SELF LEARNING MATERIAL

M.A. Sociology - 1st Semester

SOCIOLOGY

COURCE - SOC : 104

METHODS AND TECHNIQUES IN SOCIAL RESEARCH

BLOCK : 1 to 5

Directorate of Open & Distance Learning
DIBRUGARH UNIVERSITY
DIBRUGARH - 786004

M.A. Sociology - 1st Semester

SOCIOLOGY

COURSE - SOC : 104

METHODS AND TECHNIQUES IN SOCIAL RESEARCH

Contributor :

Prof. C.K. Gogoi
Department of Sociology
Dibrugarh University

Editor :

Dr. J. Borbora
Department of Sociology
Dibrugarh University

Published on behalf of the Directorate of Open & Distance Learning, Dibrugarh University by the Director, DODL, D.U. and printed at Designer Graphics, Dibrugarh-01

ACKNOWLEDGEMENT

The Directorate of Open & Distance Learning, Dibrugarh University duly acknowledgement the financial assistance from the Distance Education Council, IGNOU, New Delhi for preparation of this Self Learning Material.

M.A. Sociology - 1st Semester

SOCIOLOGY

COURSE - SOC : 104

METHODS AND TECHNIQUES IN SOCIAL RESEARCH

CONTENTS

COURSE INTRODUCTION

The application of scientific methods to understand human society is as old as the emergence of the discipline.dddd "Sociology" and therefore, the men of sociology are supposed to be the experts in research methodology. But, methodology in social sciences is under rigorous refinement and so, development of new methods and techniques in social research is an integral part of development of the subject. Sociology thus, has claimed to be a scientific discipline. The present course is oriented to disseminate higher knowledge about research methodology in social sciences. Besides conventional methods, this course intends to introduce to the students about the recent development in social science research like - use and application of computer technology in handling social data more effectively.

BLOCK – 1

METHODOLOGICAL PERSPECTIVES IN SOCIAL RESEARCH

STRUCTURE

1.0 OBJECTIVES

The objectives of this Block is two fold viz. to gain knowledge about the scientific methods of research and secondly, to acquire ideas how these methods are applied in the study of social phenomena.

1.1 INTRODUCTION

In this Unit, we shall first introduce the meaning of science and discuss some characteristics of science. In defining characteristics of social sciences, our attention will be chiefly focused on two basis points i.e. 'Objectivity' and 'Inter-subjectivity' – Similarly, 'Concept' and 'Conceptualisation' are two another important terms that are also thoroughly discussed Any scientific research, whether physical or social has to confront a dilemma whether fact is either good or bad, useful or harmful for society. Here lies the question of 'Ethics' in research. The issue of ethical considerations in social research is also discussed at the end of the unit.

1.2 SCIENCE AND ITS OBJECTIVES

1.2.1 Meaning of Science

The word 'science' connotes various meanings such as –

1. It is an objective investigation of empirical phenomena.
2. It is an accumulation of systematic knowledge,
3. It means all knowledge are to be collected by scientific methods,
4. It is united by its methodology
5. It is a method of approach.
6. It is a method of systematic rules and procedures.

1.2.2 Definition

Young and Mack defined science as a body of knowledge and it is everchanging. It is not that facts are disproved, it is that they are added to or modified. The body of scientific knowledge is cumulative, for example – Einstein's theory of relativity does not suggest that Newton's scientific facts were wrong ; it adds to them.

We encounter with science a great variety of descriptions and definitions. These two central themes i.e. description and definition involve cluster of views. About more than one hundred and fifty years ago two pioneer sociologists William J. Goode and Paul K. Hatt defined science "as a method of approach to the entire empirical world". Again, some others like to refer to the results of those methods. In terms of definition then, science emphasizes the method of gathering knowledge, while still some others focus upon the knowledge which is gathered by this method.

Mannheim and Simon tried to provide a summary definition of science as below :

> "Science is an objective, accurate and systematic analysis of a determinate body of empirical data in order to discover recurring relationship."

Each and every term used in the above definition need *Operational clarity.* By *objective,* it is meant unbiased, unprejudiced, detached, impersonal, the characteristics of viewing things as they *really* are. In this context the suggestion given by L. Gross can be added. Gross said that a statement is objective if its subject matter is something other than an event in somebody's mind. *Accurate* means that the scientist strives to be definite, precise and exact. Great care and fidelity of truth are exercised. *Systematic analysis* implies the meaning of unbreaking and logical flow of expression and ideas. Such ideas are derived after testing empirical data. Sociologists like

other *natural scientists* make systematic statements on a body of empirical data and try to discover relationships among groups.

1.2.3 Characteristics of Science

Since science is a logical system that based on knowledge of direct and systematic observation, scientific knowledge is based on empirical evidence. In simple sense, knowledge which we derive must be able to verify without own senses.

Then, one can determine the characters of science if the available definitions of science are carefully scrutinized. The fundamental character of science as introduced by Goode and Hatt is its *method of approach.*

Identifying the characteristics of science, Robert A. Nisbet held the idea that science has its *own mode of expression.* Secondly, it is the *objectives of discovery* and *systematic explanation* are universal characteristics which are to be found in any science. With regard to *explanation* Smelser asserts that explanation begins with the search for *independent variables* (or causes, or determinants or factors or conditions) to which variations in the *dependent variables* are referred. So, character of science or scientific study is to search out the *dependent* and *independent situation.* As a result, *cause-effect relationships* of two social situations can be determined. This may be stated as one important character of science. In the scientific framework it is the *research questions* that it asked by the scientist. Science, whether physical, biological or social, may be regarded as a vast collection of questions.

For a scientist the questions call for a certain type of answer. Scientific answers or explanations must be capable of *empirical verification,* that is to say that result achieved by one scientist at one time and place must be opened for empirical verification by others at other places and time. And, so verification is the very essence of science. Science, in short, is concerned with *answers, conclusions and principles.*

Science is by nature universal and this comes from the fact that the same question directed to the same body of data can produce the same results or answer to all. This is assumed that scientists follow the same procedures directed by the character of the problem. This is what we mean *by objectivity* and *universality* of science. In science, an *answer or conclusion* is valid irrespective of one's personal, political, ideological or religious like or dislikes.

Check Your Progress 1

1. List at least five characteristics of science in the space given below -

_____

_____

_____

_____

2. Define the terms objectivity, empirical verification and universality (in the space given below)

_____

_____

_____

_____

_____

_____

Try to answer

1. What is science ?
2. What do you mean by systematic knowledge ?
3. Try to understand the meaning of empirical phenomena.

Terms to be remembered

Systematic ;	empirical verification ;
objective ;	body of knowledge ;
operational	clarity principle
verification	independent variable
	dependent variable

1.3 OBJECTIVITY AND INTER-SUBJECTIVITY

1.3.1 Meaning of Objectivity

Sociologists use research techniques to study human social action and interaction. Like other natural sciences they conduct research using scientific methods and strive to reach conclusion and present findings which are considered *objectives.* Thus, *objectivity* refers to a condition of unbiased, concrete, real and not influenced by emotion or prejudices.

Knowledge of any phenomena whether it is natural or social, must be objective in nature. One of the biggest problems the social scientists comfort is the *issue of objectivity*. Sociologists are quite frequently asked – "How can a creature of society like everyone else, with feelings, moral views and prejudices, be detached and unemotional enough in investigation so that the findings are truly objectives ? Sociologists address the problems of objectivity partly through the way their discipline is organised and partly through the way they conduct research.

The ideal of objectivity has to suffer from two challenging pauses. Firstly, Sociologists and other scientists have personal opinions about whatever they study and secondly, science strives for objectivity which is a state of personal mentality in conducting research. Max Weber, an influentical German sociologist claimed

that research must be carried out dispassionately in order to study the world as it is rather than as we think it should be. This way, Weber advocated sociological research value free rather than value relevant.

We may thus conclude that the research which are influenced by researcher's personal opinion (value) are subjective in nature and so, non-scientific in handling social reality. On the otherhand any research conducted without taking care of prevailing value system is termed as objective and hence scientific in procedure. How to get rid of subjectivity ? One way to delimit subjectivity is replication i.e. repeatation of research by others in order to assess its accuracy.

Check Your Progress

1. Why objectivity is a problem in social research ?

_____

_____

_____

_____

2. Define 'value-free' and 'value-relevant' research.

_____

_____

_____

Review Question

1. What is objectivity ?
2. What is replication ?
3. Show how a natural science is objective ?

1.3.2 Objectivity and Bias

Another issue in social science research is the question of bias in research process. Sociologists claim that their study is a science of human behaviour. The research process which sociologists undertake and the conclusions which they draw are not always free from criticism. The mere presence of a researcher may distort the situation under study, may create awkward to the general public, a phenomenon often termed as 'reactivity' (Popenoe 1971).

As in other academic disciplines, knowledge in sociology is derived from the work of many researchers. The bias attitudes of any particular sociologist are kept to a minimum by the check and balance of research by other scholars on similar topics. Sociological findings are always almost reported publicity with full statements of methods applied and the sources of data. Others may look at the same evidence and try the same method to see whether they come up with the same conclusions. Thus, biasness is minimized to achieve objectivity of findings as far as possible.

Now try to understand the terms

a) Science of human behariour
b) reactivity
c) bias

1.3.3. Subjectivity

What is subjectivity ?

"Subject" refers to persons. The feelings, attachment, emotion, personal likeness or dislikeness when rooted in the understanding of phenomena may be considered subjective understanding. For example, if we put our opinion on religion that certain religious practices and rituals are good or bad to the society,

we are putting our subjective notions. Some social processes have strong influence on shaping of thinking, forming attitudes, determining values of the individual in society.

1.3.4. The idea of inter-subjectivity

The idea of inter-subjectivity in social sciences is a much-concerned debate and notion among social scientists. Social groups deeply rooted in tradition to contribute to perform a synthesis which otherwise helped to formalize social sciences. Karl Marx, Max Weber and Georg Simmel admonished researchers to be *value free in social science* and stated that there is a continous influence of social process on scientific thinking. Such notions obscure the temperament of scientific sociology.

1.3.5. Subjectivity in Social Science Research

Science first revealed the operation of the natural world. When sociologists started utilizing natural science strategy to study human world, sociologists recongnised several limitations. Some of such limitation are :-

1. Human Society is complex. It is so complex that sociologists cannot predict individuals behaviour with exact precision and accuracy.
2. Because humans respond to their surroundings, the mere presence of a researcher may effect the behaviour being studied.
3. Social patterns change constantly, what is true is one time or place may not hold in another.
4. Because sociologists are part of the social world they study, objectivity in social research is especially difficult.

Subjectivity is not totally rejected in social research. Some researchers argue that subjectivity has got importance. According to

them the logic of science does not eliminate subjectivity from research, since creative thinking is vital in sociological investigation. Nisbet (1970) points out that insight comes not from science itself, but from lively thinking of creative human beings. Secondly, it is argued that science cannot embrace the vast and complex range of human motivation and feelings, including greed, love, pride and despair. Science helps us to gather facts about how people act, but it can never fully explain, how people make sense of what they do (Berger and Kellner 1981 quoted by J.J. Macionis– Sociology 1987) Thirdly, there was an argument with regard to the position of sociology as a discipline. Data cannot speak for themselves. Sociologists and other scientists must always face the ultimate task of interpretation; creating meaning from a collection of observation. For this reason sociology is an art as well as science.

Check your Progress

Now review your reading

1. What is subjectivity ?
2. How does it differ form objectivity ?
3. What is value-free research ?
4. Define scientific sociology.
5. "Human Society is complex" Discuss.

1.4 CONCEPT AND CONCEPTUALIZATION

1.4.1 Concept : Its meaning

Science studies a series of phenomena in the natural as well as in social world. Both natural and social worlds are full of phenomena. But it is not possible to study the whole phenomena at a time. So, instead or studying whole phenomena, science deals with certain phenomena, or some aspects of phenomena. In such approach, scientists use certain terms or vocabularies or words. These are commonly called as concepts. For example – 'mass', 'weight', 'power',

social status etc. Each science has developed its owing terms or concepts for communicating its research findings or research experience. Social science has its own unique sets of concepts such as –social status, role, caste, religion, socialization etc. When a science refers to a theoretical system, essentially it refers to a conceptual system.

1.4.2 Certain qualities of concepts

(a) A meaning taken for a particular use may be spoken of as the conception. Associated with the usages of a term is a concept which may be said correspondingly to be a family of conception (Abraham Kaplan 1964).

(b) Concepts are impersonal and timeless.

(c) Concept is an abstract construction of knowledge or ideas or conception.

(d) Concepts attain the quality of theoretical expression.

(e) The construction of concept changes.

1.4.3 Concept and Construct

How to construct a concept ?

Concepts are logical construction created from sense impression, precepts and even family of complex experiences (Goode and Hatt, 1964).

Researcher usually takes into account two sets of terms such as –

(a) Terms or words which can point out to an object. For example – When we use of word 'CAR' we can point to an object, the word represents. Similarly, we can measure an area, the height, width and thickness of wall of a building.

(b) The second set of terms such as – race, religion, intelligence, achievement etc. donot have objects to point out. These terms are abstraction from observation.

Hence, we first define these set of words without observable characteristics. The terms of the first set which have direct empirical referents are referred to as concepts. A construct is a deliberately adopted concept having an added meaning for a specific purpose. For example 'acheivement' is a concept and is abstracted from the observer's viewpoint-acheiver and non-acheiver's.

1.4.4. Concept and Communication

What concept communicates ?

Concept must not a rouse feeling vaguely-it must communicate in a very special sense. The components of concepts must be known. The elements which are derived in the process of construction of concepts is to be defined and this is the basic to conceptualization.

Science develops a variety of vocafularies to communicate meanings within its own discipline. But still there are difficulties in communication. This is because of the fact that definitions of the scientific terms and concepts lack adequate clarity. The reasons are –

(a) Concepts develop from a shared experience

(b) Terms used to denote scientific concepts may have other meanings.

(c) A term may refer to different phenomena.

(d) A term may have no immediate empirical referent at all.

(e) Meaning of concept may change.

1.4.5 Concept and Conceptualisation

As the science develops we see conceptual difficulties. Since it is stated that the meaning of concept may change. The change may cause difficulty in communication. Hence, there is an added consideration in building concepts. The following considerations are added to rebuild concepts –

(a) Terms must be clear with little difficulties

(b) The context and use of language while developing concepts

(c) Avoid overlapping usages

(d) As the science develop, many conceptual problems are bypassed when the concepts themselves become irrelevant to the newer theoretical tasks (Goode and Hatt)

So, it is a procedure for clarifying for concepts in new development. This process of clarification may be called conceptualization. In this context Robert K. Merton and Paul Lazarsfeld used the words like "reconceptualisation" or "re-specification of concepts".

Review your reading

Understand the meanings of the following terms clearly Phenomena ; Natural world, Social World ; Research Findings ; Research experience ; Family of conception ; Abstract construction ; Theoretical expression ; Conceptualization ; Shared experience; Empirical referent.

1.5. ETHICS IN RESEARCH

1.5.1 Defining Ethics

Ethic is a set of moral principles or rules of behaviour. In research-we mean a set of duties and responsibilities of a researcher which are considered obligatory on the part of the researcher to follow or maintain while doing research.

What sort of duties and responsibilities a researcher has to maintain ? Are there any specific duties and responsibilities on the part of social researcher ? To answer such questions we should go through the pages to follow.

1.5.2 Ethics in Research.

Sociologists study human subjects and so as researchers they are bound to confront several ethical problems. Such problems arise in research work from the beginning to the end of research. Some of such problems are –

1. The first problem may arise on his selection of subject for research.
2. A second problem concerns with regard to subjects participation.
3. Sociologists desire to know something secret from the respondents. All aspects of social and personal life of people may not be made open under certain circumstances. The desire for the sociologists to know the privacy of the informants sometimes pause a conflicting situation.

There are certain basic things in science. A scientist demonstrates things. Such demonstration may be regardless of its goodness, rightness or beauty. It demonstrates that certain given conditions or certain events inevitably follow in the very process of its occurances.

Scientist evaluates facts in terms of its theoretical relevance. Thus, science discriminates between problems, judging one or more important than another and more desirable for solution.

About the question of 'importance' scientists usually put two sorts of questions such as –

1. What is the scientific significance of research findings ?
2. What is the practical utility of research findings ?

The first question can be answered through the use of scientific theory. In order to answer the second question science must introduce values. The two questions cannot be answered independently – one is dependent upon the answer to the next. Here comes the conviction, faith and it must be based on values.

An ethic is more than the presence of a basic value or values. In science, the basic value is simply stated as faith that '*it is better to know than not to know*'.

So, seeking after knowledge, disseminate knowledge is value. Since knowledge is believed to be better than ignorance for ordinary peoples as well as scientists, the findings of science must be made public.

Like other branches of science, sociological research has also two important aspects to be considered such as –

(a) a research may be harmful, or

(b) a research may be helpful

As a part of ethical consideration, The American Sociological Association (1984) has determined to establish some formal guidelines for the conduct of research.

These are as follows –

1. Sociologists strive for technical competence and objectivity in conducting research. Sociologists must disclose research findings in full, without ommitting significant data. Further, they must include all possible interpretation of data and they are ethically bound to present their results available to other sociologists, some of whom may wish to replicate the study.
2. Sociologists must strive to protect the right, privacy and safety of anyone involves in research projects. Sociologists are obliged to terminate any research, however promising it may seem, if they become aware of any potential danger to participants. Even if research is likely to cause subjects, discomfort or inconveniences, sociologists must ensure in advance that all participants understand and accept any risk.
3. Subjects who agree to take part in research are entitled to full anonymity, even sociologists come under legal pressure to release confidential information. Researcher should never promise such consideration, therefore, they are willing and able to carry through on their commitment.

4. Sociologists are ethically bound to present accurately the purpose of their work. They must inform to the subject, in advance, of any political organization or business that they represent and this information must also be furnished in reports of research findings. All sources of research findings must be disclosed in the publication of findings.
5. Finally, there is a global dimension to research ethics. Sociologists have been called on to collect data for government agencies. Professional ethic however, demand that sociologists never use their professional role as 'cover' while gathering intelligence information for any government.

These guidelines for research ethic framed by the American Sociological Association may be applicable to all sociologists over the globe.

1.5.3 Ethical Neutrality

In the language of natural science 'ethical neutrality' means that scientific truth should be pursued no matter whom it hurts, whom it helps or what values it serves. The advent of nuclear weapons and other means of widespread destruction has called this position into question.

The sociologist may express moral values, hold strong political opinion and speak out of public issues. Many sociologists even believe that in their work as social scientists, they are mainly guided by scientific concerns, avoid political involvement and be scholars rather than activists. That is, they must remain ethically or at least socially or politically neutral. They would believe they will stick the position of sociologist which Max Weber termed this view as *Value-Free Sociology.*

But other situation also prevailed. Many sociologists argue in favour of neutral sociology. They feel that neutral research helps to maintain existing social order. Gouldner and some others advocated that, sociologists if stricly affirm the status-quo it may hamper social changes. These critics argue that sociologists should become scholar activists and use their research to bring about social change.

1.6. LET US SUM UP

The foregoing discussion gives us the following ideas. Firstly, Science is a systematic body of knowledge which has its own characteristics. Secondly, Science follows a systematic rules and procedures. Thirdly, it is to be noted that scientific methods are objectively followed in the study of Social Phenomena. Fourthly, Social researchers have tried to be objective while studying social phenomena.

Key terms

Concepts: a logical construction or word created from sense impression, precepts and experiences.

Objectivity : a condition of unbiased, concrete, real and not influenced by emotion or prejudices.

Subjectivity : subject means person. Subjectivity is a condition when personal feelings, attachment, emotion, likeness/dislikeness are rooted in understanding phenomena.

Value-free: When value is considered in understanding phenomena and decide upon the phenomena it is termed as value-relevant.

Answer the following questions:

1. 'Science is body of knowledge' explain
2. Point out some characteristics of science.
3. 'Verification is the very essence of science'- Discuss.
4. Expand Robert A. Nisbet's description about science and scientific characteristics.
5. What do you mean by universality of science ? Give example.
6. Eleborate the meaning of intersubjectivity.
7. What are the complexities that may arise due to inter subjectivity ?
8. Define the terms 'Concept' and 'Conceptional system'.
9. Mention some concepts frequently used in sociology.
10. 'Concept are abstract construction' Discuss
11. Discuss how construction of conceptsd changes.

12. Discuss how concepts communicate ?
13. What do you mean by ethics is social research ?
14. Define clearly the meaning of 'the theoretical relevance".
15. What is practical utility or research findings ?
16. "It is better to know than not to know" Elaborate the meaning
17. Explain "global dimension of research ethics".

Suggested Readings

1. W.J. Goode & P.K. Hatt: Methods in Social Research, Tata Mcgraw Hill, 1952 (Ch. 1, 2&3).
2. Robert A. Nisbet : The Social Bond-An Introduction to the Study of Society, New York, 1970)Part-I)
3. John J. Macionis : Sociology (4th Vol.) Prentice Hall, New Jersey, 1993 (Part-I, Chap,2)
4. David Popenoe : Sociology (10th Edn.) Prentice Hall, New Jersey, 1995 (Part-I)

BLOCK – 2

FORMULATION OF PROBLEMS AND DESIGNING OR RESEARCH

STRUCTURE

2.0 OBJECTIVES

This unit aims to introduce some basic steps of research which are usually followed by the social researchers. This unit therefore, will help the reader learn the procedures adopted in conducting social research. As an initial attempt, this unit tries to provide outline as to how topic of research (problem) is selected and formulated. There are certain types of research designs/ or research strategies availably practised by social researchers. In this attempt, our discussion will confine only to four types of research designs such as – Descriptive, Exploratory, Comparative and Case Study.

2.1 INTRODUCTION

Since sociological research claims to be a scientific endeavour, the students of sociology have to learn the detail methodological steps which are long practised in the science of sociology. We want to introduce the various steps of research procedure for formation of hypothesis, research designs etc. in clear manner.

2.2 INITIATION OF RESEARCH

Simply stated research is a systematic procedure to study the inter-relationships between observable phenomena. When a scientist looks at social world, he must cope with a far greater range of variables than the natural scientist. This fact states that human social world is full of myriads of diversities and problems. Obviously, this creates a host of problems for the scientist. Scientists when takes up certain problem for study, he has to select one definite area or specific problem for systematic investigation.

A research topic is a specific kind of inquiry about *possible relationship between two social events or situations.* But to select a topic for scientific research is not an easy task. There is a wide range of potential topics for social research, since social relationship covers a wide range of social situations. Any relationship involving two or

more persons which constitute groups and human interaction, may constitute a legitimate topic for social research. In other words, social research examines not only the relationships among individuals, but also those between individuals and their families, informal social structures, communities and culture. For such reasons, the researcher may be interested to select problems that may concern to his own needs.

Thus, selection of *topic for study may vary considerably.* Sometimes, selection of a problem may be based upon *theoretical or practical desire.* The research topics of practical concerns may suggest for improvement of certain practical problems, which demand immediate solution to society. Providing facility or services, effect of settlement house, effect of campaign programme etc. are some of such example. So which one may select a topic for research.

Scientific or intellectual interest may suggest a wide ranging topics for research. Sometimes scientific or intellectual research may led general questions relating to general class of phenomena. In such investigation the researcher may explore some general subject matter about which little is known to the wider world.

Again, some researcher may be interested to know about particular phenomenon which were *already studied* by some scholars. In such attempt, phenomena may be studied in different context under which these may appear in society at different time and place.

Quest for research may originate from variety of sources and assume variety of forms. Under such variety of sources, it is indeed not easy to suggest as to how does the investigator select a problem for study. First of all, it is the *personal inclination of the researcher* to select his own topic for research. Social scientists *with different inclinations and values choose different topic* for investigation. Whatever the choices of the topic for research, it is an agreed fact among the scientists that the rationality of scientific procedures must be maintained throughout the study.

Check Your Progress

1. Discuss how a research topic can be initiated. Answer within four lines.

   _____

   _____

   _____

   _____

2. 'Scientific or intellectual interest may suggest topics for research.' Elaborate within four lines.

   _____

   _____

   _____

   _____

2.2.1 Formulation of Research Problem

Research problems may also be termed as research issue. Once a research topic has been selected for research, in academic, parlanee it is termed as *research problem,* which needs solution. Therefore, formulation of research problem is the next task of research after the topic of research has been selected.

Researcher cannot begin his research work unless the problem of study is properly formulate. It is very important to formulated a piece of scientific inquiry before taking up the step of data collection. Formulation has therefore certain steps and this involves –

(a) adoption of scientific procedure,

(b) methods of data collection

(c) use of theoretical basis, and

(d) nature of analysis of data.

Cohen and Nagel in their book "An Introduction to the Logic and Scientific Method" stated that to solve a research problem a

researcher has to face enormous difficulties in both practical or theoretical situation if the research problem is not correctly formulated. According to them, formulation is not only a systematic ordering of facts, but also proper articulation of theoretical relevance.

Check Your Progress

Cleary state what do you mean by formulation of research problem (within five lines)

_____

_____

_____

_____

_____

Any scientific inquiry is geared to the solution of problem. Seltiz and others illustrated few steps for formulation of research topic. These are –

(a) to make research problem concrete and specific

(b) to make it a task of manageable size

(c) to make research problem into a number of sub-questions

(d) to formulate hypothesis, clarification and formal definition of concepts used in the study

(e) developing of working definitions.

Each of the above points are elaborated below

(a) Concrete and Specific

The original research problem is made into specific questions – which are otherwise called research questions. The main topic is sub-divided into a number of research questions and each of such question should be concrete and specific. Each of such questions should have possibility to identify. The vital aspects of the problem

under study. The terms concrete and specific suggest the *possibility to investigate the problem area with the resources available.* In this venture, the researcher should construct some specific questions centering around his main topic of study. The exploratory studies are the examples of such research.

(b) Task of manageability

The task of manageability includes time-frame, cost and expenditure and capability to complete the research work within definite. This is a necessary step for formulation of research problem. A research problem may include diverse areas and so, it may be of diverse nature. A researcher cannot evaluate all aspects of a problem within specified period of time. It is therefore, necessary to make the research problem into a manageable size within a definite theoretical frame.

(c) Making Sub-questions

A research problem must seek answer and so research is first of all a question. The research problem is made into a number of sub-questions and each sub-question is a branch of the actual topic. That way, each sub-question is designed in such a way that it contributes to the required and related answers to the desired aspects of the topic. So, making sub-question is a purpose to understand the minute details of all aspects of the proposed research. The problem of study is thus narrowed down into a single study that can be handled effectively making into a task of manageability.

The rules for formulation of hypothesis are discussed in the subsequent pages to follow.

Remember the following statements

1. A research topic is a specific kind of inquiry searching possible relationship between two social events or situations.
2. Selection of research topic may vary.

3. Research topic is a research problem and it needs solution.
4. Research problem should be correctly formulated.
5. There are definite steps for formulation of research topic-it should be concrete and specific, manageable, making a number of sub-questions, formulation of hypothesis, developing of working hypothesis, clarification of concept etc.

2.3 HYPOTHESIS

Once we have formulated the problem for research, we proceed to formulate hypothesis. In simple meaning *a hypothesis is a tentative generalization of statement showing a relationship between two or more variables.*

2.3.1 Definitions

The available textbooks on research methodology or any introductory book on sociology provide numerous definitions of hypothesis. Some very classical definitions are discussed below.

Seltiz and others defined hypothesis "as a propositions, conditions or principle which is assumed perhaps without belief, in order to draw out its logical consequences and by this method to test its accord with facts which are known or may be determined.' Such propositions can be put to test to determine its validity. Hypothesis is a tentative generalization. A hypothesis at its initial stage may be an imagined idea or a mere guess. In this context, it is important to know what is "propositions" A proposition is a statement about relationship between two or more concepts. The propositions are constructed showing cause-effect or logical relationships between two social variables.

A hypothesis is made in order to find out correct explanation of a phenomenon through investigation. In the light of hypothesis facts are observed and collected. When by observation, the hypothesis is found to be true, a theory is obtained. A hypothesis if

found to be true, a theory is obtained. A hypothesis may thus, be defined as a propositions or principle or a condition which can be visualized, being based upon facts. A hypothesis may be proved to be correct or incorrect. In any case it requires an empirical data. It is a kind of proposition questioned in a manner that an answer of some kind is forthcoming. It is in this sense that every theory pursuits the formulation of additional hypothesis. After the theory is tested and found correct or incorrect, the theory is revised or reformulated, and it is the purpose of research.

Check Your Progress

1. Construct three hypotheses showing cause-effect relationship.

(a) _____

(b) _____

(c) _____

2.3.2 Characteristics of Hypothesis

Hypothesis must contain some specific characteristics. Goode and Hatt have outlined these as follows –

1. Hypothesis must be conceptually clear. Concepts must not be ambiguous in their meaning or connotations. Clear definition of concepts is operationally workable.
2. Hypothesis should have empirical referents. That is, all hypotheses should be based on empirical facts which can be observed, quantitatively measured or verified.
3. Hypothesis should be concrete and specific. Concrete refers to real and should indicate what it is looked up for. On the otherhand, specific carries the meaning of one aspect of situation, not the whole or general.
4. Hypothesis should be related with available techniques. Scientific researchers are accomplished with its own techniques of data collection. So, while hypotheses are

developed, a researcher should keep in mind the application of available techniques so that the designed hypothesis can be suitably tested.

5. The hypothesis should be related to a body of theory, so that in the light of available theory the phenomena can be adequately explained and analysed.

The above characteristics are the indication as to how hypothesis is formulated. If hypothesis attains those characteristics we can ascertain that such hypotheses are useable or working hypothesis.

Check Your Progress

1. 'Hypothesis should have empirical referents' Explain within three lines.

_____

_____

_____

2. Show the relationship between hypothesis and theory (within three lines)

_____

_____

_____

2.3.3 Working hypothesis

It is the aim of science to discover the cause of occurrence. Since social phenomena are by nature complex, there are multiple causative factors behind the occurrence of the phenomena. The working hypotheses are those which can be observed with empirical situation. The working hypothesis is therefore, necessary to detect the merit of incidence. In the absence of workable idea it is difficult for scientific observation. It is only the working hypothesis which guides the researcher to detect the facts which are required for explanation. The

discovery requires hunch, wit, imagination and method. Mauric Arthur noted the importance of working hypothesis thus –

"The greater the imagination and willingness to experiment, think, observe, think and think again, the more abundant and meaningful the result."

Formulation of working hypothesis gives definite guidelines for inquiry, direction and framework for looking into the problem under study. It delimits the filed of investigation. The researcher with the help of working hypothesis may determine facts which are necessary for investigation, and he may delimit the research area by eliminating the irrelevant facts.

2.3.4 Type of hypothesis

Hypothesis may be generated at any stage of research inquiry. And hypothesis may be classed or categorized. Goode and Hatt classified three types of hypothesis.

These are –

(a) **Hypothesis stating existence of empirical uniformities :** These are regularities or uniformities of behaviour of social existence like residential areas, business establishment, ethnic background of workers etc. Researcher in such condition only gather facts and donot test any relation between facts.

(b) **Hypothesis concerning with complex ideal type :** This type of hypothesis are concerned with testing relationship between empirical uniformities. The Chicago School of Sociology tested hypothesis regarding the growth of city in relation to land value, industrial concentration, types of business, ethnic groups etc.

(c) **Hypothesis of analytical variables :** This type of hypothesis refers to higher level of abstraction. Changes in one property may lead to change in another. Relationship between variables of such nature is established through this type of hypothesis.

2.3.5 Functions of Hypothesis

We have touched upon certain aspects of function of hypothesis in our proceeding pages. Some specific functions of hypothesis can be pinpointed as below –

(i) Hypothesis guides research and suggests explanation of facts.

(ii) It enables the researcher to direct inquiry along the right lines. It suggests experiments and observation. It helps in collecting necessary evidence in order to discover the order of nature.

(iii) Hypothesis determines the method of verification as well as the procedure for inquiry. Hypothesis links the scope of inquiry to a manageable area, because instead of random collection of data, it enables the researcher to search only for those facts which are relevant to the questions raised in the hypothesis. Therefore, it leads to economy of time and money.

(iv) Hypothesis leads to the discovery of laws. It explains facts and laws, and thus seeks to verify knowledge about the behaviour of phenomena studied.

(v) Hypothesis leads to conclusion which is sometimes very significant for the advancement of knowledge. The significance of an object or event can be determined by a hypothesis.

Besides, hypothesis performs some essential functions in research. It helps in the analysis of data accordingly. It is sufficient to say that a researcher cannot go beyond his pre-assumed theories when he analyses his data in accordance to his hypothesis. In this stage hypothesis guides him to reach the target of research. The selection of samples, field of study, designing of tables and data analysis, making final comments and deriving theoretical ideas etc. are solely guided by hypothesis.

In a nutshell, the function of hypothesis in research are – (1) to test theories, (2) to suggest theories and to describe social phenomena. Ram Ahuja indicated some supplementary functions of hypothesis such as – (1) it helps in formulating social policy and to solve various social problems, (2) to assist in refuting certain "common-sense" notion (e.g. men are more intelligent than women) and (3) to indicate need for change in systems and structures by providing new knowledge.

According to Pauline V. Young, hypothesis may give a vital clue for fruitful investigation and depth of problem. Usefulness of hypothesis, as Young observed depends on the researcher's (a) keen observation, (b) disciplined imagination and creative thinking which provides at best practical insight, and (c) some formulated theoretical framework.

The role of hypothesis in investigating social problems has paramount importance. A hypothesis, after testing, may be accepted or rejected. There is no question of defending the hypothesis. The researcher may be confronted with data of both favouring or disfavouring his original hypothesis. What is important here is to rely on the fact of empirical situation. The empirical facts which try to disfavour the original hypothesis may constitute a vast granary of information for finding new ideas about social phenomena.

It is not that the researcher will try to prove the correctness of his hypothesis. He should test the hypothesis and be ready to accept either the positive or the negative results of his facts and this is the true scientific spirit of inquiry.

Check Your Progress

1. List three functions of hypothesis in research.

(a) _____

(b) _____

(c) _____

Review your reading

Hypothesis is a proposition, condition or principle which is assumed perhaps without belief in order to draw its logical consequences. Hypothesis may be proved to be correct or incorrect. Hypothesis guides research and suggests explanation of facts.

Working hypothesis can be observed in empirical situation. Working hypothesis gives definite guidelines for enquiry, direction and framework for study.

Hypothesis guides research, suggests explanation of facts, suggests experiments and observation and helps in collecting necessary data.

2.4 ELEMENTS OF RESEARCH DESIGN

2.4.1 Meaning of Research Design

The term 'design' means drawing an outline or planning or arranging in detail.

A research design in a plan prepared to carry out research work. After a research topic/ problem has been formulated, the next stage of duty of a researcher is to prepare a systematic and detail plan of work, so that there is a smooth and continuous progress of research work. In this sense, research plan is a research strategy prepared with a definite target or goal to achieve. Thus, if a research plan is correctly outlined, the subsequent stages of research could be performed with no difficulty. Selltiz and Others maintained.

> "a research design is the arrangement of conditions for collection of data in a manner that aims to combine relevance to the research purpose with economy in procedure."

Various researchers put forward their definitions of research design. According to Henry Manheim "research design does not only anticipate and specify the seemingly countless decision connected with carrying out data collection, processing and analysis, but it presents a logical basis for these decisions".

Again, William Zikmund has described research design as a "master plan specifying the methods and procedures for collecting and analyzing the needed information".

To Martin Blumer

"a research design is the specification of the problem, conceptual definition, derivation of hypothesis to test and defining the population to be studied".

2.4.2. Basic Elements of Research Design

Since a research is a systematic procedure of work and thought, so a researcher must be guided by some pre conceived notions before working out of his research plan. So, to plan a research work, a researcher has to decide about certain vital questions. In other words, this is the prior condition to think about the elements of research design.

What are the elements of research design?

Sociologist P.V. Young outlined some important aspects of research design. Young suggested that some basic questions of a researcher must be taken into account while preparing a design of research. These are –

What the study is about?

Why the study is being made?

What type of data are needed?

Where the needed data can be found?

What will be the universe of study?

How much materials will be needed?

What will be the bases of selection?

What techniques will be used to collect data?

What period of time the study will cover to complete the work?

etc.

These and many other aspects of the research are to be considered while preparing the plan of research. At the sametime, any research plan is considered tentative or flexible, as because it may take up new ideas while research work is in progress.

The queries indicated above are the indications of some components of research design. A researcher keeps in mind about what, when, where, why, how much, by what means etc. These are the guiding principles for the researcher while planning his research strategy.

Whatever may be nature of research design, four general rules should be followed in planning a research design. These are –

1. Define the nature and scope of the problem.
2. Specify the related variables
3. Exclude the variables not related to the study and
4. Start from logical hypothesis.

In general, whatever the nature of topic for study, certain steps must be followed in preparing a design of study. These are discussed as under.

2.4.3. Steps in Preparing Research Design

2.4.3. (a) Define the problem

Research problem is defined in many ways. Sometimes, by way of "Introduction", or Statement made about" or sometimes as "The problem under study'. The problem selected for study should be defined clearly in operational terms so that the researcher knows positively what facts he is searching for and what is relevant to the study. Since, human behaviour is the result of various forces, it is best to delimit the scope of one's study which facilitates maximum benefits during the course of data collection. The problem selected should be practicable with regard to the monetary cost involved and time required to complete.

(b) Sources of data

Researcher should indicate the major source of data to be required for research and this is a vital part for a research design. Classical methodology suggests that library, personal documents, archives, official and unofficial sources, census publications, newspapers are the chief sources of data. In the present day sources of information are widely scattered. Network of information and communication, increasing number of research and professional journals provide large number of sources of information. So, a researcher should clearly state the sources of information in his planning of research design.

(c) Nature of study

A research may be descriptive, exploratory, experimental, comparative, historical or case study in its approach. In planning the research design the researcher should make it clear about the nature of study.

(d) Objective of Study

Social research can spelledout a large number of objectives ranging from relatively simple to complex and encompassing both basic and applied research. Research objectives include exploration, description, explanation, understanding, prediction, change, evaluation and impact assessment. A research project can pursue just one of the above objectives or perhaps a number of them in sequence. The design at this stage must explicitly make it clear about its objectives. The clarification of objectives makes clear the concepts as well as the angles of study.

(e) Universe and Sample

A research design must clearly indicate about the field of study and the units to be selected for study. The units which are to be selected are here referred to as 'sample' and the population from

where such samples are drawn is 'Universe'. In a scientific and methodical research design, it is necessary to mention the reasons behind the selection of particular universe of study, and the sampling procedure should be clearly outlined.

(f) Techniques of Data Collection

Relevant to the research design, researcher must select appropriate techniques for the collection of required data. Usually, in the social science researches, researcher apply more than one technique to study social problems or phenomena. Some of the popular techniques of data collection are observation, questionnaire, interview, interview schedule, case study methods etc.

In social science researches, observation is regarded as classical technique of data collection. Researcher observes field situation according to his pre-planned theoretical guidelines or in accordance with his assumed hypothesis. So, data collected through observation are not simply haphazard, but it is truly under the direction of hypothesis or theories. Observation is mainly of two types-participant observation and non participant observations.

Both questionnaire and the schedule are much-used techniques for collecting vast and variety of data. The questionnaire is generally sent through mail to the informants to be answered with a covering letter. The schedule on the otherhand is administered in face-to-face situation with the informants. This is called interviewing of the respondents. Interview is not merely a conversation between a researcher and the informant, it is a social process. It is an effective, informal, verbal and non verbal conversation, initiated for specific purposes and focus on certain issue and problem.

Frederic Le Play has introduced Case study method into social sciences. Through this method social scientists strive to obtain a real and enlightened record of personal experience that reveal a concrete detail of man's inner striving tensions and motivation. Thus, a comprehensive study of social unit-be that unit a person, a group, a social institution, a district, or a community is called case study.

(g) Tabulation and analysis of data

Tabulation as a statistical procedure involves variety of techniques which needs a careful design in planning for data analysis, Data are analysed with the help of tables of various design. The process of tabulation and data analysis constitute an important element for planning a research strategy.

(h) Report writing

In social science researches presenting a report of study needs a systematization of thinking in relation to its theoretical framework. Report must reflect a procedural attempt of study. An elaborate discussion on report writing is presented in Unit 3.

2.5 THE RESEARCH DESIGNS

It has been stated earlier that research design or planning of research depend upon the nature of topic of study and research purpose. Thus, social researchers adopt various research designs. Some types of research designs are ---

(a) Descriptive research design

(b) Exploratory research design

(c) Comparative research design

(d) Case study design.

The details of each design are discussed below. Each type of research design has its own specific purpose.

2.5.1. Descriptive Research Design

Descriptive study or design is basically interested in presenting detail description of the phenomena, group a community. A lot of research work has been done in social sciences describing the character of social phenomena. These include the people of a particular community, their age distribution, ethnic and social background, education, nature of earning, mental health etc.

Descriptive studies are also concerned with the study of attitudes of population towards certain issues such as political participation, educational attainment, religious activities etc. To such extent descriptive research design is concerned study the broad sprectrum of opinion of the people through larger mass of data.

The objective of descriptive study is not only confined to the *description of the phenomena and to acquire knowledge* or *to the fact finding objectives*, it also helps to a considerable extent to formulate hypothesis.

Its functions is to try out the sampling procedure, establish data collection methods, identify crucial variables, determine appropriate methods for operationalistion, assemble data, analysis techniques, formulate potential hypothesis and so on. The result is that descriptive research design becomes more explorative and less experimental.

But the main objective of descriptive research design is to acquire knowledge. There are certain areas where knowledge has not yet properly developed. Under such circumstances descriptive studies try to focus issues like – who, how, what and why of the phenomena.

(a) Preparing Descriptive Design: Steps

The following steps are usually followed in preparing a descriptive research design.

(a) Clear Formulation of objectives

Descriptive studies are concerned with wide variety of description of phenomena. Hence it is essentially important that the objectives are to be clearly specified. The clear objectives may be outlined *in form of research questions which are to be answered during investigations*. Unless the objectives are specified with sufficient precision it may be impossible to collect the relevant data. In the absence of relevant or desired information the research questions may remain unanswered.

The researcher therefore, must be able to clearly define what he wants to measure and describe. The research design much be such that the *bias can be minimized and reliability can be maximized.*

(b) Data Collection

The selection of techniques for data collection is the next important step. What data are required, what techniques will be followed, which sources of information provide maximum clues for data are to be clearly specified. Generally, *observation, interviewing questionnaire, projective techniques, examination of records and documents and combination of various other methods* are popularly used by the social researcher pursuing descriptive design.

(c) The Questions

The questions to be asked to the respondents must be *clear and straight forward. It should not be ambiguous and should be easily understandable.* The questions must aim at the desired responses. As a whole the questions should enlighten the respondents with interest.

(d) Selection the samples

In descriptive study, the researcher wants to make concluding statements about some defined groups or people. It is rather impossible to study all the people and so, *samples of representative character* are studied. The samples should be reasonably accurate. It is better to use random or matching samples. The researcher should supervise the staff of field worker who are engaged in data collection and recording.

(e) Checking the Data

The collected data must be checked properly. The respondents may leave certain questions unanswered owing to lack of understanding or personal reasons. Checking is necessary *to obtain complete, comprehensible, consistent and reliable information.*

Thus, checking of collected data is necessary to prevent any irrelevant information which may cause problem in explaining facts in an objective manner.

(f) Data Analysis and Results

In descriptive studies social phenomena are understood with wide range of data. So, to obtain the ultimate and final results, a researcher has to adopt a number of techniques to analyse data. Mainly, the following are important viz. coding of data, tabulation and use of statistics.

Coding is a process through which responses are put in numbers. Coding is required for the use of punching card. Once the responses are coded properly, it saves time to make out the tables of data which are required for analysis.

Tabulation is a process of systematic arrangement of data. Data tables are made with various purposes and objectives. In brief, tabular representation of data helps in systematic observation on the data, saves time, minimize description and makes easier for drawing conclusion.

Statistics are part and parcel for the analysis of data. Social data are enormous, complex and haphazard by nature. Hence the use of statistics is essential. A detail description with regard to the use of statistics in social science research is presented in the Unit 4.

Check Your Progress (Answer)

Q. What descriptive research aims at?

Descriptive research confines to the description of the phenomena and to acquire knowledge of the fact.

Q. What is representative sample?

Representative sample is that sample which correctly represent the characters of the universe from which the samples are drawn. It is the homogenous universe which can produce samples of representative nature.

Q. What is tabulation?

Tabulation is a systematic arrangement of data. The purpose of tabulation is to observe the data systematically, it saves time, minimize description and makes easier for drawing conclusion and generalization.

Key words to be remembered

Fact – finding objectives

Operationalization

Relevant data

Coding

Punching Card

Generalisation

Matching Sample

Representative sample

Bias

Reliability

Check Your Progress

1. Mention three aims of descriptive research design.

   (a) _____

   (b) _____

   (c) _____

2. What are the specific features of descriptive research? Mention below.

   _____

   _____

   _____

2.5.2. Exploratory Research Design

Exploratory research design aims at gaining familiarity with a phenomenon or to achieve new insight into it, often in order to formulate a more precise research problem or to develop hypothesis (Selltiz and Others 1965).

What to explore?

Much exploratory research is qualitative. Research is interested to explore the causes behind any occurrence, the issues which prompt happenings. So, the design of such inquiry has naturally to be different. According to some scholars the exploratory study is undertaken for following reasons:

1. **Feasibility of Research :** To find out whether the study on the issue in question is warranted, worthwhile and feasible.
2. **Familiarisation:** To familiarize the research with the social context of the issue i.e. detail about relationships, values, standards and factors related to the research topic.

3. **New Ideas :** To generate new ideas, views and opinion on the research issue which will help in proper understanding of the problem and to develop new insight.
4. **Formulation of hypothesis :** To show whether variable can be related to each other.
5. **Operationalisation:** To operationalise concepts by explaining their structure and by identifying indicators.

2.5.2 (a) Preparing the Design: Steps

Selltiz and his associates identified three basic steps/ phases to prepare a research plan for exploratory studies. These are

1. A review of the related and pertinent literature.
2. a survey of people who have had practical experience with the problem to be studied and
3. An analysis of "insight stimulating studies".

Review of Related Literature

A researcher before going to start his research work has to undergo a survey of all existing related studies. He has to locate the relevant literature through various means such as – books, journals, leading newspaper, proceeding of conference and symposia, documents etc. The sources of information has increased tremendously like – internet, radio, television, telephone and various other means of mass communication. Searching out information about relevant works and listing them is generally called *bibliographic survey*. The Librarians and experts in the field of information technology may be consulted for help in identifying the recent publications and availability of sources.

Usually, literature survey in the field of sociological research is done by reading large number of journals in Sociology. There are a number of international as well as national level journals in

sociology. Some of such *international journals* are – American Jr. of Sociology (AJS), American Sociological Review (ASR), Social Farces, Sociometry, Sociology of Education, International Jr. of Comparative Sociology, Urban Affairs Qtly, Rural Sociology, Public Opinion Qtly, Pacific Sociological Review, British Jr. of Sociology, British Jr. of Industrial Relation etc. Some *National level Research Journals are* – Indian Jr. of Social Research, Social Action, Indian Jr. of Social Work, Man in India, ICSSR Abstract & Reviews etc. The Indian Council for Social Science Research (ICSSR) is a leading organisation in the field of social science research which provides up-to-date information of social science research in the country.

These sources can be very fruitfully utilized while a researcher makes survey of related studies concerning to his field of inquiry. Nan Lin (1976) stated that a careful review of literature may be made in three broad directions. These are –

(a) Initial reading

(b) Summarising information

(c) Case exploration

Check Your Progress

1. What are the advantages of reviewing related literature in research? Mention three advantages.

   (a) _____

   (b) _____

   (c) _____

Experience Survey / Expert Survey

Informal interviews with the experienced people of the area of research may be very helpful in getting insight into the subject and its various facts. People who are or have been doing some work on this type of problem or who have theoretical or empirical knowledge of ideas on it, may be consulted with profit. Such people may be from

any profession, group or community. The researcher may prepare a list of such persons and may interview them informally or formally.

2.5.2. (b) Data Collection

Techniques : Interview, interview schedule, observation, case study and repeated questioning may be the basic tools for data collection, Before the actual interview, the researcher must have sufficient "Feel" of the problem. He must *know what to ask and how to ask questions.* The researcher must be able to distinguish between facts and opinion. He must discard cautiously the subjective bias of the interviewees. It is better to prepare beforehand list of question that are to be asked. The questions should be problem-raising and problem solving.

2.5.2. (c) Data Analysis and Results

An exploratory study is by nature a structured study. Questions must be sharply focused. It is important to remember that exploratory studies merely lead to insights or hypothesis, they donot test or demonstrate them. An exploratory study must always be regarded as simply a first step; more carefully controlled studies are needed to test whether the hypotheses that emerge have general applicability.

Terms to be remembered

Pertinent literature ; Bibliography ; Feasibility ; Familiarisation ; Expert Survey ; Problem-raising issues ; Controlled studies

2.5.3 Comparative Research

Quite often, experimental research finds difficulties in social sciences. Many sociologists therefore, consider comparative research as the fundamental method of their discipline. This was the opinion of the father of sociology August Comte and Emile Durkheim.

Auguse Comte defined comparative Research as –

> "a method by which the life of people living in the sametime in various parts of the world were compared so as to establish general laws of existence and development of society".

For Durkheim "comparative method is a comparison on levels of development of society."

In Comte's view comparison of animal societies with human ones, so as to show similarities and differences could also be useful. Finally, one could compare the social position of the various classes of one and the same society.

The term comparative analysis usually refers to research involving the comparison of several social systems such as – nations or major segments of nations- or the same social system at more than one point of time.

2.5.3. (a) Need of Comparative Research

For a scientific and objective assessment, there is a need of comparison. The reason is that social theory is developed in small corner – so it needs comparison for universal applicability. For example – the theory of Sanskritization developed by MN Srinivas in Southern India needs a comparative analysis of its contents to other parts of world situation to make it universal acceptance.

To make universal applicability sociologists conduct cross-cultural research or cross-societal research or cross-national research. This makes possible for understanding of comparative problems.

(b) Views of Comparative Research

According to Redcliffle Brown, a British social anthropologist,

> "Comparative method provides general explanations, while particularistic explanation is given by the historians."

Durkheim distinguishes three applications of comparative methods.

(i) Analysis of variations within one society at one point of time.

(ii) The comparison of societies generally alike, but differing in certain aspects i.e. different societies or same society at different period.

(iii) Comparision of societies generally dissimilar, yet showing some features in common.

These show intra-societal comparison. But intra-societal comparison has some shortcoming as because it is a sectional study within society, it minimizes the range of society, and its explanation is little.

G.P. Murdock, a British sociologist criticized the above type and said – "comparison gives rise to cross-cultural test or cross-cultural comparison which ultimately develops comparative sociology" –

Murdock raised a number of questions to be considered while a researcher takes up a comparative method. These are

1. What is a society?
2. What kind of data are relevant in comparative sociology?
3. Why is it so important that comparison is so explicit?
4. What kind of hypothesis is tested in cross-societal studies?
5. In making comparison can we attain universal proposition?
6. On what systematic theoretical base should comparison be made?

Check Your Progress

1. Mention Durkheim's methods of application of comparative design.

   (a) _____

   (b) _____

   (c) _____

(c) Models for Comparative Research

Comparative research has limited theoretical systematization in social science. But yet, some prevailing considerations can be shown as follows;

(a) Principle of comparative analysis

(b) Conditions for comparative analysis

(c) Type of comparison

The details of the above are mentioned below –

Principles :

Comparative analysis consists of seeking *differences* and *similarities,* existing characteristics between situations which are the object of comparison. This is achieved by interpreting the significance of the similarities and differences and by trying to discover regularities through them. Durkheim was the proponent of this kind of research.

Conditions:

To make comparative analysis valid, the comparison must pertain to comparable phenomena or situations. Situations choosed for comparison have a certain degree of resemblance. The objective of comparison is to draw and study the similarities as well as the differences existing between compared phenomena. It presupposes that similarities do exist between phenomena.

The comparative researcher has one or two options-

Either to compare the phenomena which have negligible differences than to proceed with "pseudo-comparison" or to compare those phenomena with too many differences and thereby make "artificial" or "illegitimate" comparison. In social sciences, the second risk is greater than the first, for it is rare to discover absolutely identical phenomena in this discipline.

(d) Types of comparison:

Keeping in view the above- mentioned elements, Morris Duverger proposed two terminologies for comparison such as –

I. Close comparison

II. Remote comparison

Close comparison pertains to those phenomena with fairly high degree of analogy in respect of their structure and their context. This comparison is minute, scrupulous and detailed.

Remote comparisons are characterized by their freedom as regards to requirement of analogy. The researchers take into account concern structural phenomena which are perceptively different and whose dimension and cultural context can be fairly remote.

(e) Certain Types of Comparative Research

I. Cross- national research

II. Longitudinal study

III. Cross-sectional study

Cross-National Research

It is the comparative study of two or more nations. Some areas or sections of the nations are studied and compared. Sociologist Alex Inkeles studied two important variables in his cross-national research. His research tried to set out to discover if 'Nationality is related to psychological modernity. In this approach his two studied variables are "nationality" and "psychological modernity" He defined that psychological modernity has openness to new experience, resistance to traditional authority interest in the news and community affairs, ambition and belief in the importance of science. He obtained *samples* of young men from six countries and matched according to their education and occupation. Next, he measured the degree to which they showed the traits of psychological modernity. He found widespread national variations in the presence of these traits. In his

study of six countries he found out that Argentina and Israel led Nigeria, Chile and India and Bangladesh trailed far behind.

Longitudinal Study

Sociologist studies social change by means of longitudinal studies and cross-sectional studies. A longitudinal study follows the same people over a period of time. A comparison is made of the same society at different points of time, which in effect, is the study of social change. For example, researcher may examine the lives of men born between 1916 and 1925 and compare the degree of changing factors.

Cross-Sectional study

The longitudinal study involves a high cost. Hence sociologists often choose to study social change through cross-sectional studies. Cross-sectional studies compare the responses of people of different age, education, economic status and ethnic background at a single point of time.

Through cross-sectional study, the researcher can study the opinion of two different sections of population on a particular opinion-how they react, how they contrast or how they are similar in attitudes.

Review your reading and answer the following

1. What is comparative study?
2. What is the scientific validity of comparative study?
3. How longitudinal study differs from cross-sectional study? Explain.
4. "Comparative study examines the similarities as well as the differences". Examine.

2.5.4. Case Study Design

Meanwhile, we have defined earlier in this block about the case study in a nutshell while dealing with the techniques of data collection in social research.

Case study is a record of personal experience of a person which would reveal in concrete detail a man's inner strings, tensions, motivations that drive him to action, the barriers that frustrate him or challenge him, the forces that direct him to adopt a certain patterns of behaviour and to live according to a certain scheme and philosophy of life (Parterfield 1941). Social scientists, in their study of human behaviour strive to obtain a fundamentally real and enlightened record of human personal experience with the help of case study method.

Case study is an intensive study of cases which may be an individual, an institution, a system, a community, an organisation, an event, a geographical region, a country or even the entire culture. Case study is also defined as "an empirical inquiry that investigates a contemporary phenomenon within its real-life context. (Lin 1991). Kromerey (1986) holds that "case study involves studying individual cases, often in their natural environment and for a long period of time. It is thus a "kind of research design which usually involves the qualitative method of selecting the source of data. It presents a holistic account that offers insight into the case under study. When attention is focused on the development of the case, it is called "case-history". Few important aspects of case study design are discussed below -

(a) It is a design

Case study is not a method of data collection; rather it is a research strategy, or an empirical inquiary that investigates a contemporary phenomenon by using multiple sources of evidence.

(b) Design of study

The research design of case study involves four components such as -

Designing initial question: These pertain to who, what where, when and how. For example, the case study of a drug addict focuses on questions like *what* type of drugs are taken, *how* often these are taken, *when* was initial step in starting taking drug, *what* was the sources of getting drugs, how much money is spent on drugs in a day/week/month and so on.

Study proposition: There may be two types questions – *General and specific.* The initial questions are general in nature, while the specific questions need to be asked for seeking specific evidence. In the above example, the specific questions could be – in the last one week what drugs were taken by the addict, from whom did he obtain drugs etc.

Unit of analysis : This is concerned with defining the actual case. The case may be a person, an event, and the system that is to be studied. If the unit of study is a group, the persons to be included in the group must be established. So, the researcher has to collect data only from the persons who may included in the group. He is not allowed to collect data from person outside the group.

Linking data to propositions and criteria for interpreting findings. This component relates to data analysis step.

(c) Methodological Principles

There are some principles which are followed for the advancement of case study. These are –

Use of multiple sources: Use of one single source of data collection is not considered adequate evidence. Information from several sources like – interviewing, observing, analyzing documents etc. is regarded as strength of case study approach to obtain the reliability and validity of the findings.

Maintaining a sequence of evidences.

Conclusion derived from one evidence is not considered isolate, since one evidence, is linked with other evidences and so, researcher has to maintain a chain of evidences from one to another. This helps in drawing valid conclusion.

Recording Data : The data may be recorded either in the form of notes, or it may be tape-recorded in minute detail. A full notes may be written latter on as soon as possible immediately after the completion of interview.

(d) Aims of Case Study

Conducting case study has certain purposes. Some of such purposes are mentioned below -

1. Case study method is considered as a tool *firstly as preliminary inquiry, gradually moving to major investigation.* It brings to light the variable, processes and relationship that deserve more intensive investigation. This process may fulfill the aim for detecting a source of hypothesis for future research.
2. To probe the phenomenon deeply and analyse it intensively with a view *to establishing generalising about the wider population to which the unit belongs.*
3. To get anecdotal evidence that illustrates more general findings.
4. *To refute universal generalization.* A simple case can represent a significant contributions to the theory building and assist in focusing the direction for future investigation in the area.
5. To use it as a unique, typical and as interesting case in its own right.

Bergar and Others (1988) illustrated some more reasons for employing case study method in social research.

These are –

- to get intimate and detail information about the structure, process and complexity of the research object
- to formulate hypothesis
- to conceptualise
- to operationalise variables
- to expand qualitative findings and
- to test the feasibility of the quantitative findings

(e) Source of Data for case study

For primary data, interviews and observations are regarded as two main sources, while secondary data are collected through a variety of sources like – *reports, records, newspapers, magazine, books, files, diaries* etc. The secondary sources may not so accurate or may be biased. But *they specify events and issues in greater detail than interview can.*

Interview may be both structured and unstructured, most commonly, it is the unstructured interview which is used by the investigators. The questions are usually open – ended with a conversational tone. However, at times, structured interview is also used as part of case study.

(f) Skills to collect Data

To collect data for case study research, the investigator must has certain skills. These are –

He should be able to formulate precise and relevant questions to extract full information from the respondents. Sometimes, the unexpected responses prompt the investigation to dig deeper.

He should be a good listener i.e. he should pay attention to all clues, moods and words used by the respondents.

He should be flexible and adaptive because data collection does not proceed exactly as planned. Even the focus of inquiry can change a little.

He should try to grasp the responses in the context of the respondents perspective. Sometimes, the responses may be at variance with one another and lead to the need for more evidence.

He should not have any bias in recording information or in analyzing it.

(g) Advantage of case study

There are some advantages of case study method. Some of such advantages are :-

- It makes indepth study possible
- It is flexible with respect to using methods for collecting data e.g. Questionnaire, Interview, Observation etc.
- It could be used for studying any dimension of the topic i.e. it could study one specific aspect and may not include other aspects.
- It can be conducted in practically any kind of social setting.
- Case studies are inexpensive.

(h) Disadvantages

Case study method/research is not entirely free from criticism. Criticism usually goes as follows :-

- Case study may have subjectivity and bias. Investigators subjectivity in collecting data for supporting or refuting particular explanation.
- Little evidence for scientific generalization. Since case study is a study of single case, it cannot be easily

generalized. The evidences collected from case study is to expand theory not to undertake statistical generalization.

- Case study is time consuming
- Case study collects a lot of data, so selectivity has naturally a tendency to be bias.

Doubtful reliability: When we talk of reliability, we talk of measurement which refers to the consistency or repeatability of the measurement of some phenomena. When we speak of reliability we are not speaking of individuals, we are actually talking about scores. Since, case study data are basically qualitative, the measuring instruments frequently used in quantitative data cannot be applied in assessing the final result of case studies.

Missing validity: Sometimes investigator fails to develop a sufficiently operational set of measures. Sometimes situation appears like – what 'Seems' to be true is more important than what is 'true'.

Check Your Progress

1. Mention three points how case study differs from other designs? Answer within the space below.
   a) _____
   b) _____
   c) _____
2. Case study method has little scope for scientific generalization. Mention three points with regard to above in the space below.
   a) _____
   b) _____
   c) _____

2.6 LET US SUM UP

From the foregoing discussion, it becomes clear that any systematic scientific endeavour has some methods to follow. The methodical procedures take shape of a design. Research in social sciences also follows some methodical designs and such designs take various shapes. The various shapes of research designs are named as – descriptive, exploratory, comparative or case study design.

Try to answer the following questions

1. What is case study?
2. What do you man by 'general' and 'specific' questions?
3. Mention some sources of secondary data
4. How do you define 'in-depth' study?
5. Define and give examples of 'open-ended' question.
6. 'Case-study is time-consuming' Explain.

Terms to remember these terms

holistic account
case history
unit of study
reliability
validity
intensive investigation
generalization
structured interview
unstructured interview
bias
feasibility

Question to be answered

1. Discuss the meaning of 'exploration' in social science research.

2. How does review of literature help in formulating research problem?
3. What specific cares should be taken by the researcher to collect data in exploratory research?
4. Differentiate between descriptive and exploratory research design.
5. Discuss how a descriptive research design in prepared.
6. Discuss the aims of descriptive research design.
7. Why various techniques of data collection are applied in descriptive research? Explain.
8. Briefly describe coding, tabulation and statistics.
9. Write briefly on observation, Interview, Questionnaire and projective techniques.
10. Case study is not only a method, it is a design. Explain
11. Explain how case study contributes theory building.
12. Case study uses multiple sources of data. Explain
13. Discuss the skills to be adopted by case study investigators.
14. Discuss the advantages and disadvantages of case study method.

Suggested Reading

1 Claire Selltiz & Others: Research Methods in Social Relations, Methuen, 1965

2. Nan Lin : Foundation of Social Research, McGraw Hill. 1976.

3 Ram Ahuja : Research Methods, Rawat Pub, Delhi, 2001.

4. Young, P.V. (1976) : Scientific Social Survey and Research, Prentice Hall.

5. Norman, Blaikie : Designing Social Research Blackuel Publishers.

6. Yin, R.K. (1989) : Case study Research Design and Method, Sage.

7. Robert Marsh : Comparative Sociology (1967)

BLOCK – 3

TECHNUQUES OF DATA COLLECTION AND ANALYSIS

STRUCTURE

3.0 OBJECTIVES

The aim of this unit is to introduce the available techniques of data collection popularly used by social researchers. Some commonly used techniques, Schedule etc. are discussed in detail in this unit. Besides, an attempt is made to discuss various methods of data analysis and methods of preparing reports of research.

3.1. INTRODUCTION

Social science researchers, more specifically sociologists, since the days of Auguste Comte have displayed variety of techniques to gather information from human society. In this Block we will try to analyse the processes for applying those techniques. Readers will find interest and gain knowledge about the application of those methods.

3.2 OBSERVATION

"Science begins with observation and must ultimately return to observation for its final validation." (W.J. Goode and P.K. Hatt).

3.2.1 Meaning of Observation

Observation may be defined as systematic viewing coupled with consideration of the seen phenomena. That is consideration must be given to the larger units of activity in which the specific phenomena occur.

Sociologists conduct systematic observation to carry out promising research. Therefore, it is possible to work out careful procedure for observing and recording human behaviour. Systematic observation may be conducted either in the laboratory or in the field. We are concerned here mainly to discuss a very important techniques of data collection called observation. Observation is carried out by sociologists either in the field situation or on the data collected for the purpose of research, and so observation is a continuous and ongoing process for any careful researcher.

Social scientists may observe in homes, playground, streets or office. Sometimes they watch formal groups, committees, therapy groups or professional societies-in action.

Observation is the basic method of obtaining information about person or social phenomena under investigation. All of us are constantly engaged in observation. It is probably the oldest method used by men in scientific investigation. If we go through the history of development of science we find that the theories of physics, chemistry, biology, astronomy etc. were developed at first through observation and still the same method has been used by so many scientific researches. In social science also, the earliest method of investigation was probably the observation. In the contemporary period also the anthropologists, sociologists, psychologists and other social scientists have been using observation as very effective means of scientific enquiry.

Observation does not however, means casual or general observation without any specific purpose. As a research technique observation means a carefully planned observation with defined goal. It involves pursuance of questions – what is going to be observed and how to observe.

Now try to answer the question (in the space given below)

1. How observation is a basic technique of data collection in all fields of research ?

   _____

   _____

   _____

3.2.2 Types of Observation

There are several types of observation. The types of observation sometimes depend on the degree of its structuredness i.e. *structured* or *unstructured.* Accordingly, we get two observational procedures-structured observation and unstructured observation. On the otherhand, observation may be classified in terms of role played by

the researcher. Accordingly, there may be two types of observations i.e. participant observation and non-participant observation.

(a) Structured Observation

Structured observation usually refers to a study of clear and definite problem. Researcher is clear about the definition of a particular unit of study to be observed and data to be recorded. This is possible only when the problem of study is accurately formulated. Such situation is possible only in the case of exploratory study.

(b) Unstructured Observation

Observation procedure cannot be planned always in advance. When a researcher goes to the field situation, he tends to confront a variety of changing situations which may alter his original plan for research. Thus, the researcher does not have enough clues to structure his observation. Under such condition, the observation becomes flexible since it changes from time to time. So, the very character of observation is its flexibility.

(c) Participant observation

Participant observation is a method in which a researcher systematically observes people while joining in their routine activities. Participant observation requires personal involvement. Researcher actually becomes one of the members of the groups they study, in order to see the world through the eyes of that group's members. This technique is widely used in cultural anthropology and has undergone great refinement over the years. Cultural anthropologists refer to participant observation as field work, their description of an unfamiliar culture is termed as ethnography. (Maciouis 1987 : 49).

The participant observer shares to lesser or greater degree the life of the observed group. But the degree of participation depends on largely on the nature of the study and the practical demands of the

situation. Sometimes investigators identify themselves closely with the groups they study because the subject matter needs intensive study under close proximity.

Participant observation may be both descriptive as well as exploratory. Researcher usually begins with a few specific hypothesis to test, since they may not know initially what the important questions will turn out to be. Initially, a researcher is likely to be concerned with gaining entry into what may be an unfamiliar social setting. In time, observation provokes many specific questions and answers are organized into a detailed description of a way of life.

Participant observation has two sides. On the onehand, it aims at gaining '*insider's'* look attempting to act, think or even feel the way they do. Unlike other method, participant observation requires a researcher to become immersed in a social setting, not for a week or two, but usually, for a period of years. On the otherhand, during the period of study, the researcher must maintains the role of an 'observer' standing back from the action applying sociological perspective of study.

Participant observation is typically a qualitative research, meaning *inquiry based on subjective impressions*, and so participant observation involves little quantitative techniques (investigation based on the analysis of numerical data). Because, personal impressions play a central role in participant observation.

(d) Non-participant observation

'Non-participant' as the term implies not coming in close contact to any group or community. In non-participant observation the researcher has a detached role and try to record information without any attempt to participate in the interaction process with the group being observed.

Purely non-participant observation is difficult. Because it is difficult to evolve a standards set of role pattern for the non-member who is present but never participating. Both the group and the

outsider (observer) are likely to feel uncomfortable. One has to take part during the process of data collection. In this respect non-participant observation carries the quality of *'quasi-participant' observation.*

3.2.3 Hypotheses and Observation

In research, observation is solely guided by hypothesis. Researcher visits his field with his pre-assumed idea and such pre-assumed idea is the guiding element in the immediate observation. If the anticipated idea is not proved by available data, then another hypothesis is formulated.

3.2.4 Observation and Insight

Sociologists train themselves to observe social phenomena carefully. If he can become a good observer he can investigate with more data at his disposal. Sociologist takes conscious note on social behaviour. During the period of observation he takes note on what he sees and write a complete report on his trip. He analyses his data carefully to discover how much of it is concerned with social behaviour. He needs to learn the socio-cultural context of the observed phenomena. The continuity and meaning of events are the important elements in observation.

3.2.5 Field Observation as a Research Technique

Observations as a research technique slowly involves the research or researchers making observations. There are many positive aspects of the observational research approach. Namely, observations are usually flexible and do not necessarily need to be structured around a hypothesis. Remember – a hypothesis is a statement about what you expect to observe. For instance before undertaking more structured research, a researcher may conduct observation in order to form a research questions. A research questions may be regarded as a bridge between the research topic and hypotheses.

The findings obtained from field observation are considered strong in terms of validity because the researcher is able to collect a depth of information about particular behaviour. However, there are negative aspects also in observation. There are problems in *reliability and generalisability. Reliability* refers to the extent that observations can be replicated. Setting behaviour occur over and over again may be a time consuming task. *Generalisability* or external validity is described as the extent that the study's findings would also be true for other people, in other places, and at other times. In observational research, findings may only reflect a unique population and therefore cannot be generalized to others. There are also problems with researcher bias. Often it is assumed that the researcher may "see what they want to see", Bias, however can often be overcome with training or electronically recording observations. Hence, overall, observation is a valuable tool for the researchers.

Key terms

Cause and effect relationship : a relationship between two variables in which change in one (the independent variable) causes change in another (the dependent variable).

Participant observations : a method in which researcher systematically observes people while joining in their routine activities.

Qualitative research : inquiry based on subjective impressions.

Reliability : The quality of consistent measurement.

3.3. QUESTIONNNAIRE AND SCHEDULE

Questionnaire and Interview schedule are the important tools for collecting data in social science research. Questionnaire refers to a device for securing answers to questions by using a form which the respondent fill in himself. *Schedule* is the name usually applied to a set of questions which are asked and filled in by an interviewer or in a face-to-face situation with another persons. Questionnaire and

schedule obviously have much in common particularly the fact that in both cases the making of the questions is the same for all respondents. However, there are little differences between these two techniques. These are as follows ;

1. Questionnaires are usually sent to the respondents by post or through other means to get it filled up and with a request to return. A note along with the questionnaire is also sent requesting the respondents that he/she is selected as one of the respondents for a research project. While schedules are applied with the respondents and the researcher himself take note of the desired answers counting the respondents personally and with face-to-face interview.
2. There is physical distance between the research and the respondents when data are collected with the help of questionnaire. But with the help of schedule, face-to-face contact between the researcher and the respondents is possible. With the help of schedule researcher conducts interview with his respondents. This makes possible face-to-face contact between the researcher and the respondent.
3. In the questionnaire method, there is every possibility for mistakes in replying to the questions. Respondents may not understand the content of questions or the methods of answering questions desired by the researcher. As such, many questions may be left unanswered or wrongly answered. Respondents may not be interested to pay attention to answer the questions with care. This may result incompleteness of the questionnaire and hence, incomplete questionnaire cannot provide the desired information. Since, collecting data through schedule is accompanied by face-to-face interview the shortcomings appeared through questionnaire are minimized.
4. Interview schedule is administered personally by the investigator. Hence all questions listed in the schedule are properly filled up on the spot. Respondents are interviewed

and observed simultaneously and so, Interview schedule as a tool of data collection serves two purposes-interview as well as observation. Both the purposes are not fulfilled through the method of questionnaire.

5. Many respondents may not return the questionnaire or pay little interests to do that. This creates inconvenience to the researcher because he looses the desired number of respondents. This shortcoming is minimized when data collected through interview schedule.

3.3.1 Constructing a Questionnaire

In this stage we will pay attention as to how questionnaire or schedule is prepared in terms of format or design. Both schedule and questionnaire carry similar patterns of questions, hence some specific research procedures commonly applicable to both the methods while preparing them.

Questionnaire or schedule must contain a set of related items. A topic for research is considered a central problem. The central problem is made into sub-topics and each sub-topics may be considered as an item. All items are related to one another. Hence the questions framed for each item are logically interrelated and such logical relations must be maintained from beginning to the end.

3.3.1 (a) Items and Central Problem

It is important to know that every item in a questionnaire ideally constitutes a hypothesis, or a part of hypothesis in itself. So, the inclusion of item has logical connection and the answer to each item is logically expected since it bears the significance to the central problems. Each item should be defensible one the ground of such logical connection. The accurate answer to each item fulfils requirement to solve the central problem (the subject of study).

(b) The Questions

Constructing a question in a schedule, meant for the purpose of research is a complex process. It requires careful and patient effort. The researchers must know at the beginning – what are the important questions to be asked.

Methodologists Goode and Hatt stated – "developing a questionnaire can be thought of as moving from the 'inside' to 'outward'. At this stage researcher should lay out tentatively the logical implication of his problem. Secondly, he should draw upon his own experience on the problem. Thirdly, the researcher should consult the colleagues, experts, friends to get their thinkings on the problem. Then a draft schedule is outlined and the list of areas to be covered and formulation of rough questions can be tentatively designed.

After such consultation there may be some possible changes in the areas like –

- the list of possible questions grow
- the number of areas which are of interest increase
- the number of areas which the research can cover must be decrease, *ambiguties, biases and poor phrasing etc.* are corrected gradually,
- and closer logical relationship developed between various part of the schedule.

(c) The schedule as unit

The entire schedule or questionnaire is a unit which means each item of the schedule is logically arranged. This logical arrangement is seen in certain qualities much as- (a) from one item to the next, the respondents are drawn into by a awakening their continuous interest, (2) Questions should be from simple to those of complex, (3) not to request for personal of subjective information immediately at the beginning, (4) never ask embarrassing questions,

(5) beginning as smoothly as possible from one frame of reference to the next.

These are the principles of arranging items of the schedule which make the *schedule as a total unit.* The questions formulated in each item are so constructed that response to one question is seen having logical connection from one to the next.

The above points are elaborated thus –

Awakening interest from one item to the next : The general introductory questions and attention-catching questions like – age, sex, marital status, place of residence, education etc. are usually placed at the beginning of the questionnaire.

Simple to complex : placing simpler items first and gradually complex questions which require thought are put step by step. If any item requires considerable thought on the part of the respondents, they should not be placed too early in the schedule.

Avoid personal/subjective information : Some issues are 'delicate' in the senses that respondent's personal issues, feelings sentiments are involved in it. Such questions should be intelligently avoided from the schedule or the special case and intelligence are required to put such questions.

Embarrassing questions : Some issues like voting in election, talking part in strike, dharnas etc. naturally evoke embarrassment on the part of the respondents. If a respondent did not vote he may react otherwise because of certain post-election scenario. Such situations should not be the part of direct questioning to the informants.

Moving smoothly from item to item : merely means that questions of general nature should be put together so that respondents could grasp the questions with clarity.

(d) Type of Questions

There is a variety of questionnaire which can be classified in various ways. We will discuss some types of questions below –

I. structured question

II. Unstructured question

III. Question

I. Structured questions are those which pose definite, concrete and preordained questions. They are prepared in advance and not constructed on the spot during the questioning period. The form of particular questions may require responses which are either close or open. The close end questions are used when categorized data are required. Respondents are asked to answer the questions within the choice answers given against each question. For example -

Q. How do you spend your leisure time ?

(a) engaged in extra earning ( )

(b) assisting parents ( )

(c) playing and gossiping ( )

(d) enjoying movie ( )

(e) watching television ( )

(f) any other (specify) ( )

The open-end questions are free and informations are opened to express opinion. Informants are not kept restricted to the definite limits of responses as it is done in close-end questions. For example –

Q. What is your opinion about the functioning of present government in your state ?

In this type of question respondents can express his opinion without any limitation of alternative responses. Open-end responses are used chiefly for intensive studies of a limited number of cases or for preliminary exploration of new problems and situations. At times the respondent is even asked to write a detail description and express his viewpoints, describe his relationship, attitudes, indicate his problems and reports on detail and events, without restrictions imposed as in the case of closed questions.

Activity 3.A (1)

Prepare 2 structured questions on a research topic titled "Educations and Occupational Change in a Tribal Village".

Q. 1 ……. (a)

(b)

(c)

(d)

(e) any other

Q. 2. …….. (a)

(b)

(c)

(d)

(e) any other

Use of structured questions

Structured questionnaires are used in a wide range of projects, sometimes to initiate formal inquiry and sometimes to supplement and check data. They may pertain to studies of economic or social problem, measurement of opinions on public issues or events, studies of administrative policies and changes, studies on the cost of living, consumer expenditure, child welfare, public health and numerous other issues.

II. Unstructured questionnaire

This type of questionnaire is also referred to as Interview guide. It aims at precision and contain definite subject-matter area ; the coverage of which is required during the interview. It is designed to obtain viewpoints, opinions, attitudes and to show relationship and interconnections between data. Such data may be noticed under more mechanical types of interrogation. The object is to give the respondent maximum opportunity to reveal how arrived or developed his world of experience. No check lists with predetermined responses

are provided. Rather free responses are solicited. For example question like –

Q. How would you describe the activities in which your group engage ?

This form of questioning assumes insight, articulateness and possession of facts by the respondents. It is used for intensive studies, but generally, for a limited number of selected cases.

Activity 3. A(2)

Prepare 2 unstructured questions on a research topic "Political Participation among the college students of Assam

Q. 1 ……

……

Q. 2 ……

……

III. Preference type questions

Social researchers quite often apply a type of question called preference question. Respondents are asked to choose their responses from a lists of close answers on the basis of their preferences. In certain social situations respondents have to express their opinion on more than one responses. Under such conditions it becomes difficult to identify the crucial factor that determines or shapes the attitude of the respondents. At this stage, respondents are asked to choose their responses in order of preferences. The question below may illustrate such example –

Q. On what considerations you would like to vote in election ?

(in order of preference)

(a) caste …….. ( )

(b) religion ……. ( )

(c) locality ……..……. ( )

(d) party ideology ……. ( )

(e) merit of the candidate ……. ( )

(f) education ……..…… ( )

There are various reasons for making answers on preferences. People take into consideration various factors while forming opinion. When several factors operate simultaneously, it is difficult to determine which single factor plays the important role in making decisions. In such situations responses are arranged in order of preference.

Activity 3.A (3)

Prepare two preference types questions on any research topic

Q. 1 …….. (a)

(b)

(c)

(d)

(e)

Q. 2 ……. (a)

(b)

(c)

(d)

(e)

(e) Pilot studies and pretest

Mere construction of questions is not sufficient. The questions so constructed needs *empirical checking.* In this stage the researcher should keep in mind certain important points. Firstly, the formulation of items must be supported by adequate literature. Secondly, items included in the schedule should be capable of detecting sources of

information. To verify the first point a procedure is maintained which is usually called pilot study through which answer is sought. And pretest is done to make the schedule final. Pretest is a sort of preliminary application of the scheduled of few cases.

Key Terms to remember

Questionnaire : a series of questions presented to the selected respondents usually sent by mail.

Schedule : is a set of questions which are asked and filled in by an interviewer in a face-to-face situation.

Interview Guide : is a list of points or topics which an interviewer must cover during the interview.

Open-end question : a question when the responses to the items cannot be anticipated in detail.

Close-end question : a question when the responses can be classed together and anticipated earlier.

Pilot study : a type of survey adopted in formulating items in the schedule specially when literature is inadequate.

Pretest : selection of items in the final schedule.

3.4 INTERVIEW

Like other tools, interview is also an important tool of data collection in social research. Simply stated, it is a form of communication and exchange of ideas between two persons. But in terms of a research tool interview illicits information. Interview is a conversation which develops precision, focus, reliability and validity of social acts. P.V. Young asserted that interview is "not simply a two-way conversation between two persons. It exhibits gestures, glances, facial expression, pause or even reveal subtle feelings."

3.4.1 Purpose of Interview

There are many types of interview and their purposes, are also many. Interview develops skills, people can learn to improve his techniques by learning to avoid certain types of error.

Modern social investigators develop their skills, it is the very increasing interest of the social scientists that they have turn *from books to social phenomena* in an effort to build the science. The personal interview has been variously defined. It may be seen as an effective informal, verbal and non-verbal conversation, initiated for specific purposes and focused on certain planned content areas. The objectives of interview may be exchange of ideas and experiences, eliciting information pertaining to a wide range of data in which the interviewee wish to rehearse his past, define his present and canvass his future possibilities (Goode and Hatt 1945).

Check Your Progress

1. Mention two points showing interview as technique of data collection.

(a) _____

(b) _____

3.4.2 Major objectives of research interview

Societies are of varying levels of development. Whether it is a simple tribal, peasant, urban, high level urban-industrial nature, societies are adorned with various qualities and characteristics. Modern societies are secondary societies with varied experiences. Attitudes and values are quite divergent. With the growth of urban life face-to-face contacts are decreasing, primary groups are receding and social distance widens. In such circumstances, it becomes the task of interview *to penetrate into the inner life of persons.* It becomes essential to as certain opinion, attitudes and values that are on the surface which social scientists believe to be the role of interviewer.

According to anthropological viewpoints, the aim of interview is two-fold –

First, a description of the situation as the interviewer sees it, looking from *outside in* and

Second, description of the situation as he sees it, looking from the *inside out.*

The first comprises of the visible world of object and actions, and the second the subjective frame of reference embraces the world view of the people. It is the social interviewer who can act with both the perspectives.

In fact, there is no empirical study in which interview did not play during some aspects of investigation. Thomas and Zneniecki's study of *personal adjustment of immigrants, personality* study by Adorna and his associates and Stouffer's study of the **American Soldiers** are the finest examples of sociological research conducted with the help of interview.

3.4.3 Types of Interview

Researcher P.V. Young outlined few types of interview such as –

(a) No-directive Interview

(b) Directive Interview

(c) Focussed Interview

(d) Repeated Interview, and

(e) Depth Interview

Non-*directive interviews* are basically designed as uncontrolled or unguided or unstructured. Interviewers ordinarily do not follow a definite system of lists of predetermined questions. It is a free-flowing interview.

Directive interviews on the otherhand, are highly based on standardized technique and a set of predetermined questions. It is generally useful for administrative and market research of various types.

Focused interview as the term implies to have a detail reactions, feelings, emotion one a particular situation, event or occurrences. The interview takes place with the persons who are directly involved to such situations. The interviewee (respondent) is given complete freedom to express his opinion. The focused interview is based on the assumptions that through it, it is possible to secure precise detail of personal reactions, specific emotions called into play, definite mental associations provoked by a certain stimulus and the like.

The repeated interview aims at the study of continuous development of a social or psychological process. The progressive actions, factors, attitudes etc. may determine the development of certain behaviour pattern or social situation. Use of repeated interview technique is done to study the attitude of the voters when elections campaign progresses.

The Depth interview is designed to encourage free expression of respondents. It is a lengthy process of interviewing respondents aiming at detecting unconscious and other materials dealing primarily to personality dynamics and motivations. To carry out successfully, this type of interview needs special, skill and training on the part of the researcher.

3.4.4 Rapport Establishment

The term "rapport" means a relationship between the researcher and the informants specially during data collection. In terms of research, rapport is a state of relations exist between the researcher and the respondents. It is not simply a relation, its deeper meaning goes to the extent that the respondent must realize the goal and purpose of the investigator, and actively seek to help the researcher in obtaining necessary information. So, mere friendliness between the researcher and the respondent does not carry the meaning of rapport.

It is equally important to note that rapport establishment is not an easy process. There may exist enumerable differences between the research and his field of study and persons. Language, culture, general acceptance, behaviour etc. may sharply differ between them.

The interviewer has to approach his informants with fear, anxiety, doubt and suspicion at the beginning. He needs time to overcome all such hindrances and. as such to get convince of his new situation. But he must approach with confidence that the informants, though new to him, will gradually come to his fold. Confidence is derived from an assurance on the part of the researcher that the respondents will find pleasure in associating him.

A researcher must be frank enough to disclose his purpose of research. He must introduce himself with greet and smiles and with professional maintenance.

Some respondents may feel dissociation at the beginning since the purpose of research may be unknown to them. In this stage some types of conversations is often useful in establishing friendly relation with the respondents. Researcher must assure that under no circumstance, the name of the respondents will be made public at any place or at any time. He has to make statement that the information supplied by the respondents are very useful for the purpose of research. This will enhance confidence on the part of the respondents. This constitutes an important part of ethics is social research which we discussed earlier.

Thus, the researcher may carry out his interview forward, firstly by small, simple and friendly question and gradually with difficult and probing questions.

Check Your Progress

1. How do you access that establishment of rapport helps in conducting good research ? Mention three points.

(a) _____

(b) _____

(c) _____

3.4.5 Recording the Interview

The content of the Interview must be recorded adequately and with sufficient care. If the researcher during interview realizes that informants are trying to hide information, a care must be taken in recording. As far as possible the exact words used by the respondents should be recorded. They should not be edited in grammar or meaning. The researcher should develop skill in writing quickly and legibly, with some attention to symbols for some common short phrases. In order to make perfect recording of interview, researcher immediately after completing interview should have the practice to go the his desk or table for full recording while the materials are still fresh in mind.

3.4.6 Closing of Interview

The increased popularity of research techniques in social sciences tell us many new things. Once a rapport is established between the researcher and the field situation, there happens to be an enduring relationships. So, quite often most of the researchers while noting a paragraph on *difficulties and experiences* in fieldwork expressed that fieldwork gives many new insights on the issues of hitherto unexplored areas of society. Knowing things of real situation of society is never stopping process and therefore, interview with the respondents cannot be cut short if the researcher is fully trained with social science methodology. This was what observed by Goode and Hatt saying as "*closing of interview is more difficult than the opening"*.

To close an interview however, the researcher should not forget to tell 'thanks' to his respondents. It is obviously a courtesy to say 'goodbye' to his respondents. The good relations which each interviwer develops will, directly or indirectly, help both to his own later research and that of others. This should be thankfully acknowledged by the researchers to his respondents.

Key Terms

Interview : a series of questions administered personally by a researcher to respondents.

Outside in : description of the situation as the interviewer sees it – looking from outside.

Inside out : description of situation as the interviewer sees it from inside.

Rapport : is a relationship between the researcher and the informant specially during data collection.

3.5 ANALYSIS, INTERPRETATION AND REPORT WRITING

After the completion of data collection the researcher has to go for certain very important steps in research. They are –

(a) systematic and careful planning of data

(b) classification and ordering of data

(c) tabulation

(d) use of statistics, and finally

(e) writing the report.

A social scientist assumes that behind his accumulated data there is something more important and revealing than the facts and figures themselves. Therefore, he has to consider that how the collected data are systematized and planned so as to reach the desired goal for analysis. P.V. Young stated that carefully thought out, well marshaled facts and figures when related to whole body of data have significant meaning from which valid generalization can be drawn. Since analysis is a continuous process throughout the research undertaking, some basic considerations at the beginning must be taken into consideration, such as – which data are to be secured, what techniques are to be used in securing them, what sources to tap, what hypothesis to formulate and test them and how to test them, what necessary classifications are to make etc. For all these considerations

it is necessary to make a perfect plan of processing and analysis of data.

For a careful and systematic planning of data sometimes researcher needs the assistance of computer. It is now desirable that the researcher himself should be involved in processing his data according to his desired aims of analysis.

3.5.1 Data processing

Data processing refers to certain operational work such as – editing, coding, computing of scores, preparation of master charts etc. After collecting the filled in questionnaire it is necessary to check the entry whether all questions are filled up completely or not. Sometimes researcher may miss some entries by mistake or some entries may be illegible. So, the discrepancies remained in the schedule if any, should be meticulously edited. An incomplete interview schedule may create problems at the time of tabulation of data. Many a times, some categories of responses such as – "not applicable "any other" etc. need specific care in editing.

Coding of data involves assigning some symbols, usually numerical to each response of the question. For example, when a response "strongly agree" is coded as 5, the subsequent codes would be in order. If the subsequent responses are 'agree', 'undecided', 'disagree' and 'strongly disagree', then may have to be coded as 4,3,2 and 1.

Example

Question : Do you think that industrialization has an evil effect on family system ?

	Code
Strongly agree …….	5
agree ……..	4
Undecided …….	3
disagree …….	2
strongly ……	1

Preparation of Master Chart

After a code book in prepared, this can be transferred either to a master chart or directly to computer through a statistical package. A muster chart contains all codes of the questions of the schedule. The master is a plan where all responses collected from the respondents are charted together. The preparation of master chart helps in easy tabulation of data.

3.5.2. Classification and ordering of Data

Classification is an essential step before making any ordering of data. *Classification* is a method of grouping together the data on the basis of similar characteristics. For example, researcher may collect information regarding educational qualifications of respondents which vary from the category of 'illiterate' to any higher education such as illiterate, primary, read upto class VIII standard read, upto class X standard, matriculate, graduate, post-graduate etc. Thus, education levels of populations are numerous and these are to be ordered on certain categories or classes Researcher, according to his need of analysis, such levels of education may be made into certain specific categories like 'illiterate' 'low education', 'middle education', high education etc. Thus is called, summerisation' or grouping' information. One basic aim of classification is to reduce innumerable variety into certain categories of types. The system of making classification exists in all branches of science.

So, before making tabulation or arranging data in a systematic and logical order, it is necessary to work out a detail system of classification. Classification is fundamental to any kind of scientific analysis. A general scheme of classification is determined before the data are made into order.

3.5.3 Criteria of classification

There are certain bases of statistical classification. However, such bases are sometimes determined by the nature of data or sometimes

by the problem at hand. But some criteria are invariably used in statistical classification. These are –

i. geographical area-wise e.g.. state, region, city, country, ward or census tract.
ii. chronological i.e. on the basis of time e.g. days, week, month or year.
iii. qualitative i.e. according to some attributes e.g. sex, colour, nativity, occupation, marital status etc.
iv. Quantative i.e. in terms of magnitudes of occurrence.

Geographical classification

Data are classified in terms of geographical or vocational differences between various items. See example below.

State – wise estimates of production of foodgrains

Name of States	Total Foodgrains (in thousand tons)
Andhrapadesh	1093.7
Bihar	12899.0
Haryana	11934.7
Punjab	21149.9
Uttar Pradesh	41828.0
All India	1,92,433.6

Chronological classification

When data are observed over a period of time, such type of classification is known as chornological classification. We may present the figures of population as follows

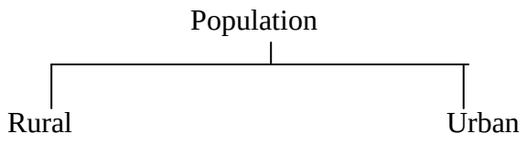
Sex-ratio of India, 1901-1951

Census Year	Sex ratio
1901	972
1911	964

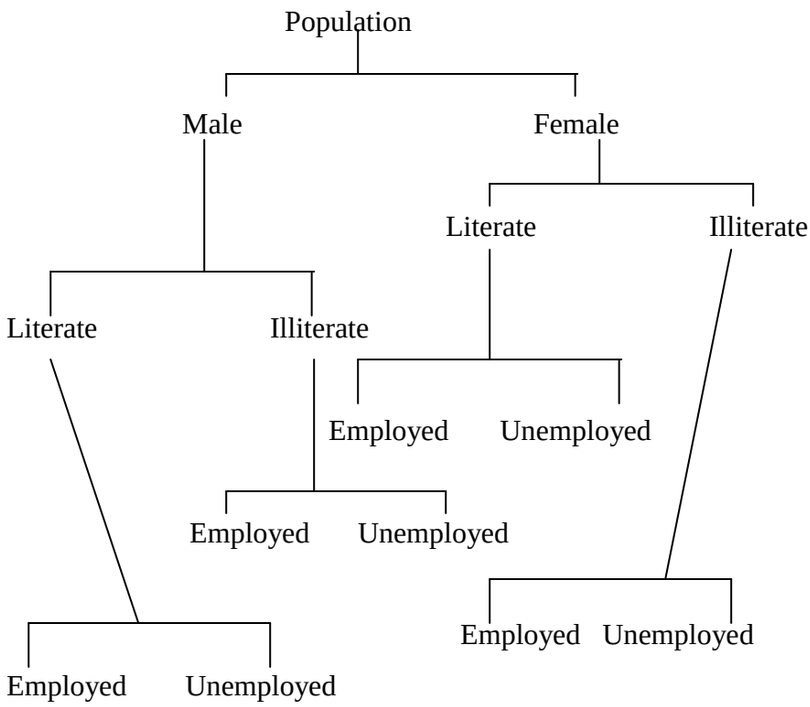
1921	955
1931	950
1941	945
1951	946

Qualitative classification

Data can be classified on the basis of qualities or attributes such as-sex, colour, religion etc. For example, the population under study may be classed as under,



Further, classification can be made on the basis of some other attributes as shown below –



Quantitative Classification

Quantitative classification refers to the classification of data according to some characteristics such as – height, weight, income etc. For example, the students of a college may be classified according to their weight as follows –

Weight of Students (in kg)	Number of Students
50-55	50
55-60	200
60-65	260
65-70	360
70-75	90
Total	**70**

Thus, classification is considered essential step before making any sort of tabulation of data. In this context, it is important to know the structural parts of a statistical table.

But before going to next items for discussion, now let us check your progress. Do you want to say that classification is a scientific progress ? If you think so, mention 2/3 points in support of your answer in the space provided below.

1. _____
2. _____
3. _____

3.5.4 Tabulation

There are some vital parts of a statistical table. These are commonly discussed as follows – table number, title of the table caption, stub, body, footnote, sources etc. The format below shows an example of such format.

Table No.

Title — **School Enrolment Rate by age, Title Selected years, 1910-1957**

Year	**Age (in Years**				
	5-6	**7-13**	**14-17**	**18-19**	**20**
1910					
1920					
1930					
1940					
1950					
1960					

Caption — **Stub** — **Body**

Source : Footnote

Table number : Each table should be numbered. The practice of numbering may vary. The number of the table is usually given at the centre of the title or at the left above the table.

Title of the table : Every table must be given a very suitable title. The title should be very precise, short, clear and self-explanatory. A title should be very clearly worded.

Caption : Caption refers to the column heading. It explains what the coloumn represents. It may consist of one or more column heading. Under a column heading there may be sub-heads.

Stub : Stub is the designations of rows or row heading. They are at the extreme left of the table.

Body : The body of the table contains numerical information. This is the most vital part of statistical table.

Footnote : Anything in a table which the reader may find difficult to understand from the title, caption or stub should be explained in the footnote given below the table. Providing footnote is specially most essential for the tables developed from secondary sources of data.

Sources : In tabulating secondary data collected from the records of Governmental, non governmental, department or offices, the sources of those data must be indicated clearly at the bottom of the table.

It is therefore, noted that a through and clear understanding of table construction is necessary. There are certain advantages of presenting statistical data through tabular form. These are –

- statistical tables conserve space and reduce explanatory and descriptive statements.
- table facilitates comparison of data.
- tabular arrangement facilitates summarizing of items and detection of errors and missions
- statistical tables provides basis for computations.

Tabulation of Data :

One of the most simplest and revealing devices for summarizing data is the presentation of data in meaningful fashion which is termed *statistical table.* A table is a systematic arrangement of statistical data in columns and rows. Rows are horizontal arrangement whereas columns are vertical ones. The basic purpose of tabulation is to simplify the presentation and to facilitate analysis and comparisons.

The first step of tabulation is to plan a detail classification of data as we have discussed earlier. After the classification has been determined, various types of tables are designed starting from simple to complex and more complex tables. Initially, sketched or designs of the tables are constructed. Statistically such designs are called *dummy tables,* Dummy tables for all questions in the questionnaire are to be checked by the statistical experts to examine whether the arrangements are efficient for logical interpretation of data.

Social researchers and data analysts have adopted variety of statistical tables in analysing data in their researches. Some types of tables are as follows –

(i) Simple or one way table

(ii) Two-way table

(iii) Cross tables between variables

(iv) Preference table.

Simple or one way table

In this type of table only one item is shown and presented in a tabular form. The following is the illustration of such table.

Number of employees in Factory according to Age-group

Age (in years)	No. of Employees
Below 25	----------
26-35	----------
36-45	----------
46-55	----------
Above 56	----------
Total	

Two-way table

Two-way table shows two characteristics or variables. Two way tables are designed when either the stub or the caption is divided into two coordinate parts. Below is an example of such table.

Number of employees in Factory according to Age and Sex

Age (in years)	Employee		Total
	Males	Females	
Below 25			
26-35			
37-45			
46-55			
Above 56			
Total			

Cross tables between variables

Social researchers quite often try to examine the association between two or more variables. This is done when relationships are supposed to be existed between social situations, events or conditions. Thus, social scientist examines logical connections between two occurrences. This is how cause - effect relationships are determined.

A *variable* may be an object or an event which can vary in successive observations either in quantity or in quality.

Thus, we may classify variable into two types – *quantitative and qualitative.* A quantitative variable is one which may take on various magnitudes, i.e. may exist in greater or smaller amounts. Examples of quantitative variables are – age, height, income, size of population, size of family, length of service, birth rate and numerous other characteristics. These variables can be measured and each resulting magnitudes is called **variate**.

A qualitative variable may also vary in successive observations, not in magnitudes, but rather in quality or kind. Such qualities are customarily called **attributes**. Thus, sex will vary according to the attributes of male and female, a person's marital status may be single, married, widow or divorced. Similarly, nationality may be American, French or Italian. The attributes of a given qualitative variable cannot be scaled, or arranged in order of magnitudes. Because of such limitations, statistical methods would have little to contribute to the analysis of qualitative data. But there is a large and growing body of statistical techniques for the treatment and manipulation of qualitative variables. (For detail see "Statistical Reasonings in Sociology" by Mueller and Schuessler.)

Cross tables are prepared making arrangements for two or more variables. Below is an example of such table.

Example – 1

Table No 1

Sex and preference for Professional wages

Sex	**Professional courses**			**Total**
	Medical	**Engineering**	**Law**	
Male				
Female				
Total				

Example – 2

Table No

Occupations and Social Status

Occupations	Social Status			Total
	Upper	Middle	Lower	
Medical Engineering Legal Business Teacher				
Total				

Thus, we can observe from the above two examples about the cross tables and its method of preparation.

Preference Table

This type of tables are made when data are collected in order of preference. When people's opinion are sought from among variety of factors, one single factor cannot be taken for granted for final opinion. In such context respondents are asked to put their opinion on

the structured set of alterative responses on the basis of their preference like – 1st choice, 2nd choice, 3rd choice etc. The opinions thus collected can be arranged in form of specially designed table which lateron made into weightage form and found out the ranks of preferences on the basis of the total scores. A design of such format of table is presented below –

Example

Table No

Consideration for casting votes in election

Bases for Consideration	**Preferences**					**Total Scores**	**Rank**
	1st	**2nd**	**3rd**	**4th**	**5th**		
Caste							
Locality							
Language							
Religion							
Education							
Efficiency							
Ideology							

3.5.5. Analysis and Interpretation

In social science research one can adopt variety of skills to analyse and interprete data. This is to say that in contemporary social science research, there are an innumerable development of methods, techniques and procedures of data from a number of individuals and these are called case. The data themselves consist of attributes, behaviours, attitudes and beliefs. These become our variable.

The analysis of data involves description, explanation, interpretation and comparison of social events or phenomena. These ultimately lead to arrive at certain conclusions.

Social researchers then, commonly employ certain procedures to analyse the collected data. Step-wise such procedures can be understood as mentioned below –

a) Calculating percentages for each cell of the table to ascertain the distributions of cases under study. This is done in the analysis of simple frequency table or one-way table.

b) Calculating both row and column percentages of two-way tables or cross tables. This is done to examine the two variables to examine the relationship exist between the variables.

c) To present a simple statistical tests to demonstrate the probability of findings.

Usually, analysis is done in percentages. In simple tabular forms, percentages are calculated only on the distributions of the data. Hence, a critical analysis is not possible in those tables. In case of two-way or cross tabulation, analysis can be done more vigorously, since percentages in such tables can be calculated from both sides, i.e. rows and columns thus making possible for critical analysis. On the basis of cross tabulation, a varied interpretation can be made on the trend of data.

The research report is not usually presented to lay audience. For a social researcher, the audience is the person who are professionally equipped to go through such reports. So, he needs to clarify all concepts and framework of his research which must be aimed at his targeted audience. This includes the maintenance of professional or ethical issues of research report. This covers the parameters of report such as- use of language, technical terms, style of writing etc.

The outline

There are many forms that the outline of report might take. There may be sponsored research projects for which clear outline is predetermined. Again, there are researches carried out for individual

research interest. These are professional or academic researches leading to the achievement of academic degrees associated with institutions. The report writer should be familiar to such requirements. However, in the contemporary social science researches of academic interest the guidelines which are quite often followed are indicated below –

Report writing

The writing of the report is the final stage of research. Its purpose is to convey to the interested person about the result of the study. The presentation of result should be in sufficient detail and should be so arranged that the data be comprehended to the interested readers and the conclusions valid for the purpose of academic interest.

How to go ahead?

The topic under study is placed in its proper context. Explain the importance of topic, theoretical importance, practical utility etc.

Outline of the study

Write down the outlines of study, define the terms, state the hypotheses. Then write down the rough draft of the report. Revise the draft and put everything in proper format. It's is called polishing. It errors are corrected, then the mater is given for typing.

Report writer has to take into account some important aspects before writing the report.

These are –

Who is the audience?

How to maintain professional or ethical problems?

What language is to be used

How to present the report? etc.

Introduction

Research report begins with an introduction. It gives a clear statement of the problem under study. The author tells about the theoretical importance and practical significance of study. There should be a clear indication about the aims and objectives. In this section the author should also define the terms used and meaningful background of study. The introductory chapter also contains, the major hypothesis to be tested in the study. A series of research questions relating to the study be placed at proper place in the introduction.

The review of related literature concerning to the study is major part of introduction chapter. The researcher should follow a pattern of review of literature that are accepted in contemporary social science research.

Methodology

Description about methodological steps followed in research is an integral part of social science research. There are various styles of presenting "methodology part" in the research report. Sometimes, it is clubbed with the "Introductory Chapter" and sometimes it is placed as a separate chapter. Whatever the arrangement may be, a good description on methodological steps followed in the study, is essential for a good research report.

Researcher should clearly mention the reasons for selections of universe for study, methods of sampling followed, size of samples, techniques of data collection (observation, questionnaire, case study etc.), procedures of tabulation and the process of data analysis. besides, the contemporary reports writers quite fashionably mention about the time and duration of field work, difficulties and experiences gathered, techniques for establishing report in field situation etc. One of the vital components of methodology chapter is the introduction about the use of statistics in the report. Usually, the methodology chapter tells us about the chapter scheme of the entire study.

Findings and Discussion

Report writer should ensure that the findings of study are systematically and lucidly presented in the report. In such presentation researcher is aware of the fact that some chapters in the report are definitely the "Core Chapters" since the studied topic/problem and its relevant empirical data are discussed in those chapters. Further, the hypotheses developed for the study are also discussed and tested in those "core-chapter".

The data tables should be suitably placed in the chapters. There should not be repetition or overlapping of statements. Each level of analysis must be guided by the theoretical. Outlines introduces at the introductory chapter. If there are any negative or non-conforming findings, these should not be discarded or concealed. Such negative findings may suggest some important clues for future research and these must be highlighted.

Findings should be suitably compared with other similar studies conducted at various time and place. This entails a comparison of data and thereby a researcher can achieve a character of universality of his research.

Summary and conclusions

Towards the end of the report, major findings of the study are summarized. The summarises are drawn in accordance with the theoretical premises made at the beginning. The main points and the major trends of data are brought together in this section. New data are not brought in this section. A summary is not an abstract. It may refer to the tables and figures incorporated in the text of the report.

Finally, an attempt is made in the conclusions which are emerged out of the findings and discussions. Conclusions derived must be guided by theoretical context or in the light of hypothetical statements planned earlier. Thus, concluding remarks may call for two aspects firstly – a researcher may conclude his report with the

results affirmative to his problems or secondly, negative or contradictory evidences may also cause reformulation of the theories. Here lies the typical character of research study. Such typical characterstics should be well guarded which may contribute to new kind of investigation for future research.

Acknowledgement

This is a place to acknowledge thanks to all who have directly or indirectly helped the researcher to smoothly carry out his work. Several persons besides the respondents, organistions, institutions, authors are to be included while offering thanks. The researcher can thank to the sponsors who have assisted his work financially. Thanks are due to research supervisor, friends, colleagues who inspired the researcher at various phases of his research work.

Footnotes

There are several ways of using footnotes. Conventionally, footnotes are used with symbols or numbers at the bottom at the relevant page. Since methodology in social science research is constantly changing, most researchers like to give footnotes at the end of each chapter, or even at the end o the report, citing author's name in the text. For example-Srivastava 1985; Boruah 1962 etc.

Bibliography

This is the list of all documents contributed in the study. Stirctly speaking, a bibliographer is a list of published works, although by common usages both published and unpublished materials are listed in a bibliography. The accepted method of preparing bibliography is show below"

Examples for books

Lal Das, D.K. (2000): Practice of Social Research, Social Work Prespective, Rawat Publications, Jaipur.

Goode, W.J. ; P.K. Hatt (1952) : Methods in Social Research, New York, Mc. Graw Hill.

Sethi J.D. (1991) : A General Theory of Voluntary Action in a Decentralised Policy, Man and Development, Vol. 13, No. 20, June, p.p. 41-43.

Appendix

The report generally contains the statements or original documents on the bases of which interpretations are made. These are generally placed at the end of the report in form of appendix. The questionnaire or schedule or other forms and statements are included in the appendix. If help has been taken from any report, written documents or any other paper, it is good to give its relevant portions in the appendix. In short, all those facts that need a special elaboration, but cannot be given in the main report for the fear of killing the lucidity and sequence should be given in the appendix.

Style and Language

A research report is essentially a scientific document and hence must be clear, accurate and precise. Confusion, ambiguity and duplicity must be avoided while explaining facts, concepts etc. The very sensitive aspect of a report is its language. The emphasis should be given on clearity, correct exposition, expression and simplicity. Typical terminology or jargons should be avoided.

Scientists are not supposed to take sides. They are expected to write *impartially and impersonally.* They should avoid using ornamental language. The should write clearly and precisely in matter of fact way. Not that they have to use telegraphic language. They must proceed systematically. The report should be free from subjective basis. The results should be objectively presented in a logical and systematic manner. Repetition of facts and statements must be avoided.

Words, Sentences and Paragraphs

We have to use proper words at the light place and at the right time. We should not use any word without understanding its full meaning, including its connotations and denotations. Do not use difficult words where a simple one will do. Author should take care in removing spelling mistakes and grammatical errors.

Sentences should be logically framed and so, these should be a consistent relations from one sentence to the next. Similarly, such logical flow should be maintained from one paragraph to the next.

Each paragraph contains one central idea. The discussion in one paragraph should be kept under control. The proceeding sentences should smoothly flow into the succeeding ones. Jumping here and there, from one idea to another creates problems and confusions. Long paragraph are broken up into shorter ones.

Different paragraphs are combined together with chains of reasoning, into sub-sections and sections. The use of sub-headings in this context is a mechanical aid. Sub-headings are created in order to call the attention of the reader to the materials being presented. The use of clear headings and sub-headings will keep the writer within the organization of the report. With sub-headings, rewriting is considerably easier and to delete sections or to combine sections without destroying the basic plan of the report.

Key terms

Classification : is the grouping of related facts into class

Qualitative Classification: is a classification on the basis of some attributes or qualities such as – age, sex, colour, religion etc.

Quantitative Classification: It refers to the classification of data according to some characterstics that can be measured such as-height, weight, income, profit, production etc.

Editing: to scrutinize the data, with care and attention and to check possible errors and irregularities.

Coding : Coding of data is a method of assigning symbols, usually numerical to each response of the items o the schedule.

Questions to be answered (15 marks each)

1. "Observation is carried out with purpose". Bring out the significance of the statement.
2. What are the various types of observation? Discuss.
3. Discuss the advantages of participant observation in social research.
4. How do you establish the relationship between hypothesis and observation? Discuss.
5. What are questionnaire and schedule.
6. In there any differences between questionnaire and schedule? Discuss.
7. Discuss how a good questionnaire is prepared.
8. What are structured and unstructured Interview schedules? Discuss with examples.
9. Prepare a model of Interview schedule on a research topic "students" participation in politics: A study of college students Assam:.1
10. Explain interview as a tool of data collection in social research.
11. Explain " Interview as a social process".
12. Discuss how report is established in social research.
13. Highlight some techniques of recording interview.
14. Discuss how a research maintains impression in the field situation.
15. "Social interview can delineate past, present and future" Explain.

Answer the following questions:

16. Explain why classification is a necessary pre-step for tabulation of data.
17. Discuss with sketch the various parts of a statistical table.
18. What are the merits of cross-tabulation in analysing data? Explain.
19. Discuss the essential qualities of a good research report.
20. Throw light on the latest trend of bibliography and references in social science research.

Answer briefly the following questions : (3 marks each)

1. What is classification of data?
2. What are the criteria of classification of data in social research?
3. What is dummy table? Why dummy tables are essential?
4. Define with examples
   (a) One-way table
   (b) Two-way table
   (c) Preference table
5. What is footnote?
6. What is bibliography?

Suggested Readings

1. W.J. Goode and P.K. Hatt: Methods in social Research, (chapter 10, 11, 12 & 13) Macgraws Hills.
2. P.V. Young: Scientific Social Survey and Research. (Part II chapter 7, 8, 9 & 10), Prentice Hall 1966.

****

BLOCK – 4

STATISTICAL APPLICATIONS IN SOCIAL RESEARCH

STRUCTURE

4.0 OBJECTIVES

In this Block, an attempt is made to understand the need of statistics in social sciences. The impact of computer in present day social life cannot be denied. A discussion is also added in this chapter with regard to the application of computer device as to how statistical data are computerized for easy interpretation of facts and phenomena.

4.1 INTRODUCTION

Statistics is regarded as a basic science in man's life. There are reasons as to why a student goes through statistics. The application of statistics in all spheres of human life is so vast and ever-expanding. And so, interest in the application of statistical methods to all types of problems has grown rapidly.

4.2 NEEDS AND VALIDITY IN STATISTICS

4.2.1 Meaning and Definition

The term 'STATISTICS' is used in several ways. It denotes the collection of data such as – those found in Labour Gazette or say Labour Statistics of the Labour Bureau published annually by the Government of India. Secondly, the term 'Statistics' refers to the statistical principles and methods employed in the collection, processing, analyzing and interpretation of any kind of data. In this sense, it is a branch of applied mathematics and helps us to know the complex social phenomena better and provide precise ideas.

Note: Understand the meaning of the term 'Statistics'

Definitions of Statistics

There are indeed many definitions of statistics. Webster defined statistics as- "The classified facts representing the conditions of people in a state ……. Specially those facts which can be stated in numbers or in tables or in any tabular or classified arrangement".

For Yule and Kendall “Statistics means quantitative data affected to a marked extent by multiplicity of causes”

Horace Secrist defined statistics as -

“An aggregate of facts affected to marked extent by multiplicity of causes, numerically expressed, enumerated or estimated according to reasonable standard of accuracy, collected in a systematic manner for a pre-determined purpose and placed in relation to each other.”

These definitions have provided scopes to learn some basic characteristics of statistics. These are –

i) Statistics are aggregate of facts.

ii) Statistics are affected to a marked extent by multiplicity of causes.

iii) Statistics are numerically expressed.

iv) Statistics are enumerated of estimated according to reasonable standard of accuracy.

v) Statistics are collected in a systematic manner.

vi) Statistics are collected for pre-determined purposes.

vii) Statistics should be placed in relation to each other.

Check Your Progress

1. How do you define the meaning of Statistics ? Answer in the following space.

_____

_____

_____

2. What specific characteristics of Statistics your observe ?

a) _____

b) _____

c) _____

4.2.2 Statistics in Daily life

When we call upon to treat large aggregate, it is not easily identifiable. In the aggregate, individualization becomes mentally impossible. Consequently, we can view them only in collectively in terms of one or more common characteristics. The individual will lose their identity in the total anonymous mass. So, far purpose of description or prediction, various devices of summarization have been evolved, or a limited number of individuals (or samples) are selected from the aggregate to represent the total, about which generalization may be made with varying degrees of precision. Such procedures have been in use for centuries, but during the last 200 years, these procedures have become formalized and have come to be know as “Statistics”.

This evolution of statistics is not arbitrary, but rather a matter of gradual compulsive social adaptation. This has become possible with the expanding social organizations which have become increasingly immense and complex. Statistics, therefore, is an instrument for viewing the mass of happening around us. To be sure, in our private life we are concerned with personal opinions, marriage, death and many other individual transactions. But in our social collective relations, we are concerned with public opinion polls, population and vital statistics on the basis of which society pursues its collective interest.

Statistics have become a personal and social necessity, reflecting our need for understanding the past and anticipating the future.

4.2.3 Need of Statistics

We find today that there is hardly a phase of human activity which does not come across with statistics. Statistics convey a variety of meaning to the people. For some, statistics is an imposing form of mathematics, while to others, it suggests tables, chart and figures,

which are commonly, find in newspaper, journals, books, various reports, speeches, classroom lectures, T.V etc. Each day we are exposed to wide assortment of numerical information which often has a profound impact in our lives. For example, we come across statement like there are 932 females for 1000 males in India, whereas in Russia there are 1170 females per 1000 males. The density of population has gone up from 216 to 267 persons in 1991 Mumbai is the most populous city in India with urban population 12.57 million, Calcutta taking the second place with 10.92 million, Delhi ranks third with 8.38 million followed by Chennai with 5.36 million. The numbers of universities have witnessed a considerable increase from 110 to 146 between 1981 and 1991.

The above are the examples of numbers placed in certain time or situation sequences. Therefore, numbers play an essential role in statistics. The numbers provide the raw materials of statistics. These materials must be processed to be useful. The study of statistics involves methods of refining numerical or non-numerical information into useful forms. The need of statistics is ever-growing and it is due to the functions it performs. Some of the most important functions of statistics are –

1. Statistics make complex mass of data simple and understandable with help of statistics methods.
2. To interpret various characteristics of data, classification is done by the application of improved techniques of statistics.
3. To study relationship between two or more phenomena, statistical methods are useful.
4. Statistics provide various information for the analysis of data that enhances individual experiences.
5. Statistics helps in formulating policies in different fields of life.
6. Statistical methods are highly useful tools in analyzing past data and predicting future trends.

7. Statistics help deriving correct and dependable conclusions.
8. In most governmental and non-governmental plan formulation, statistics helps in decision making.
9. One important function of statistics is to provide techniques for making comparison.

Civilised man is fully accomplished with statistics. Many prevailing quantitative life of modern people are engulfed with statistics. Public opinion polls, cost of living index, population data and vital statistics, games of chance and many other evidences are the view of life. Side by side, the common man in his unsophisticated way of life employs crude statistical concepts when he speaks of average, hunches and hypothesis, probabilities, chance, run, samples etc.

Check Your Progress and answer

1. 'Statistics help in deriving correct and valid conclusions' Explain

_____

_____

_____

Thus, statistics involves skill in thinking and therefore, not a redundant academic accomplishment, but is an integral element in the current thought ways of our civilization.

Now, it can be easily imagined that statistics has very important role in man's life. Some of the salient points emerged out of above discussion can be summed up as follow –

1. Statistics provides various devices for summarization.
2. We can view the happenings around us through statistics.

3. Statistics, for man, have become personal and social necessity.

Now you should remember these important points –

(a) Statistics are aggregate of facts

(b) It summarises the complex data for precise understanding

(c) Statistics are useful in analyzing past data and predicting future trends

(d) We can compare things with the help of statistics

Check Your Progress

(1) Mention three points showing relations between Statistics and human life.

a) _____

b) _____

c) _____

4.2.4 Validity of Statistics

The statistical method is –

(a) a set of practical technique, and

(b) also an ideology which validates the use and application of those techniques.

There are of course, views and counter views with regard to the validity of statistics. **Validity** refers to the quality of measuring precisely what one intends to measure. When statistics is applied to understand human behaviour, there appears opposite notion regarding validity of statistics. Moralists are of the view that application of statistics to measure human behaviour is an endeavour to reduce human behaviour to statistical regularities, and it was condemned as

materialistic and as denial of the axiomatic free will of man. To their views – quantitative regularities of inanimate nature can be positively accepted, but to reduce the behaviour of men's soul to mechanical laws is to undermine the very foundation of personal responsibility and morality.

A Belgium statistician *Quetelet* (1796 – 1874) was the first man to apply quantitative procedures to human behaviour. *Quetelet* could successfully predict the crime rate by using statistical data on crime.

With regard to the validity of statistics in social behaviour, there has also been a secular version. At one stage, it was thought that social science should not quarrel with mechanical method of physical sciences by treating social behaviour quantitatively. Social sciences should restrict itself to the operation of mental phenomena such as – understanding, insight, empathy, sympathetic introspection and other subjective techniques as advocated by some German theorist.

For many scholars, statistics, it is still thought somehow debases over observations and dehumanizes society. It is still thought as an unnatural imposition of external techniques upon social realities. These were the treatment with negative notion on statistics to human sciences.

Further, it was also advocated that there are "dangers and fallacies" of statistical methods. It is said that statistics can offer only 'probabilities', not certainty. **Probability** refers to an anticipation that a particular thing may happen, whereas **certainty** rests upon assumed statements. It is to the fact that statistics apply only to the mass, not to the individual.

But there are some commonly agreed principles with regard to the application of statistical techniques in social sciences. These are as follows

a) The assertion of statistics should be viewed as characteristics of the method, rather than limitations.

b) The statistical method, rather than replacing “insight” reinforces insight. No person can become less human by employing statistics or more human by avoiding them.

As a matter of fact, there is a continuity between common sense and statistics which is not only a more formal and precise version of knowledge, but is also of more extended scope.

4.3. REASONING AND APPLICATION IN STATISTICS

The application of statistics is vast and ever expanding. Hence, it becomes difficult even to define. Statistics pervades all subject matters, its uses have permeated almost every facts of our life. It is an indispensable tool to search an intelligent judgement. In all fields of study such as-trade, industry or commerce, economics, biology, botany, astronomy, physics, chemistry, education, medicine, sociology, psychology or meteorology, statistics are widely applicable. It may be pointed out that in the last three four decades there has been a virtual explosion in the use of statistical methods.

4.3.1. Statistics in Social Research

Statistics is widely employed as a tool and valuable means in the analysis of problem in natural, physical and social sciences. Sociologists collect huge amount of data for understanding social reality. Thus, they are in the midst of complex human data. Complex data predict complex set of behaviour. But social researcher seeks for a general trend of data. Out of the vast and confusing variety of individual character, social scientists search for underlying group character.

Here lies the place of statistics in social research. There are obviously two reasons why statistics become so popular and useful in social research. These are –

(a) Researcher collects huge amount of data. They need simplification of data and to make it useful for easy understanding without much difficulties, and

(b) Secondly, more important reason is the increasingly quantitative approach currently being employed in social research.

The need of statistics is felt only at the stage of data analysis of research process. At the initial stage, when sample is to be drawn, data are collected and analysis is made. But this does not mean that a social researcher can plan his research work without any knowledge of statistics.

Statistical procedures are increasingly becoming part of social research. The procedures are used to enhance the effectiveness and efficiency of services. Consequently, in recent years, there has been a significant increase in the use of computers by social scientists in understanding and handling of data. The social workers who are actively engaged in social services are much more benefited by the use of computers while they engage in computing data with regard to their beneficiaries/ clients. Micro computers are also used to improve social agency and administrative functioning.

We characterize ‘population’ (universe of study), collect data, tabulate the data and observe the data. Tabulation though is a part of statistics, yet we have to ascertain the reliability of those data. Reliability refers to the degree to which a study or research instrument provides consistently accurate results. Statistics as a method fulfils two important functions-tabulating the data and to measure the reliability of data.

Statistics may be called a branch of applied mathematics since it observes the data. In fact every research steps-collection of data processing, analysis and interpretation of numerical data belongs to the domain of statistics. If a social researcher hands over his data to a statistician, the results of entire research process will be disappointing if statistical methods may be involved throughout.

Truly, statistical method is widely used in sociology. Statistics, as a science of numbering and measuring phenomena objectively is widely applied to human facts.

Most social researches are carried out on the basis of theoretical notions of society. Therefore, there were arguments and counter-arguments as to how statistics can be used to proper study of human behaviour. Similarly, social variables are complex and use of statistical manipulation to control social variables raises doubts. Statistics, it is still thought, somehow debases our observation and dehumanize society. It is also considered unnatural imposition of external techniques upon social realities, since social reality is culture –loaded and therefore essentially unquantifiable.

But even than there are some characteristics of the methods of statistics. The statistical method rather than replacing insight, reinforces insight. No person becomes less human by employing statistics or more human by avoiding them – observed Mueller and Schuesler.

4.3.2. Functions of statistics

Statistics has two broad functions –

1. The first is description, the summarization of information in such a manner as to make it more useable.
2. The second function is induction which involves either making generalisation about some population on the basis a sample drawn or formulating general laws on the basis of repeated observations.

For example, it is desired to study the problem of labour unrest in particular areas. The first thing to be done is to analyse the various causes of labour unrest and to study its impact on the various categories of labour viz. male workers and female workers or skilled workers and unskilled workers. This kind of analysis will give us an insight into the problem and we will be able to know many things

from such analysis such as – involvement of male workers is much higher than the female workers, and that labour unrest in big industries is much higher than in small industries. Such analysis leads us to the conclusion that – male workers of big industries are responsible for labour unrest in our country.

Thus, the former is the process of descriptive statistics whereas the latter is that inferential statistics.

Check Your Progress and try to answer

1. What are the uses of Statistics in social research? Mention three points in the s pace provided below.

a) _____

b) _____

c) _____

2. Define Descriptive Statistics and Inferential Statistics.

a) _____

_____

b) _____

_____

S.P. Gupta has listed few important function of statistics. These are –

- ❖ It represents facts in a definite form
- ❖ It simplifies mass of figures.
- ❖ It facilitates comparison.
- ❖ It helps in formulating and testing hypothesis.
- ❖ It helps in prediction.
- ❖ It helps in the formulation of suitable policies.

Each of the above functions can be understood in more detail as follows :

Definiteness

One of the important functions of statistics is to present general statement in a precise and definite form. The character of definiteness helps proper comprehension of data.

Simplification

Statistics not only help in presenting data in a definite form, it also helps in condensing mass of data into a significant figures. Thus, it is difficult to remember the income figures of the people of India individually, but the figures of per capita income can be easily remembered by every one.

Comparison

Comparison of figures with one to another gives immediate meaning of social situation. For example – if we say that “There is a negligible growth of urban population in Assam during the census year 1961 to 1971”, the reader will not get the clear picture of urban situation. But, if we say that “urban growth in Assam during 1961 – 1971 is merely 2 pc”. It conveys a definite meaning.

Formulating and Testing Hypothesis

Statistical method are extremely helpful in formulating and testing hypothesis and to develop new concepts and theories.

Prediction

Statistical methods provide helpful means for predicting future trend of events. For example. In India, the rate of literacy among the females during 1991-2001 has increased in a significant way. If the government adopts many more effective programmes for

educational improvement for the women in rural areas, we can definitely predict much more literate people among women in the next decades.

Formulation of Suitable policies

Statistics provide the basic materials for framing suitable policies. For example data about population, its distribution by age and sex and other socio-economic characteristics, the rate of growth of population, migration, area etc. help in determining the future needs such as – food, clothing, housing, education, recreational facilities, electricity, transport system etc.

4.2.3. Application of Statistics

Statistics is indispensable in research work. Advancement of knowledge has taken place because of experiments conducted with the help of statistics. When data are not considered reliable, mechanical application of statistics is the best operation.

Researcher may carry out various statistical operations, but he must understand that the underlying design of study should not be affected.

The application of statistics in the present day social life is vast and ever expanding. It is a tool of all sciences and is indispensable to search an intelligent judgement and has become a recognized discipline in its own right. There is hardly any field of knowledge where statistical tools are not applicable. Whether it is trade, industry, commerce and management, all physical sciences, medicine, engineering, all discipline of social sciences, meteorology; statistics has become a vital tool for data analysis, drawing important conclusions and making prediction.

As already stated that in the past few decades there has been a virtual explosion in the use of statistical methods. This has been

particularly true in recent years with the advent and accessibility of high speed digital computers which have the capacity to process large amount of information.

4.3.3. Limitations of Statistics

Meanwhile we have discussed earlier about the positive as well as negative effect of statistics in social sciences. It should not be an impression that statistics are like magical devices which always provide the correct solution to problems. Unless the data are properly collected and critically interpreted there is every likelihood of drawing wrong conclusions. It is by far correct in social science research too. It is necessary to outline some possible drawback of the science of statistics. Briefly, the following can be mentioned.

a. Statistics does not deal with individual measurement.
b. Statistics deal only with quantitative characteristics.
c. Statistical results are true only on average.
d. Statistics is only one of the methods of studying a problem.
e. Statistics can be misused.

One obvious limitation of statistics is that it is readily applicable to problems that are amendable to quantitative expression and treatment. Qualitative attributes may sometimes be subjected to statistical analysis. The qualitative information have to be translated into quantitative indices.

It is understandable that statistical laws are applicable to average of aggregates. This may ignore the individual as solitary unit. Such neglect may, in certain investigations lead to merely superficial findings. On this context, there is likelihood that statistics may be misused, so much so that researcher may be in a position to twist them to suit his conclusion or hypothesis.

What is most important for us as potential users of statistical methods, therefore, it is to bear in mind that statistics has power that can only serve and not direct. Much would depend on the use we want to put it to and whether the situation justifies it. So, what is given out by statistical computation may not be the final, since it is basically an approximation, and it is by nature probabilistic.

Check Your Progress

1. What are the preconditions to draw correct conclusions when statistics are used? Mention three points.

   a) _____

   b) _____

   c) _____

2. Draw your visualization and drawbacks in the use of statistics?

   _____

   _____

   _____

Key terms to remember

Validity : refers to the quality of measuring precisely what one intends to measure.

Probability : refers to an anticipation that a particular thing may happen.

Reliability: degree to which a study or research instrument provides consistently accurate results.

4.4 COMPUTATERISATION : INTRODUCING THE WORLD OF COMPUTER

You are perhaps aware of the fact that computers are the greatest achievements of our time. When computers were first introduced in India, very few institutions could afford to have one, because of the cost and maintenance involved. Only the top educational institutions such as – Indian Institute of Technology (IIT) and the Indian Institute of Management (IIM) had computer laboratory. In addition to these, only government-sponsored research laboratories such as – Indian space Research Organisation (ISRO) and Bhaba Atomic Research Centre (BARC) could afford them.

In the late 1980's the country witnessed a sweeping change. Educationists and educational administrators realized that they had to keep up with the changing scenario in the field of education.

Today, at least every educational institution can boast of at least one computer. Computers have infiltrated in most of the activities around us. They are increasingly becoming the backbone of our information system.

We are now crossing the threshould of information era in which we are developing tools that permit us to amplify human intelligence and acquire informations to explore new systems. In the human social life some of the vital areas of exploration are health care, education, science, manufacturing and business. The new information era is nothing but "Computer Era".

A computer is an electronic device that can perform activities involving mathematical, logical and graphical manipulation.

Most people know that computer is a fact calculator, but it is much more than that. It is a machine which performs fast calculations and performs other activities such as – choosing, copying, moving, comparing and performing other operations on alphabetic, numeric and other symbols.

Thus, computer involves a vast systems of learning and the entire systems combined together is called *computer science.* Therefore, computer science is the study of operating principles of computers, computer programming and language for solving theoretical as well as practical problems. It involves the development and use of devices for processing informations. Information is of two types – one form is presented to the device (input information) and the other form is required from it (output information). The input information is raw material and the output is finished product. That is **input** refers to data and instructions given to the computer and the **output** refers to the results produced by the computer when it has processed data. This process of conversion of raw material into finished product is nothing but **data processing**. Then data processing may be defined as the process of transforming data (raw material of information) into useful information by the computer.

In the ear of information technology, computers are well used in almost all walks of life. Computers are widely used in various fields of education, communication, entertainment, banking, business, medicines, weather forecasting and scientific researches. The term **information technology** is widely used now-a-day to refer to the subjects related to creating, managing, processing and exchanging information. Sociology, being an advanced discipline is social science, has of late, opened enough scope to use technological devices in handling and manipulating social data.

Computer is social sciences is an application field of computer sciences and is specially concerned with the application of computer science in handling social problems. The computer application in the social science is a new academic subject. Upto few decades ago, computer uses were only the domain of scientific research laboratories. Now, it has entered into all levels of learnings including social sciences. Computer application in social science have lagged behind other area of studies. However, social scientists have shown more and more interest in the use of computers during

last few years. As such a number of new writings, books etc. have appeared or published on the necessity and uses of computers in social science research. Thus, the branch of social sciences in the science of societal relations of automated informations processing.

4.4.1. Need for computer in Social Science Research.

Computers are found nearly everywhere in our personal life and world of work. In the era of information technology, effect of computer is felt. This is how the meaning of computer literacy goes. **Computer literacy** means having a general knowledge about computers knowing who uses them, what kind of functions they perform, how they are used by others, how they are affecting society etc. Even some experts think that eventually a person who does not know how to use a computer will be just as handicapped in performing his or her jobs as the person today who cannot read.

4.4.2 Uses of Computer in Social Research

Computer application in social sciences is not a manual. It includes sample programs and printed for illustration. A program is not a set of instruction that enables a computer to perform a given task. It emphasizes the need for the importance of computer literacy and offer specific advice on how computers can improve efficiency and assist in creativity. It is creatively manipulating and presenting social science research.

But there is a prerequisite for reading and using computer application. The reader should be slightly familiar with the approaches and methods of social science research. Since research methodology is a compulsory course for all students of sociology, it may not be much difficult to use computer for the researcher in social sciences. This is particularly helpful when data are handled and manipulated for purpose of analysis.

The social sciences overlap with many other disciplines. These include Library Sciences, public administration, planning, architecture, applied statistics, economics and so forth. The mutual interest of the social sciences with these disciplines became specially apparent from the vantage of computing. There are practicing specialists in numerous fields and they will definitely find easier in writing and rewriting and graphing.

The application of computer in social sciences is of great help for social researchers, or who aspires to do so. As everyone knows, computers are increasingly employed by government at all levels of business and other organizations to secure, employ and store information. This means the records of society are stored in computer in readable form. The important point is that we should be familiar ourselves with the opportunities as well as advantages of these machine-readable records.

Sufficient to say that we are living through an information revolution of unprecedented scope and magnitudes. There are some important elements in such revolution such as – massive increase in the volume of information that is recorded and employed, change in the purpose for which information is used, greatly enhanced capacity to employ information and increased emphasis upon detailed quantitative and mathematically manipulable information etc.

These are the great changes in technologically influenced society. One can think for the further of a paper-less society, or we are witnessing a society of progressively increased use of electronic media instead of papers and other conventional modes of recording and disseminating information.

Further, computer technology has widely been used to automation of the processes that produce and record information. Automation is a production process monitored and controlled by machines with only minimal supervision from human beings. Unlike papers, most electronic information storage devices are easily readable and reuseable.

4.4.3. Computer and Research Value

In the last few decades, computers have brought about radical changes in our lives. Almost every profession makes use of computers. You will now at this stage observe how computers are used extensively in research activities.

The use of computers in research is particularly significant when some of the direct benefits of computers are taken into consideration. As preliminary knowledge one should know some characteristics of computers as listed below –

(a) High speed

(b) Accuracy

(c) Storage capacity

(d) Automation

(e) Diligence

(f) Versality

(g) Cost effectiveness

(a) Computers have the ability to perform routine task at a greater speed than human beings. They can perform complex calculations in second. A computer can add and substract members, compare letters to determine alphabetic sequence, move and copy numbers and letters. What is significant is the speed with which computers carry out these operations.

(b) When a task is performed manually, there is always a possibility of human error. Computers can be used to perform task in a way that ensures accuracy. However, if the input given in the computer is incorrect, the output will also be incorrect.

(c) Storage capacity

Computer can store a large amount of information. After the information is stored, it can be retrieved as needed. It can store millions of information in a condensed form. One can use the stored information to perform various kinds of analysis for example, you might want to track the correlation between respondents attitudes and their educational or economic background.

(d) Automation

Computers can be instructed to perform complex task automatically. For example, if you want to generate report showing the average age, income etc. of the respondents, you can use the computer to accomplish this efficiently.

(e) Diligence

Computers can perform the same task repeatedly and with the same accuracy, without getting tired. For example, you can use the computer to print your interview schedule. The computer will print each schedule with same quality in the same time.

(f) Versatility

Versatility refers to qualities of doing many things. Computers can be used to perform variety tasks of both simple and complex. For example, you can use them to create an invitation letter, play music, create a picture, or a diagram and variety of data tables as you need.

(g) Cost effectiveness

Computer reduces the amount of paper works and human effort, thereby reducing cost. For example, you can create and edit data easily in a computer or you can send electronic message via e-mail. E-mail is a way of sending messages and data to other people by means of computers connected together in a network.

Besides these above, there are other qualities of computers. Computer output is generally very reliable, subject to the condition that the input data entering the computer should be correct and the programme of instruction should be reliable and correct.

Thus, computer, one of the most powerful forces today is being put to use everywhere. Social science researchers have shown more and more interest in computers during the last few years. But the use of computer continuous to be low compared to outer subject areas. However, social researchers have utilized computers with its immense research value.

As already stated earlier that computer has compact storage media. This has helped automatically to record larger information of society which we could not assumed earlier. Because of computer technology, capacity to employ information is also much greater. A massive collection of detailed and quantitative information can be stored in computer and this is a great advantage of using computer in social research.

The other research value of computer is the capacity for manipulation of data and as such new range of research questions can be addressed. Those questions could not be pursued using the conventional method of handling data. When recorded in conventional paper form, large scale data or records cannot be effectively used for research. When recorded in computer-readable form, the utilities of these materials in quite different. In that form, they can be organized and reorganized. Further, specific items and categories of information can be selectively retrieved, and they can be subjected to massive volume of statistical reduction, summaries and analysis. In these terms, computer technology leads research value where it would otherwise be lacking.

Computer application in social science offers direction towards practical experience, using software to manipulate text, numbers, graphics, sound, video, animation, archived data and other

digital media. Researchers are also encouraged to engage in critical analysis with these technologies.

Computer technology continuous to grow and to involve at a rapid pace, offering social scientists a wealth of new computing methods and techniques. This facilitates research in social sciences in a more innovative way. Scholars and researchers in the social sciences have become increasingly sophisticated in their use of computer technology. Even today, many computational resources and tools remain virtually untrapped. Computer networking in one such resource.

4.4.4. Consequences

As a result of uses of computer technology in social science research one can take note of innumerable consequences. It brought about the possibilities of new research opportunities. Computer technology helps in data perception and data sources in a kind of detailed and precise manner which was never before imaginable. There was a time when it was not possible to examine different aspects of society either because of

(a) necessary source materials were not available, or

(b) the existing materials were too voluminous to be employed effectively.

Now, we can imagine a virtual paradise of abundant and diverse source material as a consequence and unprecedented opportunity to trace and examine the growth and evolution of society. With this opportunity, now we can establish a genuine base for a scientific study of society.

The pervasive effect of computer technology in social sciences can be best observed if we categorically observe certain characteristics of contemporary source materials.

The first category is composed of detailed quantitative information to support research for purpose of programme and policy evaluation and to meet needs of social monitoring.

The second category is made up of information bearing upon the operation of government, business and other organizations such as – budgetary, personnel, purchasing, transactional etc.

The third category includes reports, memoranda, letters and other forms of personal and organizational communications that constitute the privacy sources for traditional historical inquiry.

In recent, years, on various occasions an innumerable survey researches are carried out. These researchers have sought extensive uses of methodologies such as – drawing and designing of samples, framing of questions and use of tabulations. A large number of computer fields are used to handle such research steps and this minimizes researcher's labour.

Computer technology has been widely applied in providing and recording information. It has an increasing power and capacity to record.

4.5 Let us sum up

The foregoing illustration can be summed up with few basic points as follows;

1. Statistics refer to the statistical principles and methods employed in the collection, processing, analyzing and interpretation of any kind of data.
2. There are differences between individual and group life. Individuals lose their identity in the total anonymous mass. For purpose of description or prediction various devices have been involved, where the application of statistics is one basic device.
3. Statistics make complex mass of data in a simple and understandable form.

4. Statistical methods are highly useful tool in analyzing past data and predicting future trends.
5. Statistics can offer only probabilities not certainty.
6. There was arguments and counter-arguments as to how statistics can be used to proper study of human behaviour.
7. Statistics have both positive and negative effects in social sciences. Statistics deal only in quantitative measurement and its results are true only on average.
8. There has been a sizeable increase in the use of computers by social scientists in analyzing data.
9. In the era of information technology computers have infiltrated in most of the activities around us. Computer is becoming the backbone of our information system.
10. Computer has entered into all levels of learning including social sciences.
11. The computer users must have familiarity with the approaches and methods of social science research.
12. As a result of increasing advancement of computer technology, vis-à-vis computer-dependent society, one can visualize a society of paper-less culture.
13. Computer has immense research value. It can address new horizon of research range, since computer is gradually replacing conventional methods of handling data.

Key terms

Computer : Computer is an electronic device that can perform activities involving mathematical, logical and graphical manipulation.

Input: refers to data and instructions given to the computer.

Output: refers to the result produced by a computer when it has processed data.

Information Technology: is the term used to refer to the subjects related to creating, managing, processing and exchanging information.

Computer Literacy: means having a general knowledge about computer knowing who uses them, what kind of functions they perform, how they are used by others, where they are, how they are affecting society etc.

Program : a program is a set of instruction that enables a computer to perform a given task.

Automation : is a production process monitored and controlled by machines with only minimal supervision from human beings.

Question to answer

1. What is Webster's definition of Statistics?
2. What are the important contents of Hoarce Secrist's definition of Statistics?
3. Define 'aggregate of facts'?
4. Explain how numbers play important role in statistics.
5. What do you mean by statistical probability?
6. "Statistics minimize description" discuss
7. Discuss the need and application of computer in social research.
8. Discuss some of the basic characteristics of computer.
9. Discuss how application of computer in society research minimize human labour.

Suggested Readings.

S.C. Gupta : *Statistical Methods*, Sultan Chand and Sons. New Delhi, 2001

J.H. Mueller.

K.F. Schuessler : *Statistical Reasoning in Sociology*, Oxford and JBH, New Delhi, 1961.

D.K. Lal Das : *Practice of social Research*, Rawat Publishers, Jaipur, 2000.

S. Jaiswal : *Information Technology Today*, Galgotta Publication, New Delhi, 1999. Edward Brent Jr.

Ronald E. Anderson: *Computer Application in the Social Science*, 1990 (through Internet)

Jerome M. Clubb : *Computer Technology and the Source Materials of Social Science History*, Summer 1986, (through Internet).

BLOCK – 5
STATISTICAL METHODS

STRUCTURE

5.0 OBJECTIVES

The objectives of this section is primarily two-fold viz. to acquaint the learners with some basic statistical procedures popularly used in social science and secondly, to gather ideas about the utilities of those statistic in social science researches.

5.1 INTRODUCTION

Statistics have been widely used in social sciences. Social researchers besides describing social phenomena, tried to examine social significance of reality by applying various statistical techniques. There are available statistical techniques which are popularly used by the social scientists while analyzing their data. In this unit we will discuss certain selected statistics to the beginners.

5.2 MEASUREMENT OF CENTRAL TENDENCY

Any value between two extremes i.e. highest and lowest could be selected which is known as locational average. In statistical analysis the calculation of such single value is important because it determines the entire mass of data. The calculation of such common value is average. This is generally a central value around which the distribution seems to cluster. This apparent tendency of many statistical aggregates to concentrate around a centre is often termed *"central tendency"*, the value of this centre being the *measures of central tendency* more commonly called *average*.

Like common usage, and as in statistical language the concept "average" connotes the typical, the ordinary and expected. Every lay man in their daily life calculate average, casually or in common experience like average voter, average family, average age of student, a batting average etc. For example, the average American family prefers nuclear family, average weight of school students 50 kg etc. is most important quality of daily life calculations.

Whatever manner he employs the concept, the layman unconsciously implies what every statistician explicitly recognize. The statisticians take it that average is a kind of norm around which the values tend to vary. Here lies the difference between the layman and professional. The laymen use informal folk usage and the professionals require more precise and exact understanding. The professional therefore, derives mathematical procedures to measure the average. There are various types of central tendencies and so, scientists formulate various averages to answer respective problems of all the various averages. We will here discuss three types of averages the mean, the median and the mode.

5.2.1. The "Mean" or "Arithmetic Average"

The concept of Mean

When a series of figures or items are added together and divided the sum by the number of items, this is usually called *mean*. This is an operational definition of mean often used synonymously for "average". Mean is calculated with the following formula

$$a = \frac{\sum m}{N}$$

Where a = the mean

$\sum$ = the sum

(read : 'Summation X') (Greek letter S or sigma)

$\sum m$ = a the sum of values

N = total frequency, or number of items.

Illustration 1

The following table gives the monthly income of 10 employees in an office.

Income (in Rs.)

1780, 1760, 1690, 1750, 1840, 1920, 1100, 1810, 1050, 1950.

Calculate the arithmetic mean of incomes

Let income be denoted by the symbol x.

Solution :

Calculation of Arithmetic Mean	
Employee	**Monthly income (in Rs.)**
1	1780
2	1760
3	1690
4	1750
5	1840
6	1920
7	1100
8	1810
9	1050
10	1950
N=10	$\sum X = 16650$

$$a = \frac{\sum m}{N}$$

$$= \frac{16650}{10} = 1665$$

Hence the average income is Rs. 1665.

In discrete series

In discrete series arithmetic mean may be computed by applying two method i.e. (a) direct method (b) short cut method

Direct method

The formula for computing mean is

$$a = \frac{\sum mf}{N}$$

Where f = frequency

Mf = the total of the product of the frequency

N = total number of observation i.e. $\sum f$

Illustration 2

From the following data of marks obtained by 60 students of a class, calculate the arithmetic mean

Marks	**No. of students**
20	8
30	12
40	20
50	10
60	6
70	4

Let the marks be denoted by X and the number of students by f.

Calculation of Arithmetic mean.

m marks	**f** **no. of students**	**mf**
20	8	160
30	12	360
40	20	800
50	10	500
60	6	360
70	4	280
	N = 60	$\sum mf$ =2460

$$a = \frac{\sum mf}{N} = \frac{2460}{60} = 41$$

Hence the average mark = 41

Short cut method

According to this method

$$a = x + \frac{\sum fdx}{N}$$

Where dx = assumed mean

$\sum fdx$ =the total of the products of the deviations

Calculation of Arithmetic mean in Continuous Series

In continuous series, arithmetic mean may be computed by applying any of the following methods.

Direct method

Shortcut method

In direct method the following formula is used (the same formula is used as is used for discreate series)

$$a = x + \frac{\sum fdx}{N}$$

Where mv=midpoint of various classes (m.v)

f = the frequency of each classes

N = total frequency

dx = assumed mean

$\sum fdx$ = the total of the products of the deviations from the assumed mean and the frequencies of the items.

Steps : i) obtain the midpoint of each class and denote it by m.v.

ii) multiply these midpoints by the respective frequency of each class and obtain the total $\sum mf$

iii) divide the total obtained by the sum of the frequency i.e. N.

Illustration 3

From the following data compute arithmetic mean by direct method.

Marks--	**0 – 10**	**10 – 20**	**20 – 30**	**30 – 40**	**40 – 50**	**50 – 60**
No. of students	**5**	**10**	**25**	**20**	**20**	**10**

Steps

i) take an assumed mean suppose 40

ii) take the deviation from the variable x from the assumed mean and denote the deviations by “dx.”

iii) Multiply these deviations with the respective frequencies and take the total $\sum fdx$.

iv) Divide the total obtained in third step by the total frequency.

Illustration – 4

Calculate arithmetic mean by the short cut method using frequency distribution of illustration 2.

Solution :

Calculate of Arithmetic mean

Marks	**No. of students**	**dx – 40**	**fFdx**
20	8	-20	-160
30	12	-10	-120
40	20	0	0
50	10	+10	+100
60	6	+20	+120
70	4	+30	+120
	N=60		$\sum fdx = 60$

$$a = \frac{\sum fdx}{N} = 40 + \frac{60}{60} = 40 + 1 = 41$$

Solution :

Calculation of Arithmetic mean by direct Method

Marks	**Midpoint(m.v.)**	**No. of students(f)**	**Mf**
0 – 10	5	5	25
10 – 20	15	10	150
20 – 30	25	25	625
30 – 40	35	30	1050
40 – 50	45	20	900
50 – 60	55	10	550
		N=100	$\sum mf$ =3300

$$a = \frac{\sum mf}{N} = \frac{3300}{100} = 33$$

Hence, arithmetic mean is 33.

To calculate mean by adopting short-cut formula from the continuous series following method is adopted.

Marks	**Mid value (mv)**	**No. of students(f)**	**Deviation from dx assumed 25 mean fdx**	
0 – 10	5	5	-20	-100
10 – 20	15	10	- 10	- 100
20 – 30	25	25	0	0
30 – 40	35	30	+10	+300
40 – 50	45	20	+20	+400
50 – 60	55	10	+30	+300
	N=	100		+1000
			$\sum fdx = \frac{-200}{+800}$	

$$a = x + \frac{\sum fdx}{N}$$

$$= 25 + \frac{800}{100}$$

$$= 25 + 8$$

$$= 33$$

Hence arithmetic mean is 33.

Exercise:

1. calculate the value of the mean wage of the labourers from the following data.

Wages in (Rs.)	**60 – 70**	**50 – 60**	**40 – 50**	**30 – 40**	**20 – 30**
No. of labourers	5	10	20	5	3

2. Compute the mean family size from the following data (one hundred family)

Family size	**f**
1	20
2	24
3	17
4	15
5	9
6	5
7	4
8	3
9	1
10	2

5.2.2. Characteristics of mean

1. The mean is the best known and frequently used average. Often, it may be advantageous to use the mean because it is so well understood.
2. From an algebraic standpoint the mean is superior to either the median or the mode.
3. It is the simplest average to understand and easiest to compute.
4. It is affected by the value of every item in the series.
5. It is defined by a rigid mathematical formula with the result that everyone who computes the average gets the same answer.
6. It is relatively reliable in the sense that it does not vary too much when repeated samples are taken from one and the same population.
7. The mean is typical in the sense that it is the centre of gravity balancing the value of either side of it.

5.2.3. Median or Position Average

The median is another simple average or measure of central tendency. Many statisticians define median as the size of middle item when the items are arranged in order of magnitudes. By simple definition it is the middle value in a distribution. The median is just the 50^{th} percentile value below which 50 pc of the values in the sample falls. It splits the observation into two halves. The ‘M’ denote Median.

The median is called a positional average. The term ‘position’ refers to the place of a value in a series. The place of the median in a series is such that an equal number of items lie on either side of it.

For example, if the age of five students is 22, 23, 24, 25 and 27 years, the median would be 24.

22

23

24 – value at middle position of the arrary.

25

27

These are *odd numbers* of observation and the calculation of median was simple. When an even numbers of observation are listed, there is no single middle position value and the median is taken to be the arithmetic mean of two middlemost items. For example, if in the above case we are given the age of six students as 22, 23, 24, 25, 26 and 27, the median age will be

22

23

24 } there are two middle position values.
25 }

26

27

$$.M = \frac{24 + 25}{2} = \frac{49}{2} = 24.5$$

Hence, in case of even numbers of observation median may be found by averaging two middle position values.

5.2.4. Steps in calculating median

1. Arrange the data in ascending or descending order of magnitudes. (both arrangement would give the same answer).
2. In a group composed of odd number of values say 7, add 1 to the total number of values and divide by 22. Thus, N + 1 would give 8 which is divided by 2 gives 4.

Illustration 4

From the following data of the wages of 7 workers compute the median wage.

Wages – 1100, 1150, 1080, 1120, 1200, 1160, 1400

(in Rs.)

Solution:

Calculation of Median

Sl. No.	Wages arranged in ascending order
1	1080
2	1100
3	1120
4	1150
5	1160
6	1200
7	1400

$$.\mathrm{M} = size\ of\ \frac{N+1}{2} th \quad item = \frac{7+1}{2}$$

$$= 4th \quad item$$

Size of the 4^{th} item is 1150. Hence the median wage is Rs.1150.

In a group of even numbered

The median value for a group composed of an even number of items is estimated by findings of the arithmetic mean of the two middle values – i.e. adding the two values in the middle and dividing by two-

Formula is

$$.\mathrm{M} = size \quad of\ \frac{N+1}{2} th\ item.$$

Thus, we find that it is both when N is odd as well as even that 1(one) has to be added to determine median value.

Illustration 5

Find out the value of median from the following data.

391, 384, 591, 407, 672, 522, 777, 753, 2488, 1490.

Solution :

Calculation of Median

Sl.No.	Data arranged in ascending order
1	384
2	391
3	407
4	522
5	591
6	672
7	753
8	777
9	1490
10	2488

$$.\text{M} = size \quad of \ \frac{N+1}{2} th \ item. = \frac{10+1}{2} = \frac{11}{2}$$

= 5.5th item

Size of the 5.5th item = $.\frac{5th\ item + 6th\ item}{2}.$

$= .\frac{591+672}{2}.$

$= .\frac{1263}{2}.$

= 631.5

Calculation of Median in continuous series

In a continuous series, the following procedure is adopted for the calculation of median.

1. covert the frequencies into cumulative frequencies by giving the progressive total.
2. find out the middle item by using formula n/2.
3. locate the middle number in the cumulative frequencies and thus find out the median group.

   The following formula to locate the median may be used.

   M = Median

   1_1 =lower limit of the size of the median group.

   i = class interval of median group

   M = middle number

   C = cumulative frequency of the previous group

Illustration 6

Calculate Median from the following data.

Measurement	f
0 – 5	5
5 – 10	7
10 – 15	10
15 – 20	18
20 – 25	20
30 – 35	8
35 – 40	6
40 – 45	4
45 – 50	1

Solution:

Arranging in ascending order

Measurement	**Frequencies (f)**	**Cumulative frequency(cf)**
0 – 5	5	5
5 – 10	7	12
10 – 15	10	22
15 – 20	18	40
20 – 25	20	60
25 – 30	12	72
30 – 35	8	80
35 – 40	6	86
40 – 45	4	90
45 – 50	1	91

Middle item (m) $= .\frac{91}{2}. = 45.5$

By looking at the cumulative frequencies we find that this item lies in the group 20 – 25. It is thus the median group.

The median is –

$$M = i_1 + \frac{i}{f}(m - c)$$

$$= 20 + \frac{5}{20}(45.5 - 40)$$

$$= 20 + \frac{5 \times 5.5}{20}$$

$$= 20 + 1.3.75$$

$$= 21.375$$

Exercise :

1. Find out the median from the following data.

Class interval	**frequency (f)**
10 – 15	8
15 – 20	23
20 – 25	97
25 – 30	224
30 – 35	417
35 – 40	242
40 – 45	112
45 – 50	57
50 – 55	16
55 – 60	4

5.2.5 Merits of calculating median

1. Median is easy to locate
2. It is not affected by extreme items and thus possess greater stability. For example, the median of 10, 20, 30, 40 and 50 would be 30, whereas the mean is 50. Hence, very often when extreme values are present in a set of observations, the median is a more satisfactory measure of the central tendency than the mean.
3. The value of median can be determined graphically whereas the value of the mean cannot be graphically ascertained.

5.2.6. Characteristics of Median

1. It is the value of the middle point of the array such that half the items are above and half below it.
2. The value of the median is fixed by its position in the array and does not reflect the individual values.
3. The aggregate distance between the median point and all the values in the array is less than from any other point.
4. Each arrary has one and only one median

5. It cannot be manipulated algebraically ; medians of subgroups cannot be weighted and combined.
6. It is stable in that grouping procedures donot appreciably affect it.
7. Values must be ordered and may be grouped for computation.
8. It can be computed when ends are open.
9. It is not applicable to qualitative data

5.2.7. The Mode or Probability Average

The mode is simply the most frequently occurring value in an ordered distribution, it is the value where the concentration of items is most dense. Etymologically, it is related to the notion of the prevailing fashion of dress or etiquette to which a majority of a given social class would be expected to conform. Hence, the mode (Mo) may also be defined as *the most probable value*, and therefore, distinctively labeled the probability average.

Technically speaking the mode is the value of the item that has highest frequency. It is the measurement or size that the directly applicable to largest number of cases. It is the most common type of average. Thus, mode is a value that is repeated most often in the data set. For example, in the series of values – 71, 73, 74, 75, 75, 78, 78, 80 and 82, the mode is 75, because 75 occurs more than any other value (three times). In grouped data, the mode is located in the class where the frequency is greatest.

Calculation of Mode

Since the mode is the most frequently occurring value, it is obviously necessary to count the number of occurrences of each value. There are two distinct steps in the determination of the mode –

a) locate the predominant or modal frequency, and
b) find the value corresponding to the total frequency.

The following table will give the example

Table

Items arranged in order of magnitudes

Gain (in k.g) of weight	No. of children (f)	Cumulative frequency (cf)
1	3	3
2	4	7
3	6	13
4	9	12
5	12	34
6	11	45
7	9	54
8	7	61
9	4	65

In a discrete series the value against which the frequency is the largest would be the modal value. In the above table we find that 12 is the largest frequency. Thus, 5 being the value against which the frequency is the largest, is the mode of the series.

The calculation of mode is very easy. It depends upon the frequencies. The data should be grouped in continous or discrete series and the item value with higher frequency would be the mode see the illustration below

Illustration 7

Calculate Mode from the following data regarding number of children per couple.

Couple	No. of children
A	0
B	3
C	5
D	2
E	1
F	2
G	4

H	2
I	1
J	0
K	3
L	2

Now, arrange in discrete series

No of children	**No of couple**
0	2
1	2
2	4
3	2
4	1
5	1
Total	**12**

The modal number of children is thus 2, as it has the highest frequency

Illustration 8

Following are the data regarding income of 90 families. Find out the average income by means of Mode.

Income (in Rs)	**No of families**
Upto 100	8
100 – 150	10
150 – 200	15
200 – 250	25
250 – 300	12
300 – 350	11
350 – 400	7
Above 400	2
Total =	90

If we see the above distribution, we find the highest frequency is 25. The modal group therefore can be easily located as 200-250. But *Mode is a single number*, it should not be a group. Therefore, actual place of mode must be located. This is possible with the help of frequencies before and after the modal group. This is called interpolation. For this purpose, the following two formulae are generally used.

First formula

$$.Z = 1_1 + \frac{f_1 + f_0}{21 f_2 - f_0 - f_2} \times i$$

Second formula

$$Z = 1_1 + \frac{f_2}{f_0 + f_2} \times i$$

Where

Z = Mode, generally indicated by letter “Z”

1_1= lower limit of the size of modal group

f_1= frequency of the modal group

f_0= frequency of the group prior to modal group

f_2= frequency of the group subsequent to the modal group

i = class interval

The calculation of the illustration 8 will be as follows -

Formula – 1

$$Z = 1_1 + \frac{f_1 - f_0}{2 f_1 - f_0 - f_2} \times i$$

$$= 200 + \frac{25 - 15}{2.25 - 15 - 12} \times 50$$

$$= 200 + \frac{25 - 15}{50 - 15 - 12} \times 50$$

$$= 200 + \frac{10 \times 50}{23}$$

$$= 200 + 21.7 = 221.7$$

Formula – 2

$$Z = 1_1 + \frac{f_2}{f_0 + f_2} \times i$$

$$= 200 + \frac{12}{12 + 15} \times 50$$

$$= 200 + \frac{12 \times 50}{27}$$

$$= 200 + 22.2 = 222.2$$

Exercise

Calculate Mode from the following data (marks obtained by 80 students)

Marks	No of students
0 – 10	3
10 – 20	5
20 – 30	7
30 – 40	10
40 – 50	12
50 – 60	15
60 – 70	12
70 – 80	6
80 – 90	2
90 – 100	8

5.2.8 Characteristics of Modes

Why should we calculate mode ?

Mode, as another form of average has certain characteristics. Since it is a positional average, it is not affected by extreme items of the observed series. Therefore, mode is frequently difficult to locate. The mode is not susceptible to algebric manipulation.

Some characteristics of Mode can be listed as follows :

1. It is the most frequent value in the distribution ; it is the point of greatest density
2. The value of the mode is established by the predominant frequency and not by the values in the distribution.
3. It is the most probable value and hence the most typical.
4. A given distribution may have two or more modes. On the otherhand, there is no mode in a rectangular distribution
5. The mode does not reflect the degree of modality
6. It cannot be manipulated algebraically; modes of subgroups cannot be combined.
7. It is unstable in that it is influenced by grouping procedures.
8. Values must be ordered and grouped for its computations
9. It can be calculated when table ends are open
10. It is the only average which can be applied to qualitative variables.

5.2.9 Measures of Central Tendency : A Summary

We have discussed various measures of central tendency as above. These measures provide different kinds of information

The mean (arithmetic mean) is thought of as the point on the scale around which two extreme cases of distribution may be taken into consideration. If the distribution are not symmetric, the mean may be misleading is a pointless information. Its usefulness will depend on what purposes to do with it. The median implies a concept of balance. In fact, it is possible to have more than one mode in a distribution. For example, we may find two equally large groups of

students, one studying between 1 and 2 hours, and the other studying between 3 and 4 hours. In such distributions there will thus be two modes i.e. the distribution is bimodal There may be also multimodal distributions. In such cases, it is perhaps to speak of central tendency.

Key Terms

Average : Any value between two extremes i.e. highest and lowest could be selected known as average.

Mean: The arithmetic average of a series of numbers.

Median: The value that occurs midway in a series of numbers arranged in order of magnitude or simple the middle case.

Mode: The value that occurs most often is a series of numbers.

5.3 MEASURES OF DISPERSION

The homogeneity and heterogeneity of social data are the subject of statistical understanding. To be more precise, any set of social data is characterized by its heterogeneity of values. In fact, the extent to which they are heterogenous or vary among themselves is of basic importance to statistics. The measures of central tendency as we have discussed earlier, describe one important characteristic of a set of data, such as central point, mean, position etc. but they do not tell us anything about other character, i.e. heterogeneity. In simple meaning heterogeneity refers to *"consisting of many different kinds of people or things"*.

Point to note

Measures of central tendency tells one character of a set of data, but do not tell the heterogenous character of data.

Social researchers then, want to study one interesting aspect of data i.e. *"heterogeneity.* We need to measure heterogeneity. The extent to which data are dispersed and the measure which provides us the description are called *measures of dispersion or variability.*

Here we find few important concepts such as-average, dispersion, variations etc.

Average is defined in various ways –

- average is an attempt to find out one single figure to describe whole figures, or
- average is sometimes described as a number which is typical of the whole group.

It is clear then, average is a single value that represents a group of values. Such a value is of great significance because it depicts the characteristics of the whole group. Since an average represents the entire data, its value lies somewhere in between the two extremes i.e. the largest and the smallest items. For this reason an *average is frequently referred to as a measure of central tendency.*

Dispersion or variability is a concept which gives a higher meaning than the 'average'. Average gives one single figure that represents the entire data. But average alone cannot adequately describe a set of observation, if all the observations are dissimilar or heterogenous. So dispersion or variability takes the measurement about the extent to which items vary from on central value. In other words, dispersion gives an average of the differences of various items from an average. For this reason, dispersions is also known as *scatter, spread or variation.*

An average becomes more meaningful when it is examined in the light of dispersion. For example –

Average wage of workers of factory A- is Rs. 3885 and average wage of workers of factory B- is Rs. 3900.

We cannot immediately conclude that the workers of factory B are better off than the workers of factory A. Because there may be greater dispersion of wages among the workers in the factory B.

Here is an example

The wages of 5 worker in 3 different factories are –

Factory A	**Factory B**	**Factory C**
100		1
100	105	489
100	102	2
100	103	3
100	90	5
Total = 500	500	500
Average = 100	100	100

In the above example the total income as well as a average income of all five persons is equal, but we cannot say that they are equal in all respects. There is difference in the regularity or certainty. A's income is perfectly regular. B's income is changing but slightly. C's income is extremely irregular and uncertain. Thus we come to conclusion that besides the measure of central tendency we have also to find a measure of variability, degree of deviation from the central tendency or scatter of the items within the group. This measurement is know as *dispersion* in statistical terminology. The dispersion is therefore, the statistical measure to compute the scatteredness of the items within the group or their tendency of deviation from the average.

Dispersion has generally been called the *average of second order*. The first kind of averages are the measures of central tendency ad we have described earlier. The measures of dispersion tells us the degree of average deviation from the central tendency. The individual items of the group may deviate from the average in varying degrees. We have to find the average deviation of all the units.

5.2.1. Why to measure variation ?

While stating significance of calculating variation, Gupta refers four basic reasons such as –

- To determine the reliability of an average.
- To serve as a basis for the control of the variability.
- To compare two or more series with regard to their variability.
- To facilitate the use of other statistical measures.

5.3.2 Standard deviation.

The most useful and frequently used measure of dispersion is standard deviation. Standard *deviation represents a mean of deviation items.* The standard deviation concept was introduced by Karl Pearson in 1823. Standard deviation is also know as "**root-mean square deviation**" which means that it is the square root of the mean of the squared deviation from the arithmetic mean. Standard deviation is denoted by small Greek letter δ (read as sigma)

Calculation of Standard Deviation

Let us calculate standard deviation from a series of data given below.

Weights of 265 male students

Class interval (weight)	**f**	**d**	**fd**	**fd^2**
90 – 99	1	-5	-5	25
100 – 109	1	-4	-4	16
110 – 119	9	-3	-27	81
120 – 129	30	-2	-60	120
130 - 139	42	-1	-42	42
140 – 149	66	0	0	0

150 – 159	47	1	47	47
160 – 169	39	2	78	156
170 – 179	15	3	45	135
180 – 189	11	4	44	176
190 – 199	1	5	25	25
200 - 209	3	6	18	108

N= 265 $\sum fd = 99$ $\sum fd^2 = 931$

STEPS

1. See the data-the first column contains the class intervals and the second the frequencies. The total number of cases (N) is 265.
2. Select an interval group in which mean most likely to occur. Let us select the class interval 140 – 149 and was considered as the zero interval and mark of the intervals above (+) and below (-) This is assumed mean. This is third column.
3. Multiply the frequency of each class by its corresponding deviation (fd) and enter the product in the fourth column.
4. In the fifth column put the products of f x d^2 i.e. 2^{nd} column x square of third column.
5. Obtain a total of fifth column i.e. …….. it will be seen that there are never any minus quantities in the fd^2 column. The $\sum df^2$ comes as = 931

Apply the following formula to calculate standard deviation (0)

$$\sigma\left(\sqrt{\sum fd^2/N - \left(\sum fd/N\right)^2}\right) \times i$$

where, σ = sigma (std. deviation)

$\sum$ = read as summation

f = frequency

d = deviation

N = total

i = interval

Solution :

$$\sigma = \left(\sqrt{\sum fd^2 / N - \left(\sum fd / N \right)^2} \right) \times i$$

$$= \sqrt{931/265 - (99/265)^2} \times 10$$

$$= \sqrt{3.5132 - 1396 \times 10}$$

$$= \sqrt{3.3736 \times 10}$$

$$= 1.8367 \times 10$$

= 18.27 or 18.4

Hence standard deviation or σ is = 18.4

Now attempt the following exercise

Exercise 1

Calculate standard deviation from the following data.

Weekly income of 50 families

Weekly income **(in Rs.)**	No. of families **(f)**
100-200	5
200-300	6
300-400	15
400-500	10
500-600	5
600-700	4
700-800	3
800-900	2

N = 50

Exercise 2

Calculate standard deviation from the following series

Size of land holding (in bighas)	No of families (f)
1 – 5	5
6 – 10	8
11 – 15	12
16 – 20	10
21 - 25	5

5.3.3. Significance of standard deviation (S.D)

The standard deviation measures the absolute dispersion. The greater the standard deviation, greater will be the magnitudes of deviation of the values for the mean. A small standard deviation means a high degree of uniformity among the observed units as well as homogeneity of the series. A large standard deviation means just opposite. The standard deviation provides a more refined and statistically important measure of variability.

The use of measures of dispersion is sometimes made to calculate the probable limits of various measurements. It this respect S.D has been found to be a very useful measure.

Key Terms to remember

Homogeneity : The quality of being homogenous consisting of things or people that are all the same or all of the same type.

Heterogeneity : Consisting of many different kinds of people or things e.g. the heterogenous population of the United States.

Reliability : The degree to which a study or research instrument provides consistently accurate results.

5.4. MEASURES OF CORRELATION

In our day-to-day life we talk about relations between two or more social situations. In the language of research we talk about relationship between variables. We have in our earlier chapter, defined and discussed the term “variable”. Ordinarily, when we use the term “correlation” we mean to understand relationship between two or more objects. For example, we make comments like-there is relation between suicide and mental illness, poor housing and disease, poverty and infant mortality, urbanization and economic development, population growth and unemployment and others such variables and their assumed relations. In social situation, such phenomena occur constantly. Correlation then indicate *a regular, recurrent relationship between variables in which a change in one variable is associated with a change in another variable.*

Boddington remarked “whenever some definite connection exists between two or more groups, classes or series of data there is said to be correlation. It doesn’t matter whether data in one section changes in the same or reverse direction”. From the above we can derive the following –

As student of social research, we would be interested to know the relationship that may exist between such variables. Sociologist want to examine and measure the relationship. In the science of statistics we can find various devices to study relationship between variables.

With heavy reliance on empiricism and increasing dependent on empirical data, sociologist have started using a number of statistical devices to measure relations between social variables. In this section we will chiefly discuss two types of correlations i.e. i) Karl Pearson’s method of correlation and ii) Rank difference correlation.

5.4.1. Positive and Negative Correlation

Correlation may be either positive or negative. It means the variables may be positively related or negative related. The positive correlation is also known as *'direct correlation'* and the negative correlation is known as *'inverse correlation'*.

Positive correlation is established if both the variables vary in the same direction. If one variable (x) on an average has its increasing affect on other variable (y), or if the variable 'x' is decreasing other variable 'y' on average, such correlation is said to be positive correlation.

On the otherhand, if the variables are varying in opposite directions i.e. the increasing effect of one variable has its decreasing effect on the other or vice-versa, correlation is said to be negative.

The following samples will illustrate the difference between positive and negative correlation.

1. Positive Correlation

x- 10, 12,15,18,20	x – 80,70,60,30
y- 15,20,22,25,37	y- 50,44,30,20,10

2. Negative Correlation

x- 20,30,40,60,80	x- 100,90,60,40,30
y- 40,30,22,15,10	y- 10,20,30,40,50

So, from the above we can derive some ideas of correlation

a) Correlation exists when a change in one series is followed by a change in other series.

b) The change may be in the same direction of in the opposite direction.

c) The change is not casual, but based on cause-effect relationship

5.4.2. Karl Pearson's Method

Pearson's method is also known as product-moment correlation. Of the several mathematical methods of measuring correlation, Karl Pearson's method is popularly known as Pearson's

coefficient of correlation. This method is widely used in practice. The Pearson's coefficient of correlation is denoted by "r". Symbolically, it may be expressed as below –

$$r = \frac{\sum d \times dy}{n \times a_x a_y}$$

Where

R= Co-efficient of correlation

d_x = Deviation of item values of x series from its arithmetic average

d_y = deviation of item values of y series from its arithmetic average

$d_x d_y$ = Product of d_x and d_y

n = Number of pairs

σ_x = S.D of x series

σ_y = S.D of y series

In this method, deviation has been calculated from the actual arithmetic average, which entails a lot of calculation work. A shortcut method is also used in which deviations are calculated from the assumed average.

The shortcut method is –

$$r = \frac{\sum d_x d_y - n\left(\frac{\sum d_x}{n}\right)\left(\frac{\sum d_y}{n}\right)}{n \times a_x \times a_y}$$

If the standard deviation is not to be calculated separately the same formula would be as under –

$$r = \frac{\sum d_x d_y - n\left(\frac{\sum d_x}{n}\right)\left(\frac{\sum d_y}{n}\right)}{n \times \sqrt{\left(\frac{\sum d_x^2}{n}\right) - \left(\frac{\sum d_x}{n}\right)^2} \times \sqrt{\left(\frac{\sum d_y^2}{n}\right) - \left(\frac{\sum d_y}{n}\right)^2}}$$

Illustration 1

The heights of 10 father and their sons are given below. Find out if there is any correlation between the two.

Height of father – 63, 64, 65, 67, 68, 69, 70, 71, 71 (in inches)

Height of sons – 65, 63, 63, 65, 67, 67, 71, 68, 69 (in inches)

Solution :

Height of father x	Dev. From ass. av 68 (d_x)	Sq. of dev (d_x^2)	Height of sons (y)	Dev. From ass. av 67 (d_y)	Sq. of d $(d_y^{y_2})$	Product of $(d_x)^2$ (d_y) $(d_x.d_y)$
63	-5	25	65	-2	4	10
64	-4	16	63	-4	16	16
65	-3	9	63	-4	16	12
67	-1	1	65	-2	4	2
68	0	0	67	0	0	0
69	+1	1	67	0	0	0
69	+1	1	68	+1	1	1
70	+2	4	71	+4	16	8
71	+3	9	68	+1	1	3
71	+3	9	69	+2	4	6
	-3	75		-4	62	58

$$r = \frac{\sum d_x d_y - n\left(\frac{\sum d_x}{n}\right)\left(\frac{\sum d_y}{n}\right)}{n \times \sqrt{\left(\frac{\sum d_x^2}{n}\right) - \left(\frac{\sum d_x}{n}\right)^2} \times \sqrt{\left(\frac{\sum d_y^2}{n}\right) - \left(\frac{\sum d_y}{n}\right)^2}}$$

$$= \frac{58 - \left(10 \times \frac{-3}{10} \times \frac{-4}{10}\right)}{10 \times \sqrt{\left(\frac{75}{10}\right) - \left(\frac{-3}{10}\right)^2} \times \sqrt{\left(\frac{62}{10}\right) - \left(\frac{-4}{10}\right)^2}}$$

$$= \frac{58 - 10 \times .3 \times -.4}{10 \times \sqrt{7.5 - .09 \times \sqrt{6.2 - .16}}}$$

$$= \frac{56.8}{10 \times \sqrt{7.41} \times \sqrt{6.04}}$$

$$= \frac{56.8}{66.9}$$

$$= 0.85$$

Probable error

In order to understand the significance of the correlation, probable error is also calculated. The formula for calculating probable error is-

$$P.E. = .6745 \frac{1 - r^2}{\sqrt{n}}$$

$$= .6745 \frac{1 - (85)^2}{\sqrt{n}}$$

$$= .6745 \frac{1 - .7225}{\sqrt{n}}$$

$$= .6745 \times .2775 = .06$$

Interpretation of Correlation

The Co-efficient of correlation measures the degree of relationship between two sets of figures. The reliability of estimates depends upon the closeness of the relationship. Therefore, it is imperative that utmost care can be taken while of co-efficient of correlation is calculated, otherwise fallacious conclusion can be drawn.

It is to note that the interpretation of the co-efficient of correlation depends very mush on experience. The full significance of 'I' can only be grasped after working out a member of correlation

problems and seeing the kind of data that give rise to various values of *r*. The investigator must know his data thoroughly in order to avoid errors of interpretation and emphasis. However, the following general rules are given which would help in interpreting the value of '*r*'

- When r = + 1, it means there is perfect positive relationship between the variables.
- When r = - 1, it means there is perfect negative relationship between the variables.
- When r = 0, it means there is no relationship between the variables i.e. the variables are uncorrelated.

Review Reading

1. What do you mean by correlation in the language of research ?
2. What is variable ?
3. Define empiricism
4. What are the specific characteristics of social variable ?
5. What do mean by statistical association ?

Important Terms

Direct correlation
Inverse correlation
Cause-effect relationship
Arithmetic average
Assumed average
Error in statistics
Subjectivity, objectivity
Quantitative data
Qualitative data

Answer the following questions

1. Explain what do you mean by positive correlation and negative correlation. Give examples.

2. Define Karl Pearson's co-efficient of correlation. What does it measure ?
3. "Correlation is based on cause effect relationship" Explain
4. Discuss how with the use of statistics reliable and valid knowledge could be achieved.

5.4.4. Method of Rank Difference

The Rank Difference method was introduced by British Psychologists Charles Edward Spearman in 1904. This method was specially useful when factors such as-leadership, ability, performance, beauty etc. are measured in terms of quantitative data. In this method individual items of rank measurement is not taken into consideration, but the views of the whole group is taken. The rank difference methods of correlation may be found useful for relatively small sample-usually less than 30 cases.

From mathematical point of view, rank difference method is less satisfactory than the product –moment method. This method is particularly useful in case of social phenomena where actual measurement cannot be given but rank can be fixed. But it is agreed that for measuring qualitative phenomena, this method is suitable.

The Greek letter rho,p is used in computing co-efficient of correlation by Rank difference.

5.4.5 Measurement of Ranks

A very simple ordering of data is in the form of rank. The very process of ranking may be subjective for quite sometime such ranking may not be very precise, but very useful. For this reason, many sociological concepts cannot be manipulated in true sense of objectivity. But a considerable volume of social research is founded on subjective base. Thus, occupation may be ranked by prestige; fellow students may be ranked in order of preference; race and

nationality may be ranked by favorable or unfavourable prejudice; picture may be ranked by aesthetic value etc.

On many condition, units of measurement are available such as-size of city, birth rate etc. But even than ranking is required to know the exact precision. Rank order is based on the concept of "more or less" even though the "unit of measure" is subjective. We may conceive of the collective items ranks as a quantitative, rather than qualitative variable. When two sets of ranks vary together, we speak of rank-order correlation.

Rank difference Method of Correlation

Social phenomena are by nature qualitative, and qualitative measurement of the phenomena is not easily possible. For example- efficiency, intelligence etc. in such situation, social scientists apply rank difference method of correlation for finding out the degree of correlation.

The following formula is used for calculating the coefficient of correlation by method of rank difference.

$$r = 1 - \frac{6\left(\sum d^2\right)}{n\left(n^2 - 1\right)}$$

Where, d^2 = square of the differences of the rank of individual pair of two series

n= number of items

Here are some illustrations

Illustration 1

Following are the data of ages of husbands and wives. Calculate the co-efficient of correlation by method of rank difference.

Age of husbands – 18, 19, 25, 27, 30, 35, 36, 40, 42, 44

Age of wives - 16, 17, 24, 23, 22, 30, 28, 36, 40, 41

Solution :

Age of husband	Rank R_1	Age of wives	Rank R_2	Difference of rank $R_1 R_2$	Square of difference (d^2)
18	1	16	1	0	0
19	2	17	2	0	0
25	3	24	3	2	0
27	4	23	4	0	0
30	5	22	5	2	0
35	6	30	6	1	0
36	7	28	7	1	0
40	8	36	8	0	0
42	9	40	9	0	0
44	10	41	10	0	0
					$\sum d^2 = 10$

Formula $$r = 1 - \frac{6\left(\sum d^2\right)}{n\left(n^2 - 1\right)}$$

$$= 1 - \frac{6 \times 10}{10\,(100 - 1)}$$

$$= 1 - \frac{60}{990}$$

$$= 1 - .06$$

$$= .94 \text{ (approx)}$$

Illustration 2

10 competition in a beauty contest are ranked by two judges as follows –

1st judge – 1, 6, 5, 10, 3, 2, 4, 9, 7, 8

2nd judge - 6, 4, 9, 8, 1, 2, 3, 10, 5, 7

Find out by method of rank difference how far the opinions of two judges are similar.

Solution :

Rank by judge 1	**Ranks by judge 2**	**Difference (d)**	**Square of differences (d^2)**
1	6	5	25
6	4	2	4
5	9	4	16
10	8	2	4
3	1	2	4
2	2	0	0
4	3	1	1
9	10	1	1
7	5	2	4
8	7	1	1
Item 10	$\sum d^2 = 60$		

Formula $r = 1 - \frac{6\left(\sum d^2\right)}{n\left(n^2 - 1\right)}$

$= 1 - \frac{6 \times 10}{10\,(100 - 1)}$

$= 1 - \frac{60}{990}$

= 1 - .36

= .64 (approx)

The opinion of the two judges are fairly similar.

You can now take the exercise below

Exercise – 1

1. Two different judges ranked the efficiency of social worker as follows -

Social worker	Judged by A	Judged by B
A	3	4
B	9	7
C	6	
D	5	8
E	1	
F	2	3
G	4	2
H	7	5
I	8	10
J	10	9

Find out the rank difference if any between the judgements of two judges by the method of rank correlation.

5.4.6 Utility of Calculating Rank Difference

Though data are quantitative in its character, its subjective elements cannot be ignored. Social scientists who deal with rough and coarse quantitative data, may use rank difference method quite effectively in finding out the subjective character of social variables. Ranking of preferences, aesthetic judgement of electors and judges – all be correlated to determine the degree of agreement. Occupation may also be ranked by socio-economic levels.

Key terms

Correlation : a relationship between two or more variables.

Variable: a concept whose value changes from case to case.

5.5. TESTING OF STATISTICAL HYPOTHESIS : CHI-SQUARE (χ^2)

In the proceeding Block 2 we have discussed at length about hypothesis in social research. So to repeat it again, hypothesis is an assumption, made on the basis of probabilities, shrewed guesses and profound hunches. When such hunches penetrate to a possible explanation, it becomes possible for fruitful investigation, it becomes a hypothesis.

Sometimes, in many situations, it is not possible to make any rigid assumption about the population from which samples are drawn. Because of this limitation, we find a development of a group of alternative techniques which are known as non-parametric tests. It is *non-parametric* because no clear assumption about the parameters of population from which samples are drawn are made. Chi-*Square test* is one such example of non-parametric test.

The theory of non-parametric test was initially developed in the middle of 19^{th} century, but it was only after 1945, that non-parametric test came to be widely used. Firstly, it was originated in sociological and psychological researches. Non-parametric test today, are very popular in behavioural sciences. These are two general types of statistical inference.

a) Estimation and

b) Hypothesis testing

Estimation, which begins without any state assumption about the value of the parameter and merely seeks to estimate on the basis of description.

Hypothesis testing, which begins with a hypothesis about the parameter and then use the sample data to check the plausibility of the statement. We formulate our hypothesis is about a parameter in advance of the collection of sample data, which is often used to test that hypothesis.

For example – we hypothesize that "average voting rate of rural area does not differ from that of urban area". We begin with that supposition and then we take an appropriate sample from rural and urban area, compare the two means in the prescribed manner and finally reach a decision whether in the light of sample difference, the hypothesis should be *accepted or rejected.*

5.5.1 χ^2 defined

The χ^2 test (pronounced as chi-square test) is one of the simplest and most widely used non-parametric tests in statistical work. The symbol χ^2 is the Greek letter chi, which is pronounced as '*ki*' as in 'kite'. The χ^2 test was first used by Karl Pearson in the year 1890. The quantity χ^2 describes the magnitudes of the discrepancy between theory and observation. Chi-square is a test which provides us with a method to evaluate whether the frequencies which have been *empirically observed* differ significant form those which would be *expected* under a certain set of theoretical assumption. In a very simple meaning it can be illustrated thus-

- Any social situation can be observed firsthand and gather information about the situation.
- There are certain functional situation which are to be expected
- Such expectations are guided by certain ideal existance.

For example – in a certain region of a country, 65% respondents cast their votes in favour of a national political party which were not supposed to be the real expectation of responses. In the context of prevailing political situation (ideal situation), the casting of votes from those respondents were expected somewhat differently. Under such condition, a difference may arise between the *observed situation (observed frequency)* and the *expected situation (expected frequency).* This is to be investigated in the procedure of chi-square test. So, the description on the magnitude of the

discrepancy between the theory and observation can be said to be chi-square test.

χ^2 defined as

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

Where O = observed frequencies

E = expected frequencies

Steps

To determine the value of x^2, the steps are:

- Calculate the expected frequencies for each cell of observed frequencies. Generally, the expected frequency for any cell can be calculated from the following equation:

$$E = \frac{RT \times CT}{N}$$

Where, E = Expected frequency

RT = Row total for the row containing the cell

CT = Column total for the column containing the cell

N = total number of observation

The illustration below will describe the steps for calculating χ^2

Illustration 1

In an anti-malaria campaign in a certain area, quinine was administered to 812 person out of a total of 3248. The number of fever cases were found as shown below –

Table A

Treatment	Fever	No, Fever	Total (row)
Quinine	20	792	812
No Quinine	220	2216	2436
Total column)	240	3008	3248

Discuss the use usefulness of quinine in checking malaria

Step 1

First let us calculate expected frequencies for each cell of the observed frequencies.

Step 2

Let us take the hypothesis that, - 'Quinine is not effective in checking malaria'

The expected frequency to the first row first column is 60.

Calculation procedure is

$$E = \frac{RT \times CT}{N}$$

$$= \frac{812 \times 240}{3248} = 60$$

Thus, the expected frequencies for the observed data will be –

60 752 812

180 2258 2436

220 3008 3248

Step 3

Observed O	**Expected E**	**$(O-E)^2$**	**$(O-E)^2/E$**
20	60	1600	26.667
220	180	1600	8.889
792	752	1600	2.128
2216	2256	1600	0.709

$$\sum (O-E)^2 / E = 38.393$$

Hence

$$\chi^2 = \sum \frac{6(O-E)^2}{E}$$

$$= 38.393$$

Calculation of Degrees of Freedom (df)

In the table A above, the degree of freedom will be –

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

$$= (2 - 1)(2 - 1)$$

$$= 1$$

There are 2 rows and 2 columns in the table A.

Hence-degrees of freedom of this table is 1

There are various sets of calculated Chi-square with its degrees of freedom (df) and showing probability levels. The tabulated value of χ^2 at 1 df and at 0.05 p.c. probability level is 3.84. The calculated value of χ^2 is 38.393 which is greater than the table value i.e. 3.84

So, the hypothesis presumed earlier is rejected.

Hence, conclusion is "Quinine is useful in checking malaria".

The hypothesis presumed earlier is call *Null hypothesis.*

5.5.2 The Null Hypothesis (H_0)

In the science of statistics, it has become a convention that there should be a null hypothesis to open an investigation. The null hypothesis is symbolized as H_o (pronounced as 'H not'). In its most current usages, the *hypothesis hold that two or more given samples have come from statistically identical populations and that therefore, any observed difference between such samples is merely a chance variation.* The aforesaid hypothesis that rurality and urbanity don't necessarily differ in their average voting rates would therefore be a typical null-hypothesis.

Essentially, the null hypothesis is set up to be nullified, but even then, null hypothesis is usually launched with an expectation, though the term implies that it will be nullified. Because of the diversities of social situation,, it is quite often difficult to obtain elaborate explanation. Hence, the researcher proceeds with the

possibility that there is no linear relation between variables. But, it is not suggestive that exploration about social interaction would never be initiated for the purpose of testing *null hypothesis of zero correlation* that is, to demonstrate the essential unrelatedness of things. Quite the contrary, the social analyst is moved by an interest in establishing functional relationship.

The Null Hypothesis : Some issues

The null hypothesis has undergone some shifts in both meaning and implementation since it was first launched by Ronald Fisher in 1920. There is some diversity in the employment of the term in the standard texts, resulting in some confusion among the readers. The confusion, in fact has its principle sources of two phases of the concept such as –

a) It has undergone some modification in meaning which had made "null" equivalent to "no difference" and "zero". The concept has thereby, lost its original inclusiveness, applicable to any statistical hypothesis subject to probabilistic testing.

b) It has often succumbed to the *'fallacy of mixed significance level', by* which the .01 and .05 levels are routinely and casually applied as if the traditions were legitimate for all types of decision. Such an arbitrary norm was not in the spirit of the original practice.

But, when we come into the essential nature of these controversies we see that they do not seem as destructive as might appear, because-

* they are on the level of nomenclature rather than of essence; and

* they are matter of detailed strategic implementation rather than erroneous application.

Thus, it is apparent that the concept has been in a state of flux, stimulated by the expanding use of the term. Chi-square test has several associated terminologies which must be understood clearly in computing chi-square. Some of such terms are -

Degrees of freedom (df)

Level of significance

5.5.3 Degrees of Freedom (df)

The collected data are placed in the cells of the table in rows and columns. The cell frequencies of all columns but one i.e. (c – 1) and all rows but one i.e. (r – 1) are assigned. The number of degrees of freedom for all the cell frequencies i.e. (c-1) (r-1) is calculated. Thus, in a 2 × 2 table (excluding row and column), the degrees of freedom is

$$df = (2 – 1)(2 – 1)$$

$$= 1$$

and for 3 × 3 table,

$$df = (3 – 1)(3 – 1)$$

$$= 2 \times 2 = 4$$

and if there are 2 rows and 3 columns then the df is

$$= (2 – 1)(3 – 1)$$

$$= 1 \times 2$$

$$= 2$$

The following figure will illustrate

Fig:1

	Column 1	Column 2	
Row 1	a	c	Row total 1
Row 2	b	d	Row total 2
	Col. Total 1	Col. Total 2	

5.5.4 The level of significance

When we calculate expected frequency we find a difference between the observed and the expected frequencies. The difference arises due to various reason – Sampling fluctuation or some other reasons. The chi-square test is used to evaluate whether the difference

between observed and expected frequencies is due to sampling fluctuation. If it is due to sampling fluctuation then the difference is considered insignificant or if it is due to some other reasons-it is significant.

Usually, null hypothesis states that there is no relationship between variables. The null hypothesis assumes that there is sampling error and the difference is due to chance above. Then the probability of occurrence of such a difference is determined. The *probability indicates the extent of reliance* that we can place on the inference drawn.

The table values of chi-square are available at various probability levels. These levels are called Levels of Significance. We can find out from the table value of Chi-square at certain Level of Significance. Usually, the value of Chi-Square at 0.05 or 0.01 levels of significance from the given degree of freedom is seen from the table and is compared with observed value of chi-square. If the observed value of chi-square is more than the table value at .05 probability level, it means the difference is significant. Thus, the level of significance or in significance is determined.

5.6. LET USSUM UP

1. Average or locational average is the value which lies between the highest and the lowest extreme values. The calculation of average is called as measure of central tendency, statistically average is called mean.
2. Mean is the simplest average to understand and easiest to compute
3. Median is the size of middle number. It is also called positional average.
4. Heterogeneity rather than homogeneity of social data are basic to statistical calculations. Heterogeneity within the group gives rise to variation from one to another. The extent of variation is the measures of dispersion.

5. Absolute dispersion is measured through calculation of standard deviation.
6. Correlation is the measurement of relation between two or more social situations, groups or classes.
7. There may be probable difference between theoretical assumption and empirical world. The difference between theory (expected or established notion) and empiricism (observed situation) are examined with the help of Chi-square test (χ^2 Test).

Terms to be remembered

Parametric test

Non-parametric test

Observed frequency

Expected frequency

Zero correlation

Sampling fluctuation

Probability

Frequency cell

Row total

Column total

Answer the following questions

1. Explain the significance of calculating central tendency ?
2. Discuss how mean and median are different
3. Discuss the characteristics of median.
4. What are the specific advantages of calculating mode? Explain
5. The relationship between occupational choice and rural-urban background of 112 students is shown below -

Rural-Urban background	Occupational Choice			
	Medical	Engineering	Legal	Row Total
Rural	10	26	6	42
Urban	38	20	12	70
Column Total	48	46	18	112

Find out using χ^2 test how far there is a relationship between rural-urban background and occupational choice.

6. Discuss the significance of measuring dispersion.
7. Bring out the steps for calculating standard deviation.
8. "Standard deviation is superior statistical calculation than calculation of Mean, Median and mode" Discuss.
9. What are the uses of chi-square test in social science research ?
10. Discuss the functions of null hypothesis.
11. Discuss how degrees of freedom is determined.
12. Differentiate between statistical significance and explanatory significance.

Suggested for further reading

S.P. Gupta : *Statistical Methods*, New Delhi, 1980.

Mueller, J.H and K. F Schuessier : *Statistical Reasoning in Sociology*, Oxford, 1961.

Lal Das, D.K : *Practice of Social Research*, Rawat Publication, 2000.

Young, P.V. : *Scientific Social Survey and Research*, Prentice Hall, 1953.

****