hypothesis testing provides various techniques to test the hypothesis statement depending upon the variable and the data points. It finds its use in almost every field of research while answering statements such as whether the new AI tool will work, a new testing method is appropriate, or if the outcomes of a random experiment are probable or not.

For example, radio station selects the music program it plays based on the assumption that the average age of its listening audience is 30 years. To determine whether this assumption is valid, a hypothesis test could be conducted with the null hypothesis ($H_0: \mu = 30$) and the alternative hypothesis given as $H_a: \mu \neq 30$. Based on a sample of individuals from the listening audience, the sample mean age, \bar{x} , can be computed and used to determine whether there is sufficient statistical evidence to reject H_0 . Conceptually, a value of the sample mean that is “close” to 30 is consistent with the null hypothesis, while a value of the sample mean that is “not close” to 30 provides support for the alternative hypothesis. What is considered “close” and “not close” is determined by using the sampling distribution of \bar{x} , (Encyclopedia Britannica).

What is p-value? :

- The p-value or probability value tells you the statistical significance of a finding.
- A p-value is a statistical measurement used to validate a hypothesis against observed data.
- A p-value measures the probability of obtaining the observed results, assuming that the null hypothesis is true.
- The lower the p-value, the greater the statistical significance of the observed difference.
- A p-value of 0.05 or lower is generally considered statistically significant.
- P-value can serve as an alternative to—or in addition to—preselected confidence levels for hypothesis testing.

It is common practice in mass media research studies to set the probability level at .01 or .05, which means that either one or five times out of 100 significant results of the study occur because of random error or chance. There is no logical reason for using these figures but the practice has been followed for many years. In many research areas, researchers set the significance level according to the purpose of the study rather than by general convention.

5.6 HYPOTHESIS ERRORS :

There is always the possibility of making an error in rejecting or failing to reject a null hypothesis. Under these circumstances, researchers must evaluate the various consequences of making a Type I or a Type II error. There are two types of errors generally commits while testing the hypothesis such as type I and type II error. Type I error occurs when we reject a true hypothesis while as Type II error occurs when we accept a wrong hypothesis.

Balancing Type I and Type II Error :

There are different consequences for different decisions. Setting a significance level at .0001 will virtually eliminate Type I error but will dramatically increase the odds of Type II error. Suppose a researcher is testing the efficacy of a new TV Channel. Which result is more harmful: making a Type I error and claiming that a drug works when it does not, or making a Type II error and overlooking a drug that might actually work.

5.7 IMPORTANCE OF STATISTICAL SIGNIFICANCE :

Salkind (2007) offers the following considerations when evaluating the importance of statistical significance:

- Significance is not very meaningful unless the study has a sound conceptual base that lends meaning to the results.
- Significance cannot be interpreted independently of the context in which it occurs.
- Significance should not be the end all of all research. If a study is designed correctly, failing to reject the null hypothesis might be an important finding.

5.8 TYPES OF HYPOTHESIS TESTING :

Based on the data provided, there are several types of hypothesis testing tools use in research methodology. Depending on the sample size and the data given, we choose among different hypothesis testing methodologies.

5.8.1 Normality :

This type of testing is used for normal distribution in a population sample. If the data points are grouped around the mean, the probability of them being above or below the mean is equally likely. Its shape resembles a bell curve that is equally distributed on either side of the mean.

5.8.2 T-test :

This test is used when the sample size in a normally distributed population is comparatively small, and the standard deviation is unknown. Usually, if the sample size drops below 30, we use a T-test to find the confidence intervals of the population.

5.8.3 Chi-Square Test :

The Chi-Square test is used to test the population variance against the known or assumed value of the population variance. It is also a better choice to test the goodness of fit of a distribution of data. The two most common Chi-Square tests are the Chi-Square test of independence and the chi-square test of variance.

5.8.4 ANOVA :

Analysis of Variance or ANOVA compares the data sets of two different populations or samples. It is similar in its use to the t-test or the Z-test, but it allows us to compare more than two sample means. ANOVA allows us to test the significance between an independent variable and a dependent variable, namely X and Y, respectively.

5.8.5 Z-test :

It is a statistical measure to test that the means of two population samples are different when their variance is known. For a Z-test, the population is assumed to be normally distributed. A z-test is better suited in the case of large sample sizes greater than 30. This is due to the central limit theorem that as the sample size increases, the samples are considered to be distributed normally.

5.9 FORMULATION OF HYPOTHESIS :

Following is the general process to formulate the hypothesis and find the P value:

- ❖ **Identify the correct test statistic :** All hypothesis tests boil your sample data down to a single number known as a test statistic. T-tests use t-values. F-tests use F-values.

Chi-square tests use chi-square values. Choosing the correct one depends on the type of data you have and how you want to analyze it. Before you can find the p value, you must determine which hypothesis test and test statistic you'll use. Test statistics assess how consistent your sample data are with the null hypothesis. As a test statistic becomes more extreme, it indicates a larger difference between your sample data and the null hypothesis.

- ❖ **Calculate the test statistic using the relevant properties of your sample** :How you calculate the test statistic depends on which one you're using. Unsurprisingly, the method for calculating test statistics varies by test type. Consequently, to calculate the p value for any test, you'll need to know the correct test statistic formula.
- ❖ **Specify the characteristics of the test statistic's sampling distribution** : Sampling distributions are a type of probability distribution. Consequently, they allow you to calculate probabilities related to your test statistic's extremeness, which lets us find the p value. The same sampling distribution can have a variety of shapes depending upon its parameters. For this step, you need to determine the characteristics of the sampling distribution that fit your design and data.
- ❖ **Place your test statistic in the sampling distribution to find the p value** : After calculating the test statistic and determine the correct properties for its sampling distribution, we need to find the probability of values more extreme than our observed test statistic. More extreme means further away from the null value in both directions for a two-tailed test or in one direction for a one-tailed test. At this point, there are two ways to use the test statistic and distribution to calculate the p value. The formulas for probability distributions are relatively complex. Consequently, you won't calculate it directly. Instead, you'll use either an online calculator or a statistical table for the test statistic.

5.10 SUMMARY :

This chapter is introductory information about hypothesis and hypothesis testing and indeed a quite simplified for absolute beginners. Development of hypothesis in scientific research is important because the process refines the study. Researchers generally conduct a study developing either some type of research questions or hypothesis. An applicable hypothesis must be compatible with current related knowledge and it should be testable. Hypotheses must be tested for statistical significance. It involves error, particularly Type I and Type II error. Error must be considered in all research. There is a much emphasis often placed on significance testing.

5.11 GLOSSARY :

- ❖ **Hypothesis testing** : Hypothesis testing is a form of statistical inference that uses data from a sample to draw conclusions about a population parameter or a population probability distribution. It can be understood as quantitative results for a qualitative problem. It provides various techniques to test the hypothesis statement depending upon the variable.
- ❖ **Null hypothesis** : The null hypothesis is the claim that no relationship exists between two sets of data or variables being analyzed.
- ❖ **Type I and Type II errors** : Type I error occurs when we reject a true hypothesis while as Type II error occurs when we accept a wrong hypothesis
- ❖ **Probability level** : Pprobability value also called the p value tells the statistical significance of a finding.

- ❖ **Statistical significance** : Statistical significance refers to the claim that a set of observed data are not the result of chance but can instead be attributed to a specific cause. Statistical significance can be considered strong or weak.

5.12 SELF-ASSESSMENT QUESTIONS :

1. Develop three hypotheses in any mass media area that could be investigated or tested.
2. Explain the relationship between Type I error and Type II error.
3. Under what circumstances might a mass media researcher use a probability level of .001?
4. What are the types of Hypothesis Testing?
5. What is the general process to formulate the hypothesis and find the P value?
6. What is p value and explain its key features?
7. Who developed the concept of significance testing?

5.13 REFERENCES AND SUGGESTED READING :

1. Bailey, K. D. (1978). *Methods of social research (3rd ed.)*. New York: The Free Press. Grinnell, (1988). *Social work research and evaluation*. Itasca: Illinois, F. E. Peacock Publishers
2. Burns, R. B. (1997). *Introduction to research methods*. Addison Wesley Longman.
3. Drew, Chris. (2022, October 31). *The 13 different types of Hypothesis*. *Helpful professor.com*. Retrieved from <https://helpfulprofessor.com/types-of-hypothesis>
4. Farrugia, P., Petrisor, B. A., Farrokhyar, F., & Bhandari, M. (2010). Research questions, hypotheses and objectives. *Canadian Journal of Surgery*, 53(4).
5. Feldman, D. (2006). *Why do clocks run clockwise?* New York: Harper & Row.
6. Grabe, E.M. (2002). Maintaining the moral order: a functional analysis of “the jerry springer show”. *Critical studies in media communication*, 19(3), 311-328
7. Haynes, R., Sackett, D., Guyatt, G., & Tugwell, P. (2006). Forming research questions. *Clinical epidemiology: how to do clinical practice research*, 3-14.
8. Hubbard, Raymond., Bayarri, M. J. (2013, September 4). *P Values are not Error Probabilities Archived, Wayback Machine*.
9. Hulley, S. B., Cummings, S. R., Browner, W. S., Grady, D. G., & Newman, T. B. (2013). *Designing clinical research*. Lippincott Williams & Wilkins.
10. Jim Frost (2022). How to find the P value: Process and calculations. Statistics by Jim. Retrieved from <https://statisticsbyjim.com/hypothesis-testing/how-to-find-p-value>
11. Kerlinger, F. N. (1979) *Foundation of Behavioral Research*, New York: H 107, Rinehart and Winstem Inc.
12. Parahoo, K. (2014). *Nursing research: principles, process and issues*, London, Macmillan Press.
13. Roger D. Wimmer, Joseph R. Dominick (2013). *Mass Media Research: An Introduction* Cengage Learning, Retrieved from https://books.google.co.in/books/about/Mass_Media_Research.html
15. Singh, Y. K. (2006). *Fundamental of research methodology and statistics*. New Age International.
16. Tukey, J. W. (1986). *The collected works of John W. Tukey, Vols III and IV*. Belmont, CA: Wadsworth and Brooks/Cole. <https://www.britannica.com>

LESSON - 6

RESEARCH DESIGN

OBJECTIVES :

After reading this lesson, the reader should be able to gain an overview of Research Design in mass communication.

Key objectives of this lesson are :

- To define the research design and know the purpose of designing the research
- To distinguish various types of research design and their uses in the research
- To comprehend the features of good research and their characteristics
- To analyze the difference between research design and research approaches
- To get an idea of how to create a research design
- To distinguish the difference between fixed design and flexible design

STRUCTURE :

- 6.1 Introduction
- 6.2 Definition of the Research Design
- 6.3 Purpose of Research Design
- 6.4 Characteristics of Research Design
- 6.5 Types of Research Design There are four types of research designs:
 - 6.5.1 Descriptive Research Design
 - 6.5.2 The Exploratory Research Design
 - 6.5.3 The Explanatory Research Design
 - 6.5.4 The Evaluation Research Design
 - 6.5.5 Qualitative Research Design
- 6.6 Quantitative Research Design
 - 6.6.1 Descriptive Design
 - 6.6.2 Correlational Design
 - 6.6.3 Quasi-Experimental Design
 - 6.6.4 Experimental Design
- 6.7 How to create a Research Design
- 6.8 Difference Between Research Design and Approach
- 6.9 Summary
- 6.10 Glossary
- 6.11 Self-Assessment Questions
- 6.12 References and Suggested Readings

6.1 Introduction :

Research design refers to the way in which a research idea is transformed into a research project or plan that can then be carried out in practice by a researcher or research team. Research design is a well-organized and methodical procedure utilized by a researcher to carry out a scientific study. It is a comprehensive coexistence of already identified elements and data leading to a reasonable end result. The term research design also refers to and encompasses decisions about how the research itself is conceptualized, the subsequent

conduct of a specific research project, and ultimately the type of contribution the research is intended to make to the development of knowledge in a particular area. The process of developing a research design combines three broadly connected and interdependent components: *the theoretical, methodological, and ethical considerations* relevant to the specific project. The theoretical understandings and assumptions about research held by a researcher and/or research team provide an overarching frame that shapes and influences the research design at every point.

1. **Definition of the Research Design :** Research design is defined as a framework of methods and techniques chosen by a researcher to combine various components of research in a reasonably logical manner so that the research problem is efficiently handled. It provides insights about “how” to conduct research using a particular methodology. Every researcher has a list of research questions which need to be assessed. This can be done with research design. The sketch of how research should be conducted can be prepared using research design. The market research study will be carried out on the basis of research design. For Silverman (2013, p.38), research designs depend upon cool and rational assessment of alternatives.

It is the master plan specifying the methods and procedures for collecting the data and analyzing the required information in a research study. According to Kerlinger, “Research Design is a plan, structure and strategy of investigation conceived so as to obtain answers to research questions and control variance.”

- **Plan:** It includes everything the investigator will do from writing the hypothesis and their operational implications to the final analysis of the data.
- **The structure:** It is the outline, scheme, the practice of the operation of the variables.
- **Strategy:** It includes the methods to be used to collect and analyze the data.

Popular Definitions :

- “Research design is a master plan specifying the methods and procedures for collection and analysing the needed information.” **William Zikmund**
- “Research design is the plan, structure, and strategy of investigation conceived so as to obtain answers to research questions and to control variance.” **Kerlinger**
- “Research design is the planned sequence of the entire process involved in conducting a research study.” **By Miller.**
- “Research design is a catalogue of the various phases and facts relating to the formulation of a research effort. It is an arrangement of the essential conditions for collection and analysis of data in a form that aims to combine relevance to the research purpose with economy in the procedure”. **By Seltiz and others.**
- “Research designates the logical manner in which individuals or other units are compared and analysed, it is the basis of making interpretations from the data”. **By Anonymous.**
- “Also known as a market research briefing, this is a basic plan which guides the data collection and analysis phase of the research project. It acts as a framework that details the type of information to be collected, the data sources, and the data collection procedure”. **By Market Intelligence Group, India**

In short, a research design is a plan of what data to gather, from whom, how and when to collect the data, and how to analyse the data obtained.

Therefore, we can simply say that research design is planning a strategy of conducting a well-structured research study.

6.3 PURPOSE OF RESEARCH DESIGN :

Research designs are used for the following purposes;

- ❖ **To minimize the expenditure:** Research design carries an important influence on the reliability of the results attained. It therefore provides a solid base for the whole research. This makes the research as effective as possible by providing maximum information with minimum spending of effort, money and time by preparing the advance plan of all about the research.
- ❖ **To facilitate the smooth scaling:** Research design is needed because it facilitates the smooth scaling of the various research operations, thereby making research as efficient as possible yielding maximal information with minimal expenditure of effort, time and money.
- ❖ **To collect the relevant data and technique:** Research design stands for advance planning of the methods to be adopted for collecting the relevant data and the techniques to be used in their analysis, keeping in view the objective of the research and the availability of staff time and money. Poor preparation of research design upset the entire project.
- ❖ **To provide blue print for plans:** Research design is needed due to the fact that it allows for the smooth working of many research operations. It is like blue print which we need in advance to plan the methods to be adopted for collecting the relevant data and techniques to be used in its analysis for preparation of research project. Just as for better economical and attractive construction of a house need a blue print and a map of that, similarly we need a blue print or a design for the smooth flow of operation of research.
- ❖ **To provide an overview to other experts:** A research design provides an overview of all the research process and with the help of the design we can take the help and views of experts of that field. The design helps the investigator to organize his ideas, which helps to recognize and fix his faults.
- ❖ **To provide a direction:** A research design provides a proper or particular direction to the other executives and others who are helping us into the process. The researcher studies available, literature and learns about new (alternative approaches).

Need for Research Design :

Research is a scientific investigation of a problem for which we need systematic planning of research. For successful research, we need a research design because it includes

- the formulation of a strategy to resolve a particular question
- the collection and recording of information and evidence
- the processing and analysis of these data and their interpretation
- the publication of results.

A research design states the structure and process of conducting a research process. Thus, it shows a path to the researcher without which he may be lost or confused as to what next step he has to take. More so, it also takes care of the budget and time frame of the research study. All this planning can only make a research study a success story.

6.4. CHARACTERISTICS OF RESEARCH DESIGN :

There are four main characteristics of research design which include *Reliability, Neutrality, Validity as well as Generalization*. Further, a researcher should have a clear

understanding of how their project can be implemented in the research design. Apart from these, there are other characteristics of research design such as:

- ❖ A good research design ought to be flexible, effective, efficient, appropriate, reliable, economical tool
- ❖ A good research design ought to increase reliability and decrease bias of any data or information collected and reviewed
- ❖ A good research design must be less prone to an error whenever it comes to experiments and survey. The best design always gives the smallest experimental error.
- ❖ A design, which yields maximum information and provides an opportunity for considering many different aspects of a problem, is considered an efficient design.
- ❖ A good design contains a clear statement of research problem, procedures and techniques to be used for gathering information, the population to be studied and methods to be used in processing and analyzing data.

Generally, a good research design minimizes bias and maximizes the reliability of the data collected and analysed. The design which gives the smallest experimental error is reported to be the best design in scientific investigation. Similarly, a design which yields maximum information and provides an opportunity for considering different aspects of a problem is considered to be the most appropriate and efficient design.

A good research design possesses the following characteristics;

1. **Objectivity** : It refers to the findings related to the method of data collection and scoring of the responses. The research design should permit the measuring instruments which are fairly objective in which every observer or judge scoring the performance must precisely give the same report. In other words, the objectivity of the procedure may be judged by the degree of agreement between the final scores assigned to different individuals by more than one independent observer. This ensures the objectivity of the collected data which shall be capable of analysis and interpretation
2. **Reliability**: It refers to consistency throughout a series of measurements. For example, if a respondent gives out a response to a particular item, he is expected to give the same response to that item even if he is asked repeatedly. If he is changing his response to the same item, the consistency will be lost. So, the researcher should frame the items in a questionnaire in such a way that it provides consistency or reliability.
3. **Validity**: Any measuring device or instrument is said to be valid when it measures what it is expected to measure. For example, an intelligence test conducted for measuring the IQ should measure only the intelligence and nothing else and the questionnaire shall be framed accordingly.
4. **Generalizability**: It means how best the data collected from the samples can be utilized for drawing certain generalizations applicable to a large group from which sample is drawn. Thus, a research design helps an investigator to generalize his findings provided he has taken due care in defining the population, selecting the sample, deriving appropriate statistical analysis etc. while preparing the research design. Thus, a good research design is one which is methodologically prepared and should ensure that generalization is possible. For ensuring the generalization we should confirm that our research problem has the following characteristics;
 - The problem is clearly formulated.
 - The population is clearly defined.

- Most appropriate techniques of sample selection are used to form an appropriate sample.
- Appropriate statistical analysis has been carried out
- The findings of the study are capable of generalizations.

Adequate Information : The most important requirement of good research design is that it should provide adequate information so that the research problem can be analysed on a wide perspective. An ideal design should take into account important factors like;

- Identifying the exact research problem to be studied
- The objective of the research
- The process of obtaining information
- The availability of adequate and skilled manpower and
- The availability of adequate financial resources for carrying research.

Other Features : Some other important features of a good research design are flexibility, aptability, efficiency, being economic and so on. A good research design should minimize bias and maximize reliability and generalization

6.5.TYPES OF RESEARCH DESIGN :

There are four types of research designs:

- The descriptive Research Design
- The exploratory Research Design
- The explanatory Research Design
- The evaluation Research Design

6.5.1 Descriptive Research Design :

Purpose of Descriptive Research is to describe the characteristics of groups, community or people. It systematically and accurately describes the population and situation of phenomenon. It provides researchers with a vast amount of information about many social context and settings. The heuristic value of descriptive studies must be considered a major contribution. Descriptive studies can reveal potential relationships between variables, thus setting the stage for more elaborate investigation later. It is good for a large-scale survey and use archival data methods, naturalistic observation and structured test methods. Descriptive research answers to the questions of what and how. Ex: What is the role of journalists in new media? How many journalists favor to AI supported Chat GPT tool in India?

6.5.2. The Exploratory Research Design :

The exploratory research design is the short of research design that explore the subject matter of research and there by answers to the questions of what, why and how. Probing, fact-finding, searching and investigating are the key features of exploratory Design. Investigator has limited or no knowledge about the problem or situation. Since there is unfamiliarity with a particular group of people or research targets, it is more sophisticated. Qualitative research approach is commonly used in exploratory design. This type of research design generates new ideas. Exploratory research design diagnosis the situation. For example, in case of workers strikes, conducting preliminary interviews with the workers give information on problems, working conditions, wages, safety, extramural facilities, sharing profits, career opportunities etc. This method provides alternatives. In workers strike, alternatives for the workers could be frequent dialogue with decision-makers.

6.5.3 The Explanatory Research Design :

Explanation about unexplored aspects of a subject is provided along with details about what, how and why related to the research questions. It concerned more with causes or why factor about some phenomenon. Ex: why farmers are committing suicides in Maharashtra, AP. Why and how communal riots are takes place. Why the rate of child marriage is high. In exploratory research design, the researcher's ideas and thoughts are key as it is primarily dependent on their personal inclination about a particular topic.

6.5.4. The Evaluation Research Design :

Evaluation research is a common research design that entails carrying out a structured assessment of the value of resources committed to a project or specific goal. It often adopts social research methods to gather and analyze useful information about organizational processes and products. This short of research determines the productivity of the subject matter or a program. Online surveys created and administered via data collection platforms make it easier for the researcher to gather and process information during evaluation research. There are two types of evaluation research: *Quantitative research and Qualitative research.*

6.5.5. Qualitative Research Design :

Qualitative designs vary with respect to how theory is explicitly used in a particular study. In some research designs, a specific theory may provide the organizing construct for the entire research design. For example, in a study design that is overtly Marxist in orientation, Marxist theory and traditions will shape and influence that design at every point. Other qualitative research designs may use theoretical concepts, as opposed to entire theories, to shape the design. These theoretical concepts may be derived from one or a number of theoretical traditions to provide the focus for the study and/or as a vehicle to explore a topic of interest. Gender or class as concepts can provide a theoretical lens or frame in qualitative research. In addition, some research designs may employ combinations of these concepts.

Following points qualifies the qualitative research design.

1. A central question that states the objectives or the problem
2. It has pre-set procedure or sequence in order to answer research questions
3. Reviews the generated data
4. The collection and analysis of the data are done in order to draw conclusions

6.6. QUANTITATIVE RESEARCH DESIGN :

In Quantitative Research Design, a researcher examines the various variables while including numbers as well as statistics in a project to analyses its findings. The use of graphics, figures, and pie charts is the main form of data collection measurement and meta-analysis (it is information about the data by the data). Its analysis variable relationships in terms of numbers and statistical strategies to review findings and is generally divided into four main types namely: Descriptive, correlational, quasi-experimental design and experimental design.

6.6.1 Descriptive Design :

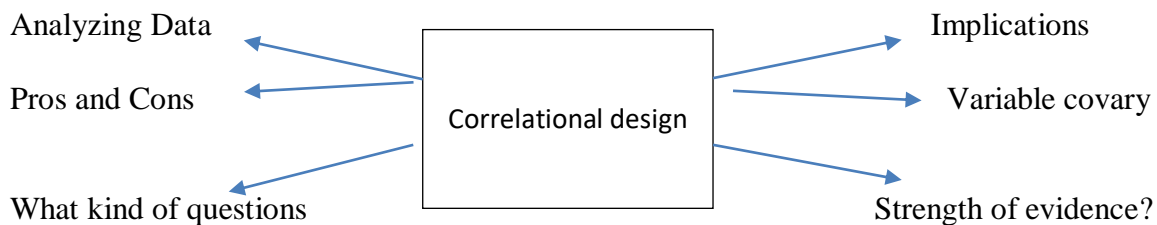
This type of design describes the present status of phenomenon or variable that not required any hypothesis for initiation. It is developed only after the data were collected. For instance, case studies, naturalistic observations and survey.

6.6.2. Correlational Design :

A correlational research design looks into correlations between variables without allowing the researcher to control or manipulate any of them. Correlational studies reveal the magnitude and/or direction of a link between two (or more) variables. Correlational study designs might have either a positive, negative or zero. It is great for swiftly collecting data from natural settings. This allows you to apply your results to real-world circumstances in an externally legitimate manner. Correlational studies research is a viable choice in a few scenarios like:

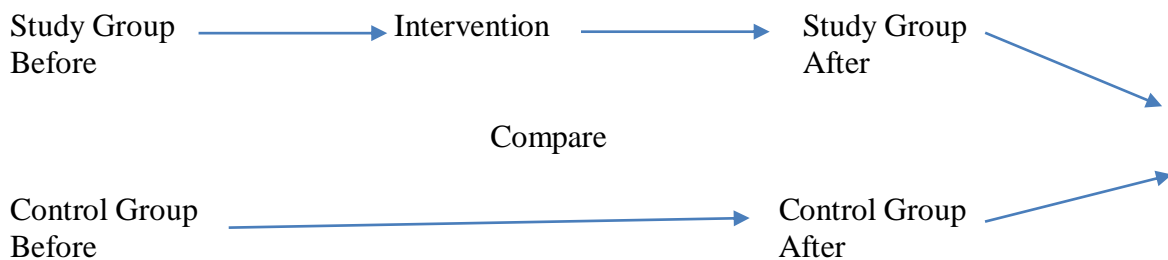
- To investigate non-causal relationships
- To explore causal relationships between variables
- To test new measurement tools

This shot of design utilizes statistical analyses in order to determine whether two variables are related or not. For instance, case control study, observational study.



6.6.3. Quasi-Experimental Design :

This shot of design resembles a true experimental design and it utilized when a standard research design is not applicable. However, it is not based upon randomized sample groups.



6.6.4 Experimental Design :

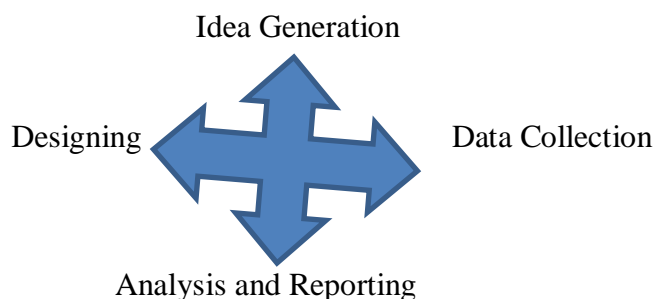
This shot of design determines cause and effect relationships amongst various variables in such a way independent variable is changed in order to observe its effect on the dependent variable easy. For instance, Experiment with random assignment.

Cause= the reason for something happening

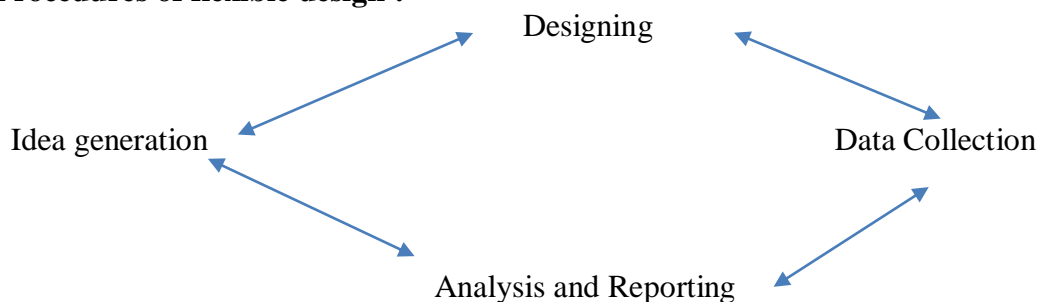
Effect= what happened

Cause and effect relationships :**What is the difference between fixed design and flexible design? :**

A fixed design generally follows pre-set or pre-determined design or a sequence before collection of the data, and is usually driven by theory. Variables are usually measured by quantitative research.

Procedures of fixed design

A flexible design on the other hand offers freedom when it comes to the collection of the data. This usually because the variables and this case is not measured quantitatively. Example, culture. In a few cases, the theory may not even be available before the actual research begins.

Procedures of flexible design :**6.7 HOW TO CREATE A RESEARCH DESIGN :**

When you identify the research problem, you need to focus on refine it by formulating various research questions/hypothesis. Once it is happened, then there is possible to develop clear aims and objectives for the study. Simplifying in the extreme, the aims and objectives of a study refer to what the study hopes to achieve. The aims and objectives of a study also provide and guide a framework or study design for choosing the specific methods and techniques for data collection.

In a tight and refined research design, it should be possible to cross-link and relate every part of the study design, including the data collection techniques, back to what the research is attempting to achieve as expressed in the stated aims and objectives. For example, if the design involves the use of interviews, then why interviews are an appropriate method to use, what type of interview is being used, and who is being interviewed should be apparent from, and congruent with, the aims and objectives of the research. Often this information is presented as a series of related and sequential points pertaining to the proposed design.

Points of research design :

1. Selection of the topic, specify the research problem, pinpoint your learning outcomes
2. Refine the research problem and formulate the research questions/hypothesis
3. Defining concepts and variables and set your research objectives
4. Develop your research design
5. Choose a sampling framework
6. Pick up suitable methods of data collection
7. Set up appropriate data collection tool/s and measurement instruments
8. Identify statistical analysis tools for using wherever it is required
9. Identify proper data analysis techniques and tools for tabulation, interpretation and generalization of the data
10. Contemplate dissemination and publication of your findings

6.8 DIFFERENCE BETWEEN RESEARCH DESIGN AND APPROACH :

You need to understand the main difference between research design and research approach. Research design is overall framework or outline or structure of research proposal whereas a research approach incorporates various tools, techniques, procedures or processes utilized to collect or review data or information.

6.9 SUMMARY :

A Research Design in mass communication is simply a structural framework of various research methods as well as techniques that are utilised by a researcher. Research design is much more than simply identifying techniques that will be used to collect data. It involves theoretical, methodological, and ethical considerations that shape both the design and what the research is aiming to achieve.

Research design also involves a degree of reflexivity on the part of the researcher with respect to acknowledging the underlying theory and/or theoretical assumptions that have shaped his or her perspectives and understandings of the research focus and process.

6.10 GLOSSARY :

- ❖ **Research design:** Research design is planning a strategy of conducting a well-structured research study
- ❖ **Descriptive Research Design:** Descriptive Research describes systematically and accurately the characteristics of the population, groups, community or people and also the situation of phenomenon.
- ❖ **Exploratory Research Design:** Explore the subject matter of research and there by answers to the questions of what, why and how. Probing, fact-finding, searching and investigating are the key features of exploratory Design.

- ❖ **Explanatory Research Design:** Explanation about unexplored aspects of a subject is provided along with details about what, how and why related to the research questions.
- ❖ **Evaluation Research Design:** Evaluation research design entails carrying out a structured assessment of the value of resources committed to a project or specific goal. It determines the productivity of the subject matter or a programme.

6.11 SELF-ASSESSMENT QUESTIONS :

1. Define the research design.
2. What are the characteristics of research design?
3. What is the purpose of research design?
4. What do you know about the digital media research?
5. What are the types of research design?
6. Write the features of good research design.
7. Explain how to create the research design.
8. What is the difference between fixed design and flexible design?
9. Difference between research design and approach

6.12 REFERENCES AND SUGGESTED READINGS :

1. Aysha.(2021). Research Design.Educational Hub Research Centre.[You Tube video]. Retrieved from <https://www.youtube.com/watch?v=LpmGSioXxdo&t=4s>
2. Cheek, Julianne."Research Design."*The Sage Encyclopedia of Qualitative Research Methods*. 2008. SAGE Publications. 4 Mar. 2009. <http://sage-reference.com/research/Article_n384.html>.
3. Klaus Bruhn Jensen. (2002). A handbook of media and communication research.Routledge Kothari, C.R. (2004). *Research Methodology: Methods and Techniques*. New Delhi: New age International Publishers.Silverman, David (2013), *Doing Qualitative Research* (4thEdition), London, Sage Publication

Dr. K. JOHN BABU

LESSON - 7

RESEARCH APPROACHES AND METHODS

OBJECTIVES :

After reading this lesson, the reader should be able to understand an overview of the research approaches and methods.

Key objectives of this lesson are:

- To define the research approaches and understand the significance of methodology in mass communication research;
- To know various types of research approaches and methods and how they are used in mass communication depending on the purpose of research;
- To comprehend the importance of survey research in mass media and to know various survey approaches and data collection methods of survey;
- To understand the significance of content analysis in mass media research and various steps to be followed while conducting content analysis research;
- To comprehend the historical method of research and how the historical data is analyzed with its various sources of data;

STRUCTURE :

- 7.1 Introduction
- 7.2 Types of Research Approaches
 - 7.2.1 Positivist paradigm
 - 7.2.2 Interpretive paradigm
 - 7.2.3 The critical paradigm
 - 7.2.4 Qualitative approach
 - 7.2.5 Quantitative approach
 - 7.2.6 Deductive Approach
 - 7.2.7 Inductive Approach
 - 7.2.8 Abductive Approach
- 7.3 Definitions of Research Methodology
- 7.4 Types of Research Methods
- 7.5 Survey Research
 - 7.5.1 Types of research
 - 7.5.2 Pilot study
 - 7.5.3 Questionnaire design
 - 7.5.4 General problems in survey research
- 7.6 Content Analysis
 - 7.6.1 Definition of Content Analysis
 - 7.6.2 Types of Content Analysis
 - 7.6.3 Steps in Content Analysis
 - 7.6.4 Cases of content analysis
 - 7.6.5 Advantages and disadvantages of Content Analysis
- 7.7 Historical Method
 - 7.7.1 Source of Data
 - 7.7.2 Types of Historical Research

7.7.3 Data Analysis and Reporting of Findings

7.8 Summary

7.9 Glossary

7.10 Self-Assessment Questions

7.11 References and Suggested Reading

7.1 INTRODUCTION :

Research approaches refer to the systematic and structured ways that researchers use to conduct research, and they differ in terms of their underlying logic and methods of inquiry. Research approach decides the methods for data collection, analysis, and interpretation. There are many factors the selection of research approach, such as, research objective, experience of research, and the audience of research study. A researcher's approach has a great influence on the specific research methods the researcher uses.

7.2. TYPES OF RESEARCH APPROACHES :

Neuman (1997) and Blaikie (1993) suggest that there are three distinct approaches to social science research: *positivism* (or objectivism), *interpretive*, and *critical*. Each of these represents a model or a paradigm for research an accepted set of theories, procedures, and assumptions about how researchers look at the world. Paradigms are based on statements that are universally accepted as true. Paradigms are important because they are related to the selection of research methodologies.

7.2.1 Positivist Paradigm :

It is the oldest and still the most widely used in mass media research. Derived from the writings of philosophers such as Comte and Mill, positivism is the paradigm most used in the natural sciences. When the social sciences developed, researchers modified this technique for their own purposes. The positivist paradigm involves such concepts as quantification, hypotheses, and objective measures.

7.2.2 Interpretive Paradigm :

Interpretive social science traces its roots to Max Weber and Wilhelm Dilthey. The aim of the interpretive paradigm is to understand how people in everyday natural settings create meaning and interpret the events of their world. This paradigm became popular in mass media research during the 1970s and 1980s and gained added visibility in the 1990s and the new century.

7.2.3 Critical Paradigm :

It draws on analysis models used in the humanities (an academic disciplines that study human society and culture). Critical researchers are interested in such concepts as the distribution of power in society and political ideology.

These three approaches have a different philosophy of reality. Let us compare the positivist and interpretive paradigms. For the positivist researcher, reality is objective. It exists apart from researchers and can be seen by all. For the interpretive researcher, there is no single reality. Each observer creates reality as part of the research process. It is subjective and exists only in reference to the observer. Example, if a tree falls in the forest and there is no one there to hear it, does it make any noise? A positivist would answer yes, reality doesn't depend on an observer. It exists independently. On the other hand, an interpretive researcher would say no noise was made, reality exists only in the observer. Furthermore, the positivist researcher believes that reality can be divided into component parts, and knowledge of the

whole is gained by looking at the parts. In contrast, the interpretive researcher examines the entire process, believing that reality is holistic and cannot be subdivided.

Apart from these three paradigms, research approaches are two types: Qualitative and Quantitative. Both the approaches increase the researchers understanding of the social phenomena. Many researchers now use a combination of the quantitative and qualitative approaches to understand fully the phenomenon they are studying. Quantitative methodology generally adopts the positive paradigm while qualitative researchers promote the critical paradigm. Those who accept the critical paradigm generally follow the methodology of the humanities.

7.2.4 Qualitative Approach :

Qualitative research uses a flexible questioning approach and employ open-ended questions which utilize direct quotations. Although a basic set of questions is designed to start the project, the researcher can change questions or ask follow-up questions at any time. Qualitative research is a useful mass media research approach only when its limitations are recognized. In most cases, qualitative research studies use small samples, respondents or informants that are not necessarily representative of the population from which they are drawn. Qualitative research can be an excellent way to collect and analyze data, but the researchers must keep in mind that the results of such studies have interpretational limits if sample sizes are small. Researchers interested in generalizing results should use large samples or consider other methods. The researcher can utilize qualitative methods in the form of semi- structured interviews.

7.2.5 Quantitative approach :

Quantitative techniques make use of advanced statistical techniques and standardized set of questions. All respondents are asked the same questions. Although follow-up questions (and skips) can be designed into a questionnaire, they must be included in the questionnaire or measurement instrument before the research project begins. Interviewers conducting the interview are not allowed to stray from the questionnaire. Quantitative approach provides an avenue for furthering our understanding about communication. It is grounded on conventionalized approach to social science research. In a research project, data can be collected from various sources like surveys through mail questionnaires, face- to-face interviews, pre-test or post-test experimental data, fieldwork data and so forth. The researcher can utilize quantitative methods in the form of statistical techniques for hypothesis testing. There are three main research approaches: Deductive, Inductive, and Abductive.

7.2.6 Deductive Approach :

The deductive approach starts with a theory or a hypothesis, and the researcher tests the hypothesis through the collection and analysis of data. The researcher develops a research design and data collection methods based on the theory or hypothesis. The goal of this approach is to confirm or reject the hypothesis.

7.2.7 Inductive Approach :

The inductive approach starts with the collection and analysis of data. The researcher develops a theory or an explanation based on the patterns and themes that emerge from the data. The goal of this approach is to generate a new theory or to refine an existing one.

7.2.8 Abductive Approach :

The abductive approach is a combination of deductive and inductive approaches. It starts with a problem or a phenomenon that is not fully understood, and the researcher develops a

theory or an explanation that can account for the data. The researcher then tests the theory through the collection and analysis of more data. The goal of this approach is to generate a plausible explanation or theory that can be further refined or tested.

7.3. DEFINITION OF RESEARCH METHODOLOGY :

The word methodology is the study of methods and the underpinning philosophical assumptions of the research process. Different research questions suggest different methodologies. Different methodologies are associated with different paradigms and approaches. Methodology deals with the question of “why” to do research in a certain way. It is a guide to what problems are worth investigating and how the research should proceed. For instance, a researcher is interested in how the social media affects copyright laws; he/she would probably choose the methodology of legal research. A researcher wants to trace how radio programming has evolved, he/she would probably choose historical methodology. A study about the effects of television on children might use scientific methodology.

In contrast, a method is a specific technique for gathering information following the assumptions of the chosen methodology. Researchers who choose the positivist paradigm will use such methods as surveys and experiments while those who choose the interpretive paradigm will rely on methods such as focus groups, ethnography, and observation.

7.4 TYPES OF RESEARCH METHODS :

There are qualitative and quantitative data collection method and techniques such as: field observations, focus groups, intensive interviews, case studies, ethnography, survey, experimental study, content analysis and other methods like historical study. Let us discuss about survey, content analysis and historical methods in this lesson.

7.5 SURVEY RESEARCH :

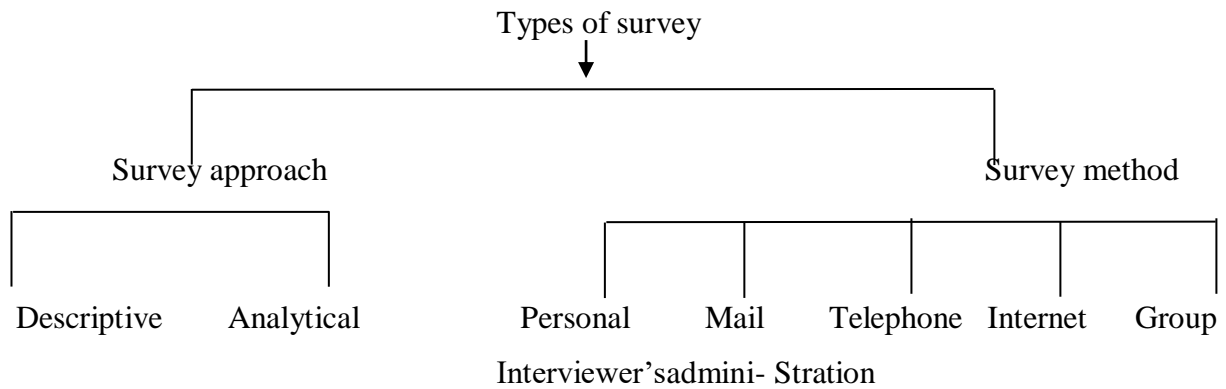
Survey research is one of the most widely used methods of media research, primarily because of its flexibility. “Survey research is a specific type of field study that involves the collection of data from a sample of elements (e.g., adult women) drawn from a well-defined population (e.g., all adult women living in the United States) through the use of a questionnaire (for more lengthy discussions,” (Babbie, 1990; Fowler, 1988).

Surveys, however, involve a number of steps, a researcher must consider in conducting a survey. The steps are defining the purpose of the study, review the available literature in the area, select the survey approach, questionnaire design, and sample, analyze and interpret the data, and finally, decide whether to publish or disseminate the results. Survey research, including online surveys, requires careful planning and execution, and the research must take into account a wide variety of decisions and problems.

7.5.1 Types of Survey :

If a decision is made to go ahead and conduct a survey, two more questions then emerge - What type of survey should be used, and should the survey be of the entire population or a sample of the population being studied? There are two major types of surveys: descriptive and analytical.

Researcher must decide whether to use a descriptive or analytical approach. Researchers must also select the most appropriate survey approach from among the following basic types: personal interview, mail, telephone, Internet, and group administration. Each approach has advantages and disadvantages that must be weighed.



- ❖ **Descriptive Survey** : A descriptive survey attempts to describe or document current conditions or attitudes, that is to explain what exists at the moment. In descriptive surveys, the interest is in discovering the current situation in the area under study. For example, the newspapers organisations regularly conduct surveys on the rate of circulation of the newspapers. Broadcast stations and networks continually survey their audiences to determine programming tastes, changing values, and lifestyle variations that might affect programming. Professional pollsters survey the electorate to learn its opinions of candidates or issues.
- ❖ **Analytical Survey** : An analytical survey attempts to describe and explain *why* situations exist. In this approach, two or more variables are usually examined to investigate research questions or test research hypotheses. The results allow researchers to examine the interrelationships among variables and to develop explanatory inferences. For example, television station owners survey the market to determine how lifestyles affect viewing habits or to determine whether viewers' lifestyles can be used to predict the success of syndicated programming. On a broader scale, television networks conduct yearly surveys to determine how the public's tastes and desires are changing and how these attitudes relate to viewers' perceptions of the three major commercial networks.
- ❖ **Personal Interviewers** : Presence of interviewers is extremely important in face-to-face personal surveys. An interviewer ideally should function as a neutral medium through which the respondents' answers are communicated to the researcher. The interviewer's presence and manner of speaking should not influence respondents' answers in any way. For example, if an interviewer shows disdain or shock over an answer, it is unlikely that the respondent will continue to answer questions in an honest manner. Showing agreement with certain responses might prompt similar answers to other questions. Skipping questions, carelessly asking questions, and being impatient with the respondent also cause problems. Therefore, adequate training and instruction can minimize the bias that the interviewer might inject into the data.
- ❖ **Mail Surveys** : Mail surveys are popular in some time ago and are common in some types of businesses. Mail surveys involve sending self-administered questionnaires to a sample of respondents through post/courier. Stamped reply envelopes are enclosed to encourage respondents to send their completed questionnaires back to the researcher.
- ❖ **Telephone Surveys** : Telephone and personal interviewing techniques have certain similarities, and much of what follows applies to both. The respondents generally do not see the actual questionnaire in telephone surveys. Telephone surveys fill a middle ground between mail surveys and personal interviews. They offer more control and higher response rates than most mail surveys, but they are limited in the types of questions that can be asked. Telephone interviews are generally more expensive than

mail surveys but less expensive than face-to-face interviews. Because of these factors, telephone surveys are a compromise between the other two techniques, and this may account for their popularity in mass media research.

- ❖ **Internet Surveys :** Collecting research data via the Internet is now commonplace. A respondent is contacted via telephone, letter, or email and asked to participate in a research project. Respondents who agree are either sent a questionnaire via email or given a link to access the questionnaire online. In most situations, respondents are given a password to access the website. When finished, respondents either click on a “submit” button or email the questionnaire back to the research company or business conducting the study. Collecting required data through Google form questionnaire is easiest way. This type of online research offers a huge list of advantages including low costs, no geographic limitations, no specific time constraints because respondents can complete the survey or measurement instrument at their convenience, flexibility in the approach used to collect data, and the ability to expose respondents to almost any type of audio or visual materials.
- ❖ **Group Administration :** When telephone and face-to-face interviews take place on a one-on-one basis, Group-administered surveys are questionnaires completed in group settings, such as at a screening clinic or in a classroom. Main advantages in this technique are group can be surveyed at one time, anonymity may provide more honest answers to sensitive questions, costs are low and less staff is needed than for personal interviews. There is a controlled situation where staff can answer questions and address concerns that also ensures a high response rate. Problems in group administration are it not practical for most populations being studied and members of the group are not independently chosen.

The type of survey used depends on the purpose of the study, the amount of time available to the researcher, and the funds available for the study. It is clear that many researchers now depend less on the face-to-face survey and more on computer-assisted telephone interviewing and Internet data collection.

7.5.2 Pilot Study :

Researchers suggested conducting one or more pilot studies to detect any errors in the approach. It will ensure that all the steps in the survey process are in harmony. Pilot studies save time, money, and frustration because an error that could void an entire analysis sometimes surfaces at this stage.

7.5.3 Questionnaire Design :

The goal in questionnaire design is to avoid bias in answers. Question wording, length, style, and order may affect a respondent’s answers. An interviewing approach will elicit a specific response. Extreme care must be taken when developing questions to ensure that they are neutral. To achieve a reasonable response rate, researchers should consider including an incentive, notifying survey subjects beforehand, and personalizing the questionnaire. Researchers should mention the response rate when they report the results of the survey.

There can be significant differences when designing questionnaires for academic use and those used for the private sector. Academic research usually requires additional explanations, procedures, and anonymity guarantees. If you are in an academic setting, you should know what rules or regulations does your institution have in reference to conducting research with humans.

Open-ended questions always include an opportunity for interviewers to ask for additional information. From experience, we have learned that interviewers should ask the respondent “What else?” instead of “Anything more?” or “Is that all?” The “What else?” approach does not give the respondent the same opportunity to say, “Nothing.”

7.5.4 General Problems in Survey Research :

Although surveys are valuable tools in mass media research, several obstacles are frequently encountered. According to Roger Wimmer and Joseph Dominick (2005) survey research has the following problems based on the experience:

- ❖ Respondents/subjects are often unable to recall information about themselves or their activities. This inability may be caused by memory failure, nervousness related to being involved in a research study, confusion about the questions asked, or some other intervening factor. Questions that are simple to researchers may cause significant problems for respondents. For example, radio station manager often want to ask respondents which radio stations they listen. People do not know which stations are programmed and also do not know how many buttons are on their radio.
- ❖ Due to respondents’ feelings of inadequacy or lack of knowledge about a particular topic, they often provide “prestigious” answers rather than admit to not knowing something. This is called prestige bias. For example, some respondents claim to watch International TV and listen to International radio when, in fact, they do not.
- ❖ Subjects may purposely deceive researchers by giving incorrect answers to questions. Almost nothing can be done about respondents who lie. A large sample may discount this type of response. However, there is no acceptable and valid method to determine whether respondents’ answers are truthful; the answers must be accepted as they are given, although one way to discover deception is to ask the same question in different ways a few times throughout the survey.
- ❖ Respondents often give elaborate answers to simple questions because they try to figure out the purpose of a study and what the researcher is doing. People are naturally curious, but they become even more curious when they are the focus of a scientific research project. In addition, some respondents use a research study as a soapbox for their opinions.
- ❖ Surveys are often complicated by the inability of respondents to explain their true feelings, perceptions, and beliefs not because they do not have any but because they cannot put them into words. The question “Why do you like to watch soap operas?” may be particularly difficult for some people. They may watch soap operas every day, but respond by saying only “Because I like them.” From a research perspective, this answer does not provide much information and probing respondents for further information may help, but not in every case.
- ❖ Conducting survey research is an exciting activity. It is fun to find out why people think in certain ways or what they do in certain situations. But researchers must continually remain aware of obstacles that may hinder data collection, and they must deal with these problems. In many areas around the world, many citizens now refuse to take part in any type of research project. Researchers must convince respondents and subjects that their help is important in making decisions and solving problems.
- ❖ The face of survey research is continually changing. One-on-one and door-to-door interviews are now difficult to conduct. The emphasis is now on research via the Internet, and it will be interesting to see how the Internet continues to change the research survey process.

7.6 CONTENT ANALYSIS :

After the war II, researchers used content analysis to study propaganda in newspapers and radio. In 1952, Bernard Berelson published *Content Analysis in Communication Research*, which signaled that the technique had gained recognition as a tool for media scholars. Since that time, content analysis has become a popular research technique in mass media research. Many of the steps followed in laboratory and survey studies are also involved in content analysis. Sampling procedures need to be objective and detailed, and operational definitions are mandatory. Content categories should be reliable, exclusive, and exhaustive. Coders must be carefully trained to gather accurate data. Interpreting a content analysis, however, requires more caution. No claims about the impact of the content can be drawn from an analysis of the message in the absence of a study that examines the audience. A content analysis should demonstrate acceptable inter coder reliability and validity.

7.6.1 Definition of Content Analysis :

- ❖ Walizer and Wienir (1978) define it as any systematic procedure devised to examine the content of recorded information.
- ❖ Krippendorf (2004) defines it as a research technique for making replicable and valid references from data to their context.
- ❖ Kerlinger's (2000) defines it as "Content analysis is a method of studying and analyzing communication in a *systematic, objective, and quantitative* manner for the purpose of measuring variables.
- ❖ Kerlinger's definition involves three concepts that require elaboration. First, content analysis is *systematic*. This means that the content to be analyzed is selected according to explicit and consistently applied rules: Sample selection must follow proper procedures, and each item must have an equal chance of being included in the analysis.

7.6.2 Types of Content Analysis :

Qualitative Content Analysis : A qualitative approach to content analysis, however, is typically inductive, beginning with deep close reading of text and attempting to uncover the less obvious contextual or latent content therein. For example, a researcher seeking to understand participants' experiences or understandings of a phenomenon of interest is likely to use such an inductive approach to analysis of interview data. When analyzing qualitative data such as interview transcripts, analyses across the whole set of data typically produce clusters or codes that translate into "themes."

Textual data include non written text, such as photographic data, equally open to content analysis. In this case, the researcher may identify content as straightforwardly as identifying objects evident in photographs or may conduct more subtle analyses of symbolic communications that can be unconsciously discerned from a physical space. The level at which content analysis occurs varies widely, from obvious surface-level groups of similar responses to a particular interview question to deeper inductive insights inferred from more sustained, iterative, and recursive interaction with textual data. For example, the varying connotations associated with particular words used by participants, or the degrees of enthusiasm expressed about an issue, are open to content analysis.

- ❖ **Quantitative Content Analysis :** In quantitative research work, content analysis is applied in a deductive manner, producing frequencies of pre-selected categories or values associated with particular variables. Either quantitative or qualitative

approaches may be combined within a single research study depending on the purpose of the analysis.

- ❖ **Descriptive Content Analysis** : Here, analysis means counting the frequency in appearance of certain elements of the research question and comparing it with other elements.
- ❖ **Categorical Content Analysis** : Where analysis involves a study of the documents by means of a set of categories, producing nominal as well as original and interval data, which are then processed statistically.
- ❖ **Contingency Content Analysis** : Which is basically a semantic analysis that usually employed to make an inference from the text about the personality of the author.(Touching, contact, connection and affinity).
- ❖ **Contextual Content Analysis** : Where sequence in which concepts appear together is examined. The systematic appearance of concepts, the thinking pattern of the author will be examined..
- ❖ **Intensity Content Analysis** : Which is processed by means of multi-step scale based on theoretical criteria.

7.6.3 Steps in Content Analysis :

A content analysis is conducted in several stages. The following steps may be used in content analysis study:

1. Formulate the research question or hypothesis.
2. Define the universe in question.
3. Select an appropriate sample from the population.
4. Select and define a unit of analysis.
5. Construct the categories of content to be analyzed.
6. Establish a quantification system.
7. Train coders and conduct a pilot study.
8. Code the content according to established definitions.
9. Analyze the collected data.
10. Draw conclusions and search for indications.

7.6.4 Cases of Content Analysis :

Content Analysis method was used in analyzing violence against Sikhs as reported in newspapers and magazines in 1984. It was used in analyzing caste massacres in Bihar during 1999, and 2000 as reported in print media. DA Patil's (2006) study *Print Media and Environment: A content Analysis of Marathi Newspapers* found that majority of environment related space covered on editorial page followed by Sunday special supplement. In a study on *Medicine on British Television: A content Analysis*, 54 British TV programmes on medical topics were viewed and analyzed during a four-month period in the winter. Psychological, behavioral and neurologic topics appeared frequently with common medical problems and preventive medicine rarely emphasized.

7.6.5 Advantages and Disadvantages of Content Analysis :

The symbols and messages contained in the mass media have become increasingly popular research topics in the academic sector. Public relations firms use content analysis to monitor the subject matter of company publications, and *The Media Monitor* publishes periodic studies of how the media treat social and political issues. Mass media are aimed at the largest possible audience, most material is simple in form and uncomplicated in content.

In this, researcher is directly involved with content. When it is quantitative research, there will be minimum chances of bias. This method is more useful where research budget is small and resources are limited. Useful for cross cultural studies and comparing. CA serves an important function by comparing the same material as presented in different nations or by comparing media content with some explicit set of standards or abstract categories.

There are some disadvantages with content analysis like problems of subjectivity in qualitative content analysis, problems of constituency and validity, problems of sampling and problems of estimation and generalization.

7.7 HISTORICAL METHOD :

Historical research is most often associated with historiography as the primary research method. Historiography goes beyond data gathering to analyze and develop theoretical and holistic conclusions about historical events and periods. It includes a critical examination of sources, interpretation of data, and analysis that focuses on the narrative, interpretation, and use of valid and reliable evidence that supports the study conclusions. Although a historian studies history or may teach history, the historiographer writes, analyzes, and interprets history. As technology has advanced, few people keep diaries or journals or write letters. Those who deliver speeches often do not write them verbatim, as in the past. Storytelling has most often been associated with the oral traditions of the undereducated and less developed cultures and was often expressed for society's entertainment rather than for leaving historical evidence. Yet, storytelling is an important historical method since it is a way of communicating values and cultural paradigms. Contemporary storytelling may now take the form of emails, blogs, and instant messaging, creating challenges for historians accustomed to depending on handwritten letters and traditional data sources. Historical research will change as more correspondence and eyewitness accounts are recorded on the internet and become available as new data sources for historical research.

7.7.1 Source of Data :

Diaries, photographs, art, literature, minutes of meetings, eyewitness accounts in newspapers or other official documents, court records, letters, maps, and other relevant sources can often be found in university and specialty collections. Government websites and collections are also excellent beginning points for locating data sources, as are special collections from museums and art galleries. Many of these can now be located on the internet. Historical research that is often overlooked includes a variety of art and media forms, such as paintings, sculpture, poetry, music, film, television, and literature. These resources can depict the shifts and changes in the social, cultural, and political context of history. Data should always be labeled and dated and authorship identified, including all critical details of the data source.

- ❖ **Primary Sources:** Primary sources refer to first-person accounts of events in original documents, letters, artwork, literature, music, observational notes, journals, and photographs. Primary sources enable the researcher to get as close as possible to what actually happened during a historical event or time period. These original sources of data hold the greatest value in the validity and reliability of historical analysis.
- ❖ **Secondary Sources:** Secondary sources are data from letters, diaries, and account descriptions of persons who were not eyewitnesses of the event or who did not personally know the person who is the focus of the study. This category of sources is significantly easier to define, understand, and access. A secondary source is any item that was created after the events it describes or is related to or is created by someone

who was not directly involved in or was an eyewitness to the events. Types of secondary sources include biographies and accounts written years after the event, scholarly or popular books and articles, reference books, biographies, and textbooks.

7.7.2 Types of Historical Research :

Historical research can take many forms, depending on the purpose of the research as well as on the availability and quality of data and resources available to the researcher. The following types of historical research data-collection methods represent the most commonly used. Researchers often combine the data collection methods as well.

- ❖ **Oral History:** It is a biographical approach in which the researcher gathers personal recollections of events from a living individual through audio and videotape recordings. Oral history can include written works of an individual who has died, but is primarily limited to a living individual. Most researcher questions and comments are unstructured, although a general interview schedule may be employed to guide the story depending on the nature of the oral history. This method provides the respondents, or storyteller, a natural and effective environment that allows a reciprocal interchange between the researcher and the respondent.
- ❖ **Autobiography Narrative :** It is an account of a person's life that has been written or recorded by the individuals themselves. A biography narrative account of a person's life can be either told to the researcher or found in archives, documents, and other sources.
- ❖ **Life History :** It is a biographical writing in the form of an extensive record of a person's life, as told to the researcher. The life history of a person involves a living individual.
- ❖ **Case Study:** It is a type of historical research that sheds light on a phenomenon through an in-depth examination of a single case exemplar of a phenomenon. The case can be an individual person, an event, a group, or an institution. Case studies take a relatively small sub-sample of research subjects as a source of in-depth, qualitative information.

7.7.3 Data Analysis and Reporting of Findings :

Interpretation occurs at the analytic stage through interpretation of meaning. Extensive examples should be used, with excerpts from documents and other artifacts. Although most historical research is based on incomplete data, the researcher must extend and derive opinion beyond what is discovered and is known from the research data collection process. A critical description of historical evidence, an evaluation of its historical significance to contemporary society, and creative narratives should be included in the written research report, including the derived inferences. The researcher should include all sources in the reference list, footnotes about data sources, and multiple references. This list should reflect the corroboration of facts, as evidence of reliability and validity of the findings.

7.8 SUMMARY :

We discussed in this lesson the differences between the positivist approach, which generally favors quantitative methods, and the interpretive approach, which favors qualitative methods. Mass media research can be influenced by the research paradigm that directs the researcher. During the past several years, technology has advanced the use of the internet for correspondence of both primary and secondary sources. The Internet has opened up new opportunities and challenges for survey, content analysis and historical methods. Survey and content analysis are the most used and important methods of data collection in mass media research.

7.9 GLOSSARY :

- ❖ **Paradigm:** Paradigms are based on statements that are universally accepted as true. Paradigms are important in research because they are related to the selection of research methodologies.
- ❖ **Qualitative Approach:** Qualitative research approach has a different philosophy of reality. It uses a flexible questioning approach and employ open-ended questions.
- ❖ **Quantitative approach:** Quantitative techniques make use of advanced statistical techniques and standardized set of questions.
- ❖ **Deductive approach:** The deductive approach starts with a theory or a hypothesis, and the researcher tests the hypothesis through the collection and analysis of data.
- ❖ **Inductive Approach:** The inductive approach starts with the collection and analysis of data. The researcher develops a theory or an explanation based on the patterns and themes that emerge from the data.
- ❖ **Descriptive survey :** A descriptive survey attempts to describe or document current conditions or attitudes, that is to explain what exists at the moment.
- ❖ **Content analysis:** Content analysis is a method of studying and analyzing the content in a systematic and objective manner for the purpose of obtaining findings.
- ❖ **Historical research:** Historical research is associated with historiography as the primary research method. A historian studies history or may teach history, the historiographer writes, analyzes, and interprets history.

7.10. SELF-ASSESSMENT QUESTIONS :

1. Define research approaches and explain various philosophical paradigms of research.
2. Explain various types of research approaches.
3. Define the research methodology and explain its significance in mass media research.
4. Define survey research and explain various types of survey research.
5. Define content analyses.
6. What are the stages for conducting a study using content analysis method?
7. Conduct a brief content analysis of one of the topics: Similarities and differences between local newspapers and national newspapers.
8. Define historical research and explain how do you analyze historical data.
9. What are the types of historical research?

7.11 REFERENCE AND SUGGESTED READING :

1. Babbie, E. R. (1990). *Survey research methods*. Belmont, CA: Wadsworth.
2. Blaikie, N. W. (1993). *Approaches to social enquiry*. Cambridge, MA: Polity Press.
3. Fowler, F. J. (1988). *Survey march methods*(2nd 4). Beverly Hills, CA: Sage.
4. Julien, Heidi. "Content Analysis."*The Sage Encyclopaedia of Qualitative Research Methods*. 2008. SAGE Publications. 4 Mar. 2009.<http://sage ereference.com/research>.
5. Neuman, W. L. (1997). *Social research methods*. Boston: Allyn& Bacon.
6. Roger D. Wimmer& Joseph R. Dominick. (2005). *Mass media research: An introduction*: Wadsworth Pub Co.
7. Saucier Lundy, Karen."Historical Research."*The Sage Encyclopedia of Qualitative Research Methods*. 2008. SAGE Publications. 4 Mar. 2009.<http://sage-ereference.com/research>

Dr. K. JOHN BABU

LESSON - 8

SAMPLING

OBJECTIVES :

By the end of this topic, you should be able to:

1. Define what is sampling and its concept;
2. Identify the advantages and disadvantages of sampling;
3. Describe sampling terminologies;
4. Identify sample size and selection method; and
5. Differentiate between probability sampling and non-probability sampling techniques.

STRUCTURE :

- 8.1 Introduction
- 8.2 Need for Sampling
- 8.3 Advantages of Sampling
- 8.4 Disadvantages of Sampling
- 8.5 Sampling Terminologies
- 8.6 Characteristics of Good Sample
- 8.7 Sampling Methods
 - 8.7.1 Probability Sampling Methods
 - 8.7.2 Non-Probability Sampling Methods
- 8.8 Bias in Sampling
- 8.9 Sampling Error
- 8.10 Summary
- 8.11 Glossary
- 8.12 Self-Assessment Questions
- 8.13 References and Suggested Readings

8.1 INTRODUCTION :

It would normally be impractical to study a whole population, for example when doing a questionnaire survey. Sampling is a method that allows researchers to infer information about a population based on results from a subset of the population, without having to investigate every individual. Reducing the number of individuals in a study reduces the cost and workload, and may make it easier to obtain high-quality information, but this must be balanced against having a large enough sample size with enough power to detect a true association.

If a sample is to be used, by whatever method it is chosen, the individuals must select are representative of the whole population. This may involve specifically targeting hard-to-reach groups. For example, if the electoral roll for a town was used to identify participants, some people, such as the homeless, would not be registered and therefore excluded from the study by default.

There are several different sampling techniques available, and they can be subdivided into two groups: probability sampling and non-probability sampling. In probability (random) sampling, you start with a complete sampling frame of all eligible individuals from which you select your sample. In this way, all eligible individuals have a chance of being chosen for

the sample, and you will be more able to generalize the results from your study. Probability sampling methods tend to be more time-consuming and expensive than non-probability sampling. In non-probability (non-random) sampling, you do not start with a complete sampling frame, so some individuals have no chance of being selected. Consequently, you cannot estimate the effect of sampling error and there is a significant risk of ending up with a non-representative sample which produces non-generalizable results. However, non-probability sampling methods tend to be cheaper and more convenient, and they are useful for exploratory research and hypothesis generation.

8.2 NEED FOR SAMPLING :

Sometimes it is not feasible to study a whole group or an extremely large group. For example

social work researchers may be interested in learning about mentally challenged children, mentally ill, prison inmates, street children, or some other large group of people. It would be difficult or rather impossible to study all members of the groups. Here comes the process called sampling, which allows studying a manageable number of people from the large group to device inferences that are likely to apply to all the people in the large group. Another reason why we would study a sample is that the results obtained from the sample are more precise and correct than the results obtained from the study of the whole group. There are a few advantages and disadvantages associated with the sampling process.

8.3 ADVANTAGES OF SAMPLING :

1. Sampling reduces the time and cost of research studies.
2. Sampling saves labour.
3. The quality of the study is often better with sampling.
4. Sampling provides much quicker results.

8.4 DISADVANTAGES OF SAMPLING :

A researcher may not find information about the population being studied especially on its characteristics. The research can only estimate or predict them. This means that there is a high possibility of error occurrence in the estimation made.

1. The sampling process only enables a researcher to estimate the actual situation instead of finding the truth. If you take a piece of information from your sampling population, and if your reasoning is correct, your findings should also be accurate to a certain degree.
2. In the absence of thorough knowledge, and sampling methods the result option may be incorrect or misleading.
3. A complicated sampling may require more labour than complete coverage.
4. A pure representation is impossible in most cases

8.5 SAMPLING TERMINOLOGIES :

In sampling, there are a few terminologies that a researcher should be familiar with. For example, let us say you are working on a research project on computing implementation for elderly and disabled citizens for a smart home system. You are supposed to find out the average age of senior and disabled citizens involved in your study.

- a) The community, and families living in the town with smart homes form the population or study population and are usually denoted by the letter N.

- b) The sample group of elderly people or senior citizens and disabled people in the vicinity of the smart home community is called a sample.
- c) The number of elderly people or senior citizens and disabled people you obtain information to find their average age is called the sample size and is usually denoted by the letter n .
- d) The way you select senior citizens and disabled people is called the sampling design or strategy.
- e) Each citizen or disabled person that becomes the basis for selecting your sample is called the sampling unit or sampling element.
- f) A list identifying each respondent in the study population is called a sampling frame. In case when all elements in a sampling population cannot be individually identified, you cannot have a sampling frame for the study population.
- g) Finally, the obtained findings based on the information of the respondents are called sample statistics.

8.6 CHARACTERISTICS OF GOOD SAMPLE :

- a) Representativeness: A sample must be representative of the population. In measurement terms as well as in quality.
- b) Accuracy: Accuracy is defined as the degree to which one must be absent in which sample.
- c) Precision: The sample must yield a precise estimate. The standard error should be minimized.
- d) Size: A good sample must be adequate in size. It should not too small or too big.

8.7 SAMPLING METHODS :

Sampling methods may be classified into two types.

- a) Probability of Random sampling
- b) Non-Probability or Non-Random sampling

Probability sampling is the following types :

This Sampling technique uses randomization to make sure that every element of the population gets an equal chance to be part of the selected sample. It is alternatively known as random sampling.

- a) Simple Random Sampling
- b) Systematic Random Sampling
- c) Stratified Random Sampling
- d) Cluster Sampling
- e) Multi-stage Sampling

Non-Probability sampling may be classified into :

It does not rely on randomization. This technique is more reliant on the researcher's ability to select elements for a sample. The outcome of sampling might be biased and makes it difficult for all the elements of the population to be part of the sample equally. This type of sampling is also known as non-random sampling.

- a) Convenient sampling
- b) Purposive sampling
- c) Quota sampling
- d) Snow-Ball sampling

8.7.1 Probability Sampling Methods :

A) Simple Random Sampling :

The most basic type of probability sampling is simple random sampling. Here, each subject or unit in the population has an equal chance of being selected. In principle, one can use this method for selecting random samples from populations of any size. But in practice, it becomes very cumbersome. In this case, everyone is chosen entirely by chance and each member of the population has an equal chance, or probability, of being selected. One way of obtaining a random sample is to give everyone in a population a number, and then use a table of random numbers to decide which individuals to include. For example, if you have a sampling frame of 1000 individuals, labelled 0 to 999, use groups of three digits from the random number table to pick your sample. So, if the first three numbers from the random number table were 094, select the individual labelled “94”, and so on.

As with all probability sampling methods, simple random sampling allows the sampling error to be calculated and reduces selection bias. A specific advantage is that it is the most straightforward method of probability sampling. A disadvantage of simple random sampling is that you may not select enough individuals with your characteristic of interest, especially if that characteristic is uncommon. It may also be difficult to define a complete sampling frame and inconvenient to contact them, especially if different forms of contact are required (email, phone, post) and your sample units are scattered over a wide geographical area.

B) Systematic Random Sampling :

The most practical way of sampling is to select every n^{th} item on a list. A sampling of this type is known as systematic random sampling. Individuals are selected at regular intervals from the sampling frame. The intervals are chosen to ensure an adequate sample size. If you need a sample size n from a population of size x , you should select every x/n^{th} individual for the sample. For example, if you wanted a sample size of 100 from a population of 1000, select every $1000/100 = 10^{\text{th}}$ member of the sampling frame.

Systematic sampling is often more convenient than simple random sampling, and it is easy to administer. However, it may also lead to bias, for example, if there are underlying patterns in the order of the individuals in the sampling frame, such that the sampling technique coincides with the periodicity of the underlying pattern. As a hypothetical example, if a group of students was being sampled to gain their opinions on college facilities, but the Student Record Department's central list of all students was arranged such that the sex of students alternated between male and female, choosing an even interval (e.g., every 20th student) would result in a sample of all males or all females. Whilst in this example the bias is obvious and should be easily corrected, this may not always be the case.

C) Stratified Random Sampling :

In this method, the population is first divided into subgroups (or strata) that all share a similar characteristic. It is used when we might reasonably expect the measurement of interest to vary between the different subgroups, and we want to ensure representation from all the subgroups. For example, in a study of stroke outcomes, we may stratify the population by sex, to ensure equal representation of men and women. The study sample is then obtained by taking equal sample sizes from each stratum. In stratified sampling, it may also be appropriate to choose non-equal sample sizes from each stratum. For example, in a study of the health outcomes of nursing staff in a county, if there are three hospitals each with different numbers of nursing staff (hospital A has 500 nurses, hospital B has 1000, and

hospital C has 2000), then it would be appropriate to choose the sample numbers from each hospital proportionally (e.g., 10 from hospital A, 20 from hospital B and 40 from hospital C). This ensures a more realistic and accurate estimation of the health outcomes of nurses across the county, whereas simple random sampling would over-represent nurses from hospitals A and B. The fact that the sample was stratified should be taken into account at the analysis stage.

Stratified sampling improves the accuracy and representativeness of the results by reducing sampling bias. However, it requires knowledge of the appropriate characteristics of the sampling frame (the details of which are not always available), and it can be difficult to decide which characteristic(s) to stratify by.

D) Cluster Sampling :

In a clustered sample, subgroups of the population are used as the sampling unit, rather than individuals. The population is divided into subgroups, known as clusters, which are randomly selected to be included in the study. Clusters are usually already defined, for example, individual GP practices or towns could be identified as clusters. In single-stage cluster sampling, all members of the chosen clusters are then included in the study. In two-stage cluster sampling, a selection of individuals from each cluster is then randomly selected for inclusion. Clustering should be considered in the analysis.

A key difference between cluster sampling and stratified random sampling is that the total population is divided into clusters or groups or categories and the sample units are randomly selected from each of these clusters. The ultimate sample consists of all units from these clusters. With cluster sampling, the researcher can divide the state into districts or PIN code areas and select groups from these areas. Cluster sampling can be more efficient than simple random sampling, especially where a study takes place over a wide geographical region.

Disadvantages include an increased risk of bias, if the chosen clusters are not representative of the population, resulting in an increased sampling error.

E) Multi-stage Sampling :

Multi-stage sampling is a further development of the principle of cluster sampling. In many nationwide surveys, researchers use a form of cluster sampling called multi-stage sampling, in which individual households or persons and not groups are selected. In multi-stage sampling, each stage is a separate sampling task. Therefore, the separate stages can be approached with different methods. For example, in a large population, stratification may be useful within the initial steps and simple random procedures in the final ones.

Let us take an example of a city, where one does not have exact data on all the residents. You can divide the city into its zones from available municipal data, divide the zones further into wards, and select clusters from wards, ensuring that all zones are represented. So, if there are 6 zones and 10 wards in each zone (total of 60), you may choose two wards from each zone, and then choose two clusters (each of 10 persons) from each ward. In this way, you can get an adequate sample of 120 respondents, representative of the whole city.

8.7.2 Non-Probability Sampling Methods :

A) Convenience sampling ;

Convenience sampling is perhaps the easiest method of sampling because participants are selected based on availability and willingness to take part. Useful results can be obtained,

but the results are prone to significant bias, because those who volunteer to take part may be different from those who choose not to (volunteer bias), and the sample may not be representative of other characteristics, such as age or sex. Note: volunteer bias is a risk of all non-probability sampling methods.

B) Judgment (or Purposive) Sampling :

Also known as selective, or subjective, sampling, this technique relies on the judgment of the researcher when choosing whom to ask to participate. Researchers may implicitly thus choose a “representative” sample to suit their needs or specific approach to individuals with certain characteristics. This approach is often used by the media when canvassing the public for opinions and in qualitative research. The basic assumption behind purposive sampling is that the subjects are selected for a specific characteristic or quality and eliminate those who fail to meet these criteria.

Judgment sampling has the advantage of being time-and cost-effective to perform whilst resulting in a range of responses (particularly useful in qualitative research). However, in addition to volunteer bias, it is also prone to errors of judgment by the researcher, and the findings, whilst being potentially broad, will not necessarily be representative.

C) Quota Sampling :

This method of sampling is often used by market researchers. Interviewers are given a quota of subjects of a specified type to attempt to recruit. The basic objective of such sampling is the selection of a sample that is like the population in terms of the proportion of certain characteristics. For example, an interviewer might be told to go out and select 20 adult men, 20 adult women, 10 teenage girls, and 10 teenage boys so that they could interview them about their television viewing. Ideally, the quotas chosen would proportionally represent the characteristics of the underlying population.

Whilst this has the advantage of being relatively straightforward and potentially representative, the chosen sample may not be representative of other characteristics that weren't considered (a consequence of the non-random nature of sampling).

D) Snowball Sampling :

This method is commonly used in social sciences when investigating hard-to-reach groups. Existing subjects are asked to nominate further subjects known to them, so the sample increases in size like a rolling snowball. Sometimes, especially in exploratory research, it is difficult to find relevant respondents. Therefore, each respondent is asked to identify one or more respondents, e.g., people with walking sticks; cancer specialists; experts in a particular field, mothers who also volunteer their services, and even other friends who are social media users. Thus, the sample grows to a predetermined number by each respondent identifying others who can be sampled.

Snowball sampling can be effective when a sampling frame is difficult to identify. However, by selecting friends and acquaintances of subjects already investigated, there is a significant risk of selection bias (choosing many people with similar characteristics or views to the initial individual identified).

8.8 BIAS IN SAMPLING :

Five important potential sources of bias should be considered when selecting a sample, irrespective of the method used. Sampling bias may be introduced when:

1. Any pre-agreed sampling rules deviated from

2. People in hard-to-reach groups are omitted
3. Selected individuals are replaced with others, for example, if they are difficult to contact
4. There are low response rates
5. An out-of-date list is used as the sample frame (for example, if it excludes people who have recently moved to an area)

9.9 SAMPLING ERROR :

There is always a need to make a sample error-free. However, often, this may not be possible. The degree to which measurements of the units or subjects selected differ from those of the population as a whole is known as sampling error. The procedure adopted to select the sample introduces the likelihood of sampling error. Sampling error is also known as standard error. In other words, sampling error occurs when measurements taken from a sample do not correspond to what exists in the population. Therefore, while selecting a sampling procedure, the researcher must ensure that the sample design causes a relatively small sampling error and helps to control the bias.

To minimize sampling error, the characteristics of a good sample design should adhere to the following criteria:

- ❖ The sample design must result in a truly representative sample.
- ❖ The sample design selected must ensure minimum sampling error.
- ❖ The sample design selected should consider budget and time constraints.

Sampling error does not depend on the type of sample. Whether convenience, judgment, or probability sampling, the possibility of error always exists when population members vary and something less than the entire population is drawn. With convenience and judgment samples, one can do nothing more than accept that possibility. With probability samples, we can estimate the rate of sampling error, thereby giving us an idea of the risk assumed in accepting the sample as true.

8.10 SUMMARY :

In this unit, we have discussed the methods employed to select a sample from a given population. To make predictions about events, concepts, or phenomena, researchers must perform detailed objective analyses. One procedure to use in such analyses is a census, in which every member of the population is studied. However, conducting a census for each research project is impractical. The most widely used option is to select a random sample from the population and make predictions from it that can be generalized to the population. If the scientific procedure is to provide valid and useful results, close attention must be paid to the methods used in selecting a sample. Probability and non-probability samples are adopted depending on the kind of research, costs, and time. The non-probability samples are accidental, available, volunteer, purposive, and quota samples while the probability methods are simple random sampling, systematic random sampling, stratified random sampling, and cluster sampling. The size of the sample required depends on many factors - the purpose of research, time, and financial constraints. To sum it up, it is important for any researcher to seriously consider the various factors discussed above to select a research sample for any study.

8.11 GLOSSARY :

- ❖ **Population or Universe:** Population or universe is the aggregate of all units possessing certain specified characteristics on which the sample seeks to draw inferences.

- ❖ **Frame:** The frame describes the population in terms of sampling units. It may be a geographical area. In essence a frame lists or maps elements of the universe.
- ❖ **Census:** Census denotes a total enumeration of individuals elements for units in defined population.
- ❖ **Sample:** A Sample is composed of some fractions or part of the total number of elements or units in a defined population.
- ❖ **Design:** Designing means the method by which the sample to choose.
- ❖ **Random:** A mathematical term 'Random' means that every element of the total population has an equal chance of probability of being chosen for the sampling.
- ❖ **Unit:** any population or universe should contain some specifications in terms of content units, extent, and time e.g. "A farmer's household in a district in Punjab in 1975" There is a unit determination in a household and time destination of the population.
- ❖ **Parameter:** The parameter is the value of a variable calculation from the population which is being studied.
- ❖ **Precision:** The precision of a sample is designated by the computation of slandered error.
- ❖ **Stratification:** It makes the segmentation of a sample. It is several data.

8.12 SELF-ASSESSMENT QUESTIONS :

1. Describe the guidelines for selecting samples.
2. Identify how bias can occur during the sample selection process.
3. Identify the importance of sampling in research work.
4. What are the sampling techniques of probability sampling?
5. Distinguish between quota and purposive sampling techniques.

8.13 REFERENCES AND SUGGESTED READINGS :

- 1) Priest H. Susanna (2010). Media Research-An Introduction, Sage Publications, New Delhi.
- 2) Ina Bertrand and Peter Hughes (2005). Media Research Methods-Audiences, Institutions, and Texts, Palgrave Macmillan Publications, New York.
- 3) Wimmer and Dominick (2003). Mass Media Research - An Introduction, Thomson Wadsworth, Singapore.
- 4) Kothari, C. R. (2004). Research Methodology: Methods and Techniques, New Age International Ltd. Publishers, India.
- 5) Wilkinson and Bhandarkar (1990). Methodology and Techniques of Social Research, Himalaya Publishing House, Delhi.
- 6) Easton, V. J. & McColl, J. H. (2007) Statistics Glossary: Sampling [Electronic Version] <http://www.stats.gla.ac.uk/steps/glossary/sampling.html>
- 7) Galloway, A. (1997). Sampling: A Workbook [Electronic version] <http://www.tardis.ed.ac.uk/~kate/qmcweb/scont.htm>
- 8) Trochim, W. K. (2007). Research method tutorials [Electronic version] <http://www.socialresearchmethods.net/kb/sampling.php>

LESSON - 9

METHODS OF DATA COLLECTION

OBJECTIVES :

When you have studied this session, you should be able to:

1. Develop research questions and link them to study designs
2. Understand the differences between quantitative and qualitative research and their application
3. Be familiar with different methods for collecting and analysing qualitative data
4. Be familiar with different methods for collecting quantitative data and basic concepts of probability sampling
5. Understand simple descriptive analyses for quantitative data
6. Interpret multiple sources of data and develop evidence-based conclusions and recommendations

These learning outcomes will equip you to better understand the data collection methods and tools that are used within the overall system. Each of the methods of data collection, as explained later in the module, has specific tools and methods for data collection – some qualitative, some quantitative.

STRUCTURE :

- 9.1 Introduction
- 9.2 Importance of Data
- 9.3 Sources of Data
- 9.4 Objective Data and Subjective Data
- 9.5 Qualitative Data and Quantitative Data
- 9.6 Main Methods of Data Collection
 - 9.6.1 Interview Method
 - 9.6.2 Importance of interview
 - 9.6.3 Characteristics of Interview
- 9.7 Types of Interviews
- 9.8 Advantages of Interview
- 9.9 Limitations of Interview
- 9.10 Questionnaire Method
 - 9.10.1 Characteristics of a Good Questionnaire
 - 9.10.2 Uses of Questionnaire
 - 9.10.3 Limitations of Questionnaires
 - 9.10.4 Types of Questionnaires
- 9.11 Summary
- 9.12 KeyWords
- 9.13 Self-Assessment Questions
- 9.14 References and Suggested Readings

9.1 INTRODUCTION :

The search for answers to research questions is called the collection of data. Data are facts, and other relevant materials, past and present, serving as bases for study and analyses. The data needed for social science research may be broadly classified into

- (a) Data about human beings,
- (b) Data relating to organizations, and
- (c) Data about territorial areas.

Personal data or data related to human beings consist of (1) Demographic and socio-economic characteristics of individuals: Age, sex, race, social class, religion, marital status, education, occupation, income, family size, location of the household, lifestyle, etc., (2) Behavioural Variables: Attitudes, opinions, awareness, Knowledge, practice, intentions, etc. Organizational data consists of an organization's origin, ownership, objectives, resources, functions, performance, and growth. Territorial data are related to geophysical characteristics, resources endowment, population, occupational pattern, infrastructure, structure, degree of development, etc. of spatial divisions like villages, cities, tehsil, districts, states, and the nation.

9.2 IMPORTANCE OF DATA :

The data serve as the bases or raw material for analysis. Without an analysis of factual data, no specific inferences can be drawn on the questions under study. Inference based on imagination or guesswork cannot answer research questions correctly. The relevance adequacy and reliability of data determine the quality of the findings of a study.

Data form the basis for testing the hypotheses formulated in a study. Data also provide the facts and figures required for constructing measurement scales and tables, which are analysed with statistical techniques. Inferences on the results of statistical analysis and tests of significance provide the answers to research questions. Thus, the scientific process of measurements, analysis, testing, and inferences depended on relevant data availability and accuracy. Hence, the importance of data for any research studies.

9.3 SOURCES OF DATA :

The sources of data may be classified into (a) primary sources and (b) secondary sources. Primary sources are sources from which the researcher directly collects data that have not been previously collected. Primary data are first-hand information collected through various methods such as interviewing, mailing, observation, etc. Secondary sources contain data that have been collected and compiled for another purpose. The secondary sources consist of readily available compendia and already compiled statistical statements and reports whose data may be used by researchers in their studies. E.g., census reports, annual reports, and financial reports. Secondary sources consist of not only published records and reports, but also unpublished records.

9.4 OBJECTIVE DATA AND SUBJECTIVE DATA :

Objective data is independent of any single person's opinion, whereas subjective data can be an individual's opinion or it can be dependent upon the researcher.

9.5 QUALITATIVE DATA AND QUANTITATIVE DATA :

Qualitative data is the description of things made without assigning numeric value. For example, facts generated from the unstructured interview. It needs the researcher's

interpretation. Quantitative data entails measurements in which numbers are used directly to represent properties of things. It is ready for statistical analysis. A larger sample is required in quantitative data, and with proper sampling design, the ability to generalize is also high.

9.6 MAIN METHODS OF DATA COLLECTION :

Most research studies collect fresh data from the respondents even though already existing data are utilized for developing the research design or supplementing the data to be collected. There are various methods of data collection. 'Method is different from a 'Tool' while a method refers to the way or mode of gathering data, a tool is an instrument used for the method. For example, a schedule is used for interviewing. The important methods are **(a) observation, (b) interviewing, (c) mail survey (D) schedule.** Observations involve gathering data relating to the selected research by viewing and or listening. Interviewing involves a face-to-face conversation between the investigator and the respondent. Mailing is used for collecting data by getting questionnaires completed by respondents. Experimentation involves the study of independent variables under controlled conditions. Experiments may be conducted in a laboratory or the field in a natural setting. Simulation involves the s creation of an artificial situation similar to the actual life situation. Projective methods aim at drawing inferences on the characteristics of respondents by presenting their stimuli. Even the method has its advantages and disadvantages.

In primary data collection, you collect the data yourself using qualitative and quantitative methods. The key point here is that the data you collect is unique to you and your research and, until you publish, no one else has access to it. There are many methods of collecting primary data. The main methods include –

- Interview
- Questionnaire
- Observation
- Case Study

9.6.1 Interview Method ;

Interviewing is one of the prominent methods of data collection. Interviewing involves asking questions and getting answers from participants in a study. Interviewing has a variety of forms including individual, face-to-face interviews, and face-to-face group interviewing. The asking and answering of questions can be mediated by the telephone or other electronic devices (e.g., computers). Interview emerged as a tool of data collection by the turn of the last century and has by now become an integral part of social research. During earlier times, interviews were conducted more like a probing conversation. Guided by a careful observer this method was used as a powerful instrument for obtaining information. It involves not only conversation, but also learning from the respondent's gestures, facial expressions and pauses, and his environment. Interviewing requires face-to-face contact or contact over the telephone and calls for interviewing skills. It is done by using a structured schedule or an unstructured guide.

9.6.2 Importance of interview :

Interviewing may be used either as a main method or as a supplementary one in studies of persons. Interviewing is the only suitable method for gathering information from illiterate or less educated respondents. It is useful for collecting a wide range of data from factual demographic data to highly personal and intimate information relating to a person's opinions, attitudes, values, beliefs, experience, and future intentions. When qualitative information is required or probing is necessary to draw out fully, and then interviewing is

required. Where the area covered for the survey is compact, or when a sufficient number of qualified interviews are available, a personal interview is feasible.

The interview is often superior to other data-gathering methods. People are usually more willing to talk than to write. Once rapport is established, even confidential information may be obtained. It permits probing into the context and reasons for answers to questions. Interviews can add flesh to statistical information. It enables the investigator to grasp the behavioural context of the data furnished by the respondents. It permits the investigator to seek clarifications and brings to the forefront those questions, that for one reason or another, respondents do not want to answer.

9.6.3. Characteristics of Interview :

Interview as a method of data collection has certain characteristics.

1. The participants- the interviewer and the respondent- are strangers. Hence, the investigator has to get him appropriately introduced to the respondent.
2. The relationship between the participants is a transitory one. It has fixed beginning and termination points. The interview proper is a fleeting, momentary experience for them.
3. The interview is not a mere causal conversational exchange, but a conversation with a specific purpose, viz., obtaining information relevant to the study.
4. The interview is a mode of obtaining verbal answers to questions put verbally.
5. The interaction between the interviewer and the respondent need not necessarily be on a face-to-face basis, because interviews can be conducted over the telephone also.
6. Although the interview is usually a conversation between two persons, it need not be limited to a single respondent. It can also be conducted with a group of persons, such as family members, a group of children, or a group of customers, depending on the requirements of the study.
7. The interview is an interaction process. The interaction between the interviewer and the respondent depends upon how they perceive each other.

9.7 TYPES OF INTERVIEWS :

The interviews may be classified into (a) structured or directive interview, (b) unstructured or non-directive interview, (c) focused interview, (d) clinical interview, and (d) depth interview.

- ❖ **Structured, directive interview** : This is an interview made with a detailed standardized schedule. The same questions are put to all the respondents and in the same order. Each question is asked in the same way in each interview, promoting measurement reliability. This type of interview is used for large-scale formalized surveys.
- ❖ **Unstructured or non-directive interview** : This is the least structured one. The interviewer encourages the respondent to talk freely about a given topic with a minimum of prompting or guidance. In this type of interview, a detailed pre-schedule is not used. Only a broad interview guide is used.
- ❖ **Focused Interview** : This is a semi-structured interview where the investigator attempts to focus the discussions on the actual effects of a given experience to which the respondents have been exposed. It takes place with the respondents known to have been involved in a particular experience.
- ❖ **Clinical Interview** : This is like the focused interview but with subtle differences. While the focused interview is concerned with the effects of a specific experience, the clinical interview is concerned with broad underlying feelings or motivations or with the course of the individual's life experiences.

- ❖ **Depth Interview** : This is an intensive and searching interview aiming at studying the respondent's opinion, emotions, or convictions based on an interview guide. This requires much more training in interpersonal skills than structured interviewing. This deliberately aims to elicit unconscious as well as extremely personal feelings and emotions.

9.8 ADVANTAGES OF INTERVIEW :

There are several real advantages to a personal interview.

- ❖ First, the greatest value of this method is the depth and detail of information that can be secured.
- ❖ Second, the interviewer can do more to improve the percentage of responses and the quality of the information received than other methods. He can note the conditions of the interview situations, and adopt appropriate approaches to overcome such problems as the respondent's unwillingness, incorrect understanding of questions, suspicion, etc.
- ❖ Third, the interviewer can gather other supplemental information like economic level, living conditions, etc. Through observation of the respondent's environment.
- ❖ Fourth, the interviewer can use special scoring devices, and visual materials, and like to improve the quality of interviewing.
- ❖ Fifth, the accuracy and dependability of the answers given by the respondent can be checked by observation and probing.
- ❖ Last, the interview is flexible and adaptable to individual situations. Even more, control can be exercised over the interview situation.

9.9 Limitations of Interview :

- ❖ First, the Interview is not free from limitations. Its greatest drawback is that it is costly both in money and time.
- ❖ Second, the interview results are often adversely affected by the interviewer's mode of asking questions and interactions, incorrect recording, and also the respondent's faulty perception, faulty memory, inability to articulate, etc.
- ❖ Third, certain types of personal and financial information may be refused in face-to-face interviews. Such information might be supplied more willingly on mail questionnaires, especially if they are to be unsigned.
- ❖ Fourth, the interview poses the problem of recording information obtained from the respondents, no foolproof system is available. Note-taking is invariably distracting to both respondent and the interviewer and affects the thread of the conversation.
- ❖ Last, the interview calls for highly skilled interviewers. The availability of such persons is limited and the training of interviewers is often a long and costly process.

9.10 QUESTIONNAIRE METHOD :

A questionnaire is a research instrument consisting of a series of questions and other prompts to gather information from respondents. Although they are often designed for statistical analysis of the responses, this is not always the case. The questionnaire was invented by Sir Francis Galton (1822 - 1911). It is a list of questions related to one topic. It may be defined as; Questionnaires have advantages over some other types of surveys in that they are cheap, do not require as much effort from the questioner as verbal or telephone surveys, and often have standardized answers that make it simple to compile data.

As a type of survey, questionnaires also have many of the same problems relating to question construction and wording that exist in other types of opinion polls. Types: A distinction can be made between questionnaires with questions that measure separate

variables, and questionnaires with questions that are aggregated into either a scale or index. Questionnaires within the former category are commonly part of surveys, whereas questionnaires in the latter category are commonly part of tests. Questionnaires with questions that measure separate variables, could, for instance, include questions on preferences (e.g. political party) behaviors (e.g. food consumption) facts (e.g. gender). Questionnaires with questions that are aggregated into either a scale or index include for instance questions that measure – latent traits (e.g., personality traits such as extroversion) attitudes (e.g. towards immigration) an index (e.g. Social Economic Status).

9.10.1 Characteristics of a Good Questionnaire :

1. It deals with a significant topic.
2. Its significance is carefully stated on the questionnaire itself or its cover letter.
3. It seeks only that data which cannot be obtained from the resources like books, reports, and records.
4. It is as short as possible, only long enough to get the essential data.
5. It is attractive, neatly arranged, and duplicated or printed.
6. Directions are clear and complete and, important terms are clarified.
7. The questions are objective, with no clues, hints, or suggestions.
8. Questions are presented in order from simple to complex.
9. Double negatives, adverbs, and descriptive adjectives are avoided.
10. Double-barrelled questions or putting two questions in one question are also avoided.
11. The questions carry an adequate number of alternatives.
12. It is easy to tabulate, summarize and interpret.

9.10.2 Uses of Questionnaire :

1. A questionnaire is a popular means of collecting different kinds of data in research. It is widely used in educational research to obtain information about certain conditions and practices and to inquire into the opinions of an individual or a group.
2. A questionnaire is administered personally either individually or to a group of individuals or is mailed to them to save a great deal of time and money in travel. In the former situation, the person administering the tool has an opportunity to establish rapport with the respondent to explain the purpose of the study to the respondents and explain the meaning of questions that may not be clear to them. In the latter situation, the mailed questionnaire is mostly used when the individuals cannot be contacted personally. The range of administration of a mailed questionnaire may be national or international.
3. Questionnaires are used both to initiate a formal inquiry and also to supplement and check data previously accumulated. They may pertain to studies of economic or social problems, measurement of opinion on public issues or events, studies of administrative policies and changes, studies on the cost of living, consumer expenditure, child welfare, and numerous other issues.

9.10.3 Limitations of Questionnaires :

1. A questionnaire cannot be used with children and illiterates.
2. The return of the mailed questionnaire is often as low as 40 per cent to 50 per cent. As a result of this poor response, the data obtained are sometimes of limited validity. The respondents who return the questionnaires may not be representative of the entire group. It will make the sample a biased one and thus vitiate the findings.
3. Sometimes respondents may not like to respond in writing to questions of intimate and confidential nature or questions involving controversial issues. For example, it has been

experienced that people avoid questions related to marriage, and government servants avoid answering questions about policy matters of the government.

4. It is sometimes difficult to formulate and phrase questions on certain complex and delicate problems.
5. There is no check on the respondent who misinterprets a question or gives incomplete or indefinite responses.
6. Sometimes the respondent may modify his/her earlier/original responses to the questions when he/she finds that his/her responses to letter questions are contradicting the previous ones.

9.10.4 Types of Questionnaires :

Questionnaires can be classified in various ways. Here we confine ourselves to structured and unstructured questionnaires.

Structured questionnaires are those which pose definite and concrete questions. They are prepared well in advance and not on the spot. Additional questions may be used only when there is a need to clarify vague or inadequate replies by respondents or when further details are needed. The form of questions may require responses that are either closed or open.

Closed-form of questionnaires is used when categorized data are required. They include a set of questions to which a respondent can reply in a limited number of ways - 'yes', 'no', 'no-opinion', or an answer from a short list of possible responses. He/she is asked to put a tick (✓) mark in a space provided on the answer sheet or is requested to underline a response. Sometimes he/she is asked to insert brief answers of his/her own. The open-ended responses, on the other hand, are free and spontaneous expressions by the respondents to the questions posed to him/her. The open-ended responses are used mainly for the intensive study of a limited number of cases or preliminary exploration of new problems and situations. At times, the respondent is asked to write a descriptive essay and express his/her viewpoints or report on details and events, without restrictions imposed as in the case of closed questions.

Unstructured questionnaires are frequently referred to as interview guides. They also aim at precision and contain definite issues that are covered while conducting an interview. Flexibility is the chief advantage of unstructured questionnaires. It is designed to obtain viewpoints, opinions, and attitudes and to show relationships between various types of information which might escape notice under more mechanical types of interrogation. No predetermined responses are provided; instead, free responses are solicited.

9.11 SUMMARY:

1. A questionnaire is a Performa containing a sequence of questions to elicit information from the interviewees.
2. A questionnaire consists of a number of questions printed in a definite order on a form.
3. Questionnaires are either personally administered or sent to the respondents by post or e-mail.
4. There are two types of questionnaire: Structured and Unstructured questionnaire.
5. There are three main aspects of questionnaire viz. (i) The general form, (ii) The question sequence, and (iii) The question wording.
6. Physical forms of the questionnaire should be attractive.
7. Pre-testing is necessary to check the effectiveness of the questionnaire.
8. Questionnaire must contain simple but straightforward directions for the respondents so that they may not feel any difficulty in answering the question.

9. Questionnaire form is essential while measuring media exposure.

9.12 KEY WORDS :

- ❖ **Questionnaire:** A questionnaire is a performa containing a sequence of questions to elicit information from the interviewees. A questionnaire consists of a number of questions printed in a definite order on a form. Questionnaire must contain simple but straightforward directions for the respondents so that they may not feel any difficulty in answering the question.
- ❖ **Questionnaire Administration:** Questionnaires are either personally administered or sent to the respondents by post or e-mail.
- ❖ **Types of Questionnaire:** There are two types of questionnaire: Structured and Unstructured questionnaire.

9.13 SELF-ASSESSMENT QUESTIONS :

1. What is questionnaire? Explain briefly how to design a questionnaire.
2. What are the main aspects of a questionnaire? Write the advantages and disadvantages of a mailed questionnaire.
3. Prepare a questionnaire to measure the media exposure of literate adults in your locality.
4. Write short notes on:
 - a. Pre-testing
 - b. Structural questionnaire
 - c. Physical form of questionnaire
5. What are the basic differences between a schedule and a questionnaire?
6. Define a target group and design questions to collect information on the following topics:
 - i. Political party affiliation.
 - ii. Attitudes towards television advertisements
 - iii. Attitudes towards newspaper editorials.
 - iv. Attitudes towards the frequency of television commercials
 - v. Public television viewing habits.

9.14 REFERENCES AND SUGGESTED READINGS:

1. Introduction to Communication Research by *John C. Reinard*.
2. Mass Media IV by *Ray Eldon Hiebert et al*
3. Mass Media Research by *Roger D. Wimmer & Joseph R. Dominick*
4. Doing Media Research - An Introduction by *Susanna Houring Priest*
5. Introduction to Communication Research by *John C Reinard* (Benchmark, 1994)
6. Practical Marketing Research by *Jefferey L Pope* (Amacom, 1993)
7. Introduction to Survey Research by *Pamela L Alreck & Robert B Settle* (Irwin, 1995)

Dr. Tarakeswara Rao Ippili

LESSON - 10

OBSERVATION AND CASE STUDY METHODS

OBJECTIVES :

Observation becomes scientific when it

- serves a formulated research purpose,
- is planned deliberately,
- is recorded systematically, and
- is subjected to checks and controls on validity and reliability

Validity refers to the extent to which the recorder observations accurately reflect the construct they are intended to measure. Validity is assessed by examining how well the observations agree with alternative measures of the same construct. Reliability entails consistency and freedom from measurement error.

STRUCTURE :

- 10.1 Introduction
- 10.2 Use of Observational Method
 - 10.2.1 Characteristics of the Observation Method
 - 10.2.2 Types of Observation
 - 10.2.3 Observation Tools and Recording Devices
 - 10.2.4 Advantages of Observation
 - 10.2.5 Limitations of Study
- 10.3 Case Study Method
 - 10.3.1 Definition of Case Study
 - 10.3.2 Characteristics of Case Study Method
 - 10.3.3 Design of Case Study
 - 10.3.4 Category of Case Study
- 10.4 Summary
- 10.5 Glossary
- 10.6 Self-Assessment Questions
- 10.7 Reference and Suggested Readings

10.1 INTRODUCTION :

Observation is a basic method of getting information about the world around us. Observation is part and parcel of our daily life but many types of data required as evidence to support social research are also obtained through the observational method. The greatest asset of the observational technique is that it is possible to record the actual occurrence of social events. While many research techniques depend mainly if not entirely on recalling past events, observational method yields such as are related to real-life situations. A trained researcher can even observe and record all the minor details of a community with the help of this technique which to others might seem insignificant.

Observation means viewing or seeing. Most of such observations are just causal and have no specific purpose. But observation in a method of data collection is different from such causal viewing. Observation may be defined as a systematic viewing of a specific phenomenon in its proper setting for the specific purpose of gathering data for a particular

study. Observation as a method includes both 'seeing' and 'hearing'. It is accompanied by perceiving as well.

Observation is a classical method of scientific inquiry. Observation also plays a major role in formulating and testing a hypothesis in social sciences. Behavioural scientists observe interactions in small groups; political scientists observe the behaviour of political leaders and political institutions.

Observation may serve a variety of research purposes. It can be used in exploratory research to develop a preliminary understanding of social phenomena. It can be applied to study real-life situations as well as to conduct experimental research. Again, it can simply be used to collect supplementary data in support of other tools of data collection. Observation includes the most causal and uncontrolled experiences as well as exact recording as is done in experimentation. Observation is useful for studying simpler as well as complex research problems.

10.2 USE OF OBSERVATIONAL METHOD :

There are a variety of reasons for collecting observational data. Some of these reasons include

- i. When the nature of the research question to be answered is focused on answering a how- or what-type question.
- ii. When the topic is relatively unexplored and little is known to explain the behaviour of people in a particular setting.
- iii. When understanding the meaning of a setting in a detailed way is valuable.
- iv. When it is important to study a phenomenon in its natural setting.
- v. When self-report data (asking people what they do) is likely to be different from actual behaviour (what people do). One example of this is seen in the difference between self-reported versus observed preventive service delivery in health care settings.
- vi. When implementing an intervention in a natural setting, observation may be used in conjunction with other quantitative data collection techniques.

Observational data can help researchers evaluate the fidelity of an intervention across settings and identify when 'stasis' has been achieved.

10.2.1 Characteristics of the Observation Method :

Observation as a method of data collection has certain characteristics.

1. It is both a physical and mental activity. The observing eye 'catches' many things which are slighted, but attention is focused on data that are pertinent to the given study.
2. Observation is selective. The researcher does not observe anything and everything but selects the range of things to be observed based on the nature, scope, and objectives of his study
3. Observation is purposive and not casual. It is made for the specific purpose of nothing things relevant to the study.
4. It captures the natural social context in which a person's behaviour occurs.
5. It grasps the significant events and occurrences that affect the social relations of the participants.
6. Observation should be exact and be based on standardized tools of research such as observation schedules, social-metric scales, and precision instruments if any.

10.3.2. Types of Observation :

Observation may be classified in different ways. Concerning the investigator's role, it may be classified into

- ❖ participant observation, and
- ❖ non-participant observation, in terms of mode of observation, it may be classified into
- ❖ direct observation and
- ❖ indirect observation. Concerning the rigor of the system adopted, observation is classified into
- ❖ controlled observation, and
- ❖ uncontrolled observation.

Participant Observation : In this observation, the observer is a part of the phenomena or group which is observed and he acts as both an observer and a participant. The persons who are observed should not be aware of the researcher's purpose. Then only their behaviour will be 'natural'. The concealment of the research objective and researcher's identity is justified on the ground that it makes it possible to study certain aspects of the group culture which are not revealed to outsiders makes it possible to study certain aspects of the group's culture which are not revealed to outsiders.

The advantages of Participant Observation are :

1. The observer can understand the emotional reactions of the observed group, and get a deeper insight into their experiences.
2. The observer will be able to record context which gives meaning to the observed behaviour and heard statements.

Disadvantages :

1. The participant observer narrows his range of observation.
2. To the extent that the participant observer participates emotionally, objectivity is lost.

Because of the above limitations, participant observation is generally restricted to those cases where non-participant observation is not practical.

- ❖ **Non-participant Observation :** In this method, the observer stands apart and does not participate in the phenomenon observed. Naturally, there is no emotional involvement on the part of the observer. This method calls for skill in recording observations in an unnoticed manner.
- ❖ **Direct Observation :** This means observation of an event personally by the observer when it takes place. This method is flexible and allows the observer to see and record subtle aspects of events and behaviour as they occur. He can be free to shift places and change the focus of the observation. A limitation of this method is that the observer's perception circuit may not be able to cover all relevant events when the latter moves quickly, resulting in the incompleteness of the observation.
- ❖ **Indirect Observation :** This does not involve the physical presence of the observer, and the recording is done by mechanical, photographic, or electronic devices. This method is less flexible than direct observation, but it is less biased and less erratic in recording accuracy. It also provides a permanent record for an analysis of different aspects of the event.
- ❖ **Controlled Observation :** This involves standardization of observational technique and exercise of maximum control over extrinsic and intrinsic variables by adopting experimental design and systematically recording observations. Controlled

observation is earned either in the laboratory or in the field. It is typified by clear and explicit decisions on what, how, and when to observe. It is primarily used for inferring causality, and testing causal hypotheses.

- ❖ **Uncontrolled Observation** : This does not involve control over extrinsic and intrinsic variables. It is primarily used for descriptive research. Participant observation is a typical uncontrolled one.
- ❖ **Planning of Observation** : The use of the observation method requires proper planning.
 - First, the researcher should carefully examine the relevance of the observation method to the data needs of the selected study.
 - Second, he must identify the specific investigative questions which call for use of the observation method. These determine the data collected.
 - Third, he must decide the observation content, viz., specific conditions, events, and activities that have to be observed for the required data. The observation content should include the relevant variables.
 - Fourth, for each variable chosen, the operational definition should be specified.
 - Fifth, the observation setting, the subjects to be observed, the timing and mode of observation, recording, procedure, recording instruments to be used, and other details of the task should be determined.
 - Last, observers should be selected and trained. The persons to be selected must have sufficient concentration powers, strong memory power, and intrusive nature. Selected persons should be imparted both theoretical and practical training.

10.2.3 Observation Tools and Recording Devices :

Systematic observation requires the use of an observation schedule (or observation area), a diary, and various mechanical recording devices.

- **Schedule**: The data requirements are identified by analysing the core of the problem, the objectives of the study, the investigative questions, the hypothesis, and the operational definition of concepts and out of the data requirements, items of data to be collected through observation are identified. A schedule is then constructed, covering those items of data.
- It should be constructed in such a manner as to make it possible to record the necessary information easily and correctly. Enough space should be provided for recording observations each time. The item should appear in logical groupings and in the order in which the observer would observe them.
- **Field Observation**: This may take the form of a diary or cards. Each item of observation is recorded under the appropriate sub-heading. At the time of observation, rough noting may be made, and at the end of the day, the full log may be made. The card system is flexible and facilitates the arrangement and rearrangement of items in any desired order.
- **Mechanical Devices**: These may include cameras, tape recorders, videotapes, and electronic devices. Still, motion, sound, color, and time-lapse cameras give a permanent record of events, The microscopic and telescopic lenses may be used in cameras.

10.2.4 Advantages of Observation :

Observation has certain advantages

1. The main virtue of observation is its directness; it makes it possible to study behaviour as it occurs. The researcher need not ask people about their behaviour and interactions, he can simply watch what they do and say.
2. Data collected by observation may describe the observed phenomena as they occur in their natural settings. Other methods introduce elements or artificiality into the researched situation.
3. Observation is more suitable for studying subjects who are unable to articulate meaningfully.
4. Observation is less demanding of the subjects and has a less biased effect on their conduct than questioning.

10.2.5 Limitations of Study :

1. Observation is of no use in studying past events or activities. One has to depend upon documents or narrations by people for studying such things.
2. It is not suitable for studying opinions and attitudes.
3. Observation poses difficulties in obtaining a representative sample.
4. Observation is a slow and expensive process. Requiring human observation and/or costly surveillance equipment.

10.3 CASE STUDY METHOD :

Case studies are in-depth investigations of a single person, group, event, or community. Typically, data are gathered from a variety of sources and by using several different methods (e.g. observations & interviews). The case study research method originated in clinical medicine (the case history, i.e., the patient's personal history). The case study method often involves simply observing what happens to, or reconstructing 'the case history' of a single participant or group of individuals (such as a school class or a specific social group), i.e., the idiographic approach. Case studies allow a researcher to investigate a topic in far more detail than might be possible if they were trying to deal with many research participants (nomothetic approach) with the aim of 'averaging.'

In a case study, nearly every aspect of the subject's life and history is analysed to seek patterns and causes for behaviour. The hope is that learning gained from studying one case can be generalized to many others. Unfortunately, case studies tend to be highly subjective and it is difficult to generalize results to a larger population.

In a case study information (commonly known as case data) may be gathered exhaustively of an entire life cycle of a social unit or a definite section of it. Whether a section or the whole of life is studied, the aim is to ascertain the natural history, that is, an account of the generic development of a person or group, or whatever constitutes the social unit in a particular study, revealing the factors that modelled the life of the unit within its cultural setting (Young, 1966). Because of its aid in studying behaviour in specific, precise detail, Burgess (1949) termed the case study method, "the social microscope".

10.3.1 Definition of Case Study ;

The case study method enables a researcher to closely examine the data within a specific context. In most cases, a case study method selects a small geographical area or a very limited number of individuals as the subjects of study. Case studies, in their true essence, explore and investigate contemporary real-life phenomena through detailed

contextual analysis of a limited number of events or conditions, and their relationships. Yin (1984:23) defines the case study research method “as an empirical inquiry that investigates a contemporary phenomenon within its real-life context; when the boundaries between phenomenon and context are not evident; and in which multiple sources of evidence are used.”

In some case studies, an in-depth longitudinal examination of a single case or event is used. The longitudinal examination provides a systematic way of observing the events, collecting data, analysing information, and reporting the results over a long period. For instance, studies on child language development can be conducted using this longitudinal case study method. Data collected through observations are recorded to ascertain the language development of a child. In another example, a researcher conducting a case study may examine the reading processes of only one subject over some time. In other words, a case study is a unique way of observing any natural phenomenon which exists in a set of data.

10.3.2 Characteristics of Case Study Method :

- i. Case study research is not sampling research. Selecting cases must be done to maximize what can be learned in the period available for the study.
- ii. The unit of analysis is a critical factor in the case study. It is typically a system of action rather than an individual or group of individuals. Case studies tend to be selective, focusing on one or two issues that are fundamental to understanding the system being examined.
- iii. Case studies are multi-perspectives analyses. This means that the researcher considers not just the voice and perspective of the actors, but also the relevant groups of actors and the interaction between them. This one aspect is a salient point in the characteristic that case studies possess. They give a voice to the powerless and voiceless.

10.3.3 Design of Case Study :

Since the case study method receives criticism in terms of its lack of robustness as a research tool, crafting the design of case studies is of paramount importance. Researchers can adopt either a single-case or multiple-case design depending on the issue in question. In cases where there are no other cases available for replication, the researcher can adopt the single-case design. For instance, a social study on the effects of the collapse of the Super tech Twin Towers in Noida in 2022, or the effects of the tsunami in Chennai in 2004 can be conducted using a single-case design, where events are limited to a single occurrence.

The multiple-case design, on the other hand, can be adopted with real-life events that show numerous sources of evidence through replication rather than sampling logic. According to Yin (1994), the generalization of results from case studies, from either single or multiple designs, stems from theory rather than from populations. By replicating the case through pattern-matching, a technique linking several pieces of information from the same case to some theoretical proposition (Campbell, 1975), the multiple-case design enhances and supports the previous results. This helps raise the level of confidence in the robustness of the method. For instance, research on dyslexic children with reading problems requires several replications that can be linked to a theory before conclusive results are generalized.

10.3.4 Category of Case Study :

There are several categories of the case study. Yin (1984) notes three categories, namely exploratory, descriptive, and explanatory case studies.

First, exploratory case studies are set to explore any phenomenon in the data which serves as a point of interest to the researcher. For instance, a researcher conducting an exploratory case study on an individual's reading process may ask general questions, such as, "Does a student use any strategies when he reads a text?" and "if so, how often?". These general questions are meant to open the door for further examination of the phenomenon observed.

Second, descriptive case studies are set to describe the natural phenomena which occur within the data in question, for instance, what different strategies are used by a reader and how the reader uses them. The goal set by the researcher is to describe the data as they occur. The challenge of a descriptive case study is that the researcher must begin with a descriptive theory to support the description of the phenomenon or story. If this fails there is the possibility that the description lacks rigour and that problems may occur during the project.

Third, explanatory case studies examine the data closely both at a surface and deep level to explain the phenomena in the data. For instance, a researcher may ask the reason why a student uses an inference strategy in reading. Based on the data, the researcher may then form a theory and set to test this theory

a). Advantages of Case Study :

1. The examination of the data is most often conducted within the context of its use that is, within the situation in which the activity takes place.
2. The variations in terms of intrinsic, instrumental, and collective approaches to case studies allow for both quantitative and qualitative analyses of the data.
3. The detailed qualitative accounts often produced in case of studies not only help to explore or describe the data in a real-life environment but also help to explain the complexities of real-life situations which may not be captured through experimental or survey research.

b). Disadvantages of Case Studies :

1. Case studies are often accused of a lack of rigour.
2. Case studies provide very little basis for scientific generalization since they use a small number of subjects, some conducted with only one subject.
3. These are often labelled as being too long, difficult to conduct, and producing a massive amount of documentation.
4. A common criticism of the case study method is its dependency on a single case exploration making it difficult to reach a generalizing conclusion.

10.4 SUMMARY :

Research is a systematic effort to secure answers to certain 'questions'. In the case of communication or media research, the questions are related to media. It involves the description, analysis, and interpretation of media or communication events and messages related. Most of the methods applicable to social research are used for communication research. While fields like history, literature, etc. use only specific methods, communication research uses all kinds of methods.

Also individual studies in the field of communication research use multiple methods. behaviour. This session has taken you through the process of identifying research questions and selecting appropriate methodologies. You now hopefully have a better understanding of the difference between quantitative and qualitative data collection methods and associated

benefits and limitations. We also introduced you to some common methods and techniques of data analysis for both quantitative and qualitative research. This session is useful and will draw on it to develop systematic investigations that can be used to improve the quality, impact, and accountability of programs.

10.5 GLOSSARY :

- ❖ **Assessment:** Often used as a synonym for evaluation. The term is sometimes recommended for processes that are focused on quantitative or testing approaches.
- ❖ **Central tendency:** A representative or typical value for the dataset. It can be represented by the mean (or average), median (a middle point that divides the data set into two halves), or mode (most frequent value) depending on the data and your needs.
- ❖ **Closed-ended questions:** A broad category of questions that provide a set of possible responses from which to choose.
- ❖ **Data analysis:** The process by which meaning, themes, and useful information are extracted from raw quantitative or qualitative data.
- ❖ **Descriptive data:** Information and findings expressed in words; unlike statistical data, which are expressed in numbers.
- ❖ **Evaluation:** The systematic collection of information to answer important questions about activities, characteristics, and outcomes of a program.
- ❖ **Frequency distributions:** A tabulation of the possible values for a variable (e.g., levels of satisfaction) and the number or range of observations that fall into each of the possible value categories.
- ❖ **Findings:** Summaries, impressions, or conclusions reached after an examination or investigation of data.
- ❖ **Generalise:** The ability to make statements and draw conclusions that can have a general application.
- ❖ **Indicators:** Measurable elements that tell you, or indicate, that the program efforts are successful. Indicators help to define what information must be collected to answer evaluation questions.
- ❖ **Inputs:** Resources, such as costs, materials, and personnel, required to carry out the program
- ❖ **Observation:** A method to gather information about things that can be observed. For example, by visiting a participant's workplace, you can directly collect information on the physical surroundings.
- ❖ **Open-ended questions:** Questions that stimulate free thought by asking people to write their answers in their own words rather than choosing from a predetermined set of response options.
- ❖ **Outcomes:** Outcomes are the positive differences the program makes in the lives of people and communities. Outcomes are changes in beliefs, attitudes, knowledge, and action the program produces.
- ❖ **Outputs:** What the program is intended to produce, such as the number of services provided or the number of people reached, and other results you can count, observe, or measure.
- ❖ **Quantitative data:** Information that can be counted or expressed numerically. Examples of quantitative data are age, personal income, and amount of time.
- ❖ **Qualitative data:** Thoughts, observations, opinions, or other data expressed in words. Qualitative data typically come from asking open-ended questions to which the answers are not limited by a set of choices or a scale.

- ❖ **Questionnaires:** An instrument useful in gathering focused, limited information from a specific population. Questionnaires ask questions in a standardized format that allows consistency and the ability to aggregate responses.
- ❖ **Reliability:** The consistency of measurement or the degree to which results obtained by an evaluation instrument can be reproduced.
- ❖ **Research:** To study (something) systematically, gathering and reporting on detailed and accurate information.
- ❖ **Response rate:** The percentage of people who respond to a survey. For example, if 100 people receive a survey questionnaire and 73 people complete the questionnaire, the response rate is 73/100, or 73%.
- ❖ **Validity:** The degree of accuracy of a measurement. For survey instruments, validity refers to the degree to which the instrument measures what it is intended to measure.

10.6 SELF-ASSESSMENT QUESTIONS :

- 1) Enumerate the different methods of collecting data.
- 2) Which one is the most suitable for conducting an inquiry regarding family welfare programs in India? Explain its merits and demerits.
- 3) How would you schedule observations?
- 4) Write a detailed note on the various methods of media research.
- 5) Describe how you might use public records to answer the same research question. What might be some limitations of this approach?
- 6) Discuss the interview method in detail.
- 7) Write a detailed note on the observation methods of media research.

10.7 REFERENCE AND SUGGESTED READINGS :

- 1) Bailey, Kenneth, D. (1978), *Methods of Social Research*, The Free Press, London.
- 2) Baker, L. Therese (1988), *Doing Social Research*, McGraw Hill, New York.
- 3) Black, James A., and Champion, Dean J. (1976), *Methods and Issues in Social Research*, John Wiley, New York.
- 4) Burgess (1949), "Research Method in Sociology," in Georges Gurvitch and W.E. Moore (Eds.) *Twentieth Century Sociology*.
- 5) Festinger, L. and Katz., D. (Eds.) (1953), *Research Methods in the Behavioral Sciences*, The Dryden Press, New York.
- 6) John C. Reinard (2005). *Introduction to Communication Research* McGraw-Hill.
- 7) Roger D. Wimmer & Joseph R. Dominick (2015). *Mass Media Research: An Introduction*, Cengage Learning India Private Limited.
- 8) Susanna H. Priest, (1996). *Doing Media Research: An Introduction*, Sage.

Dr. I. Tharakeswara Rao

LESSON - 11

STATISTICAL METHODS IN MEDIA RESEARCH

OBJECTIVES :

The objectives of this lesson are as follows :

- To describe Measures of Central Tendency in Media Research
- To elaborate Mean comprehensively
- To analyses Mode in details
- To explain Media largely and widely
- To enlist the formulae of Mean, Median and Mode.

STRUCTURE :

- 11.1 Introduction
- 11.2 Statistical Methods (Summary Measures)
 - 11.2.1 Measures of Central Tendency
 - 11.2.2 Mean
 - 11.2.3 Median
 - 11.2.4 Mode
- 11.3 Summary
- 11.4 Keywords
- 11.5 Self-Assessment Questions
- 11.6 References/Suggested Readings

11.1 INTRODUCTION :

Measures of central tendency are considered to be the statistical averages which are continuously growing indispensable in the emerging research areas of mass communication. As per the basic concept, the word -Average is commonly used in media research and in our day-to-day activities. But it has been differently defined in statistics.

1. According to Clark - An average is an attempt to find one single figure to describe whole of figures.
2. Leabo defines -Average as -The average is sometimes described as a number which is typical of the whole group.
3. A.E. Waugh defines it as -An average is a single value selected from a group of values to represent them in some way - a value which is supposed to stand for whole group, of which it is a part, as typical of all the values in the group
4. According to Ya-Lun-Chou -An average is a typical value in the sense that it is sometimes employed to represent all the individual value in a series of a variable.
5. Croxton and Cowden define average as -An average value on a single value within the range of the data that is used to represent all of the values in the series. Since the average is somewhere within the range of the data, It is also called measures of central tendency.

Thus, the main objectives of average are to get single value that describes the

characteristics of the entire group, and to facilitate comparison. Since the comparative and evaluative measurement concept of unexplored variables of communication and media are more in demand these days due to the growing professionalism in this area, so getting a single value of the mines of media data and then using such data in various comparisons is becoming the crying need of the hour.

Types of average: There are three types of average

1. Mean
2. Median
3. Mode.

11.2 STATISTICAL METHODS (SUMMARY MEASURES) :

1. Quantitative data show certain general characteristics, which can be briefly put together as follows:
2. They show a tendency to concentrate at certain values, usually somewhere in the centre of the distribution. Measures of this tendency are called measures of central tendency or averages.
3. The data vary about a measure of central tendency. This refers to the variability characteristics of the data, often called dispersion. A measure of variability is concerned with quantifying the magnitude of spreadness of individual observations around the central value.
4. The data in a frequency distribution may fall into symmetrical or asymmetrical patterns. The measures of this direction and degree of asymmetry are called measures of skewness.
5. Polygons in frequency distribution exhibit peakedness. Measures of peakedness are called measures or kurtosis.
6. The purpose of these measures is to discover the characteristics of a set of data and hence to facilitate comparison within a set of data or between sets of data. In this part we shall discuss the measures of first two of the above-mentioned characteristics.

Statistical Methods in Media Research are:

1. Measures of Central Tendency
2. Mean
3. Median
4. Mode

11.2.1 Measures of Central Tendency :

Measures of central tendency indicate the point about which items have a tendency to cluster. This is generally accepted as the most representative figure for the whole bunch of data. It gets a single value that describes the characteristics of entire mass of scattered and wide data. Taking such example of media research, suppose there are 35 lakhs media persons in India and each media person has some income. But we say that per capita income per media person per annum is Rs. 10 lakhs. It reflects the entire picture of income of media persons of the whole country with a single value. Hence, measures of central tendency in media research are generally the computation of mean, media and mode in the field of print media, electronic media, advertising, corporate communication & public relations, web media, development

communication & social marketing, social media and various other media.

11.2.2 Mean :

In communication research or social research only arithmetic mean is use full and applicable, while geometric mean or harmonic mean is not required at any point of research. But arithmetic mean is the most largely and widely used measures of representing the whole of media or communication data by one value. Its value is obtained by adding together all the items and by dividing this total by number of items. In media research, arithmetic mean is so popular and relevant that when we talk of mean, we generally mean "Arithmetic Mean".

1. Mean of Simple Distribution (or Ungrouped or Unclassified Distribution) :

The arithmetic mean, which is sometimes simply referred to as "mean" is the most commonly used average. It is sum of the values observed divided by the total number of observation of the data set

Formula, $\bar{X} = \frac{\sum x}{N}$ where

(i) \bar{X} (called X bar) is mean (ii) X is scores

(iii) N is number of score (iv) \sum is summation of

\sum is a letter of Greek alphabets also called sigma

e.g. In a media awareness test, 10 students secured the following marks:
10, 18, 16, 15, 19, 15, 14, 25, 16, 18. Find the mean of the said data.

No.of students	Scores(x)
1	10
2	18
3	16
4	15
5	19
6	15
7	14
8	25
9	16
10	18
N=10	$\sum X=166$

$$\bar{X} = \frac{\sum x}{N} = \frac{166}{10} = 16.6$$

Hence, the average score is 16.6 marks.

2. Mean of Unclassified Data with Frequency Distribution:

Formula,

$$\bar{X} = \frac{\sum fX}{N}$$

Where \bar{X} = mean, X is scores, N is number of scores. fx= Frequency multiplied by X value.

e.g. Calculate the arithmetic mean of the unclassified data with the following frequency distribution offive print journalists and their no. of investigative stories:

X (Print journalists) 1,2,3,4,5

F (Investigative stories) 2,9.9.7,5

Variable (x)	Frequency (f)	fx
1	2	2
2	9	18
3	9	27
4	7	28
5	3	15
	$N = \sum f = 30$	$\sum fx = 90$

$$\bar{X} = \frac{\sum fX}{N} = \frac{90}{30} = 3$$

Hence the arithmetic mean of investigative stories of five journalists is 3.

3. Mean of classified data with Frequency Distribution

Formula

$$\bar{X} = \frac{\sum fX}{N}$$

Find the mean of the following frequency distribution of media persons and their exclusives:

Class(Media persons)	0-8	8-16	16-24	24-32	32-40	40-48
Frequency(No.of exclusives)	10	20	14	16	18	22

Class	Frequency(f)	Mid value of class(X)	Fx
0-8	10	$\frac{0+8}{2} = 4$	40
8-16	20	$\frac{8+16}{2} = 12$	240
16-24	14	$\frac{16+24}{2} = 20$	280
24-32	16	$\frac{24+32}{2} = 28$	448
32-40	18	$\frac{32+40}{2} = 36$	648
40-48	22	$\frac{40+48}{2} = 44$	968
	$N = \sum f = 100$		$\sum fx = 2624$

$$\bar{X} = \frac{\sum fX}{N} = \frac{2624}{100} = 26.24$$

Hence, the mean of the media persons and their successful exclusives is 26.

4. Mean by Short-cut method:

1. Simple Distribution
2. Frequency Distribution

Formula,

1. Simple Distribution

$$\bar{X} = A + \frac{\sum d}{N}$$

(Where as Assumed Mean and d is Deviation of scores from Assumed Mean)

For assuming (A), any single value can be selected from \bar{X} values.

2. Frequency Distribution

$$\bar{X} = A + \frac{\sum d}{N}$$

e.g. (i) In a media creativity test, 10 students secured the following mark: 10,18,16,15,19,15,14,24,25,16,18. Find the mean of this simple distribution by short-cut method:

Solution: Let us take the assumed mean =19

x	d= X-A
10	10-19=-9
18	18-19=-1
16	16-19=-3
15	15-19=-4
19	19-19=0
15	15-19=-4
14	14-19=-5
25	25-19=6
16	16-19=-3
18	18-19=-1

$$N= 10 \quad \sum d = -24$$

$$\bar{X} = A + \frac{\sum d}{N} = 19 + \frac{-24}{10} = 19 - 2.4 = 16.6$$

Hence mean value is 16.

11.2.3 Median :

Median is the middle value in the entire distribution of data or scores. It refers to the central position of the data or the score. When all the observations are arranged in order of their size, the median is the value of that item, which has equal number of observations on either sides. In other words, median is the central value of the distribution or the value that divides the distribution into two equal parts. Hence, for the computation of median, it is the first condition that items be arranged in ascending or descending order. That is why it is also called the positional average or average of position, as it occupies a definite place in the given set of observation.

J. Wert, C. Neidt and J. Ahmmann has aptly defined median as -The Median is that point in the distribution above which and below which 50 percent of class lie'.

1. Median of Simple Distribution (or Ungrouped or Unclassified Distribution) :

The median of a collection observations is defined as the middle most observation in the sense that one half of the number of observations is less than median while the other half of observation is larger than the median. To determine the median, it is necessary to first array the observation either in a descending or ascending order of magnitude and then count their number. If the number of observation is odd, media is equal to $(n+1)/2$ the observation. If the number is even, median is arithmetic mean of $n/2$ and $n/2+1$ the observations.

Formula, $M = 1 + \left[\frac{N+1}{2} \right]$ th term

Where M is Median and N is the no. of scores. When N is an odd number, then deriving median is an extremely simple process.

e.g., Find the Median of 5, 8, 6, 11, and 9 of a Advertising campaigning by different advertising agencies.

First, the scores are arranged in ascending or descending order as follows:

5, 6, 8, 9, 11. According to the formula,

$$\text{Median} = \left[\frac{N+1}{2} \right] \text{th term}$$

Here $N=5=5+1/2=6/2=3^{\text{rd}}$ term

$$\therefore \text{Median} = 8$$

Thus, the median of ad campaigning by different corporate organizations is 8.

2. Median of Frequency Distribution:

i) Unclassified ii) Classified

Median of unclassified frequency distribution: first of all, the value of $\frac{N}{2}$ will be derived.

Here Median will be the value of the variable having $cf \geq \frac{N}{2}$

(Note: when $N/2$ is not equal to any cumulative frequency, then cf which is just greater than $N/2$ is taken)

e.g., Find the median of the following unclassified frequency distribution of marks secured by PG Diploma students of Journalism:

Marks(X)	11	12	13	14	15	16	17	18
No. of students	5	7	11	9	8	7	3	5

Solution

X	f	cf
11	5	5
12	7	12
13	11	23
14	9	<u>32</u>
15	8	40
16	7	47
17	3	50
18	5	55
	N=55	

$$N/2 = 55/2 = 27.5$$

(cf is calculated by just adding previous frequency/frequencies. Like $5+7=12$, $12+11=23$, $23+9=32$ and so on.)

Here cf just greater than 27.5 is 32 and the value of the variable corresponding to 32 is 14. $M = 14$.

Thus, the Median of marks of PG Diploma in Journalism is 14.

i) Median of the Classified Frequency Distribution:

$$\text{Formula, } M = l + \left(\frac{\frac{N}{2} - cfb}{fm} \right) \times h$$

Where l = lower limit of median class (if total frequency be N , then class $wcf \geq N$ is called median class)

cfb : cumulative frequency just below median class

fm : frequency of median class

h : height of the class

It is important to mention here that the lower limit of the median class is the same as given in the class in overlapping or continuous class. But 0.5 less in non-overlapping or discrete class.

Note : Instead of length of the class, height of the class is preferred because it is always the same whether the class is overlapping (continuous) or non-overlapping (discrete). In height, we have to just subtract the lower limit of any class from the lower limit of the nearest higher class. Or we have to just subtract the higher limit of any class from the higher limit of the nearest higher class. In case of use of height, there is no need to add +1 in non-overlapping (discrete) class unlike length of the class).

e.g. 1 Find the Median of the following class distribution of no. of investigative stories covered by journalists of different age groups:

Age group of journalists	30-40	40-50	50-60	60-70	70-80
No. of investigativestories	50	54	85	45	30

Class	Frequency	cf
30-40	50	50
40-50	54	104
<u>50-60</u> ←	85	189
60-70	45	234
70-80	30	264
	$N = 264$	

$$h = 50 - 40 = 10$$

$$\frac{N}{2} = \frac{264}{2} = 132$$

Here, cf 189 is just greater than $N/2$, i.e.; 132. Hence the corresponding class, i.e., (50-60) is the median class and 104 is Cfb , i.e., cumulative frequency just below median class. Further h is 70-60=10, 60-50=10, 50-40=10 or 40-30=10 i.e. subtraction of the lower limit of any class from the lower limit of the nearest higher class as shown above in the class column.

As per formula

$$M = l + \left(\frac{\frac{N}{2} - cfb}{fm} \right) \times h$$

$$= 50 + \left(\frac{132-104}{85} \right) \times 10$$

$$= 50 + \left(\frac{280}{85} \right) = 53.29$$

Thus, the median of no. of investigative stories covered by journalists of the said age group is 53.

e.g. 3. Find the median of the following distribution:

Marks Obtained Less than 10	No. of students of MMC students 25
Less than 20	40
Less than 30	60
Less than 40	75
Less than 50	95
Less than 60	125
Less than 70	190
Less than 80	240

First of all, the frequency table will be prepared in the following manner:

Classes	Cumulative Frequency	Frequency
0-10	25	25
10-20	40	40-25=15
20-30	60	60-40=20
30-40	75	75-60=15
40-50	95	95-75=20
50-60	125	125-95=30
60-70	190	190-125=65
70-80	240	240-190=50

N=240

Here c.f 125 is just greater than $N/2$, i.e, 120. Hence, the corresponding class, i.e, cumulative frequency just below median class. Further, h is $10-0=10$, i.e., subtraction, of the lower limit of any class from the lower limit of the nearest higher class as shown above in the class column.

$$\text{Now } M = 1 + \left(\frac{\frac{N}{2} - cfb}{fm} \right) \times h$$

$$= 50 + \left(\frac{120-95}{8530} \right) \times 10$$

$$= 50 + \left(\frac{25}{3} \right) = 50 + 8.33$$

$$= 58.33$$

11.2.4 Mode :

Mode is the score of the greatest frequency. It indicates that point in the distribution

which occurs most frequently. According to J.P. Guilford & B. Fructer, -The mode is defined as the point on the scale of measurement with maximum frequency in a distribution.

Croxtton and Cowden define it as: -The mode of a distribution is the value at the point around which the items tend to be most heavily concentrated. It may be regarded of the most typical of a series of values’.

1.Mode of a Simple Distribution :

It is the value in a series of observations which has the highest or greatest or maximum frequency. There are many situations in which arithmetic mean and median fail to reveal the true characteristics of data. When we talk of average consumer average student, average size of shoes or ready-made garments we have in mind mode and not the arithmetic mean or median. Here average student means the type of student who is met most frequently with regard to some quality. In statistics the value of variable, which occurs most frequently, is called mode.

e.g. Find the value of mode in the following data: 3, 5, 6, 6, 7, 6, 8.

In this Simple Distribution of data, mode is ‘6’ as it has the highest frequency ‘3’.

2.Mode of Frequency Distribution :

i) Mode of unclassified Frequency Distribution

ii) Mode of Classified Frequency Distribution.

i).Mode of Unclassified Frequency Distribution: Here, again it is the value of the highest frequency.

e.g. Find the mode of the following Unclassified frequency distribution of scores secured by MA students of Public Relations:

Scores(X)	11	12	13	14	15	16	17	18
No.of students	5	7	11	9	8	7	3	5

Here Mode is 13’as it has the highest frequency, i.e, 11.

ii). Mode of Classified Frequency Distribution: Mode of a classified data is normally the mid-point of the class interval of the highest frequency. But for more accuracy and precision, the following formula is used:

$$M_o = l + \left(\frac{f_a}{f_a + f_b} \right) X h$$

Where M_o is Mode,

l is lower limit of modal class (class of highest frequency). It is the same in overlapping or continuous class and 0.5 less in non-overlapping or discrete class.

f_a is frequency of pre-modal

class f_b is frequency of post-modal class

h is height of the class (height is the subtraction of the lower limit of any class from the lower limit of any class from the lower limit of the nearest higher class.

11.3 SUMMARY :

Measures of central tendency in social research, especially communication research

which is described as statistical averages, gets a single value of largely and widely scattered data of mass media which describes the characteristics of the entire data. There are three types of averages mean, median and mode. In media research when we talk of mean, we generally mean arithmetic mean. Median is the average of position, which refers to the middle value in the distribution.

Hence median is the central value in the distribution or the value that divides the distribution into two equal parts. Mode is the score of the highest frequency. Choosing a measure of central tendency in media research out of the mean, median and mode, mean is the most widely and largely used descriptive statistics in media research or descriptive media metrics because so many statistical methods relating to media are based on means. So if we use something other than mean, we would need a lot of extra knowledge on the application of inferential statistics of mass communication or inferential media metrics

11.4 KEYWORDS :

- ❖ **Central Tendency :** A single value which represents the large and widely scattered data, either in the form of a mathematical average or positional average or concentration of the highest frequency, mean, median and mode
- ❖ **Mean:** It is the mathematical average of the given data.
- ❖ **Median:** It is the positional average of the given data.
- ❖ **Mode:** It is the score having the greatest value

11.5 SELF-ASSESSMENT QUESTIONS :

1. Write the formulae for calculating median?
2. Write the formulae for calculating mode?
3. Find the arithmetic mean of following scores secured by the students of bachelors of mass communication in reporting: 50, 40, 30, 45, 25, 42, 44, 38, 48, and 36.
4. Ten working journalists of a regional daily have the following monthly income in rupees: 14780, 15760, 26690, 27750, 24840, 24920, 16100, 17810, 27050, 26950. Find out the arithmetic mean of their income.
5. 50 students of Master of Mass Communication secured the following marks in the paper of editing. Find their mean, median and mode: 32, 30, 45, 75, 35, 33, 51, 61, 44, 33, 45, 48, 56, 71, 70, 73, 80, 34, 46, 44, 48, 33, 31, 46, 61, 63, 64, 68, 69, 76, 77, 79, 71, 75, 55, 56, 34, 36, 38, 40, 32, 49, 51, 58, 55, 31, 62, 66, 45, 62.

Answers: (3) A.M=39.8 (4) A.M=Rs.22,265

11.6 REFERENCES /SUGGESTED READINGS :

1. Berger, Arthur Asa .1998. Media Research Techniques‘ (Second Edition) USA: SagePublications.
2. David, P. 2000. ‘Precision Journalism: A practical Guide’ USA: Sage Publications.
3. Dayal, M. 2010. ‘Media Shodh‘ (3rd edition), Panchkula: Haryana Sahitya Akademi.
4. Dayal, M. 2017. ‘Media Metrics: An Introduction to Quantitative Research in MassCommunication’ .pp 73-85, New Delhi: Sage Publications.
5. Elhance, D.N. & Elhance, Veena .1990. ‘Fundamentals of Statistics’ pp. 8.1-8.134,

- New Delhi: Kitab Mahal, Delhi.
6. Gupta, Ashit Das & Thakur, S.N. 1998. 'Modern Approach to Intermediate Statistics and Probability' pp.14-101, Patna: Bharti Bhawan.
 7. Gupta, S.P. 2011. 'Statistical Methods' pp 177-238, New Delhi: Sultan Chand & Sons.
 8. Gupta, S.P & Gupta, Alka .2008. 'Statistical Method in Behavioral Sciences' pp.62-86, Allahabad: Sharda Pustak Bhawan.
 9. Husain, Shamshad .1986. 'Elementary Statistics in Psychology & Statistics' pp 24-38, New Delhi: Motilal Banarasidas.
 10. Kerlinger, F. N.1983. 'Foundations of Behavioural Research New Delhi: Surjeet Publications.
 11. Kothari, C.R .2000. 'Research methodologies- Methods and Techniques', New Delhi: WishwaPrakashan
 12. Krishnaswami, O.R & Ranganatham, M.2005. 'Methodology of Research',
 13. New Delhi: Himalaya Publications.
 14. Mangal, S.K .1987. 'Statistics in Psychology and Education' pp. 36-48, New Delhi: McGrawHill.
 15. Sarangi, Prasant. 2010. 'Research Methodology' Delhi: Taxman Publication.
 16. Singh, Arun Kumar .1986. 'Basic Statistics In Psychology and Education' pp 20-57, Patna: Novelty & Company.
 17. Wilkinston, T, S, & Bhandarkar, P.L. 1996. 'Methodologies and Techniques of Social Research' Mumbai: Himalaya Publishing House.
 18. Wimmer, Roger D & Dominick, Joseph R .2003. 'Mass Media Research: An Introduction' (7th edition) pp. 247-249 Belmont, USA: Thomsons & Wadsworth

Dr. J. Madhu Babu

LESSON - 12

STANDARD DEVIATION AND VARIANCE

OBJECTIVES :

The objectives of this lesson are as follows:

- To describe the meaning and concept of measures of dispersion in media studies
- To elaborate methods of dispersion
- To analyze mean deviation
- To explain meaning and concept of standard deviation
- To describe variance
- To explain range, inter-quartile range & quartile deviation
- To enlist the formulae

STRUCTURE :

- 12.1 Introduction
- 12.2 Meaning and Concept of Measures of Dispersion in Media Studies
 - 12.2.1 Methods of Dispersion
 - 12.2.2 Mean Deviation
 - 12.2.3 Meaning and Concept of Standard Deviation
 - 12.2.4 Variance
 - 12.2.5 Range, Inter-quartile Range & Quartile Deviation
- 12.3 Summary
- 12.4 Keywords
- 12.5 Self-Assessment Questions
- 12.6 References and Suggested Readings

12.1 INTRODUCTION :

In the previous chapter we have discussed measures of central tendency in which mean, median and mode were comprehensively discussed. In this chapter, we will describe and analyze the measures of dispersion. In media research or communication research, the word dispersion is described as variability or scatter or spread. It refers to the degree or extent of stretched or squeezed of the available data. Measures of dispersion in media research are generally the mean deviation, coefficient of mean deviation, standard deviation, variance, range, semi-range or inter-quartile range in the field of print media, electronic media, advertising, corporate communication & public relations, web media, development communication & social marketing, social media and various other media. Thus measures of dispersion in media research or communication research generally measure as to how spread out a set of media.

12.2 MEANING & CONCEPT OF MEASURES OF DISPERSION IN MEDIA STUDIES :

The measures of central tendency give a single value of largely and widely spreaded data.

But the average cannot adequately describe a set of observations, unless all the observations are the same. In two or more distributions, the central value may be the same, but yet there may be huge disparities in the formation of the distribution. Measures of dispersion are instrumental in finding out this important characteristic of a distribution in media studies due to variability and fluctuation in the various areas of mass communication. Different experts have defined measures of dispersion differently :

- According to Simpson and Kafka, "The measurement of the scattered ness of the mass of figures in a series about an average is called measure of variation or dispersion".
- A.L. Bowley defines it as, "Dispersion is the measure of the variation of the items".
- Spigel Opines, "The degree to which numerical data tend to spread about an average value is called the variation or the dispersion of the data".

Hence, it is very obvious that dispersion (also described as scatter, spread or variation) is the amount of variability in a set of scores and it measures the extent to which the items vary from some value.

Since measures of dispersion give an average of the differences of various items from one average, they are also called averages of the second order.

For example, let there be two small groups of students of new media and traditional media scores in a test are such as the following:

Scores of New Media Students	37,41,36,17,20,19,18,3,5,4
Scores of Traditional Students	18, 21, 22, 18,21,23,17,20,22,18

Mean of New Media students 'scores

$$= 37+41+36+17+20+19+18+3+5+4=200/10=20$$

Mean of Traditional Media students 'scores

$$= 18+21+22+18+21+23+17+20+22+ 18=200/10=20$$

Hence, the value of mean in both the above situation is 20 and hence, so far as the mean goes, there is no difference in the performance of the two groups. Now, a valid question arises, can we take both sets of scores as identical? No doubt, there is a lot of difference between the performances of the two groups.

Whereas the scores of students of mass communication are found to range from 3 to 41, the scores of Advertising range from 18 to 23. The first group is composed of students who have wide individual differences. It consists of either very intelligent students or very dull students. The second group of students of advertising, on the other hand, is composed of average students. Students in this later group are less variable than those in the former. Thus, there is a great need of paying attention to the variability or dispersion of scores in the sets of scores or series if we want to describe or compare them.

Thus, we may conclude that there is a tendency for data to be dispersed, scattered or to show variability around the average consequently, the variability or dispersion may also be defined as the tendency of the measures of the attributes of a group to deviate from the average or central value. Variability, thus, refers to the spread of scores from their central tendency.

12.2.1 Methods of Dispersion ;

The degree of spreading, scattering, or dispersion, i.e, variability is measured by different methods. Of them, the following are popular :

- 1) Mean deviation,
- 2) Standard deviation,
- 3) Variance,
- 4) Range,
- 5) Interquartile range & quartile deviation.

12.2.2 Mean Deviation :

Mean deviation or average deviation is an integral part of measures of dispersion. Mean deviation of a distribution is the arithmetic mean of the absolute deviations of the values of the variable from a measure of their either mean. The mean deviation is also known as average deviation and is generally denoted by A.D or M.D or δ (del). It is the average difference between the items in a distribution and the median or mean or mode of that series. It is the mean of the deviations of all the separate scores in the series taken from their mean (occasionally from their median or mode) (Garnett, 1971 p-48). It is the simplest measure of variability that takes into account the fluctuation or variation of all the items in a series. Theoretically, there is an advantage of taking the deviations from median merely because the sum of deviations of items from median is minimum when signs are ignored.

But in practice the mean is more frequently used in computing the value of average deviation and this is the reason why it is more commonly called mean deviation. This is also called first moment dispersion. For mean deviation, we write δ , for mean deviation from mean, we write δ_x for mean deviation from median, we write δ_M and from mean deviation from mode, we write δ_{M0}

1. Mean Deviation of Simple Distribution or Unclassified or Ungrouped Distribution: Formula:

$$\delta \text{ (M. D. or A. D.)} = \frac{\sum[d]}{N}$$

(Where = Mean Deviation

$$d = (X - \bar{X})$$

$$|d| = \text{Mod}$$

X is score, N is no. of scores

\bar{X} = Mean of score)

e.g.(i) Find the mean deviation of the following data of media test : 12,9,8,7,4

X	d = (X - \bar{X})	d
12	12 - 8 = 4	4
9	9 - 8 = 1	1

8	8 - 8 = 0	0
7	7 - 8 = -1	1
4	4 - 8 = -4	4
$\sum X = 40, N = 5$		$\sum d = 10$

$$\bar{X} = \frac{\sum X}{N} = \frac{40}{5} = 8$$

$$\delta = \frac{\sum |d|}{N} = \frac{10}{5} = 2$$

2. Mean Deviation of Frequency Distribution:

(a) Unclassified frequency distribution

(b) Classified frequency distribution

Formula: $\delta = \frac{\sum |d|}{N}$ where f is the frequency of scores

e.g. (i) Calculate the mean deviation of the following data of reporters filling stories per day:

No. of reporters(X)	2	3	4	5	6	7
No. of stories perday(f)	5	4	7	6	3	2

X	f	fx	d = (X - \bar{X})	d	fd
2	5	10	2 - 4.15 = -2.15	2.15	10.75
3	4	12	3 - 4.15 = -1.15	1.15	4.6
4	7	28	4 - 4.15 = -.15	0.15	1.05
5	6	30	5 - 4.15 = .085	0.85	5.1
6	3	18	6 - 4.15 = 1.85	1.85	5.55
7	2	14	7 - 4.15 = 2.85	2.85	5.7
	N = 27	$\sum fx = 112$			$\sum fd = 32.75$

$$\bar{X} = \frac{\sum fx}{N} = \frac{112}{27} = 4.15 \quad \delta = \frac{\sum |fd|}{N} = \frac{32.75}{27} = 1.21$$

Hence, mean deviation is 1.21.

3. Advantages and Disadvantages of Mean Deviation :

Advantages :

1. Simplicity in calculation: The formula for computing Mean Deviation is simple. Hence it is very useful for communication issues and media analysis.
2. Freedom of choice of mean: M.D. can be computed from any of the averages like Mean,
3. Median or Mode. Thus it is highly applicable in media research
4. Importance of every term: The value of M.D. is dependent on each and every term. Hence, it is very relevant in media metrics.
5. Distribution structure: The evaluation of M.D. gives us an idea about how the terms are situated in the series of distribution. This type of distribution structure is very relevant in media analysis.

Disadvantages :

1. Avoidance of negative sign: Since the formula of mean deviation carries only absolute value, the negative sign is ignored, neglected or avoided. It tremendously affects a suitable media analysis.
2. Lack of certainty: Since it is calculated from any of the measures of central tendency, it lacks certainty. That is why, it is sometimes avoided in the analysis of media issues. But mean deviation from the median is the least and considered to be the best in macro and micro media analysis.

Therefore, mean deviation is highly practically relevant and useful in media research because of its simplicity in computation. It is particularly useful and effective in reports presented to such group or group of public of communication studies which is not aware of statistical methods. Mean deviation is highly useful in media and communication studies for small samples with no elaborate analysis required. It is important to mention that the National Bureau of Economic Research has meticulously observed that the mean deviation is the most practical measures of dispersion especially in social sciences including communication studies.

Co-efficient of Mean Deviation : It is basically the ratio of the mean deviation and the central value from which the deviation is measured.

Thus, Co-efficient of mean deviation from mean = $\frac{\delta\bar{X}}{\bar{X}}$

Co-efficient of mean deviation from median = $\frac{\delta M}{M}$

Co-efficient of mean deviation from mode = $\frac{\delta M_o}{M_o}$

12.2.3 Meaning and Concept of Standard Deviation :

The meaning and concept of standard deviation was first given by Karl Pearson in 1894. It is also called root mean square deviation or second moment dispersion (unlike mean deviation as first moment dispersion). No doubt, it is the most relevant, important and useful measure of study of dispersion. It is the most commonly used measure of dispersion in media research. It satisfies most of the properties laid down for an ideal measure of dispersion in communication and media studies. In social sciences including communication studies, standard deviation is denoted by the small Greek letter (sigma). The greater the standard deviation, the greater will be the magnitude of the deviations of the values from their mean. Similarly, a small standard deviation means a high degree of uniformity of the observations as well as homogeneity of a series. Hence, if there are two or more comparable series with the smallest standard deviation that has the most representative mean. Thus, it is largely and widely used in judging the representativeness of the mean.

Standard deviation is quite different from mean deviation. Negative signs are avoided while computing mean deviation whereas standard deviation is calculated without ignoring any algebraic sign. Further, mean deviation can be calculated either from median, or mean or mode whereas standard deviation is always calculated from the arithmetic mean as the sum of the squares of the deviation of items from arithmetic mean is the least. Since it is root mean square deviation, it is the square root of the mean of the squared deviation from the arithmetic mean.

1. Standard Deviation of Simple Distribution or Unclassified or Ungrouped Distribution:

Formula, $\sigma = \sqrt{\frac{\sum d^2}{N}}$ Where σ (Sigma) = Standard deviation

$$d = X - \bar{X}$$

X = Scores

\bar{X} = Mean of scores, N = No. of scores

e.g., Calculate the standard deviation of the following data: 12, 9, 8, 7, 4

X	d = (X - \bar{X})	d ²
12	12 - 8 = 4	16
9	9 - 8 = 1	1
8	8 - 8 = 0	0
7	7 - 8 = -1	1
4	4 - 8 = -4	16

$$\sum X = 40$$

$$\sum d^2 = 34$$

$$\bar{X} = \frac{\sum X}{N} = \frac{40}{5} = 8 \quad \sigma = \sqrt{\frac{\sum d^2}{N}} = \sigma = \sqrt{\frac{34}{5}} = \sqrt{6.8} = 2.60$$

Thus, the standard deviation is 2.60.

3. Standard Deviation of the Frequency Distribution :

$$\text{Formula: } \sigma = \sqrt{\frac{\sum f d^2}{N}}$$

(i) Unclassified or ungrouped frequency distribution:

e.g. Compute the Standard deviation of the following frequency distribution of a editors writing stories per month:

No. of editors(X)	0	1	2	3	4	5	6	7
No. of stories filed per month(f)	14	21	25	43	51	40	39	12

X	f	fx	d = (X - \bar{X})	d ²	fd ²
0	14	0	0 - 3.76 = -3.76	14.14	197.96
1	21	21	1 - 3.76 = -2.76	7.62	160.02
2	25	50	2 - 3.76 = -1.76	3.1	77.5
3	43	129	3 - 3.76 = -0.76	0.58	24.94
4	51	204	4 - 3.76 = 0.24	0.06	3.06
5	40	200	5 - 3.76 = 1.24	1.54	61.6
6	39	234	6 - 3.76 = 2.24	5.02	195.78
7	12	84	7 - 3.76 = 3.24	10.5	126
	N = 245	$\sum fx = 922$			$\sum fd^2 = 846.86$

$$\bar{X} = \frac{\sum fX}{N} = \frac{922}{245} = 3.76 \quad \sigma = \sqrt{\frac{\sum fd^2}{N}} = \sigma = \sqrt{\frac{846.86}{245}} = \sqrt{3.456} = 1.86$$

Thus, Standard deviation of the above said frequency distribution of a newspaper correspondents filing news stories per month is 1.86.

4. Advantages of Standard Deviation :

- ❖ **Mathematically useful** : Since both positive and negative signs are taken into consideration, so it is mathematically more correct, accurate and useful. As a result, media researchers often rely on this method for measuring dispersion.
- ❖ **Dependence on every item**: It is more reliable as it depends on all the terms. Hence this method can be more dependable for data analysis of media issues.
- ❖ **Effect of sampling is minimum**: If the sample slightly changes, it will have the least effect on its computation. That is why its reliability is very high in communication and media studies.
- ❖ **Certainty**: There is no ambiguity in its measurement. Due to its certainty attribute, it is emerging as one of the most dependable methods of measuring dispersion of media variables.
- ❖ **Useful for more critical studies**: As it is mathematically more convincing, it is more useful for critical analysis. That is why is considered to be a standard device of dispersion in mass communication research.

Disadvantages of Standard Deviation:

- ❖ **Lengthy calculation**: It is difficult to compute as compared to other measures. It is longer and lengthier. In media research, earlier it was sometimes avoided due to this demerit. But now with the rapid growth of information technology, lengthy computation is no more a problem. Since media researchers are generally more than average in the use of communication and information technology, so this method is often adhered to.
- ❖ **Effect of extreme value**: it is normally more influenced and affected by extreme values. Yet one of the most preferred device in media and communication analysis as it is automatically nullified in the process of squaring the data.

5. Differences between Mean Deviation and Standard Deviation :

Both these measures are based on all the items of a variable but they differ a lot in the following measures:

1. Mean deviation does not take into consideration the algebraic signs (plus or minus) in its calculation which is unreasonable and illogical.
2. Mean deviation can be calculated from median, mode or mean and its value differs in these cases (unless the distribution is normal). The standard deviation is always calculated from arithmetic mean.
3. The mathematical properties possessed by standard deviation are much greater than these possessed by Mean deviation.

Co-efficient of Standard Deviation and Co-efficient of Variation :

It is important to mention here that the standard deviation is an absolute measure of dispersion and is expressed in terms of the units of the variable. For the purpose of comparative

studies, a relative measure of dispersion is computed. It is co-efficient of standard deviation which is equal to standard deviation divided by mean.

Thus, co - efficient of SD = $\frac{\sigma}{\bar{X}}$

This value will be often in fraction and as such not ideal for any comparison and hence co-efficient of variation is calculated by multiplying the co-efficient of standard deviation by 100.

Thus, co – efficient of Variation or C. V. = $\frac{\sigma}{\bar{X}} \times 100$

This measure for the first time was evolved by Karl Pearson and is very useful for studying dispersion in more than one series.

A series in which co-efficient of variation is higher would have greater dispersion than the one in which it is lower. Similarly, when C.V. is high, the series is less consistent or less variable. This measure is largely, widely and most commonly used for a comparative study of the variability or consistence of two or more series.

e.g. (i) Find the mean deviation, standard deviation, co-efficient of standard deviation and co-efficient of variation of the following data: 5,6,9,12,13:

X	d = (X - \bar{X})	d	d ²
5	5 - 9 = -4	4	16
6	6 - 9 = -3	3	9
9	9 - 9 = 0	0	0
12	12 - 9 = 3	3	9
13	13 - 9 = 4	4	16
$\sum X = 45$ N=5		$\sum d = 14$	$\sum d^2 = 50$

$$\bar{X} = \frac{45}{5} = 9 \quad \delta = \frac{|d|}{N} = \frac{14}{5} = 2.8$$

$$\sigma = \sqrt{\frac{\sum d^2}{N}} = \sqrt{\frac{50}{5}} = \sqrt{10} = 3.162$$

$$\text{Co-efficient of SD} = \frac{\sigma}{\bar{X}} = \frac{3.162}{9} = 0.351$$

$$\text{Co-efficient of variation or C.V.} = \frac{\sigma}{\bar{X}} \times 100$$

12.2.3 Variance :

Variance is very important and relevant statistical tool in the field of media research and communication research. The term 'variance' is used to describe the square of standard deviation. This term was first used by R.A. Fisher in 1913. The relevance and importance of the term 'variance' lies in the fact that it is capable of a very exhaustive type of quantitative analysis. The concept of variance is very significant in advanced media research work where it is possible to split the total into several parts, each attributable to one of the factors causing variation in their original series. When we deal with a phenomenon affected by a number of variables in communication and media studies, the analysis of variance helps us in isolating the effects of various factors. Thus, variance is the sum of squared deviations of scores taken from mean and divided by their number (N).

$$\text{Symbolically, Variance} = \frac{\sum (X - \bar{X})^2}{N} = \sigma^2$$

$$\text{Thus, } \sigma = \sqrt{\text{Variance}}$$

In a frequency distribution, where assumed mean is taken and step-deviations are calculated:

$$\text{Variance} = \left(\frac{\sum f d_1^2}{N} - \left(\frac{\sum f d_1}{N} \right)^2 \right) \times h^2$$

Where $d = \frac{x-A}{H}$ and $h =$ height of class

e.g. (i) 5 students of advertising scored the following marks in their dissertations: 20, 22, 25, 18, and 30. What is the variance and standard deviation?

X	$d = (X - \bar{X})$	$ d^2 $
20	$20 - 23 = -3$	9
22	$22 - 23 = -1$	1
25	$25 - 23 = 2$	4
18	$18 - 23 = -5$	25
30	$30 - 23 = 7$	49
$\sum X = 115$ $N=5$		$\sum d^2 = 88$

$$\bar{X} = \frac{\sum X}{N} = \frac{115}{5} = 23$$

$$\text{Variance} = \frac{\sum d^2}{N} = \frac{88}{5} = 17.6$$

$$\sigma = \sqrt{17.6} = 4.2$$

e.g. (ii) Compute the variance of the following marks obtained by 80 students of Communication as shown in the following table:

Marks obtained	10-14	14-18	18-22	22-26	26-30	30-34	34-38	38-42	42-46	46-50	50-54	54-58
No. of students	2	4	4	8	12	16	10	8	4	6	2	4

Marks	f	Mid Value (X)	$d' = (X - A)/h$	$ fd' $	$ fd'^2 $
10 - 14	2	12	-5	-10	50
14 - 18	4	16	-4	-16	64
18 - 22	4	20	-3	-12	36
22 - 26	8	24	-2	-16	32
26 - 30	12	28	-1	-12	12
30 - 34	16	32	0	0	0
34 - 38	10	36	1	10	10
38 - 42	8	40	2	16	32
42 - 46	4	44	3	12	36
46 - 50	6	48	4	24	96
50 - 54	2	52	5	10	50
54 - 58	4	56	6	24	144
	$N =$			$\sum fd' = 30$	$\sum fd'^2 = 562$

$$\begin{aligned}\text{Variance} &= \left(\frac{\sum fd'^2}{N} - \left(\frac{\sum fd'}{N} \right)^2 \right) \times h^2 \\ &= \left(\frac{562}{80} - \left(\frac{30}{80} \right)^2 \right) \times 4^2 \\ &= 6.884 \times 16 = 110.144\end{aligned}$$

12.2.4 Range, Inter-Quartile Range & Quartile Deviation :

The relevance and importance of Range cannot be denied in media research. Range is the simplest possible measure of dispersion. It is calculated by subtracting the lowest score in the series from the highest. But it takes only extreme values into account and remains silent about the variation of different items.

Symbolically, Range (R) = L – S where L is the largest value and S is the smallest value in a series.

If range is divided by the sum of the extreme items, the resulting figure is called, 'Ratio of range' or 'Co-efficient of range'.

Symbolically,

$$\text{Co-efficient of Range or Ratio of Range} = \frac{L-S}{L+S}$$

e.g. (i) The range of the data 2,4,6,8,50 is 50-2=48

e.g. (ii) Calculate the Range and Co-efficient of Range of the following data:

Class	11-13	13-15	15-17	17-19
Frequency	3	4	5	26

$$\begin{aligned}R &= L - S \\ &= 19 - 11 = 8\end{aligned}$$

$$\text{Co-efficient of Range} = \frac{L-S}{L+S}$$

$$\frac{19-11}{19+11} = \frac{8}{30} = 0.266$$

(Note: In the calculation of Range, only the values of the variables are taken into account and the frequencies are totally ignored).

- Advantages and Disadvantages of Range:** A good measure of dispersion should be easily computed, rigidly defined and readily understood. Moreover, it should be capable of further mathematical treatment and should not be much affected by fluctuations of the sampling.

Advantages :

- Range is easily calculated, readily understood and hence the simplest and the easiest method of studying dispersion. While carrying on economic journalism research, range is useful in the study of fluctuations in share prices, money, rates, exchange rates, gold prices etc.
- It takes the least time to compute the value of range. Thus, if we are interested to get a quick rather than a very accurate picture of variability, we may calculate range. Hence, in the applied research of communication and media analysis this is often used as it is quickly computable.

Disadvantages :

- It is not based on all the observations of the series. It has nothing to do with each

and every item of the distribution as far as quality control is concerned. Since, it is subject to fluctuations of considerable magnitude from sample to sample, so we may say it is largely affected by fluctuations of sampling.

- (ii) It is impossible to use it in case of open ended distributions. In such distributions, the lower limit of the lowest class and the upper limit of the highest class are quite unknown.

Despite several demerits of range of fluctuations of considerable magnitude from sample to sample, yet it is very useful in financial and commercial journalism research, especially in the study of fluctuations in share prices, money, rates, exchange rates, gold prices etc. It is very relevant in overall mass communication research, in weather forecasts and overall in our everyday life. For example, if the salary of the highest earning journalist and the lowest earning journalist remain unchanged and if the salaries of all other journalists are changed, range will remain unaffected. The range of a symmetrical and an asymmetrical distribution can be identical. Two such distributions can never have the same dispersion. Thus it is relevant and useful in media research.

2. Inter-quartile Range & Quartile Deviation : Inter-quartile range is also a measure of dispersion. It has an advantage over range because it is not affected by the values of the extreme items. For this purpose, a measure called Inter-quartile range has been developed. This type of range includes the middle 50 per cent of the distribution, i.e., one quarter of the observations at the lower end, another quarter of the distribution are excluded in calculating the Inter-quartile range. Thus, Inter-quartile range represents the difference between the third quartile and the first quartile.

Symbolically

$$\text{Inter - quartile Range} = Q_3 - Q_1$$

When the value of inter-quartile range is divided by 2, it becomes Semi-Inter-quartile Range or Quartile Deviation.

$$\text{Thus, Semi Inter- quartile Range or Quartile Deviation} = \frac{Q_3 - Q_1}{2}$$

Where Q_3 and Q_1 stands for the upper and lower quartiles respectively.

In a symmetrical distribution, median lies half-way on the scale from Q_1 to Q_3 , i.e., median - $Q_1 = Q_3$ - median and as such the difference can be taken as a measure of dispersion. The median \pm Q. D. covers exactly 50 per cent of the observations. But mass media, generally, such data is not available as we have more possibility of the asymmetrical distribution. In an asymmetrical distribution, Q_1 and Q_3 are not equidistant from the median. Thus, an asymmetrical distribution includes only approximately 50 per cent of the observations.

If Quartile Deviation is very less, it indicates high uniformity or small variation of the central 50 per cent items and a high quartile deviation means that the variation among the central item is large.

Thus, Quartile Deviation is described as an absolute measure of the dispersion. When it is divided by the average value of the two quartiles, a relative measure of the dispersion is obtained. It is called Co-efficient of Quartile Deviation.

Symbolically,

$$\text{Co - efficient of a Quartile Deviation} = \frac{\frac{Q_3 - Q_1}{2}}{\frac{Q_3 + Q_1}{2}} = \frac{Q_3 - Q_1}{Q_3 + Q_1}$$

(Note: it is important to mention here that Co-efficient of Quartile Deviation may be used to compare the degree of variations in different distribution.)

e.g. Find the Quartile Deviation and its Co-efficient from the following data relating to the monthly income of seven freelance journalists.

Monthly Income (Rs.): 5000, 7000, 8000, 6000, 6500, 4000, 9000.

Here, the incomes are arranged in ascending order, i.e., 4000, 5000, 6000, 6500, 7000, 8000, 9000.

$$Q_1 = \text{the value of } \left(\frac{N+1}{4}\right)^{\text{th}} \text{ or } \left(\frac{7+1}{4}\right)^{\text{th}} \text{ or } 2^{\text{nd}} \text{ item} = 5000$$

$$Q_3 = \text{the value of } 3 \left(\frac{N+1}{4}\right)^{\text{th}} \text{ or } 3 \left(\frac{7+1}{4}\right)^{\text{th}} \text{ or } 6^{\text{th}} \text{ item} = 8000$$

$$\text{Quartile Deviation} = \left(\frac{Q_3 - Q_1}{2}\right)^{\text{th}} \text{ or } \frac{8000 - 5000}{2} = \frac{3000}{2} = 1500$$

$$\text{And, Co-efficient of Quartile Deviation} = \frac{(Q_3 - Q_1)}{(Q_3 + Q_1)} = \frac{8000 - 5000}{8000 + 5000} = \frac{3000}{13000} = 0.23$$

e.g. (ii) Calculate the value of Quartile Deviation and its Co-efficient from the data of 7 students of digital media.

Roll No.	1	2	3	4	5	6	7
Marks	20	28	40	12	30	15	50

Marks arranged in ascending order: 12, 15, 20, 28, 30, 40, 50

$$Q_1 = \left(\frac{N+1}{4}\right)^{\text{th}} \text{ term } \frac{7+1}{4} = 2^{\text{nd}} \text{ item} = 15$$

$$Q_3 = 3 \left(\frac{N+1}{4}\right)^{\text{th}} \text{ term} = 3 \left(\frac{7+1}{4}\right)^{\text{th}} = 6^{\text{th}} \text{ item} = 40$$

$$\text{Quartile Deviation} = \frac{Q_3 - Q_1}{2} = \frac{40 - 15}{2} = \frac{25}{2} = 12.5$$

$$\text{And, Co-efficient of Quartile Deviation} = \frac{Q_3 - Q_1}{Q_3 + Q_1} = \frac{40 - 15}{40 + 15} = \frac{25}{55} = 0.455$$

2. Advantages and Disadvantages of Quartile Deviation: Advantages :

1. It is simply calculable and easily understandable.
2. As a measure of dispersion, it is superior to range in certain aspects
3. It is relevant, significant and useful in studying dispersion in open and series.
4. It is also relevant, significant and useful in erratic or badly skewed distributions, where the other measures of dispersion would be warped by extreme values. The Quartile Deviation is not affected by the presence of extreme values.

Disadvantages:

1. Quartile Deviation is not based on all the observations and it ignores 50% items, i.e., the first 25% and the last 25%.
2. It is not capable of further algebraic manipulation and mathematical treatment.
3. It is in a way a positional average.
4. Its value is very much affected by sampling fluctuations. A change in the value of one item, may affect its value considerably.

Thus, due to considerable variation in the value of various items, it is not a suitable measure of dispersion. As its value does not depend upon every item of the series, it cannot be considered as a favorable method of measuring dispersion. Because of such limitations, it is not often useful for statistical inference. But communication and media researchers do use the term standard deviation and variance in measurements or differences between research units on what is being measured. And the knowledge of others measures are important for proper understanding of standard deviation and variance.

12.3 SUMMARY :

Measures of dispersion is so important to be discussed in the suitable analysis of media research. Dispersion, which is often described as scatter, spread or variation is the amount of variability in a set of scores and it measures the extent to which the item vary from some central value. There are five methods of dispersion: (i) Mean Deviation, (ii) Standard Deviation, (iii) Variance, (iv) Range, (v) Inter-quartile Range & Quartile Deviation. Mean deviation is the average difference between the items in a distribution and the median or mean or mode of that series. Standard deviation is always calculated from the arithmetic mean as the sum of squares of deviation of items from arithmetic mean is the least. Variance is described as the square of standard deviation. Range is calculated by subtracting the lowest score in the series from the highest. Inter-quartile range represents the difference between the third and the first quartile. Most of such measures may not be practically so pertinent, but standard deviation and variance are highly relevant pertinently applicable. Communication and media researchers do use the term standard deviation and variance in measurements or differences between research units on what is being measured.

12.4 KEYWORDS :

- ❖ **Dispersion:** It is often described as scatter, spread or variation. It is the amount of variability in a set of scores and it measures the extent to which the item vary from some central value.
- ❖ **Mean Deviation:** It is the average difference between the items in a distribution and the median or mean or mode of that series.
- ❖ **Standard Deviation:** It is always calculated from the arithmetic mean as the sum of squares of deviation of items from arithmetic mean is the least.
- ❖ **Variance:** It is described as the square of standard deviation.
- ❖ **Range:** is calculated by subtracting the lowest score in the series from the highest.

12.5 SELF-ASSESSMENT QUESTIONS :

1. Explain the various methods of dispersion comprehensively.
2. Describe the meaning and concept of measures of dispersion.
3. What is Mean Deviation? Discuss the merits, demerits and usefulness of Mean Deviation.
4. What is Standard Deviation? Discuss the merits, demerits and usefulness of Standard Deviation.
5. Describe the significance of Standard Deviation in mass communication research.
6. Discuss the role of Mean Deviation in media and communication research.
7. Throw light on co-efficient of Mean Deviation, Co-efficient of Standard Deviation and Co-efficient of Quartile Deviation.

8. Differentiate between Mean Deviation and Standard Deviation.
9. What is Variance? Discuss the significance of Variance in media research.

12.6 REFERENCES AND SUGGESTED READINGS :

1. Anderson, J.A. 1987. *Communication Research: Issues and Methods*, New York: Mc Graw-Hill
2. Berger, Arthur Asa .1998. *Media Research Techniques* (Second Edition) USA: SagePublications.
3. David, P. 2000. *Precision Journalism: A practical Guide* USA: Sage Publications.
4. Dayal, M. 2010. *Media Shodh* (3rd edition), Panchkula: Haryana Sahitya Akademi.
5. Dayal, M. 2017. *Media Metrics: An Introduction to Quantitative Research in MassCommunication* .pp 100-119, New Delhi: Sage Publications.
6. Elhance, D.N.S & Elhance, veena .1990. *Fundamentals of Statistics*, pp. 9.1-9.54, New Delhi:Kitab Mahal.
7. Gupta, Das & Thakur, S.N.1998. *Modern Approach to Intermediate Statistics and Probability*,pp. 102-144, Patna: Bharti Bhawan, Patna.
8. Gupta, S.P. 2011. *Statistical Method* pp. 273-334, New Delhi: Sultan Chand & Sons, Delhi.
9. Hussain, Shanshad .1986. *Elementary Statistics in Psychology and Education* pp. 39-59, NewDelhi: Motilal Banarasidas.
10. Kerlinger, F.N..1983. *Foundations of Behavioural Research* , USA: Rinehart and Wintson .
11. Kothari, C.R .2000. *Research methodologies- Methods and Techniques*, New Delhi: WishwaPrakashan
12. Krishnaswami, O.R & Ranganatham, M.2005. *Methodology of Research*, New Delhi:Himalaya Publications.
13. Mangal, S.K .1987. *Statistics in Psychology and Education* pp. 61-69, New Delhi: McGrawHill.
14. Sarangi, Prasant. 2010. *Research Methodology* Delhi: Taxman Publication.
15. Singh, Arun Kumar. 1986. *Basic Statistics in Psychology and Education* pp.86-111. Patna:Novelty &Co., Patna.
16. Wilkinston, T, S, & Bhandarkar, P.L. 1996. *Methodologies and Techniques of Social Research* Mumbai: Himalaya Publishing House.
17. Wimmer, Roger D & Dominick, Joseph R .2003. *Mass Media Research: An Introduction* (7th edition) pp. 249-254 Belmont, USA: Thomsons &Wadsworth.

Dr. J. Madhu Babu

LESSON - 13

CHI- SQUARE TEST

OBJECTIVES :

The objectives of this lesson are as follows :

- To describe the meaning and concept of Chi Square Test in Media Research
- To elaborate applications of Chi Square Test
- To analyze additive property of Chi Square
- To explain misuse and precaution of Chi Square Test
- To describe illustration of various applications of Chi Square Test

STRUCTURE :

- 13.1 Introduction
- 13.2 Meaning and Concept of Chi Square Test in Media Research
 - 13.2.1 Applications of Chi Square Test,
 - 13.2.2 Additive Property of Chi Square,
 - 13.2.3 Misuse and Precaution of Chi Square Test,
 - 13.2.4 Illustration of various Applications of Chi Square Test.
 - 13.2.5 Usage of 'Z' and 'T' Tests
- 13.3 Summary
- 13.4 Keywords
- 13.5 Self-Assessment Questions
- 13.6 References

13.1 INTRODUCTION :

For testing hypothesis or test of significance we use both parametric tests and nonparametric or distribution free tests. Parametric tests assume within properties of the population, from which we draw samples. Such assumptions may be about population parameters, sample size, etc. In case of non-parametric tests, we do not make such assumptions. Here we assume only nominal or ordinal data.

Important parametric tests used for testing of hypothesis are :

- i) X^2 test
- ii) z-test
- iii) t-test

Chi Square Test is often used in media or communication research. It is one of the easiest, simplest, and most popular non-parametric tests. It has a vital role in statistical analysis and statistical inference. But these tests in communication and media studies are still not as reliable as parametric tests. If both parametric and non-parametric tests can be applied under any situation, then parametric tests are considered to be more suitable, powerful and effective. Chi-square is slightly parametric also.

13.2 MEANING AND CONCEPT OF CHI SQUARE TEST IN MEDIA RESEARCH :

Chi Square which is widely and largely used in media research is denoted by the Greek letter X^2 which was first applied by British statistician *Karl Pearson* in the year 1900. It is given by

the following formula: $X^2 = \sum \frac{(f_o - f_e)^2}{f_e}$

Where f_o is observed frequencies and f_e is expected frequencies. This formula is applied by the following steps :

- ❖ First of all expected frequency is generally calculated by simply averaging all the observed frequencies. In a situation of multiple expected frequencies, the expected frequency is calculated by the following formula: $f_e = \frac{RT \times CT}{N}$
- ❖ Where RT is Row Total, CT is Column Total and N is grand total number of observations.
- ❖ Then we will take the difference between observed frequencies (f_o) and expected frequencies (f_e), i.e. ($f_o - f_e$).
- ❖ After that we will find out ($f_o - f_e$)² and divide each value of ($f_o - f_e$)² by f_e , i.e., to find out the sum of all the values $\frac{(f_o - f_e)^2}{f_e}$
- ❖ After totaling all the $\frac{(f_o - f_e)^2}{f_e}$ we will get the value of X^2 In other words
- ❖ Then we will compare the calculated value of X^2 with the table value (or, critical value) for the given degrees of freedom and at the desired level of significance.
- ❖ If the calculated value of X^2 is more than the table value, then the difference between observed frequencies and expected frequencies is significant. Similarly if the calculated value of X^2 is less than the table value, then the difference between the observed frequencies and expected frequencies is not significant and it may have arisen due to sampling fluctuations.

1. Degrees of Freedom : During the comparison between the calculated value and the table value of X^2 we are supposed to find out the degrees of freedom. Here the degrees of freedom (d_f or v) is calculated by the following formula:

$$d_f = (c-1) (r-1)$$

Where c is the number of columns and r is the number of rows.

In addition, if we fit a binomial distribution, the number of degrees of freedom is one less than the number of classes. If we fit a Poisson distribution, the degrees of freedom are 2 less than the number of classes, and if we fit a normal curve, the number of degrees of freedom is small by 3 than the number of classes. The critical or table values of X^2 are available only up to 30 degrees of freedom. Beyond that, the distribution of $\sqrt{2X^2}$ approximates the normal distribution. For d_f beyond 30, the approximation is acceptably close. The mean of distribution of $\sqrt{2X^2}$ is $\sqrt{2d_{f-1}}$ and $s = 1$. Thus, the application of the test is simple. Deviation $\sqrt{2X^2}$ from $\sqrt{2d_{f-1}}$ Here z is like 't' value for larger sample.

2. Constant of X^2 Distribution : The constants of X^2 distribution with d_f degrees of freedom are

$$i. \quad \text{Mean} = 2 d_f, \text{ Mode} = d_f - 2, \text{ Variance} = 2 d_f$$

ii. Moments = $m_0=0, m_2=2df, m_3=8df$ and $m_4=48df+12df^2$

iii. $b_1 = \frac{m_3^2}{m_2^2} = \frac{64df^2}{8df^2} = \frac{8}{df} \quad \backslash \quad g_1 = \sqrt{8/df}$

iv. $b_2 = \frac{m_4}{m_2^2} = \frac{48df+12df^2}{4df^2} = 3 + \frac{12}{df} \quad \backslash \quad g_2 = b_2 - 3 = \frac{12}{df}$

3.Alternative Method of X²: When the data are arranged into 2 × 2 consistency table as shownbelow :

A	B	A + B
C	D	C + D
(A + C)	(B + D)	N

Then $X^2 = \frac{N(AD-BC)^2}{(A+B)(C+D)+(A+C)(B+D)}$

When in a four cells of 2 × 2 consistency table, any one of the frequency is less than 5 then yate’s correction is applied and the value of C² is determined by the following formula

$X^2 = \frac{(\backslash AD-B \ /-N/2)^2}{(A+B)(C+D)+(A+C)(B+D)}$

X² = can also calculated by the following formula: $X^2 = \frac{\sum f_o^2}{f_e} - N$

This formula is helpful if there are fractions in the deviations and the use of earlier direct formulabecomes tedious.

13.2.1 Applications of Chi Square Test :

Applications of X² Test : This test is widely used and a very popular and prevalent method in mass communication today as it is applicable under a variety of conditions and situations. Since it is assumption –free *non-parametric tests*, so it can be easily applied to several problems of communication and media studies. Most important applications of X²are as follows.

1. Test of Independence
2. Test of Goodness of Fit
3. Test of Homogeneity
4. Test of population variance.

1. As a Test of Independence : With the application of X², we can easily test whether two or more attributes like the benefits of new media in education and readers’ opinion about the same are independent of each other or they are associated or related. We normally take the null hypothesis that there is no association (or relation) in attributes or in other words, –the two attributes are independent. If the calculated value of X²is more than the table value at a particular level of significance, then we say that the difference is significant at that level and the null hypothesis is rejected. That is to say that there is some association (or relation) between the attributes or in other words the two attributes are not independent. Similarly, if the calculated value of X² is less than the table value at a particular level of significance, then we say that the difference is not significant at that level and the null hypothesis is accepted. That is to

say that there is no association (or relation) between the attributes or in other words the two attributes are independent.

2. **As a Test of Goodness of Fit :** It is also known as a test of concordance which enables us to understand how well the assured theoretical distribution fits to the observed data. If some probability distribution (theoretical distribution) is fitted to the given data, we can understand how well this distribution fits with the observed data. When the calculated value of X^2 is less than the critical value (or tabulated value) at a particular level of significance i.e., null hypothesis is accepted and the fit is considered to be good one which means the difference between the observed and expected frequencies (or hypothesized value) is due to sampling fluctuations. But if the calculated value of X^2 is greater than its critical value, the fit is not considered to be a good one. Thus, this test also helps in judging whether the distribution is continuous or discrete.
3. **As a Test of Homogeneity:** It is basically an extension of test of independence. Test of homogeneity tries to understand whether two or more independent random samples are drawn from the same universe or from different universe. It is important to mention here that in both tests of independence and tests of homogeneity, we are related to cross-classified data. Hence, both tests are performed similarly. Test of independence is related to the problem of whether one attribute is independent of other or not? But the test of homogeneity is related to the problem whether different samples come from same universe. Moreover, test of independence involves a single sample taken from one universe. But test of homogeneity involves two or more independent samples one from each of the probable universe under the study.
4. **As a Test of Population Variance :** This type of test is applied to find out whether the *variance* in the population could be a specified numeric value. This has a vital role in scientific accuracy of media analysis. Hence, the value of Y^2 can be derived by the following formula := $\frac{NS^2}{s^2} = \frac{\sum(X-X)^2}{s^2}$
 Where N is the size of the sample,
 S is the standard deviation of the sample,
 s is the standard deviation of the universe.

Hence the degree of freedom is taken as (N-1). Thus, by comparing the calculated value and the critical value at about d_f and at a particular level of significance, the null hypothesis may be accepted or rejected.

It is important to mention here that this test is applicable only when the universe is normal.

13.2.2 Additive Property of Chi Square :

X^2 as an unique characteristics of addition which is very useful in the field of communication and media. When a number of sample studies are conducted in the same field of communication and media, then the results can be combined together for achieving an accurate idea about the actual position. It must be noted that during addition of the values of X^2 , two important aspects must be taken care of :

- i. The pooled result in a single inclusive test is proper if the samples are independent.
- ii. If the values of X^2 are to be combined together, Yate's corrections should be avoided because the addition theorem is applicable only for uncorrelated

constituent elements

13.2.3 Misuse and Precaution of Chi Square Test :

This test is the most widely, largely and frequently used technique. But at the same time, this test is also described as the most widely, largely and frequently misused procedure in the field of communication and media studies because of underexposure of this statistical device. It is very easy to learn its formula and computation, but very difficult to apply this test in a scientific, reasonable and judicious manner.

Its most common misuse in mass communication is the violation of independence between measures. This assumption of independence should not be confused and misunderstood as a test of independence. Regarding the misuse of X^2 test in the field of behavioural communication and media studies has been suitably manifested in the *Journal of Experimental Psychology* as follows:

- (i) It is often applied even when the frequencies are very small. This is nothing but misuse of this test. Normally no frequency should be less than 10. If the frequency is less than 5 then either it should be with the adjacent frequency or yate's corrections should be used.
- (ii) Frequency of non-occurrence is often neglected. But it should never be omitted.
- (iii) X^2 is often used even if the original data is not available. This can be described as a misuse. For example, when we have only proportions, percentages or rates, this test of X^2 should not be applied.
- (iv) Often the repeated measurements are made on the same units. This should be strictly avoided. For example, if the frequencies are given in percentage, then it does not reflect the size of N. If it is mentioned that out of 20 journalists, 8 are taken or out of 100 advertisers, 40 are taken, then the percentage is 40%. But this cannot be treated as size of N. Such percentages must be multiplied by N/100 for computations.
- (v) Another misuse of this test is that there are often failures to equalize the sum of observed frequencies and the sum of expected frequencies. This should be taken care and precautions should be there to equalize the same.
- (vi) Many a time the expected values are not calculated on a rational basis.
- (vii) Often the hypothesis is not properly set up. It can destroy the entire calculations and even the final results. So precautions needed here too.
- (viii) Similarly the calculations of degree of freedom are often mistaken. This is very important and conclusive aspect. So it should not be avoided.
- (ix) Often there are confusions between critical value and calculated value. This is again a serious matter. The calculated value is derived and the table value is available. The table value is also called the tabulated value or the critical value. The table must be seen a particular level which is generally at 0.05 level or at 0.01 level of significance.
- (x) Further, the misuse of indeterminate theoretical frequencies, incorrect or questionable categorizing, incorrect calculations, and use of non-frequency data should be avoided.
- (xi) It is highly surprising that the rate of growth of misuse of X^2 test is increasing faster than that the rate of growth of its use. Joseph Stegar in his pioneering work "*Readings in Statistics*" and further Lewis and Burke in their joint research paper

have strictly cautioned, “*One cannot simply use a statistic because one knows how to calculate it. One must understand the rationale behind its development and its limitations on its applications imposed by its assumptions underlying it.*” Hence, precautions must be taken carefully and meticulously to avoid the various misuses of X^2 as discussed above.

13.2.4 Illustration of Various Applications Chi Square Test :

Illustration of Various Applications of X^2 Test :

In media metrics, all the uses, misuses, precautions and suitability of the various applications will be further exemplified and simplified by the following problem-solving illustrations.

e.g. (i) In a classroom of 60 students of mass communication, 40 said ‘yes’ to a proposal, 10 said ‘no’ to the proposal and 10 remained neutral about the proposal. Find whether the opinion of the students is significant?

Solution : Let us take the hypothesis that the opinion of the students is not significant. Now, let us put the question into the following table:

	<i>Yes</i>	<i>No</i>	<i>Neutral</i>	<i>Total</i>
f_o	40	10	10	N=60
f_e	20	20	10	
$f_o - f_e$	20	-10	-10	
$(f_o - f_e)^2$	400	100	100	
$(f_o - f_e)^2 / f_e$	20	5	5	

$$X^2 = \sum \frac{(f_o - f_e)^2}{f_e} = 6.67 + 0.60 + 3.26 + 10.52$$

$$d_f = (r - 1)(c - 1) = (2 - 1)(3 - 1) = 2$$

The table value of X^2 with $d_f = 2$ at 0.01 level is 9.21 and at 0.05 level is 5.99. Since the calculated value with $d_f = 2$ at both the levels is much more than the table value, so the difference is very significant. Hence the null hypothesis is rejected. Thus, the opinion of the students is significant.

e.g. (iii) In a media school, 90 students were asked to give their preference for specialization, in which *electronic media*, *print media* and *advertising* preferences were given by 28, 35 and 27 students respectively. Is this difference of preference significant?

Solution : Let us take the hypothesis that there is no significant difference of preference of specialization among the students.

Now, let us put the question into the following table :

	<i>Electronic Media</i>	<i>Print Media</i>	<i>Advertising</i>	<i>Total</i>
f_0	28	35	27	N=90
f_e	30	30	30	
$f_0 - f_e$	-2	5	-3	
$(f_0 - f_e)^2$	4	25	9	
$(f_0 - f_e)^2 / f_e$	0.133	0.833	0.3	

$$X^2 = \sum \frac{(f_0 - f_e)^2}{f_e} = 0.133 + 0.833 + 0.30 = 1.266$$

$$d_f = (r - 1)(c - 1) = (2 - 1)(3 - 1) = 2$$

The table value of X^2 with $d_f = 2$ at 0.01 level and 0.05 level is 9.21 and 5.99 respectively. Since the calculated value is less than the table value with $d_f = 2$ at both the levels, so the difference is not significant. Hence the null hypothesis is accepted. Thus, the difference of preference of specialization among students is not significant.

e.g. (v) On 50 viewers of television advertisement, the colour preference was experimented in which 14 viewers preferred red, 8 preferred blue, 12 preferred yellow, 9 preferred green and 7 preferred orange. Find whether this difference of colour preference is significant?

Solution : Let us take the hypothesis that there is no difference of colour preference among viewers.

Tabular form :

	<i>Red</i>	<i>Blue</i>	<i>Yellow</i>	<i>Green</i>	<i>Orange</i>	<i>Total</i>
f_0	14	8	12	9	7	N=50
f_e	10	10	10	10	10	
$f_0 - f_e$	4	-2	2	-1	-3	
$(f_0 - f_e)^2$	16	4	4	1	9	
$(f_0 - f_e)^2 / f_e$	1.6	0.4	0.4	0.1	0.9	

$$X^2 = \sum \frac{(f_0 - f_e)^2}{f_e} = 1.6 + 0.4 + 0.4 + 0.1 + 0.9 = 3.4$$

$$d_f = (r - 1)(c - 1) = (2 - 1)(5 - 1) = 4$$

At $d_f = 4$, the table value of X^2 at 0.05 level and at 0.01 level is 9.48 and at 0.01 level is 13.27. Since the calculated value of X^2 is less than the table value at both the levels, so the difference is not significant. Hence, the null hypothesis is accepted. Thus, there is no difference of colour preference among viewers of television advertisements.

e.g. (vi) Whether new media is helpful in education or not? 100 users of *new media* were asked to give their views in which 24 strongly agreed, 17 agreed, 18 remained neutral, 23 disagreed and 18 strongly disagreed. Is this difference of views significant?

Solution : Let us take the hypothesis that there is no significant difference of views of users of

new media:

Now, converting the problem into a table:

	<i>Strongly Supported</i>	<i>Supported</i>	<i>Neutral</i>	<i>Opposed</i>	<i>Strongly opposed</i>	<i>Total</i>
f_0	30	25	10	17	18	N=100
f_c	20	20	20	20	20	
$f_0 - f_c$	10	5	-10	-3	-2	
$(f_0 - f_c)^2$	100	25	100	9	4	
$(f_0 - f_c)^2 / f_c$	5	1.25	5	0.45	0.2	

$$X^2 = \sum \frac{(f_0 - f_c)^2}{f_c} = 0.8 + 0.45 + 0.2 + 0.45 + 0.2 = 2.10$$

$$d_f = (r - 1)(c - 1) = (2 - 1)(5 - 1) = 4$$

At d_f 4, the table value of X^2 at 0.05 level is 9.48 and at 0.01 is 13.27. Since the calculated value of X^2 at both the levels is less than the table value, so the difference is not significant. Hence, the null hypothesis is accepted. Thus, there is no significant difference of among users of new media as far as education is concerned.

e.g. (vii) when the opinions of 100 students of mass communication about a political party were collected, 30 students *strongly supported*, 25 *supported*, 10 remained *neutral*, 17 *opposed* and 18 *strongly opposed*. Find whether the difference of opinion is significant?

Solution : Let us take the hypothesis that the difference of opinion of students is not significant. Now putting the question in a tabular form:

	<i>Strongly Supported</i>	<i>Supported</i>	<i>Neutral</i>	<i>Opposed</i>	<i>Strongly Opposed</i>	<i>Total</i>
f_0	30	25	10	17	18	N=100
f_c	20	20	20	20	20	
$f_0 - f_c$	10	5	-10	-3	-2	
$(f_0 - f_c)^2$	100	25	100	9	4	
$(f_0 - f_c)^2 / f_c$	5	1.25	5	0.45	0.2	

$$X^2 = \sum \frac{(f_0 - f_c)^2}{f_c} = 5 + 1.25 + 5 + 0.45 + 0.2 = 11.90$$

$$d_f = (r - 1)(c - 1) = (2 - 1)(5 - 1) = 4$$

At d_f 4, the table value of X^2 at 0.05 level and 0.01 level is 9.48 and 13.27 respectively. The calculated value of X^2 at 0.05 level is more than the table value and 0.01 level is less than the table value, so the difference is significant at 0.05 level and not significant at 0.01 level. Hence the null hypothesis is rejected at 0.05 level and is accepted at 0.01 level. Thus, the difference of opinion of students at 0.05 level is significant and at 0.01 level is not significant.

e.g. (viii) The number of attacks on journalists per month in a certain town was as follows: 12, 8, 20, 2, 14, 10, 15, 6, 9, 4. Use chi-square test to determine if these frequencies are in

agreement with the belief that attacks on journalist in that town were the same during 10-month period.

Solution: Let us take the hypothesis that attacks on journalists were the same during the period we should expect:

$$\frac{12 + 8 + 20 + 2 + 14 + 10 + 15 + 6 + 9 + 4}{10} \\ = 10 \text{ attacks per month applying } X^2 \text{ test,}$$

f_0	12	8	20	2	14	10	15	6	9	4
f_e	10	10	10	10	10	10	10	10	10	10
$f_0 - f_e$	2	-2	10	-8	4	0	5	-4	-1	-6
$(f_0 - f_e)^2$	4	4	100	64	16	0	25	16	1	36
$(f_0 - f_e)^2 / f_e$	0.4	0.4	10	6.4	1.6	0	2.5	1.6	0.1	3.6

$$X^2 = \sum \frac{(f_0 - f_e)^2}{f_e} = 26.60$$

$$df = (n - 1) = (10 - 1) = 9$$

$$\text{At } df = 9 \quad X_{0.05}^2 = 16.92$$

Since the calculated value of X^2 is more than the table value, so the difference is significant. Hence, the null hypothesis is rejected. Thus, the frequencies are not in agreement with the belief that theno. of attacks were the same as during the last 10 months period.

13.2.5 Usage of 'Z' and 'T' Tests :

The z-Test is used to measure the difference between any variable value (x) and the mean of all variable values or 'x' values, which is indicated by ' μ ', divided by the standard deviation(s). It is based on the normal probability distribution. In following cases, we find the usage of z-test:

- To judge the significance of statistical measures, particularly the mean. This is done by comparing the observed value (test statistic) with the probable value (table value) at a specified level of significance.
- It is used to compare the mean of a sample with some hypothesized mean of the population.
- It is also used to judge the significance of difference between means of two independent samples.
- It can also be used for judging the significance of difference between sample and population proportion or proportions of two independent samples.
- Finally this test can also be used for measuring the significance of medium, mode, coefficient of correlation and other measures.

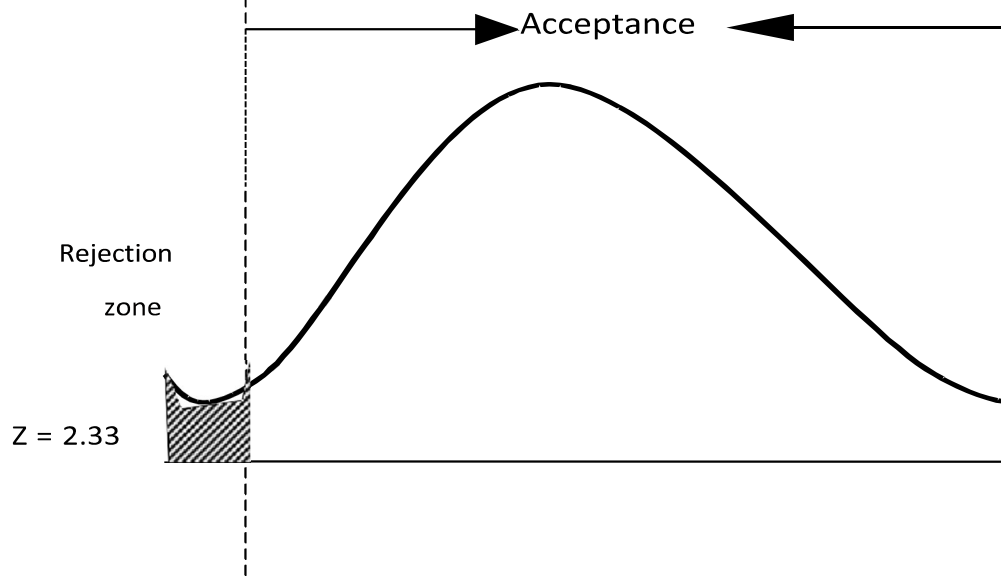
We shall now illustrate usage of z-tests in testing hypothesis with some real life examples. In case of small sample and in cases when sample variance is taken as population variance (population variance, being not known) we use t-test based on t-distribution to judge the significance of sample mean or significance of difference between two sample means. Normally when sample size is less than 30 we use t-test.

e.g. The supporters of a party, claim that their party has 60% following in Delhi. To test this claim a city paper undertakes a test sample of 400 persons on the Internet, 220 people support this claim. Test at 1% level of significance whether the supporters claim is correct or not.

Solution: it is a one tailed test $H_0 P = 0.6$

$H_1 p < 0.6$

Sample Proportion $P_x = \frac{220}{400} = 0.55$



$$\begin{aligned} SE &= \sqrt{\frac{p(1-p)}{n}} = \sqrt{\frac{pq}{n}} \\ &= \sqrt{\frac{0.6(0.4)}{400}} = 0.0245 \\ Z &= \frac{P_x - P}{\sqrt{pq/n}} = \frac{0.55 - 0.6}{0.0245} = -2.04 \end{aligned}$$

$|z|$ tabulated = -2.33 at 1% level of significance.

As, calculated value < tabulated, null hypothesis is accepted.

∴ Claim of the party is valid

13.3 SUMMARY :

Though parametric tests are more reliable, suitable, powerful and effective than the non-parametric tests, yet chi-square test is considered to be the most popular and widely used non-parametric tests for statistical analysis and statistical inference. Applied by British Statistician Karl Pearson in 1990 and denoted by Greek Letter χ^2 , is given by the following formula:

$$\chi^2 = \sum \frac{(f_o - f_e)^2}{f_e}$$

In a situation of multiple expected frequencies, the expected frequency is calculated by the following formula

$$f_c = \frac{RT \times CT}{N}$$

The Degree of Freedom is denoted by df or v is given by the formula: df (or v) = $(n-1)$ $(r-1)$. When data are arranged into 2×2 consistency table, then

$$X^2 = \frac{N(AD-BC)^2}{(A+B)(C+D)+(A+C)(B+D)}$$

When in a four cells of 2×2 consistency table, any one of the frequency is less than 5, then yate's correction is applied and the value of c^2 is determined by the following formula:

$$X^2 = \frac{(\sqrt{AD-BC} - N/2)^2}{(A+B)(C+D)+(A+C)(B+D)}$$

Most important applications of X^2 are : Test of Independence, (ii) Test of Goodness of Fit, (iii) Test of Homogeneity, (iv) Test of Population Variance. As a Test of Independence one can test whether two or more attributes are independent of each other or they are associated or related. As a Test of Goodness of fit, one can test how well the assumed theoretical distribution is fitted to the given data. Test of homogeneity, an extension of test of independence, involves two or more independent samples, one from each of the probable universe under the study. Test of population variance is applied whether the variance in the population could be specified numeric value. The rate of growth of misuse of c^2 its use.

13.4 KEY WORDS :

- ❖ **Chi-Square:** It is a specific type of cross-tabulation between observed frequencies and expected frequencies which is followed by a certain formula.
- ❖ **Test of Independence:** As a Test of Independence one can test whether two or more attributes are independent of each other or they are associated or related.
- ❖ **Test of Goodness of Fit:** As a Test of Goodness of fit, one can test how well the assumed theoretical distribution is fitted to the given data.
- ❖ **Test of Homogeneity:** This test is an extension of test of independence which involves two or more independent samples, one from each of the probable universe under the study.
- ❖ **Test of Population Variance:** It is applied whether the variance in the population could be specified numeric value. c^2 test called test of goodness

13.5 SELF-ASSESSMENT QUESTIONS :

1. What is X^2 test ? Under what conditions is it applicable? Why is of fit?
2. Describe the X^2 test of significance and state the several uses to which it can be put?
3. What is 3^2 test of goodness of fit? What precautions are necessary.
4. What are the basic conditions for the application of X^2 - test ?
5. 3^2 test is a test of independence, homogeneity and goodness of fit. Discuss briefly.
6. Why is X^2 test so popular despite being a non-parametric test?
7. Describe the meaning and concept of 3^2 . Explain the conditions for using X^2 test.
8. Explain the applications of X^2 test : (a) As a test of Independence, (b) As a Test of Goodness of Fit, (c) As a Test of Homogeneity, (d) As a Test of

Population Variance.

9. Explain the alternative method and additive property of χ^2 test.
10. It is said that χ^2 test is more widely misused test, even though it the most popular parametric test. Comment.
11. How is χ^2 test used if the degree of freedom is more than 30?
12. Explain the constants of chi-square distribution.
13. Enumerate the various formulae for measuring the value of χ^2 .
14. Which formula will you use for measuring the value of χ^2 if there are more fractions in the deviations and the use of earlier direct formula becomes tedious?

13.6 REFERENCES AND SUGGESTED READINGS :

1. Anderson, J.A. 1987. „*Communication Research: Issues and Methods*“, New York: Mc Graw-Hill
2. Berger, Arthur Asa .1998. „*Media Research Techniques*“ (Second Edition) USA: Sage Publications.
3. David, P. 2000. „*Precision Journalism: A practical Guide*“ USA: Sage Publications.
4. Dayal, M. 2010. „*Media Shodh*“ (3rd edition), Panchkula: Haryana Sahitya Akademi.
5. Dayal, M. 2017. „*Media Metrics: An Introduction to Quantitative Research in Mass Communication*“ .pp 339-345, New Delhi: Sage Publications.
6. Elhance, D.N.S & Elhance, Veena .1990. „*Fundamentals of Statistics*“, pp. 24.1-24.54, New Delhi: Kitab Mahal.
7. Gupta, S.P. 2011. „*Statistical Methods*“ pp. 954-1003, New Delhi: Sltan Chand & Sons.
8. Gupta, S.P. & Gupta, Alka. 2008. „*Statistical Methods in Behavioural Sciences*“ pp 412 – 436, Allahabad: Sharda Pustak Bhawan.
9. Hussain, Shamsad .1986. : „*Elementary Statistics in Psychology and Eduation*“, pp. 120-136, New Delhi: Motilal Banarasidas, Delhi, pp. 120-136.\
10. Kothari, C.R. 2000. „*Research Methodology: Methods & Techniques*“ pp 277-299, New Delhi: Wishwa Parkash.
11. Krishnaswami, O.R & Ranganatham, M. 2005. „*Methodology of Research*“, New Delhi: Himalaya Publications.
12. Karlinger, F.N .1983. „*Fundamental of Behavioural Research*“ (2nd Edition) pp 162 – 172. USA: Holt, Rinehart and Winston.
13. Mangal. S.K. (1987) : *Statistics in Psychology and Education*, Tata McGrad Hill Publishing Comapany Limited, Delhi pp. 169-191
14. Sarangi, Prasant. 2010. „*Research Methodology*“ Delhi: Taxman Publication.
15. Wilkinston, T,S, & Bhandarkar, P.L. 1996. „*Methodologies and Techniques of Social Research*“ Mumbai: Himalaya Publishing House.
16. Wimmer, Roger D & Dominick, Joseph R .2003. „*Mass Media Research: An Introduction*“ (7th edition), pp 280 – 284, USA Thomson & wadsworth.

Dr. J. Madhu Babu

LESSON - 14

ANOVA AND MULTIPLE REGRESSIONS

OBJECTIVES :

After going through this lesson you will be able to:

- Describe meaning and application of ANOVA ;
- Describe meaning and application Multiple Regression Analysis.

STRUCTURE :

- 14.1 Introduction
- 14.2 Basic ANOVA concepts
 - 14.2.1 Hypotheses of ANOVA
 - 14.2.2 One-Way ANOVA
 - 14.2.3 Application of ANOVA
- 14.3 Multiple Regression Analysis
- 14.4 Summary
- 14.5 Keywords
- 14.6 Self-Assessment Questions
- 14.7 References/Suggested Readings

14.1 INTRODUCTION :

Analysis of variance (ANOVA) enables us to infer whether population from which we have drawn more than two samples are having the same mean values. By testing the significance of difference between more than two sample means we make such inferences. To test the equality of variances, we use 'F' test. By comparing the observed 'F' value with table value of 'F' and analyzing the significance of differences (if any) we test our hypothesis. Before elaborating on the 'F' test, it is important for us to understand the principles and techniques of analysis of variance (ANOVA). Only after analysis of variance, we can submit it to 'F' test to make necessary inferences.

14.2 BASIC ANOVA CONCEPTS :

The Setting :

Generally, we are considering a quantitative response variable as it relates to one or more explanatory variables, usually categorical. Questions which fit this setting:

- i. Which academic department in the sciences gives out the lowest average grades?
(Explanatory variable: *department*; Response variable: *student GPA's for individual courses*)
- ii. Which kind of promotional campaign leads to greatest store income at Christmas time?
(Explanatory variable: *promotion type*; Response variable: *daily store income*)
- iii. How do the type of career and marital status of a person relate to the total cost in annual

claims she/he is likely to make on her health insurance. (Explanatory variables: *career* and *marital status*; Response variable: *health insurance payouts*)

14.2.1 Hypotheses of ANOVA :

These are always the same.

H₀: The (population) means of all groups under consideration are equal.

H_a: The (population) means are not all equal. (Note: This is different than saying “they are all unequal”!)

❖ Basic Idea of ANOVA :

Analysis of variance is a perfectly descriptive name of what is actually done to analyze sample data acquired to answer problems. Side-by-side boxplots like these in both figures reveal differences between samples taken from three populations.

❖ Assumptions of ANOVA :

Like so many of our inference procedures, ANOVA has some underlying assumptions which should be in place in order to make the results of calculations completely trustworthy. They include:

- i. Subjects are chosen via a simple random sample.
- ii. Within each group/population, the response variable is normally distributed.
- iii. While the population means may be different from one group to the next, the population standard deviation is the same for all groups.

Fortunately, ANOVA is somewhat robust (i.e., results remain fairly trustworthy despite mild violations of these assumptions). Assumptions (ii) and (iii) are close enough to being true if, after gathering SRS samples from each group, you:

- (ii) Look at normal quartile plots for each group and, in each case, see that the data points fall close to a line.
- (iii) Compute the standard deviations for each group sample, and see that the ratio of the largest to the smallest group sample S.D. is no more than two.

14.2.2 One-Way ANOVA :

When there is just *one* explanatory variable, we refer to the analysis of variance as *one-way ANOVA*.

1. Notation :

Here is a key to symbols you may see as you read through this section.

k = the number of groups/populations/values of the explanatory variable/levels of treatment

n_i = the sample size taken from group i

x_{ij} = the j th response sampled from the i th group/population.

\bar{x}_i = the sample mean of responses from the i th group = $\frac{1}{n_i} \sum_{j=1}^{n_i} x_{ij}$

s_i = the sample standard deviation from the i th group = $\sqrt{\frac{1}{n_i-1} \sum_{j=1}^{n_i} (x_{ij} - \bar{x}_i)^2}$

n = the (total) sample, irrespective of groups = $\sum_i^k n_i$

\bar{x} = the mean of all responses, irrespective of groups = $\frac{1}{n} \sum_{ij} x_{ji}$ Splitting the Total Variability into Parts

Viewed as one sample (rather than k samples from the individual groups/populations), one might measure the total amount of variability among observations by summing the squares of the differences between each x_{ij} and \bar{x} :

$$\text{SST (stands for sum of squares total)} = \sum_{i=1}^k \sum_{j=1}^{n_i} (x_{ij} - \bar{x})^2$$

This variability has two sources :

1. Variability between group means (specifically, variation around the overall mean \bar{x})

SSG: = $\sum_{i=1}^k n_i (\bar{x}_i - \bar{x})^2$ and Variability within groups means (specifically, variation of observations about their group mean \bar{x}_i): $= \sum_{i=1}^k \sum_{j=1}^{n_i} (x_{ij} - \bar{x}_i)^2 = \sum_{i=1}^k (n_i - 1) s_i^2$ It is the case that $\text{SST} = \text{SSG} + \text{SSE}$.

2.The Calculations :

If the variability between groups/treatments is large relative to the variability within groups/treatments, then the data suggest that the means of the populations from which the data were drawn are significantly different. That is, in fact, how the F statistic is computed: it is a measure of the variability between treatments divided by a measure of the variability within treatments. If F is large, the variability between treatments is large relative to the variation within treatments, and we reject the null hypothesis of equal means.

If F is small, the variability between treatments is small relative to the variation within treatments, and we do not reject the null hypothesis of equal means. (In this case, the sample data is consistent with the hypothesis that population means are equal between groups.) To compute this ratio (the F statistic) is difficult and time consuming. Therefore we are always going to let the computer do this for us. The computer generates what is called an ANOVA table:

Source	SS	df	MS	F
Model/Group	SSG	$K - 1$	$\text{MSG} = \frac{\text{SSG}}{k-1}$	$\frac{\text{MSG}}{\text{MSE}}$
Residual/Error	SSE	$n - k$	$\text{MSE} = \frac{\text{SSE}}{n-k}$	
Total	SST $n-1$			

What are these things?

- The *source* (of variability) column tells us SS=Sum of Squares (sum of squared deviations):

SST measures variation of the data around the overall mean \bar{x}

SSG measures variation of the group means around the overall mean

SSE measures the variation of each observation around its group mean \bar{x}_i

Degrees of freedom :

$k - 1$ for SSG, since it measures the variation of the k group means about the overall mean

$n - k$ for SSE, since it measures the variation of the n observations about k group means

$n - 1$ for SST, since it measures the variation of all n observations about the overall mean

- **MS = Mean Square = $\frac{SS}{df}$:**

This is like a standard deviation. Look at the formula we learned back in Chapter 1 for sample standard deviation. Its numerator was a sum of squared deviations (just like our SS formulas), and it was divided by the appropriate number of degrees of freedom. It is interesting to note that another formula for MSE is

$$MSE = \frac{(n_1-1)s_1^2 + (n_2-1)s_2^2 + \dots + (n_k-1)s_k^2}{(n_1-1) + (n_2-1) + \dots + (n_k-1)}$$

which may remind you of the pooled sample estimate for the population variance for 2-sample procedures (when we believe the two populations have the same variance). In fact, the quantity MSE is also called S_p^2 .

- **The F statistic = $\frac{MSG}{MSE}$:**

If the null hypothesis is true, the F statistic has an F distribution with $k - 1$ and $n - k$ degrees of freedom in the numerator/denominator respectively. If the alternate hypothesis is true, then F tends to be large. We reject H_0 in favor of H_a if the F statistic is sufficiently large.

As with other hypothesis tests, we determine whether the F statistic is large by finding a corresponding P -value. For this, we use Table E. Since the alternative hypothesis is always the same (no 1-sided vs. 2-sided distinction), the test is single-tailed (like the chi-squared test). Nevertheless, to read the correct P -value from the table requires knowledge of the number of degrees of freedom associated with both the numerator (MSG) and denominator (MSE) of the F -value.

Look at Table E. On the top are the numerator d.f., and down the left side are the denominator df. In the table are the F values, and the P -values (the probability of getting an F statistic larger than that if the null hypothesis is true) are down the left side.

Example : Determine

$$P(F_{3,6} > 9.78) = 0.01$$

$$P(F_{2,20} > 5) = \text{between } 0.01 \text{ and } 0.025$$

14.2.3 Application of ANOVA :

To test the equality of variances for two sample data, we use 'F' test, the formula for which is

$$F = \frac{\sigma_{s1}^2}{\sigma_{s2}^2} = \text{with d.f. } (n_1-1) \text{ for sample 1.}$$

(n_2-1) for sample 2.

$$\sigma_{s1}^2 = \text{variance of sample 1}$$

$$\text{Calculated as } \frac{\sum(X_{2i} - \bar{X}_1)^2}{(n_1-1)}$$

we can also use n_1 when sample size is 30 or more.

$$\sigma_{s1}^2 = \text{variance of sample 2}$$

Calculated as $\frac{\sum(X_{2i} - \bar{X}_2)^2}{(n_2 - 1)}$

X_{1i} = is the observed frequencies of sample 1.

X_{2i} = observed frequencies of sample 2. \bar{X}_1 and \bar{X}_2 respective mean values of samples 1 and 2.

But ANOVA technique is important to compute variances when we want to compare more than two samples to infer whether the said samples have been drawn from population having the same mean values. Without going into the trouble of considering all possible combinations of populations, ANOVA technique enables the researcher to investigate the mean differences of all the populations simultaneously.

We have introduced the term dependent variable and independent variable with an example earlier. ANOVA enables a researcher to investigate factors which influence a dependent variable. Let us assume productivity rate as a dependent variable. We know that productivity usually depends on multiple factors like level of technology used in production, workers' skill, incentives and gain sharing, training, supervision and industrial relations. For us all these factors are independent variables. Now a researcher may be interested in knowing the differences amongst these factors, i.e. which one of the above factors is more effective in raising productivity (by studying their respective differences at various level of significances) or he may be interested in investigating the differences amongst various categories of same factor (for example various methods on incentive and gain sharing) which influence productivity. When we investigate various categories of one factor only, we carry out one-way ANOVA.

But when we are interested in investigating two factors at a time, we carry out two-way ANOVA. Likewise with the help of ANOVA, we can investigate, in general any number of factors and also various categories within each factor to study their respective influences. Thus we find following steps are usually involved in analysis of variance.

- i. Estimation of population variance based on variance between the sample means.
- ii. Estimation of population variance based on variance within the sample means.
- iii. The third step is to compare the obtained F value with the respective F table value at given degrees of freedom and to judge the differences at desired levels of significances. The formula for computing 'F' value takes the following shape.

$$F = \frac{\text{MS between}}{\text{MS within}}$$

Now unless the above steps are explained with a practical example, it would be difficult for readers to understand F test in its true spirit. Let us first have an example of one-way ANOVA.

14.3 MULTIPLE REGRESSION ANALYSIS :

When there are two or more than two independent variables, the equation describing such a relationship is the multiple regression equation. In this situation, the results are interpreted as shown below.

Multiple regression equation assumes the form:

$$\hat{Y} = a + b_1X_1 + b_2X_2$$

Where X_1 and X_2 are independent variables being the dependent variable. The constants, a , b_1 and b_2 can be solved by solving the following three normal equations:

$$\begin{aligned}\sum Y_i &= na + b_1\sum X_{1i} + b_2\sum X_{2i} \\ \sum X_{1i}Y_i &= a\sum X_{1i} + b_1\sum X_{1i}^2 + b_2\sum X_{1i}X_{2i} \\ \sum X_{2i}Y_i &= a\sum X_{2i} + b_1\sum X_{1i}X_{2i} + b_2\sum X_{2i}^2\end{aligned}$$

- ❖ Please note the fact that the number of normal equations would depend upon the number of independent variables. If there are two independent variables, then three equations, if there are three independent variables then four equations and so on, are used.
- ❖ In multiple regression analysis, the regression coefficients (viz., b_1 , b_2) become less reliable as the degree of correlation between the independent variables (viz., X_1 , X_2) increases. If there is a high degree of correlation between independent variables, we have what is commonly described as the problem of multicollinearity. In such a situation we should use only one set of the independent variable to make our estimate.
- ❖ In fact, adding a second variable, say X_2 , which is correlated with the first variable, say X_1 , and distorts the values of the regression coefficients. Nevertheless, the prediction for the dependent variable can be made even when multicollinearity is present, but in such a situation care should be taken in selecting the independent variables to estimate a dependent variable so as to ensure that multicollinearity is reduced to the minimum.
- ❖ With more than one independent variable, we may make a difference between the collective effect of the two independent variables and the individual effect of each of them taken separately. The collective effect is given by the coefficient of multiple correlations

$R_{y. x_1 \times x_2}$ defined as under:

$$R_{y. x_1 \times x_2} = \sqrt{\frac{b_1 \sum Y_i X_{1i} - n \bar{X}_1 \bar{Y}_1 + b_2 \sum Y_i X_{2i} - n \bar{X}_2 \bar{Y}_2}{\sum Y_i^2 -}}$$

Alternatively, we can write

$$R_{y. x_1 \times x_2} = \sqrt{\frac{b_1 \sum X_{1i} Y_i + b_2 \sum X_{2i} Y_i}{\sum Y_i^2}}$$

14.4 SUMMARY :

- ❖ In case of population from which two means are drawn having the same mean values, the significant difference between these two can be found with the help of analysis of variance technique i.e. ANOVA. F-test is applied to check significant difference between the means. The value of F is the ratio of variance of the two samples. ANOVA is important in computing variances when there is comparison between more than two samples to inform whether the said samples have been drawn from population having the same mean values.
- ❖ The first step in ANOVA is the estimation of population variance based on variance between the sample means then the estimation of population variance based on variance within the sample means is done. The next step is to compare the obtained value of F with respect to F tabular value at given degree of freedom and to judge the differences at desired levels of significance.
- ❖ In multiple regression technique, two or more than two independent variables describe a relationship. In problem of multi collinearity, there exists a high degree of correlation between independent variables or measures.

14.5 KEYWORDS :

- ❖ **Factor** is an underlying dimension of several related variables.
- ❖ **Factor loadings** are the values that explain how closely the variables are related to each other.
- ❖ **Explanatory variable** is causal or independent variable and is also called extraneous variable.
- ❖ **Criterion variable** is resultant or dependent variable and is also called endogenous variable.
- ❖ **Observable variable** is directly observable.
- ❖ **Latent variable** is unobservable variable which may influence criterion variable.
- ❖ **Discrete variable** takes only integer value when measured.
- ❖ **Continuous variable** can assume any real value.
- ❖ **Cluster analysis** is a technique of measuring some measure of similarity.
- ❖ **Discriminant analysis** is used primarily to identify variables that contribute to differences in the priority defined groups with the use of discriminant functions.
- ❖ **ANOVA** is analysis of variance, where the comparison between means of samples drawn from some population having same mean values by testing the significance of difference between more than two sample means and inferences are made.
- ❖ **F-Ratio** is the ratio between mean square between columns and mean square of residual.
- ❖ **Multiple Regression** describes the relationship between two or more independent variables.

14.6 SELF-ASSESSMENT QUESTIONS :

1. Explain the meaning of the analysis of variance. Write down the one way analysis of variance table for testing the homogeneity of k-groups.
2. Explain the significance of ANOVA. What is the role of F-distribution in this analysis?
3. Explain the two-way ANOVA with example.
4. Give the multiple regression equation. Discuss its importance in research studies.

14.7 REFERENCES AND SUGGESTED READINGS :

1. Churchill, Gilbert. A (1983) Marketing Research: Methodological Foundations, The Dryden Press, New York.
2. Kothari C.R. (1990) Research Methodology: Methods and Technique. Wishwa Prakashan, New Delhi.
3. Mahalotra N.K.(2002) Marketing Research: An Applied Orientation. Pearson Education Asia.
4. Mustafi, C.K.1981.Statistical Methods in Managerial Decisions, Macmillan: New Delhi.
5. Raj,D.(1968),“Sampling Theory, “McGraw-Hill Book Company, New York.
6. Singh, D.and F.S. Chaudhary, 1986.Theory and Analysis of Sample Survey Designs, Wiley Eastern: New Delhi
7. Yates,E(1960),“Sampling Methods for Censuses and Surveys, “Charles Griffin & Company, Ltd., London.

Dr. J. Madhu Babu

LESSON - 15

DATA PROCESSING AND ANALYSIS

OBJECTIVES :

This lesson will elaborate on the meaning and utility of data analysis. It will provide a brief understanding of data processing, analysis, and interpretation in the research process. The major focus of this lesson is to guide data analysis - how to plan, collect, and manage the data so collected in quantitative research for a meaningful research outcome.

- able to understand the editing of data
- able to convert raw data into useful information
- able to classify data
- able to prepare tables and graphs using raw data

STRUCTURE :

- 15.1 Introduction
- 15.2 Defining data processing and analysis
- 15.3 Difference between Data Analysis, Processing, And Interpretation
- 15.4 Data Interpretation
- 15.5 Diagrammatic Presentations.
- 15.6 Summary
- 15.7 Glossary
- 15.8 Self-Assessment Questions
- 15.9 References and Suggested Readings

15.1 Introduction :

The data collected from the survey tool, observation, and interview is raw and is of no value unless and until it is presented in a usable manner. The data collected from the samples are arranged in a meaningful way by editing, and coding, and presented in tabular form for drawing useful inferences. In this unit, we will be learning the various ways by which raw data is converted into important information. The processing of data/information is an essential dimension of streamlining the facts and writing of a field report. A separate account of processing is given here.

15.2 DEFINING DATA PROCESSING AND ANALYSIS :

Processing and analysing data involves many closely related operations which are performed to summarize the collected data and organize these in a manner that answer the research questions (objectives). In any research, the step of analysis of the data is one of the most crucial tasks requiring proficient knowledge to handle the data collected as per the pre-decided research design of the project.

Analysis of data is defined by Prof Wilkinson and Bhandarkar as “Several closely related operations that are performed to summarize the collected data and organize these in such a manner that they will yield answers to the research questions or suggest hypothesis or questions if no such questions or hypothesis had initiated the study”.

15.3 DIFFERENCE BETWEEN DATA ANALYSIS, PROCESSING, AND INTERPRETATION :

The general understanding is that data analysis and processing are the same. However, some researchers and authors believe that both are two very distinct steps in the research process where data processing leads to data analysis. Let us understand the difference between the two in more detail. According to Francis Rommel, “the analysis and interpretation of data involve the objective material in the possession of the researcher and his subjective reaction and desires to derive from the data the inherent meaning in their relation to the problem. To avoid making conclusions or interpretations from insufficient or invalid data, the final analysis must be anticipated in detail when plans are being made for collecting information”.

A. Data Processing : The primary data collected from the field remains in the raw form of statements, digits, and qualitative terms. The raw data contains errors, omissions, and inconsistencies. It requires corrections after scrutinizing the completed questionnaires. The following steps are involved in the processing of primary data.

- i. Editing of data :** It is a process of examining the collected raw data to detect errors and omissions and to correct these when possible. It is also defined as the process relating to the review and adjustment of collected survey data to control the quality of the collected data. Data editing can be performed manually, with the assistance of a computer, or using a combination of both methods.

Data editing is crucial as it helps take full advantage of the available data to be converted into useful data, ensuring that the errors arising during collection, entry, and assimilation are omitted or minimized. It also assures that the consistency is coherent and consistent since such characteristics have a constructive impact on the final analysis and outcomes. The editing of data can be done at two stages: field and post-field editing. Field editing is a review of reporting by the investigator for completing what has been written in an abbreviated form during interviewing the respondent. The post-field editing is carried out when the field survey is completed and all the forms of the schedule have been collected. This type of editing requires the review of all forms thoroughly.

- ii. The coding of data :** The purpose of data coding is to bring out the essence and meaning of the data that has been collected from the respondents. To make sense of the data, it must be analysed. The analysis begins with the labeling of data as to its source, how it was collected, the information it contains, etc. When we have received hundreds of questionnaires, forms, and formats containing the data it seems impossible to figure out any outcomes just by looking at the quantum. E.g. if the Hotel guest's feedback is received in letter forms with no specific format it would be nearly impossible to assess the satisfaction levels, and major complaint areas or just find out who has been recommended by most of the guests as the best employee at the hotel.

Coding facilitates the researcher to reduce the bulk of information and data to a form that is easily understandable and can be interpreted soon either manually or through software programming. For example, the injury rate at different levels of intensive physical labor demanding operations in various hotels in the city may not be sorted under the name but each of the hotels can be assigned a numeric or alphabetical code. Content analysis computer programs help researchers to code textual data for qualitative or quantitative analysis.

To keep the response within limited alternatives, we need to assign some alphabetical or numerical symbols or both to the answers. The alternatives must be mutually exclusive i.e., defined in one concept or term only. This form of processing is known as coding. For example, in a question of educational qualifications choices given are: Uneducated; Below Matriculation; Matriculation & above but below Graduate; Graduate & above; Technical Diploma; Technical Degree.

The alphabetical codes assigned to these alternatives could be A, B, C, D, E, and F. Similarly, numerical codes to these alternatives could be 1,2,3,4., and 5 respectively. It is necessary for efficient analysis. Though the coding exercise is a part of the formulation of the questionnaire yet responses to questions need to be coded and made final at the processing stage. This simplifies the transfer of data from questionnaires to the master chart. It is a two-dimensional chart in which observations are entered on one axis (X) and details of the responses on the other axis (Y). The calculations become easier and quicker if the details are coded and entered in the master chart or fed into the computers.

- iii. **Organization of Data :** The data information collected through different sources should be organized. The first task in this regard is to develop a master chart. For example, in a local area survey, we record individual households in rows and the details of population, function, facilities, amenities, etc. in columns. Thus, a large chart is prepared that contains, practically, all relevant information/data. Finally, the total rows and columns are cross-checked. The information arranged in ascending order is known as the array of data. The set of information related to a specific entity is called the field. The following illustration demonstrates the way data is organized.
- iv. **Classification of data :** It is the process of arranging data in groups or classes based on common characteristics such as descriptive or numerical. A huge volume of raw data collected through field surveys needs to be grouped for similar details of individual responses. The process of organizing data into groups and classes based on certain characteristics is known as the classification of data. Classification helps in making comparisons among the categories of observations. It can be either according to numerical characteristics or according to attributes. The numerical characteristics are classified based on class intervals. For example, monthly income up to Rs.2000 may form its group and the respondent's reported income in the range may form its frequency. Similarly, the further group can also be made like income group Rs.2000 to Rs.3000 and so on. The number of items entered against each class is known as the frequency of the class. Every class has a lower and an upper limit. The difference between the upper and lower limits is known as the range of the class. The class intervals are mostly kept equal. Sometimes when the range of the data is too large class intervals are not kept equal, instead they are based on the perceptible gaps in the array of the data. For example, settlements having less than 2000 population can be grouped as below 200 population 200-500 population, 500-1000 population, and so on. In this group as class intervals are unequal.

The data is also classified on the following bases.

1. Descriptive characteristics-example land holding, sex, caste, and so on.
2. Time, situation, and area-specific characteristics.
3. Nature of data as continuous or discrete.

- ❖ **Simple Classification:** This means that one attribute is considered and the universe is divided into two classes. With one class consisting of items possessing the given attribute and the other class consisting of items that do not possess the given attribute.
- ❖ **Class interval Classification:** This is more relevant when we use quantitative data like the number of guests, number of spa users, age groups of tourists, income levels of travelers, daily occupancy, and other statistical data.
- ❖ **Qualitative classification :** When classification is done according to attributes, such as social status, physical status, nationality, etc., it is called qualitative classification.
- ❖ **Quantitative classification :** In quantitative classification, the data are classified based on quantitative characteristics. In other words, these characteristics can be measured quantitatively.
- ❖ **Temporal classification :** In this classification, time becomes the classifying variable and data are categorized according to time. Time may be in hours, days, weeks, months, years, etc.
- ❖ **Spatial classification :** When classification is done based on place, it is called spatial classification. The place may be a village/town, block, district, state, country, etc.
- ❖ The graphic presentation of data and information offers a quick and simple way of understanding the features and drawing comparisons. Further, it is an effective analytical tool and a graph can help us in finding the mode, median, etc.

B. Data Analysis : Data analysis depends upon the nature of the research that the researcher is undertaking. Types of data analysis vary depending on whether the research is qualitative or quantitative. In the present module, as earlier stated we will be studying various types of data analysis from the standpoint of quantitative research only.

- i. **Descriptive analysis :** According to C Emory, “descriptive analysis is largely the study of the distribution of one variable. This study provides us with profiles of companies, work groups, persons, and other subjects on multiple characteristics such as size, composition, efficiency, preferences, etc. For example, the researcher is collecting data from various law colleges in India to map the job preferences of the students in the final year of LL.B. In such a research job preferences like litigation, corporate, further studies, judiciary, etc become variable. Under it, statistical tools like percentages and means are used and the data is then represented through a graph. The data analysis may be having one variable also known as one-dimensional analysis or two variables/ bivariate analysis or more than two variables also described as multivariate analysis.
- ii. **Inferential analysis :** Inferential analysis is concerned with the various tests of significance for testing hypotheses to determine with what validity data can be said to indicate some conclusion or conclusions. It is also concerned with the estimation of population values. It is mainly based on inferential analysis and the task of interpretation (i.e., the task of drawing inferences and conclusions) is performed. For example, The researcher is studying access to the justice system in India and his hypothesis beings that the Indian justice delivery system favours the haves and marginalizes the have not’s. The data collected is from various stages in the delivery system like the police station, courts of justice, litigants, etc. Once the data is collected, proceeded then the researcher does inferential analysis to test the validity of the hypotheses.

General Characteristics of Analysis of the Data :

1. The researcher should keep in mind that the analysis of data will vary depending upon the type of study i.e. qualitative or quantitative or mixed.

2. The researcher should possess a thorough knowledge of the area of research as well as the data collected by him which will help in the analysis of data.
3. The data to be analysed and interpreted should: a. Be reproducible, b. Be readily disposed to quantitative treatment c. Have significance for some systematic theories, and can serve as a broad generalization.
4. The researcher should keep a clear set of hypotheses formulated at the very start of the research which will lead to clearer actions and better data collection as well as analysis.
5. In case the data collected is from vague clues rather than according to the specific hypothesis, in such cases, the data are analysed inductively or investigated during the process and not employing any prescribed set of rules.
6. For a successful study, the task of analysis and interpretation should be designed before the data is collected.

C. Presentation of data : The presentation of data could be tabular, statistical, or cartographic. In the case of a tabular form of presentation, data related to different variables should be classified and compared. Various statistical techniques are available to derive accurate and precise results. Since techniques have a large range coupled with limitations of their own, the selection of appropriate techniques needs to be made for the purpose. The construction of graphs, charts, diagrams, and maps are the various forms of cartographic presentations. The data is transformed into the cartographic system which is used for visual presentation. As data are generally voluminous, they need to be put in a compact and presentable form. So, the voluminous data collected could be made usable readily, and easily comprehended. There are generally three forms of presentation of data:

- Textual or Descriptive presentation
- Tabular presentation
- Diagrammatic presentation.

Textual Presentation of Data : In the textual presentation, data are described within the text. When the quantity of data is not too large this form of presentation is more suitable.

Tabular Presentation : It is the process of summarizing raw data and displaying the same in a compact form for further analysis. It is an orderly arrangement of data in columns and rows. It helps in the analysis of trends, relationships, and other characteristics of a given data. Simple tabulation is used to answer the question related to one characteristic of the data whereas complex tabulation is used to present several interrelated characteristics. Complex tabulation results in two-way, and three-way tables which give information about two or three interrelated characteristics of data.

In a tabular presentation, data are presented in rows (read horizontally and columns (read vertically)). The most important advantage of tabulation is that it organizes data for further statistical treatment and decision-making.

Tabulation is essential because:

- It conserves space and reduces explanatory and descriptive statements to a minimum.
- It facilitates the process of comparison.
- It facilitates the summation of items and the detection of errors and omissions.
- It provides the basis for various statistical computations.

Tabulation may also be classified as simple and complex tabulation. Simple tabulation generally results in one-way tables which supply answers to questions about one characteristic of data only. Complex tabulation usually results in two-way tables that give

information about two interrelated characteristics of data three-way tables, or still higher order tables known as manifold tables.

Components of Data Tables :

The components of data tables are as under:

- Table Number
 - Title
 - Head notes
 - Stubs
 - Caption
 - Body or field
 - Footnotes
 - Source
-
- ❖ **Table Number:** Each table should have a specific table number for ease of access and location. This number can be readily mentioned anywhere which serves as a reference and leads us directly to the data mentioned in that particular table.
 - ❖ **Title:** A table must contain a title that tells the readers about the data it contains, the time of the study, the place of study, and the nature of the classification of data.
 - ❖ **Headnotes:** A headnote further aids in the purpose of a title and displays more information about the table. Generally, headnotes present the units of data in brackets at the end of a table title.
 - ❖ **Stubs:** These are titles of the rows in a table. Thus a stub displays information about the data contained in a particular row.
 - ❖ **Caption:** A caption is the title of a column in the data table. It is a counterpart of a stub and indicates the information contained in a column.
 - ❖ **Body or field:** The body of a table is the content of a table in its entirety. Each item in a body is known as a 'cell'.
 - ❖ **Footnotes:** Footnotes are rarely used. In effect, they supplement the title of a table if required.
 - ❖ **Source:** When using data obtained from a secondary source, this source has to be mentioned below the footnote.

Statistical Presentation of data :

The data collected through various sources need to be processed statistically for precise explanations. Very often it becomes necessary to obtain a single representative value for the whole data set. The statistical measures that enable us to work out a single representative figure for the entire data distribution, is known as a central tendency. Measures of central tendency help us to compare different distributions besides being representative for each distribution. These measures normally denote the central points of values, distance, and occurrence in a distribution. The commonly used measures of central tendency are:

- (a) Arithmetic mean or average
- (b) Median
- (c) Mode

The various types of data that can be presented are:

- Textual presentation
- Data tables
- Diagrammatic presentation
- Time Series Data

- Bar Charts
- Combo Charts
- Pie Charts
- Tables
- Geo Map
- Scorecard
- Scatter Charts
- Bullet Charts
- Area Chart
- Text & Images

Presenting and analysing data :

1. Frame the objectives of the study and make a list of data to be collected and its format.
2. Collect/obtain data from primary or secondary sources.
3. Change the format of data, i.e., table, maps, graphs, etc. in the desired format
4. Sort data through grouping, discarding the extra data, and deciding the required form to make data comprehensible
5. Make charts and graphs to help to add the visual part and analyse trends.
6. Analyse trends and relate the information to fulfil the objectives.

Presenting the results :

- ❖ The results should be presented such that a progression of arguments is in support of the study beginning with a statement defining the purpose of the study and subsequently a logical presentation making objectives clear and related to the aim of the study.
- ❖ Bigger objectives should be broken down into smaller ones i.e., define each objective as per need and outcome. Prepare a list of data to be collected, the sources of data, and the form in which data exists and needs to be obtained, and conduct a primary survey for information that does not exist.
- ❖ Form and explain the methodology adapted to carry out a study.
- ❖ Sampling methods should be clear and confirmed for ease of collecting data that results in efficient and lesser errors in the process.
- ❖ Present only the required information and skip the background research to make your point clearer.
- ❖ Credits and references should either be provided at the end or wherever obligatory.
- ❖ The presentation methods depend upon the availability of resources and the type of results expected out of the final presentation. PowerPoint, Models, Paper Charts, Smart Boards, and Analytical software e.g., Google Analytics, etc can be used to make the presentation effective and crisp.

15.4 DATA INTERPRETATION :

Once the data has been processed and analysed, the final step required in the research process is the interpretation of the data. The line between analysis and interpretation is very thin. Through interpretation, one understands what the given research findings mean and what the underlying generalization is manifested through the data collected. This can be descriptive or analytical or theoretical. The data is interpreted from the point of the research questions and the hypothesis is tested. While interpretation is being done, generalizations are drawn. Thus, interpretation consists of conclusions that the researcher has reached after the data has been processed and analysed.

The basic concept of data interpretation is to review the collected data utilizing analytical methods and arrive at relevant conclusions. There are two methods to interpret the data:

1. **Qualitative method** – This method is used to analyse qualitative data or categorical data. The qualitative data interpretation used texts instead of numbers or patterns to represent the data. Nominal and ordinal data are the two types of qualitative data. Ordinal data interpretation is much easier than nominal data interpretation.
2. **Quantitative method** -This method is used to analyse quantitative data or numerical data. Quantitative data interpretation uses numbers instead of texts to represent the data. The types of quantitative data interpretation are discrete and continuous data. The quantitative method of data interpretation requires statistical methods and techniques like mean, median, standard deviation, etc. to interpret the data.

15.5 DIAGRAMMATIC PRESENTATIONS :

A very convenient and appealing method of data representation is by using various forms of diagrams. They in a very meaningful way highlight the salient features of the data which makes them easy to understand. Diagrams are both graphical as well as geometric in nature. The processed data is portrayed through different diagrams for visual presentations. It is important to make use of diagrams based on their relative merit of visual presentation. The diagrams mostly refer to time or space or both the characteristics related to one location. Some of the diagrams used for the presentation of primary data are discussed below:

Bar Charts/Bar Graphs : The use of a column or bar has become common in representing the comparative performance of various units and the growth of an individual unit. The length of the bar is kept proportional to the size of production or the volume of change. Thus, a bar diagram is used to represent many elements at one point in time and one element across time. The compound bar diagrams are used to represent the subclasses of an element. The block of a bar is proportionately subdivided to represent the sub-classes in a compound bar diagram.

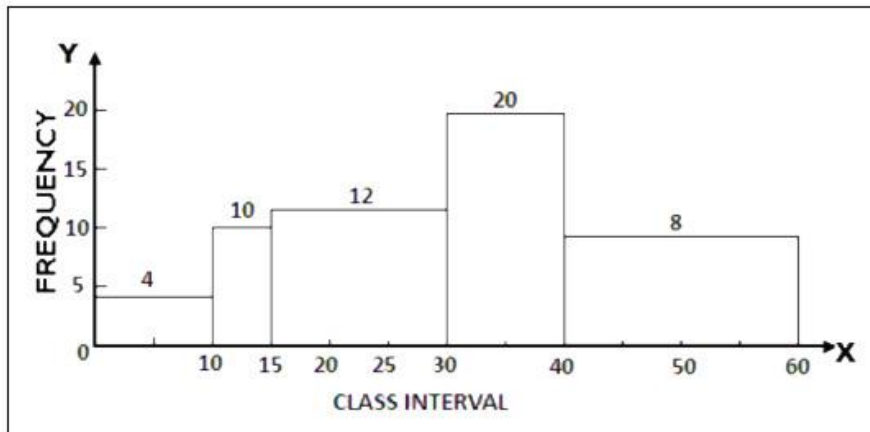
These are one of the most widely used charts for showing the growth of a company over a period. There are multiple options available like stacked bar graphs and the option of displaying a change in numerous entities. A bar graph is a way of summarizing a set of categorical data. It displays the data using several rectangles, of the same width, each of which represents a particular category. Bar graphs can be displayed horizontally or vertically and they are usually drawn with a gap between the bars (rectangles).

Pie Charts : The pie diagram is also known as a divided circle. It is used to represent the proportion of the sub-unit of a whole. The different segments of a circle represent the percentage contribution of various components of data. For drawing a pie diagram, we construct a circle of any diameter. The circle is then divided into the desired number of segments. i.e., angle 360 represents 100 per cent. A pie diagram is generally used to represent the general land use of the village, the composition of shops in a functional profile of urban areas, the social composition of the surveyed village, and the composition of the total population.

These work best for representing the share of different components from a total of 100%. For, e.g. Contribution of different sectors to GDP, the population of different states in a country, etc. A pie chart is used to display a set of categorical data. It is a circle, which is divided into segments. Each segment represents a particular category. The area of each segment is proportional to the number of cases in that category.

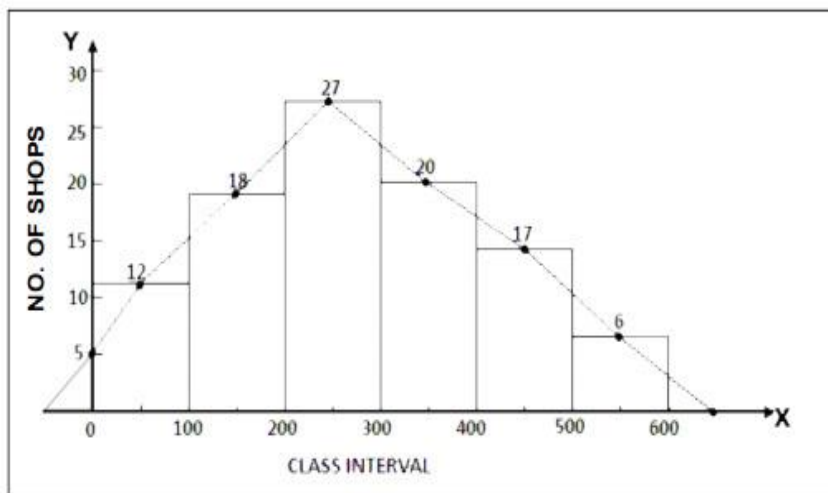
Combo Chart : As the name suggests it is a combination of more than one chart type. The one shown in the figure below is a combination of a line and a bar graph. These save space and are at times more effective than using two different charts. There can even be 3 or more charts depending on the requirement.

Histogram : A histogram is a way of summarizing data that are measured on an interval scale (either discrete or continuous). It is often used in exploratory data analysis to illustrate the features of the distribution of the data in a convenient form. In a histogram, we plot the class intervals on the X-axis and their respective frequencies on the Y-axis. Further, we create a rectangle on each class interval with its height proportional to the frequency density of the class.



Frequency Polygon or Histogram :

A frequency polygon or a Histogramy is another way of representing a frequency distribution on a graph. You draw a frequency polygon by joining the midpoints of the upper widths of the adjacent rectangles of the histogram with straight lines.



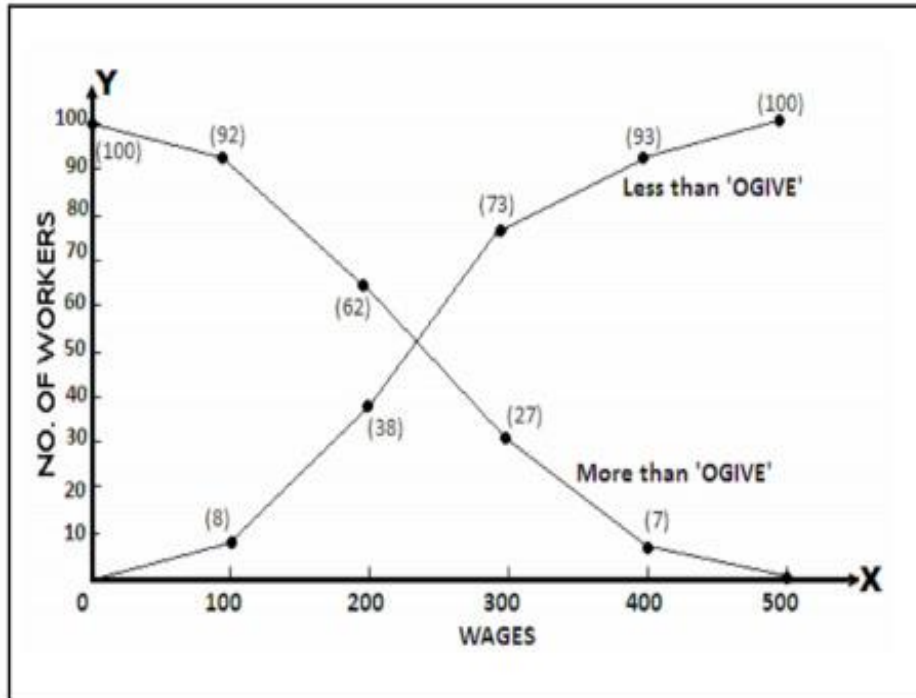
FREQUENCY POLYGON

Frequency Curve :

When you join the verticals of a polygon using a smooth curve, then the resulting figure is a Frequency Curve. As the number of observations increases, we need to accommodate more classes. Therefore, the width of each class reduces. In such a scenario, the variable tends to become continuous and the frequency polygon starts taking the shape of a frequency curve.

Cumulative Frequency Curve or Ogive :

A cumulative frequency curve or Ogive is the graphical representation of a cumulative frequency distribution. Since a cumulative frequency is either of a 'less than' or a 'more than' type, Ogives are of two types too – 'less than ogive' and 'more than ogive'.



Merits of a Graph :

- i. The graph presents data in a manner that is easier to understand.
- ii. It allows us to present statistical data attractively as compared to tables. Users can understand the main features, trends, and fluctuations of the data at a glance
- iii. A graph saves time.
- iv. It allows the viewer to compare data relating to two different periods or regions.
- v. The viewer does not require prior knowledge of mathematics or statistics to understand a graph.
- vi. We can use a graph to locate the mode, median, and mean values of the data.
- vii. It is useful in forecasting, interpolation, and extrapolation of data.

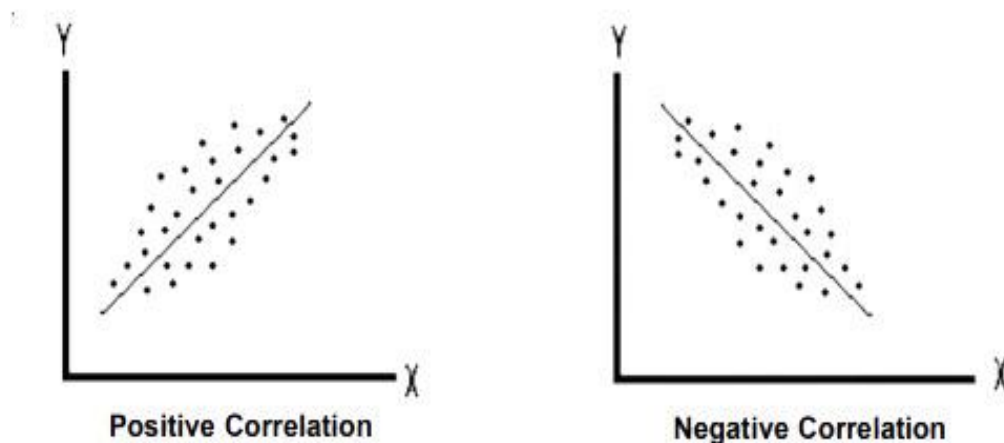
Limitations of a Graph :

- i. A graph lacks complete accuracy of facts.
- ii. It depicts only a few selected characteristics of the data.
- iii. We cannot use a graph in support of a statement.
- iv. A graph is not a substitute for tables.
- v. Usually, laymen find it difficult to understand and interpret a graph.

Typically, a graph shows the unreasonable tendency of the data and the actual values are not clear.

Scatter Diagram :

A scatter diagram or a dot chart enables us to find the nature of the relationship between the variables. If the plotted points are scattered a lot, then the relationship between the two variables is lesser.



Line Chart :

- ❖ These are best for showing the population change, i.e., for showing the trends. These also work well for explaining the growth of multiple areas at the same time.
- ❖ A line graph is particularly useful when we want to show the trend of a variable over time. Time is displayed on the horizontal axis (x-axis) and the variable is displayed on the vertical axis (y-axis). You can use a line graph to summarize how two pieces of information are related and how they vary from each other.

Advantages :

- i. You can compare multiple continuous data-sets easily
- ii. You can infer the interim data from the graph line

Disadvantages :

- i. It is only used with continuous data.

15.6 SUMMARY :

Processing and analysing data involve several closely related operations which are performed to summarize the collected data and organize these in a manner that answer the research questions (objectives). Manual Data Analysis is suitable when there are limited variables and the number of respondents is also very small. This is applicable when simple cross-tabulations are done and it is needed to calculate frequency distribution. The easiest way to do this is to code it directly onto large graph paper in columns. Each column can be given a number or a distinctive heading to identify and code information corresponding to the question. This analysis begins with manually counting various codes in a column and then decoding them. For statistical testing manual calculation is done depending on the researcher's expertise and how the results need to be communicated. Computerized Data Analysis needs the user to be familiar with the appropriate programs to be used along with an understanding of systems, statistical data, and software available. The most common software is SPSS for windows. However, data input can be a long and laborious process, and if data is entered incorrectly, it will influence the results.

15.7 GLOSSARY :

- ❖ **Body or field:** The body of a table is the content of a table in its entirety. Each item in a body is known as a cell.
- ❖ **Caption:** A caption is the title of a column in the data table. It is a counterpart of a stub and indicates the information contained in a column.
- ❖ **Characteristic:** It is the smallest unit of analysis i.e. a single item or event in a text, similar to an individual response to a variable or indicator in quantitative research.

- ❖ **Coding sorts:** Compilation of similarly coded blocks of text from different sources into a single file or report.
- ❖ **Coding:** The process of attaching labels to lines of text so that the researcher can group and compare similar or related pieces of information.
- ❖ **Computerized Data Analysis:** Computerized data analysis needs the user to be familiar with the appropriate programs to be used along with an understanding of systems, statistical data, and software available. The most common software is SPSS for windows. However, data input can be a long and laborious process, and if data is entered incorrectly, it will influence the results
- ❖ **Conjoint analysis:** Like the above method, conjoint analysis is a similar quantitative data analysis method that analyses the parameters behind a purchasing decision. This method possesses the ability to collect and analyse advanced metrics which provide an in-depth insight into purchasing decisions as well as the parameters that rank the most important.
- ❖ **Cross-tabulation:** This method used a basic table to draw inferences between different data sets available for a study. The qualities of data used are that they are either related to each other or are mutually exclusive.
- ❖ **Footnotes:** Footnotes are rarely used. In effect, they supplement the title of a table if required.
- ❖ **Gap analysis:** When it is important to understand the differential between actual and perceived values of a product or service gap analysis method is applicable. E.g., a guest may order a flashy-looking cocktail perceiving a light taste but may end up getting a drink that has a stronger taste.
- ❖ **Headnotes:** A headnote further aids in the purpose of a title and displays more information about the table. Generally, headnotes present the units of data in brackets at the end of a table title.
- ❖ **Indexing:** The process that generates a word list comprising all the substantive words and their location within the texts entered the program.
- ❖ **Manual Data Analysis:** This analysis is suitable when there are limited variables and the number of respondents is also very small. This is applicable when simple cross-tabulations are done and also it is needed to calculate frequency distribution. The easiest way to do this is to code it directly onto large graph paper in columns. Each column can be given a number or a distinctive heading to identify and code information corresponding to the question. This analysis begins with manually counting various codes in a column and then decoding them. For statistical testing manual calculation is done depending on the researcher 's expertise and how the results need to be communicated.
- ❖ **MaxDiff analysis:** This is a method that is used in studying the purchase preferences of customers and to understand why a particular factor is given more importance than other. E.g. a sun and sand tourists would prefer cleaner and less crowded beaches compared to beaches that have better availability of snacks and drinks options.
- ❖ **Source:** When using data obtained from a secondary source, this source has to be mentioned below the footnote.
- ❖ **Stubs:** These are titles of the rows in a table. Thus, a stub display information about the data contained in a particular row.
- ❖ **SWOT analysis:** Strengths, Weaknesses, Opportunities, and Threats for a subject, individual, or organization may be conducted to present a more holistic picture of competition. This is generally used when effective business strategies are to be formed.

- ❖ **Table Number:** Each table should have a specific table number for ease of access and location. This number can be readily mentioned anywhere which serves as a reference and leads us directly to the data mentioned in that particular table.
- ❖ **Text analysis:** It is an advanced statistical method where unstructured raw data is collected and then it has to be converted into a structured form for clearer understanding. Open-ended questionnaires provide data that needs conversion to statistical units for correct analysis thus Text Analysis method is appropriate as it uses intelligent tools.
- ❖ **Themes:** Clear-cut ideas that emerge from a grouping of lower-level data points.
- ❖ **Theory:** A set of interrelated concepts, definitions, and propositions that presents a systematic view of events or situations by specifying relations among variables.
- ❖ **Title:** A table must contain a title that tells the readers about the data it contains, the time of the study, the place of study, and the nature of the classification of data.
- ❖ **Trend analysis:** As the name suggests it is an interpretation of data that has been collected over a longer period thus making it easier to understand the changes that have come through. In this analysis usually one of the variables being studied remains constant.
- ❖ **TURF analysis:** Total Unduplicated Reach and Frequency Analysis are used when the researcher must find out the market reach of a product or service or a mix of both. It is helpful to develop a marketing plan when a product or service is exclusive yet has limited buyers.

15.8 SELF-ASSESSMENT QUESTIONS :

1. What are the general rules for the graphic presentation of data and information?
2. What is the importance of data interpretation?
3. What do you mean by editing data?
4. What do you mean by analysis of Data?
5. What do you mean by graphical representation of data?
6. What are the different types of Graphs? Explain each type with an example.

15.9 References and Suggested Readings :

1. Dawson, Catherine, (2002). *Practical Research Methods*, New Delhi, UBS Publishers' Distributors.
2. Kumar, Ranjit, (2005). *Research Methodology-A Step-by-Step Guide for Beginners*, (2nd.ed), Singapore, Pearson Education
3. John W. Creswell (2018). *Research Design: Qualitative, Quantitative, and Mixed Methods Approaches*, 5th Edition.
4. Norman K. Denzin&Yvonna S. Lincoln (2011).*The SAGE Handbook of Qualitative Research* (Sage Handbooks) 4th Edition.
5. Thanulingom N. (2010). *Research Methodology*, Himalaya Publishing .

Dr. Tharakeswara Rao

LESSON - 16

WRITING A RESEARCH PROPOSAL

OBJECTIVES :

After reading this unit the learner will be able to understand:

1. The purpose of writing research proposals
2. The components of research proposals
3. Tips for writing effective proposals
4. Types and steps involved in writing a report
5. The layout of the research report
6. Mechanics of writing a research report
7. Challenges of a good writing

STRUCTURE :

- 16.1 Introduction
- 16.2 What is a Research Proposal?
- 16.3 Components of a Research Proposal
- 16.4 Writing Research Report
- 16.5 Learning Outcome
- 16.6 Components of a Research Report
- 16.7 Research Report Format
- 16.8 Types and steps involved in writing a report
- 16.9 Theoretical Framework and Review of Literature
- 16.10 Editing of Reports Using External Sources
- 16.11 Components for writing effective research proposals
- 16.12 Writing style
- 16.13 Summary
- 16.14 Glossary.
- 16.15 Self-Assessment Questions
- 16.16 Reference and Suggested Reading

16.1 INTRODUCTION :

A research proposal is a document where a researcher provides all the details of a proposed research project. Research proposals are written for several different purposes such: as a part of a grant application for a research project; for Master/Ph.D. degree-related research; and, in response to calls for research proposals sent out by research institutions. While each institution/organization may have its specific formats or requirements, generally research proposals describe the process and requirements to implement a specific research project, including budgets and timelines.

16.2 What is a Research Proposal? :

A research proposal describes a research problem that is going to be investigated, the scope and rationale of this investigation, the methods used during the investigation, the implementation process, and the resources required to conduct the investigation. A research proposal outlines the entire purpose and implementation of a research project. Researchers

write research proposals for all types of research projects – those carried out based on research into written sources; others may be on research conducted in the field, and still others on experiments carried out in laboratories.

A research proposal, therefore, serves the following functions:

- i. To define and describe your research project to an external audience.
- ii. To highlight the knowledge, gap your project addresses.
- iii. To develop a plan of action for implementing your research project.
- iv. To demonstrate that a project has been conceptualized and planned in detail
- v. To define the resources (financial, human, material, and technical) that would be required to accomplish the research project.
- vi. To establish the researcher's qualifications, expertise, and credibility in the concerned area of research.
- vii. Serves as the basis for requesting grant funding and seeking approval for academic research such as Master's and Ph.D. theses

16.3 COMPONENTS OF A RESEARCH PROPOSAL :

Although the format of a research proposal may vary by the institution/organization to which you are submitting it, this section describes the key components of a research proposal. However, ensure that you review and follow the guidelines prescribed by your institution/organization carefully, as not doing so can imply a rejection of your proposal. Also, note that different institutions may label components differently – for example, some people refer to the Introduction section as the Background section. Be aware of this as you review your institution's guidelines, or read additional resources in books or on the internet. Some institutions also have fixed page number requirements – make sure you follow these.

1. **Title Page :** The title page must have the following information:
 - a) **Personal details:** Your name, your academic title or designation, date of submission, the name of your supervisor (if any), the name of your university/institution (if any), and the name of the organization/entity to whom you are submitting the proposal.
 - b) **Title of the study:** The title should be concise, relevant, and descriptive of the major focus of the study. By reading the title one should get a clear idea about what and who is studied. An effective title not only catches the reader's interest but also predisposes him/her favorably toward the proposal. Since the title reflects the nature of your entire research project, sometimes it may be useful to finalize this only once you have completed writing the entire proposal.
2. **Abstract :** An abstract is an executive summary of your research proposal and is usually written in 150-300 words. It should include a brief mention of the research question and objectives, the hypothesis if any, and the methodology to be used. Descriptions of the methodology may include the design, the sample, and the research tools. Describe your project as clearly and concisely as you can – remember you have word restrictions!
3. **Table of Contents :** The Table of Contents is a listing of all the different sections of the proposal along with their page numbers. This may not be required for short proposals with two or three pages.
4. **Introduction :** This section introduces the reader to the main area of your project. Therefore, it is important to use this section to catch your reader's attention. Provide a brief overview of the theme/area of your research study and then describe what specific goal or question your study will address within this broad area. The research question is studied with the help of specific study objectives and hypotheses (in some

cases) and they have to be spelled out. This section should also provide the justifications or rationale for your study. Also, indicate why your study is significant, or what contributions it will make to the body of literature. This part will be of special interest to reviewers and funders alike, because they may want to know in what way your research is unique and also worth granting approval and/or funds.

5. **Literature Review :** The literature review is an essential part of a proposal because it defines what information /research is available on a specific topic, and how your proposed study will address a gap in the available research. The purpose of this section is to review previous research conducted on your research topic. Remember to make note of research findings as well as research gaps. You must also explain your research question(s) and objectives within the context of these gaps as this will also help you explain how your study adds to the existing body of knowledge. You may find it useful to review Module 4 before writing this section of your proposal, as it describes the process of conducting a literature review in detail.
6. **Methods :** The methodology section is an important part of your proposal because it tells the readers and reviewers how you plan to study your research problem. It describes your work plan and the activities necessary for the completion of your project. While writing this section, maintain a focus on the research question(s), objectives, and hypotheses, and use the section to describe how you would go about collecting the data that is required to answer them. It is also important to make sure that the research questions, objectives, and methods all link up with each other, and flow smoothly. The Methods section will undoubtedly be the one that will receive the closest attention from reviewers, so it is important to demonstrate here your skills in planning and organizing research-related activities.

As you describe your chosen methods, it is important to state the reasons why you feel that your approach is the most appropriate to address your research question. This will demonstrate to the reviewers and readers that you have given considerable thought to the choice of your methods. If you have conducted any preliminary research, then it would be a good idea to mention this in this section and also describe how that has influenced your choice of methods in your current project.

Personnel or Research Team :

This section provides information on all the key staff associated with the project. One element of this section is to describe the personnel arrangements for this project. This includes a list of different positions and a brief description of the duties of the positions. You may also want to indicate what percentage of time each individual may be spending on the project. If the personnel have any specific training or qualifications that make them especially valuable to this project, be sure to mention that in this section. Finally, also include CVs/resumes of the key staff.

Budget :

The budget section includes information on the expenses required to implement the project. If you are submitting the proposal for research funding, make sure to make this section as detailed as possible. The institution you are applying for funding may have its budget template, so be sure to use the correct template. Even if you are not submitting the funding proposal, it is helpful to include the budget. Developing a budget will help you and your readers assess the expenses required for this project.

Timeline :

As you work through the various stages of your proposal, it is important to demonstrate to readers and viewers the different stages and the time it will take to complete each stage. Build in time for monitoring meetings with your supervisor or team meetings that will help you keep track of the progress made on the project. Therefore, include a timeline in your proposal. The timeline is essential as a list of steps and the approximate time it will take to complete each step. This will give readers an idea of the total time required to complete your project, and the time taken for each step. However, remember that this is a tentative plan, and it will change as you begin the project and encounter unexpected delays. This is perfectly okay. Many funding institutions usually allow for a renegotiation of the timeline.

Bibliography :

The bibliography or reference list is a list of all the texts you have referred to in your literature review and other parts of your proposal. Remember to follow the style requirements (for example, APA, MLA, etc.) of your institution, if any. If there are no requirements, you can use your choice of style. However, once you pick a style, make sure you follow it consistently throughout the document. Make sure the details provided in each bibliographic entry are accurate.

16.4 WRITING RESEARCH REPORT :

A report is a written explanation of something that one has observed, heard, done, or investigated. It is a systematic and well-organized presentation of facts and findings of an event that has already taken place somewhere or has been found out after an in-depth study has been conducted.

16.5 LEARNING OUTCOME :

After reading this unit the learner will be able to understand:

- I. Types and steps involved in writing a report
- II. The layout of the research report
- III. Mechanics of writing a research report
- IV. Challenges of a good writing

16.6 COMPONENTS OF A RESEARCH REPORT :

Scientific research articles provide a method for scientists to communicate with other scientists about the results of their research. A standard format is used for these articles, in which the author presents the research in an orderly, logical manner. This doesn't necessarily reflect the order in which you did or thought about the work. The following is a general outline for a research report.

<i>Beginning Material</i>	i.e., title page, abstract, keyword list, table of contents, list of figures and tables, acknowledgments
<i>Chapter 1</i>	Introduction – statement of the problem, hypotheses, why it is important, objectives of the work, the scope of the work

Chapter 2	Background and Literature Review – discuss related work and indicate how it relates to the report
Chapter 3	Procedure – describe the procedure used in the project, the data used, and how it was obtained
Chapter 4	Results – indicate what happened and interpret what it means
Chapter 5	Conclusions and Recommendations – summarize conclusions and what they mean (i.e., answer the question, “So what?”). What changes and further work do you recommend?

16.7 RESEARCH REPORT FORMAT :

Traditional written reports tend to be produced in the following format.

❖ Title Page :

- a) Title of the Research Project,
- b) Name of the researcher,
- c) Purpose of the research project, e.g. —A research project submitted in partial fulfillment of the requirements of Uttarakhand Open University, for the degree of in.....
- d) Date of Publication

❖ Table of Contents :

- ❖ In this section contents of the report are listed as they appear in the report, either in chapters or in subheadings e.g.
- ❖ **Acknowledgements** : In this section, the researcher may acknowledge Institution Head, Faculty Guides, research participants, friends, etc.
- ❖ **Introduction** This section introduces the research setting out aims and objectives. It includes a rationale for the research.
- ❖ Chapter I Theoretical Framework and Review of Related Literature
- ❖ Chapter II Research Design
- ❖ Chapter III Data Analysis and Interpretation
- ❖ Chapter IV Summary and Conclusion
- ❖ Suggestions/ Recommendations for Further Research
- ❖ References/ Bibliography
- ❖ Appendices
- ❖ Appendix I Questionnaire for Employees
- ❖ Appendix II Questionnaire for Managers
- ❖ List of Tables: This section includes the title and page number of all tables.
- ❖ Details about respondents (Demographic, Economical, Geographical, etc. relevant to justify the data collected for the study).
- ❖ **List of Figures** : This section contains the title and page number of all graphs, pie charts, etc.

16.9 TYPES AND STEPS INVOLVED IN WRITING A REPORT :

Reports are used as a form of written assessment to find out what has been learned from reading, researching, or experiencing an important skill that is widely used in the workplace. Good report writing is an important quality for any researcher as this also presents the findings to the readers outside your subject area with the experts in the field. It focuses on the findings, conclusions, discoveries made, efforts made and inferences drawn from the research study conducted. The report should be written in a simple but scholastic style. The language should be formal and not like the one used in media publications.

A good research report has the following functions:

1. To provide the information regarding the findings of research work i.e. methods, data analysis, conclusion, and so on in a systematic, scientific, and accepted way.
2. To elicit crucial facts for solution-derived and decision-making.
3. To prove the worth and legitimacy of the assigned research job.
4. To provide the judgment tools for the judgment of quality and talent of the researcher within and outside the academia.
5. To communicate the research findings professionally.
6. To pertain to the credibility of the research.
7. To develop an appreciation of standards, consolidate arguments, and identify the knowledge gaps.

- ❖ **Technical Report:** In the technical report the main emphasis is on the methods employed, assumptions made in the course of the study, and the detailed presentation of the findings including their limitations and supporting data. E.g, the project reports when a hotel is being conceptualized.
- ❖ **Formal or Informal Reports:** A formal report has a carefully drafted structure, and clear objectives, is organized, and has sufficient details to let the reader understand the concepts. These are written using nonperson elements whereas an informal report can be direct, and short with casual language e.g. an inter-office communication via a notice or memorandum.
- ❖ **Popular Report:** The highlights of this report are simplicity and attractiveness. The simplification is done by clear writing, minimization of technical, particularly mathematical, details, and liberal use of charts and diagrams. The attractive layout, large print, many subheadings, and even an occasional figurine is other characteristic feature of the popular report.
- ❖ **Informational or Analytical Reports:** Informational reports (annual reports, monthly financial reports, and personnel reports) carry objective information from one area of an organization to another. Analytical reports (scientific research, feasibility reports, and real estate appraisals) present attempts to solve problems.
- ❖ **Proposal Report:** The proposal is a variation of problem-solving reports. A proposal is a document prepared to describe how one organization can meet the needs of another. Most governmental agencies advertise their needs by issuing —requests for proposal or RFPs. The RFP specifies a need and potential suppliers prepare proposal reports telling how they can meet that need.
- ❖ **Vertical or Lateral Reports:** Any reports that move upward or downward the hierarchies are referred to as vertical reports; such reports contribute to management control. Lateral reports, on the other hand, assist in coordination in the organization. A report traveling between units of the same organization level (production and finance departments) is lateral.
- ❖ **Internal or External Reports:** Internal reports are circulated within the organization e.g. the reports circulated in the morning meeting of managers at a hotel. External

reports, such as annual reports of companies, are prepared for distribution outside the organization e.g. a report showing the growth, and popularity of a hotel brand in comparison to others.

- ❖ **Periodic Reports:** The reports are generated at scheduled intervals for guiding the management to exercise better control. The format is unusually pre-printed and system generated so that they are uniform.
- ❖ **Functional Reports:** This classification includes accounting reports, marketing reports, financial reports, and a variety of other reports that take their designation from the ultimate use of the report. Almost all reports could be included in most of these categories. And a single report could be included in several classifications.
 - **Pre-printed Form: These are like fill-in-the-blank** reports, relatively short (five or fewer pages), and deal with routine information, mostly numerical information.
 - **Letter:** Shorter reports that are aimed to develop an understanding of the people outside the organization. These reports include all the normal parts of a letter, but they may also have headings, footnotes, tables, and figures. Personal pronouns are used in this type of report.
 - **Memo:** Common for short (fewer than ten pages) informal reports distributed within an organization. The memo format of Date, To, From, and Subject is used. Like longer reports, they often have internal headings and sometimes have visual aids.
 - **Manuscript:** These are reports that range from a few pages to several hundred pages and require a formal approach. As their length increases, reports in manuscript format require more elements before and after the text of the report.

16.9 THEORETICAL FRAMEWORK AND REVIEW OF LITERATURE :

This section covers the sources of information and background research done by referring to the literature available. A good researcher keeps the details of all the sources and literature referred to and should ensure giving suitable references wherever necessary to avoid being accused of plagiarism i.e., breach of copyrights, unauthorized copying, or illegal use of information. The Harvard System is generally used to quote references to a particular book, a chapter in the book, or an article in the journal.

- ❖ **Research Design :** This section assists the other researchers to follow and replicate the study being presented. All the practical methods used to select the population, and area, collect the data, select the respondents, select sample sizes, and methods of analysis used for these are presented in the research design.
- ❖ **Data Analysis and Interpretation :** This is one of the most important sections of the report as it contains all the relevant information about the presentation of data in the study conducted e.g. tables, charts, graphs, statistical data charts, figures, etc when it is quantitative research and narrative prose or write up describing the findings if it's qualitative research.
- ❖ **Summary and Conclusion :** This is the chapter where the findings are summed up to conclude them, perhaps with other research or literature.
- ❖ **Recommendations :** If you have conducted a piece of research for a hotel or any other client organization, this section could be the most important part of the report. A list of clear recommendations which have been developed from the research is included- sometimes this section is included at the beginning of the report.
- ❖ **Suggestion for Further Research :** This section helps the researcher to make the readers understand the broader scope of research that may be taken up next. It also

brings forward unanswered questions, new questions about the study, newer scopes, and findings that may seem inconclusive.

E.g. while closing a report on slow down in a hotel business the report may highlight the failure in terms of innovation by the hotel and non - the inclusion of facilities offered by rival brands.

List of References /Bibliography :

- ❖ List of references contains details only of those works cited in the text.
- ❖ A bibliography includes sources not cited in the text but which are relevant to the subject. (Dissertations or thesis)
- ❖ Small research projects will need only a reference section detailing the literature referred to in the report.

The popular referencing system Harvard System lists books and periodicals in the following manner:

For Books :

1. Author's surname (alphabetically), followed by their initials,
 2. Date of publication
 3. Title of book in italics
 4. Place of publication, Publisher. E.g. Philip, T.E.; 1986, *Modern Cookery for Teaching and Trade*, Mumbai, Orient Longman.
- ❖ **For Journal Article** : The title of the article appears in inverted commas and the name of the journal comes in italics, followed by the volume number and pages of the article. e.g. Philip, T.E.; —Influence of British Raj on Indian Cuisine; *Journal of Hospitality Education*; 5:5-11
 - ❖ **Appendices** : The pilot and final questionnaires constructed for a study and the interview schedule for research etc should be included as an appendix. Appendices do not count toward your total number of pages/words. It is a useful way of including relevant material so that the examiner can gain a deeper understanding of your work by reading it.

16.10 EDITING OF REPORTS USING EXTERNAL SOURCES :

- ❖ **Substantive Editing** : This is the most comprehensive and articulated form of editing that includes evaluating the complete document and looks into the acceptability of problems of structure, organization, coherence, and logical consistency. The language used is edited, and sentences are added or removed. Paragraphs may be rewritten, condensed, or expanded. Blocks of text may be moved from one section to another.
- ❖ **Copy editing** : The editor corrects problems of grammar, style, repetition, word usage, and jargon. Copy editing includes one revision at no additional charge.
- ❖ **Proofreading** : Proofreading is the lightest form of editing. Minor errors are corrected. Minor errors include:
 - Errors of grammar and style (e.g., verb tense, units of measurements, use of numerals and words such as “10” or “ten”)
 - Errors of capitalization, and punctuation (e.g., the use of commas, semicolons, colons, periods, dashes, and apostrophes)
 - Errors in spelling and word usage (e.g., no/ know, then/than).
- ❖ **Formatting** : The editor will amend the document text to ensure that it complies with the required format, such as the format required by a specific journal.
- ❖ **References/Literature Cited** : Literature citations are checked to ensure that each citation that appears in the text is also included in the list of citations. Citations are

also checked to ensure that each citation that appears in the list of citations also appears in the text. The format of the citations is corrected so that it conforms with requirements, such as the style preferred by a specific journal.

- ❖ **Other documents** : Other documents also may be edited. For example, the editor may rework tables, figures, and figure legends to represent the data more clearly.
- ❖ **Review** : The editor may provide a one-to-two-page diagnosis of the manuscript that highlights the areas where changes might be most beneficial. Criticisms that are likely to arise during peer review, such as repetitive, ambiguous, or incomplete information, will be noted. A review includes proofreading at no charge. Turnaround time must be negotiated.

16.11 COMPONENTS FOR WRITING EFFECTIVE RESEARCH PROPOSALS :

- i. Provide an adequate context to your study by citing seminal works/landmark studies.
- ii. Present the work of other researchers accurately.
- iii. Stay focused on your research question.
- iv. Present the boundaries and limitations of your research project.
- v. Make your arguments fact-based and persuasive.
- vi. Define clearly and concisely the contributions made by your project.
- vii. Make sure all the sections of your proposal flow smoothly and strengthen the case for your project.
- viii. If you are required to include a budget, make sure all the given amounts total and tally correctly.
- ix. Once you have completed a draft version, go over it in detail and edit it. Enlist the support of a friend to read it and provide feedback, as this can help in improving the quality and clarity of your proposal.

16.12 WRITING STYLE :

- i. Do not forget to add a cover page and a table of contents
- ii. Do not forget to include page numbers – research proposals tend to be long documents and your reviewers may need to print/photocopy multiple copies. Page numbers will ensure everything is in order!
- iii. Check for spelling and grammar errors.
- iv. If you are required to submit a bibliography, make sure you use the correct citation style as per the requirements.
- v. Do not plagiarize -- make sure you cite correctly in the text.
- vi. Make sure your proposal is within the prescribed word limit. For example: if the specified limit is 5 pages or 2500 words, sending a proposal that is 8 pages long or that is about 3500 words could result in its rejection.
- vii. Use a formal font such as Times New Roman or Ariel.
- viii. Font sizes should be between 10 to 12.

16.13 SUMMARY :

A research proposal is a document describing all the key elements of a proposed research project. A research proposal helps in indicating what knowledge gaps your project addresses. It also defines the plan of action to implement it, indicates what resources are required, serves as the basis of a grant application, and establishes the expertise of the researcher and research team. Reports are used as a form of written assessment to find out what has been learned from reading, researching, or experiencing an important skill that is widely used in the workplace. Good report writing is an important quality for any researcher

as this also presents the findings to the readers outside your subject area with the experts in the field. It focuses on the findings, conclusions, discoveries made, efforts made and inferences drawn from the research study conducted. The report should be written in a simple but scholastic style. The language should be formal and not like the one used in media publications.

16.15 GLOSSARY :

- ❖ **Appendices:** The pilot and final questionnaires constructed for a study and the interview schedule for research etc should be included as an appendix. Appendices do not count toward your total number of pages/words. It is a useful way of including relevant material so that the examiner can gain a deeper understanding of your work by reading it.
- ❖ **Data Analysis and Interpretation:** This is one of the most important sections of the report as it contains all the relevant information pertaining to the presentation of data in the study conducted e.g. tables, charts, graphs, statistical data charts, figures, etc when it is quantitative research and narrative prose or write up describing the findings if it's qualitative research.
- ❖ **For Journal Article:** The title of the article appears in inverted commas and the name of the journal comes in italics, followed by the volume number and pages of the article. e.g., Philip, T.E.; Influence of British Raj on Indian Cuisine; *Journal of Hospitality Education*; 5:5-11
- ❖ **Informational or Analytical Reports:** Informational reports (annual reports, monthly financial reports, and personnel reports) carry objective information from one area of an organization to another. Analytical reports (scientific research, feasibility reports, and real estate appraisals) present attempts to solve problems.
- ❖ **Internal or External Reports:** Internal reports are circulated within the organization e.g., the reports circulated in the morning meeting of managers at a hotel. External reports, such as annual reports of companies, are prepared for distribution outside the organization e.g., a report showing the growth, and popularity of a hotel brand in comparison to others.
- ❖ **Letter:** Shorter reports that are aimed to develop an understanding of the people outside the organization. These reports include all the normal parts of a letter, but they may also have headings, footnotes, tables, and figures. Personal pronouns are used in this type of report.
- ❖ **Manuscript:** These are reports that range from a few pages to several hundred pages and require a formal approach. As their length increases, reports in manuscript format require more elements before and after the text of the report.
- ❖ **Memo:** Common for short (fewer than ten pages) informal reports distributed within an organization. The memo format of Date, To, From, and Subject is used. Like longer reports, they often have internal headings and sometimes have visual aids.
- ❖ **Periodic Reports:** The reports are generated at scheduled intervals for guiding the management to exercise better control. The format is unusually pre-printed and system generated so that they are uniform.
- ❖ **Popular Report:** The highlights of this report are simplicity and attractiveness. The simplification is done by clear writing, minimization of technical, particularly mathematical, details, and liberal use of charts and diagrams. The attractive layout along with large print, many subheadings, and even an occasional figurine is another characteristic feature of the popular report.
- ❖ **Pre-printed Form:** These are like **fill-in-the-blank** reports, relatively short (five or fewer pages), and deal with routine information, mostly numerical information.

- ❖ **Recommendations:** If you have conducted a piece of research for a hotel or any other client organization, this section could be the most important part of the report. A list of clear recommendations which have been developed from the research is included- sometimes this section is included at the beginning of the report.
- ❖ **Research Design:** This section assists the other researchers to follow and replicate the study being presented. All the practical methods used to select the population, and area, collect the data, select the respondents, select sample sizes, and methods of analysis used for these are presented in the research design.
- ❖ **Suggestion for Further Research:** This section helps the researcher to make the readers understand the broader scope of research that may be taken up next. It also brings forward unanswered questions, new questions about the study, newer scopes, and findings that may seem inconclusive.
- ❖ **Summary and Conclusion:** This is the chapter where the findings are summed up to conclude them, perhaps concerning other research or literature.
- ❖ **Technical Report:** In the technical report the main emphasis is on the methods employed, assumptions made during the study, and the detailed presentation of the findings including their limitations and supporting data. E.g., the project reports when a hotel is being conceptualized.
- ❖ **Theoretical Framework and Review of Literature:** This section covers the sources of information and background research done by referring to the literature available. A good researcher keeps the details of all the sources and literature referred to and should ensure giving suitable references wherever necessary to avoid being accused of plagiarism i.e., breach of copyrights, unauthorized copying, or illegal use of information. The Harvard System is generally used to quote references to a particular book, a chapter in the book, or an article in the journal.
- ❖ **Vertical or Lateral Reports:** Any reports that move upward or downward the hierarchies are referred to as vertical reports; such reports contribute to management control. Lateral reports, on the other hand, assist in coordination in the organization. A report traveling between units of the same organization level (production and finance departments) is lateral.

16.15 SELF-ASSESSMENT QUESTIONS :

- 1) What are the different types of the report?
- 2) Write a note on the steps in writing a report.
- 3) What challenges does one face while writing a good report?

16.16 REFERENCE AND SUGGESTED READING :

- 1) C. Rajendar Kumar, Research Methodology, APH Publishing
- 2) Galvan, J.L. (2009). Writing Literature Reviews: A Guide for Students of the Social and Behavioral Sciences, Glendale, CA: Pyczak Publishing
- 3) Slade, C., Campbell, W.G., & Ballou, Steven, V. (1993). Form and Style: Research Papers, Reports and Thesis, Orlando, FL: Houghton Mifflin.

Dr. TharakeswaraRao

Model Question Paper

M.A.DEGREE EXAMINATION

Second Year

JOURNALISM AND MASSCOMMUNICATION

Third Semester

Paper I- RESEARCH METHODOLOGY

Time: Three hours

Maximum: 70marks

Answer ALL questions

All questions carry equal marks. (5x14=70)

1. a. What is 'scientific method' what are its characteristics
Or
b. How do you identify and formulate Research problem in Communication Research.
2. a. Explain the importance of causation in preparing a hypothesis.
Or
b. Outline important Research designs in Mass Communication Research.
3. a. Examine the different types of sampling methods and their use in different research.
Or
b. When do you use case study method? Give two examples.
4. a. Explain the significance of standard deviation and correlation.
Or
b. What is importance of non-parametric tests in Mass Communication research?
5. a. Discuss the steps in report writing with suitable examples.
Or
b. How do you tabulate survey research data for presentation in report.