Research Methodology for Management Decision

This book is a part of the course by **uts**, Pune.
This book contains the course content for Research Methodology for Management Decision.

© **uts**, Pune
First Edition 2013

The content in the book is copyright of **uts**. All rights reserved.
No part of the content may in any form or by any electronic, mechanical, photocopying, recording, or any other means be reproduced, stored in a retrieval system or be broadcast or transmitted without the prior permission of the publisher.

uts makes reasonable endeavours to ensure content is current and accurate. **uts** reserves the right to alter the content whenever the need arises, and to vary it at any time without prior notice.

Published by
uts
Bavdhan, Pune - 411021

Website : www.utsglobal.edu.in
Tel : +91-20-41034800, +91 9011067684

Index

I. Content II

II. List of Figures VII

III. List of Tables VIII

IV. Case Study 175

V. Bibliography 182

VI. Self Assessment Answers 185

Book at a Glance

Contents

Chapter I **1**
Research Methodology - An Introduction **1**
Aim 1
Objectives 1
Learning outcome 1
1.1 Introduction 2
1.1.1 Defining Research 2
1.1.2 Objectives of Research 2
1.1.3 Motives in Research 2
1.1.4 Characteristics of Social Research 3
1.1.5 Limitations of Social Research 3
1.2 Types of Research 4
1.2.1 Fundamental, Pure or Theoretical Research 4
1.2.2 Applied Research 4
1.2.3 Exploratory Research 4
1.2.4 Descriptive Research 5
1.2.5 Action Research 5
1.2.6 Evaluation Research 5
1.2.7 Experimental Research 5
1.2.8 Empirical Research 5
1.2.9 Survey Research 6
1.2.10 Qualitative Research 6
1.2.11 Quantitative Research 6
1.2.12 Field Investigation Research 6
1.2.13 Ex-post Facto Research 6
1.2.14 Historical Research 6
1.2.15 Some Other Types of Research 7
1.3 Research Approaches 7
1.3.1 Descriptive vs. Analytical 7
1.3.2 Applied vs. Fundamental 7
1.3.3 Quantitative vs. Qualitative 8
1.3.4 Conceptual vs. Empirical 8
1.4 Research Approaches and Significance of Research 8
1.5 Research Methods versus Methodology 10
1.6 Research and Scientific Method 12
1.7 Research Process 13
1.8 Criteria of Good Research 20
1.9 Problems Encountered by Researchers in India 21
Summary **23**
References **23**
Recommended Reading **23**
Self Assessment **24**

Chapter II **26**
Research Problem **26**
Aim 26
Objectives 26
Learning outcome 26
2.1 Defining the Research Problem 27
2.1.1 What is a Research Problem? 27
2.1.2 Selecting a Problem 28
2.1.3 Necessity of Defining a Problem 28
2.1.4 Technique Involved in Defining a Problem 29

2.2 Research Design ... 31
2.2.1 Meaning of Research Design ... 31
2.2.2 Need for Research Design ... 31
2.2.3 Features of a Good Design ... 32
2.2.4 Important Concepts Related to Research Design ... 32
2.3 Different Research Designs ... 34
2.4 Basic Principles of Experimental Designs ... 39
Summary ... 41
References ... 41
Recommended Reading ... 41
Self Assessment ... 42

Chapter III ... 44
Sampling Fundamentals and Sampling Design ... 44
Aim ... 44
Objectives ... 44
Learning outcome ... 44
3.1 Introduction ... 45
3.1.1 Need for Sampling ... 45
3.1.2 Definition ... 45
3.1.3 Important Sampling Distributions ... 48
3.1.4 Essentials of Good Samples ... 50
3.1.5 Central Limit Theorem ... 50
3.2 Methods of Sampling ... 51
3.2.1 Random Sampling Methods ... 51
3.2.2 Non-Random Sampling Methods ... 55
3.3 Sampling Theory ... 57
3.3.1 Sandler's A Test ... 57
3.3.2 Concept of Standard Error ... 58
3.4 Sample Size and its Discrimination ... 61
3.4.1 Nature of Universe ... 61
3.5 Sampling Design ... 62
3.5.1 Census and Sample Survey ... 62
3.5.2 Implications of a Sample Design ... 62
3.5.3 Steps in Sample Design ... 62
3.5.4 Criteria of Selecting a Sampling Procedure ... 63
3.5.5 Characteristics of a Good Sample Design ... 64
3.6 Different Types of Sample Designs ... 64
3.6.1 Non-Probability Sampling ... 65
3.6.2 Probability Sampling ... 66
3.6.3 Determination of Appropriate Sampling Design ... 66
Summary ... 68
References ... 68
Recommended Reading ... 68
Self Assessment ... 69

Chapter IV ... 71
Measurement and Scaling Techniques ... 71
Aim ... 71
Objectives ... 71
Learning outcome ... 71
4.1 Introduction ... 72
4.1.1 Measurement in Research ... 72
4.1.2 Meaning of Attitude ... 73
4.2 Measurement Scales ... 73

4.2.1 Selection of Measurement of Scale ... 73
4.2.2 Measurement of Scales ... 74
4.3 Sources of Error in Measurement ... 75
4.3.1 Respondent ... 75
4.3.2 Situation ... 76
4.3.3 Measurer ... 76
4.3.4 Instrument ... 76
4.4 Tests of Sound Measurement ... 76
4.4.1 Test of Validity ... 76
4.4.2 Test of Reliability ... 77
4.4.3 Test of Practicality ... 77
4.4.4 Technique of Developing Measurement Tools ... 78
4.5 Scaling ... 78
4.5.1 Meaning of Scaling ... 79
4.5.2 Scale Classification Bases ... 79
4.5.3 Important Scaling Techniques ... 80
4.6 Scale Construction Techniques ... 81
4.6.1 Arbitrary Scales ... 82
4.6.2 Differential Scales (or Thurstone-type Scales) ... 82
4.6.3 Summated Scales (or Likert-type Scales) ... 83
4.7 Multidimensional Scaling ... 85
Summary ... 87
References ... 87
Recommended Reading ... 87
Self Assessment ... 88

Chapter V ... 90
Methods of Data Collection ... 90
Aim ... 90
Objectives ... 90
Learning outcome ... 90
5.1 Introduction ... 91
5.2 Meaning and Need for Data ... 91
5.2.1 Source of Data ... 91
5.2.2 Primary and Secondary Data ... 92
5.2.3 Methods and Techniques of Data Collection ... 92
5.2.4 Collection of Primary Data ... 92
5.2.5 Observation Method ... 93
5.2.6 Interview Method ... 94
5.3 Collection of Data through Questionnaires ... 97
5.3.1 Merits and Demerits of a Questionnaire ... 97
5.3.2 Main Aspects of a Questionnaire ... 98
5.3.3 Essentials of a Good Questionnaire ... 100
5.3.4 Functions of Questionnaire ... 100
5.3.5 Problem of Response Questionnaires ... 100
5.3.6 Pilot testing or Pre-testing the Questionnaire ... 101
5.3.7 Specimen Questionnaire ... 101
5.4 Collection of Data through Schedules ... 103
5.4.1 Purpose of Schedule ... 103
5.4.2 Essentials of a Good Schedule ... 104
5.4.3 Difference between Questionnaires and Schedules ... 104
5.5 Collection of Data through Interview ... 105
5.5.1 Purpose of Interview ... 105
5.5.2 Kinds of Interviews ... 105
5.5.3 Technique of Interview ... 106

5.6 Collection of Data through Observation 106
5.6.1 Kinds of Observation 106
5.6.2 Importance of Observation in Social Research 107
5.6.3 Limitations of Observation as a Method of Social Research 108
5.7 Modern Methods of Data Collection 108
5.8 Other Methods of Data Collection 108
5.9 Collection of Secondary Data 110
5.9.1 Reliability of Data 111
5.9.2 Suitability of Data 111
5.9.3 Adequacy of Data 111
5.10 Selection of Appropriate Method for Data Collection 111
5.11 Case Study Method 112
5.12 Guidelines for Constructing Questionnaire/Schedule 115
5.13 Choice of Suitable Method 115
Summary 116
References 116
Recommended Reading 117
Self Assessment 118

Chapter VI 120
Processing and Analysing of Data 120
Aim 120
Objectives 120
Learning outcome 120
6.1 Introduction 121
6.1.1 Processing Operations 121
6.1.2 Some Problems in Processing 125
6.1.3 Elements / Types of Analysis 126
6.1.4 Statistics in Research 127
6.2 Data Measuring 128
6.2.1 Measures of Central Tendency 128
6.2.2 Measures of Dispersion 130
6.2.3 Measures of Relationship 131
6.3 Regression Analysis 134
6.3.1 Simple Regression Analysis 134
6.3.2 Multiple Correlation and Regression 135
6.3.3 Partial Correlation 136
6.4 Other Measures 136
6.4.1 Index Numbers 136
6.4.2 Time Series Analysis 137
Summary 139
References 139
Recommended Reading 139
Self Assessment 140

Chapter VII 142
Interpretation and Report Writing 142
Aim 142
Objectives 142
Learning outcome 142
7.1 Introduction 143
7.2 Meaning of Interpretation 143
7.2.1 Essentials of Interpretation 143
7.2.2 Technique of Interpretation 143
7.2.3 Precautions in Interpretation 144

7.3 Report Writing 144
7.3.1 Purpose of Report Writing 145
7.3.2 Meaning of Report 145
7.3.3 Importance of Social Research Report 145
7.3.4 Significance of Report Writing 145
7.3.5 Different Steps in Writing Report 146
7.3.6 Layout of the Research Report 147
7.4 Types of Reports 148
7.5 Mechanics of Writing A Research Report 150
7.6 Characteristics of a Good Research Report 154
7.7 Precautions for Writing Research Reports 154
7.8 Presentation of Research Report 155
7.9 Evaluation of the Research Report 156
Summary 158
References 158
Recommended Reading 158
Self Assessment 159

Chapter VIII 161
Computer – Its Role in Research 161
Aim 161
Objectives 161
Learning outcome 161
8.1 Introduction 162
8.2 Significance of Computers 162
8.2.1 Evolution of Computers 162
8.2.2 The Computer System 163
8.2.3 Important Characteristics of Computers 164
8.3 Introduction to Binary Number System 165
8.3.1 Computations in Binary System 166
8.3.2 Binary Fractions 166
8.4 Computer Applications 167
8.5 Computers in Research 168
8.5.1 Phases of Research Process 168
8.5.2 Further Uses of Computers 170
8.6 Use of Internet in Research 170
8.7 Importance of Internet 170
Summary 172
References 172
Recommended Reading 172
Self Assessment 173

List of Figures

Fig. 1.1 Motives of doing research 3
Fig. 1.2 Research process in flow chart 13
Fig. 2.1 Research design in case of descriptive and diagnostic research studies 36
Fig. 2.2 Principles of experimental design 39
Fig. 3.1 Sampling error 47
Fig. 3.2 Methods of random sampling 51
Fig. 3.3 Methods of non-random sampling 56
Fig. 4.1 Classification of measurement of scales 74
Fig. 4.2 Technique of developing measurement tools process 78
Fig. 5.1 Types of data sources 91
Fig. 5.2 Types of control exercised in the observation 107
Fig. 6.1 Essentials of statistical table 124
Fig. 8.1 The computer system 163
Fig. 8.2 Phases of research process 169

List of Tables

Table 1.1 Methods and techniques of data collection 11
Table 2.1 Difference between research designs 38
Table 3.1 Types of sample designs 65
Table 4.1 Different scales for measuring attitudes of people 82
Table 7.1 List of abbreviations 153
Table 8.1 Table for conversion 166
Table 8.2 Important applications and uses of computers 168

Chapter I

Research Methodology - An Introduction

Aim

The aim of this chapter is to:

- define the term 'research'
- explain the objectives of research
- describe motivation in research

Objectives

The objectives of this chapter are to:

- elucidate characteristics of social research
- discuss the limitations of research
- elaborate different types of research

Learning outcome

At the end of this chapter, you will be able to:

- discuss research approaches
- understand the criteria for good research
- explain research process

1.1 Introduction

Research is an academic activity and as such the term should be used in a technical sense. According to Clifford Woody, research comprises defining and redefining problems, formulating hypothesis or suggested solutions, collecting, organising and evaluating data, making deductions and reaching conclusions and at last carefully testing the conclusions to determine whether they fit the formulating hypothesis.

1.1.1 Defining Research

Research refers to a search for knowledge. It can be defined as "a scientific and systematic search for applicable information on a specific topic". In fact, research is an art of scientific investigation.

- Redman and Mory define research as "a systematised effort to gain new knowledge."
- D. Slesinger and M. Stephenson in the Encyclopaedia of Social Sciences define research as "the manipulation of things, concepts or symbols for the purpose of generalising to extend, correct or verify knowledge, whether that knowledge aids in construction of theory or in the practice of an art."
- Research is, thus, an original contribution to the existing stock of knowledge leading to advancement in knowledge base. It is the pursuit of truth with the help of study, observation, comparison and experiment. In short, the search for knowledge through objective and systematic method of finding solution to a problem is research.
- The systematic approach concerning generalisation and the formulation of a theory is also research. As such, the term 'research' refers to the systematic method consisting of enunciating the problem, formulating a hypothesis, collecting the facts or data, analysing the facts and reaching certain conclusions either in the form of solutions(s) towards the concerned problem or in certain generalisations for some theoretical formulation.

1.1.2 Objectives of Research

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

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

1.1.3 Motives in Research

The possible motives for doing research may be one or more of the following:

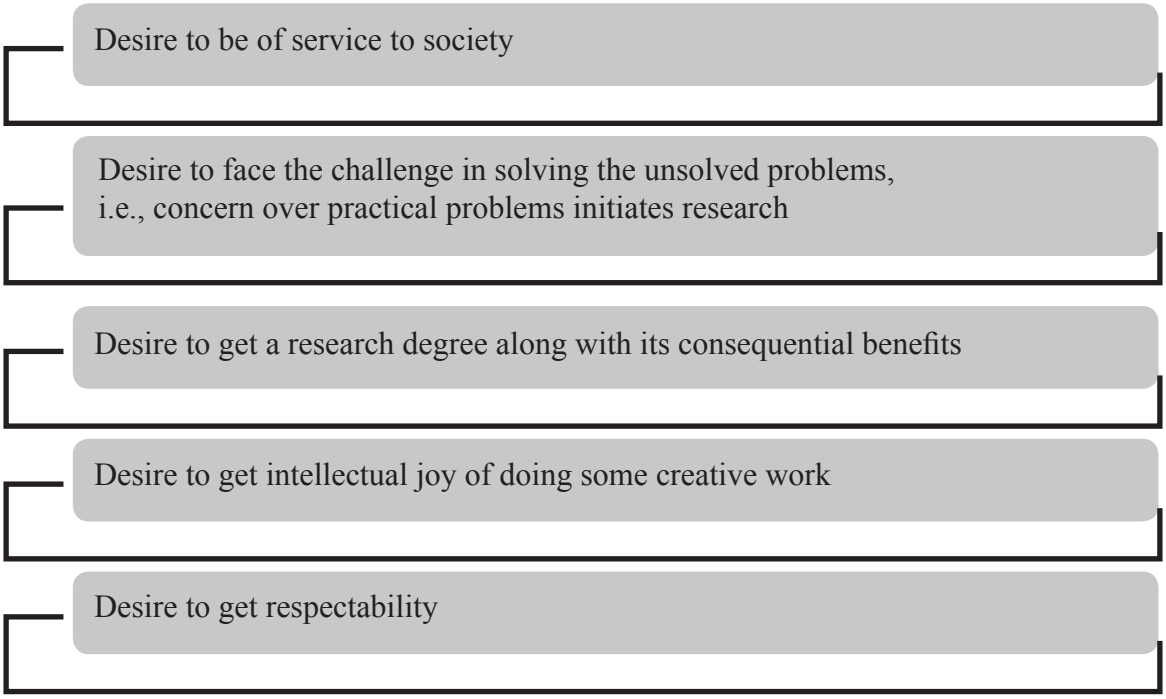


Fig. 1.1 Motives of doing research

However, this is not an exhaustive list of factors motivating people to undertake research studies. Many more factors, such as, directives of government, employment conditions, curiosity about new things, desire to understand causal relationships, social thinking and awakening and many other alike factors will motivate (or at times compel) people to perform research operations.

1.1.4 Characteristics of Social Research

Characteristics of social research are given below.

- **Study of human behaviour:** Social research deals with the social phenomena. It studies the behavior of human beings as members of society, and their feelings, response, attitudes under different circumstances.
- **Discovering new facts:** In social science, a lot of research is being carried on for both purposes and has resulted in the discovery of new facts as well as modification of old concepts. Constant verification of the old concepts is also needed, especially in case of dynamic sciences. Verification is needed because of two reasons.
  - Firstly, there may be an improvement in the technique of research and it is necessary to test the old concepts by this improved technique.
  - Secondly, the phenomena under study might have undergone a change and it may be required to test the validity of old concepts in the changed circumstances.
- **Scientific analysis:** Social research tries to establish causal connection between various human activities. It is really very interesting to note whether various complex human activities, are being performed only at random without any sequence, law or system behind them. At first causal look at varied human behavior, attitudes, moods and temperaments, the presence of any system may appear to be an impossibility, but a close and patient study of different cases, disclose the truth, that most of them are not merely jumbled as they appear but motivated by definite rules, perfect system and universal laws.

1.1.5 Limitations of Social Research

Limitations of social research are explained below.

- During the courses of conducting a research, one has to constantly guard against bias, subjectivity and inaccuracy. Yet, it is difficult to totally avoid them and a minimum of bias and inaccuracy is always present in any research. Social research happens to be no exception and it is replete with instance of problematic situation where some errors cannot be possibly avoided.

- In social survey, the traders, farmers and wage earners from whom the data were collected, may not reveal their business secrets, which also set certain limitations. The respondents are not maintaining any record of their expenses and receipts. Hence, the data provided by them from their memory may have limitations. The researcher has to carefully minimise such errors by educating the respondents about the scope of the study and with all possible verification of the data collected in the field.

1.2 Types of Research

Researches have been classified differently, depending upon the approach, the purpose and the nature of a research activity. The basic types of research are as follows:

1.2.1 Fundamental, Pure or Theoretical Research

This type of research is original or basic in character. An imaginative and particular research worker, with his qualities of honesty and integrity and his lust for the search of the truth, makes persistent and patient efforts to discover something new to enrich the human knowledge in a fundamental fashion. Such research is known as fundamental or pure. Fundamental research can take shape in two different ways:

- **Discovery of a new theory:** Fundamental research may be entirely a new discovery, the knowledge of which has not existed so far. The researcher is often born-genius, has a sharp intellect, is thirsty of knowledge and eventually has an ocean of knowledge in his possession and emerges gems of knowledge. This discovery may have nothing to do with an existing theory.
- **Development of the existing theory**: Since theory is always based on assumptions, there often exists enormous scope for altering or formulating new set of assumptions and adding new dimensions to the existing theory. There also exist the possibilities of re-interpretation of the theory that has already been developed. A researcher may, as well take off from the existing theories and come out with a new one of his own. The assumptions of a theory should always be well defined and plausible. Relaxing assumptions, altering them or making new ones altogether depends upon how a researcher views the existing theory.

1.2.2 Applied Research

This type of research is based on the application of known theories and models to the actual operational fields or populations. The applied research is conducted to test the empirical content or the basic assumptions or the very validity of theory under given conditions. Applied research contributes to social science by:

- providing the kind of convincing evidence of the usefulness to society which is necessary to continuing support
- utilising and developing techniques which can also be made to serve so called basic research
- providing data and ideas which may speed up the process of generalisation

Utility to developing countries: Applied research has practical utility in the developing countries. Developing countries may benefit by the discoveries, scientific and otherwise, that have already been made, either by direct application or by making some modifications wherever necessary.

Field investigation method: Applied research often takes the form of a field investigation and aims at collecting the basic data for verifying the applicability of existing theories and models in given situation.

1.2.3 Exploratory Research

It pursues several possibilities simultaneously, and in a sense it is not quite sure of its objective. Exploratory research is designed to provide a background, to familiarise and, as the word implies, just explore the general subject. A part of exploratory research is the investigation of relationships among variables, without knowing why they are studied. It borders on an idle curiosity approach, differing from it only in which the investigator thinks there may be a payoff in an application somewhere in the forest of questions.

1.2.4 Descriptive Research

Descriptive research, as the name suggests, is designed to describe something. For example, the characteristics of users of a given product, the degree to which product use varies with income, etc. to be of maximum benefit, a descriptive study must collect data for a definite purpose. Descriptive studies vary in the degree to which a specific hypothesis is the guide. It allows both, implicit and explicit hypotheses to be tested depending on the research problems.

1.2.5 Action Research

Action research is also a recent addition to the categories of research known to a modern social scientist. By its very definition, "it is a research through launching of a direct action with the objective of obtaining workable solutions to the given problems". This also aims at collecting information from other sources that have direct or indirect bearing on the research programme.

At the second phase, the planned action is practically launched and then at the next phase, research carries out periodical assessment of the project. At a subsequent stage, changes, modifications and other improvements are made in the functional aspect of the project and finally the whole process culminates in the evaluation of the project as a whole.

The methods used for this type of research are usually personal interview methods and the survey methods. Sometimes attitude measurement techniques are also used. Some problems associated with action research are the personal values of the individuals, lack of social scientist's interest and exclusive locations with the respondent.

1.2.6 Evaluation Research

It is a recent addition to the types of research. This type of research is primarily directed to evaluate the performance of the developmental projects and other economic programmes that have already been implemented. The objective of evaluation research is to realistically assess the impact of any such programmes. Evaluation is held to mean comprehensive concept of measurement and it is because of this definition of evaluation that project evaluations have become frequent in the recent years.

The evaluation research is of three types:

- **Concurrent evaluation**: It is a continuing process of an inspection of the project that has been launched. In this manner, such type of research not only evaluates the performance but also stimulates it and gives direction and control as and when possible.
- **Phasic or periodic evaluation**: This type of evaluation takes place at different phases or stages of performance of the project. It enables us to evaluate the performance of the completed phase and make adjustments in the subsequent phases after keeping in view the failures and successes of the previous phase.
- **Terminal evaluation**: It is the evaluation of the final phase of the project. Once the project has been completed, an overall assessment is made to see how best a project has served the objectives for which it was launched.

1.2.7 Experimental Research

Experimentation is the process of research in which one or more variables are manipulated under conditions, which permit the collection of data and which show the effects. Experiments will create artificial situation so that the researcher can obtain the particular data needed and can measure the data accurately. Experiments are artificial in the sense that the situations are usually created for testing purposes.

1.2.8 Empirical Research

This type relies on experience or observation alone, often without due regard for system and theory. It is a data based research, coming up with conclusions, which are capable of being verified by observation or experiment. In such a research, it is necessary to get the facts first hand, at their source, and actively do certain things to stimulate the production of desired information. Empirical research is appropriate when proof is sought that certain variables affect other variables in some way.

1.2.9 Survey Research

This type of research has become very popular these days as a scientific method for discovering relevant impact and inter-relationships of social psychological variables from given population. Survey research studies large and small populations by selecting and studying samples chosen from the populations – sociological and psychological variables.

The advantage of this type of research is that it links sample investigations with populations and thereby offers an easy opportunity of studying population behavior through sample survey research assessments.

Methods of survey: Survey research is approached through the methods of personal interviews, mailed questionnaires and personal discussions besides indirect oral investigation.

Advantages: This type of research has the advantage of greater scope in the sense that a larger volume of information can be controlled from a very large population.

1.2.10 Qualitative Research

It is concerned with qualitative phenomenon, i.e., phenomena relating to or involving quality or kind. This type of research aims at discovering the underlying motives and desires, using in depth interviews for the purpose.

1.2.11 Quantitative Research

It is based on the measurement of quantity or amount. It is applicable to phenomena that can be expressed in terms of quality.

1.2.12 Field Investigation Research

This is generally credited with a few virtues which are supposed to be unique to this category of research. These virtues may be listed as:

- The variables in a field experiment operate more strongly than those used in laboratory experiment. This is because of the fact that field situation takes stock of realistic natural operations.
- Secondly, field experiments have the advantage of investigating more fruitfully the dynamics of inter-relationships of small groups of variables.
- Field experimental studies are also ideal to testing of the theory and to the solution of the real world problems.

Uses: Field experimental studies are important part of the applied research which, at times, play an important role in pointing out the nature and direction of the refinements required for an existing doctrinaire.

1.2.13 Ex-post Facto Research

This is a systematic empirical inquiry in which the scientist does not have direct control of independent variables because their manifestations have already occurred or because they are inherently not manipulative. Inferences about relations among variables are made, without direct intervention, from concomitant variation of independent and dependent variables.

1.2.14 Historical Research

Historical study is a study of past records and other information sources with a view to reconstructing the origin and development of an institution or a movement or a system and discovering the trends in the past. It is descriptive in nature. It is a difficult task; it must often depend upon inference and logical analysis of recorded data and indirect evidences rather than upon direct observation.

1.2.15 Some Other Types of Research

- All other types of research are variations of one or more of the above stated approaches, based on either the purpose of research, or the time required to accomplish research, on the environment in which research is done, or on the basis of some other similar factor.
- Research can be field-setting research or laboratory research or simulation research, depending upon the environment in which it is to be carried out. Research can as well be understood as clinical or diagnostic research. Such research follows case-study methods or in-depth approaches to reach the basic causal relations.
- Such studies usually go deep into the causes of things or events that interest us, using very small samples and very deep probing data gathering devices. The research may be exploratory or it may be formalised. The objective of exploratory research is the development of hypotheses rather than their testing, whereas formalised research studies are those with substantial structure and with specific hypotheses to be tested. Historical research is that which utilises historical sources like documents, remains, etc. to study events or ideas of the past, including the philosophy of persons and groups at any remote point of time.
- Research can also be classified as conclusion-oriented and decision-oriented research. While doing conclusion oriented research, a researcher is free to pick up a problem, redesign the enquiry as he proceeds and is prepared to conceptualise as he wishes. Decision-oriented research is always for the need of a decision maker and the researcher in this case is not free to embark upon research according to his own inclination.
- Operations research is an example of decision oriented research since it is a scientific method of providing executive departments with a quantitative basis for decisions regarding operations under their control.

1.3 Research Approaches

Different research approaches are explained below.

1.3.1 Descriptive vs. Analytical

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

1.3.2 Applied vs. Fundamental

- Research can either be applied (or action) research or fundamental (to basic or pure) research. Applied research aims at finding a solution for an immediate problem facing a society or an industrial/business organisation, whereas fundamental research is mainly concerned with generalisations and with the formulation of a theory.
- Research concerning some natural phenomenon or relating to pure mathematics are examples of fundamental research. Similarly, research studies, concerning human behaviour carried on with a view to make generalisations about human behaviour, are also examples of fundamental research, but research aimed at certain conclusions (say, a solution) facing a concrete social or business problem is an example of applied research.
- Research to identify social, economic or political trends that may affect a particular institution or the copy research (research to find out whether certain communications will be read and understood) or the marketing research or evaluation research are examples of applied research. Thus, the central aim of applied research is to discover a solution for some pressing practical problem, whereas basic research is directed towards finding information that has a broad base of applications and thus, adds to the already existing organised body of scientific knowledge.

1.3.3 Quantitative vs. Qualitative

- Quantitative research is based on the measurement of quantity or amount. It is applicable to phenomena that can be expressed in terms of quantity. Qualitative research, on the other hand, is concerned with qualitative phenomenon, i.e., phenomena relating to or involving quality or kind. For instance, when we are interested in investigating the reasons for human behaviour (i.e., why people think or do certain things), we quite often talk of 'Motivation Research', an important type of qualitative research.
- This type of research aims at discovering the underlying motives and desires, using in depth interviews for the purpose. Other techniques of such research are word association tests, sentence completion tests, story completion tests and similar other projective techniques. Attitude or opinion research, i.e., research designed to find out how people feel or what they think about a particular subject or institution is also qualitative research.
- Qualitative research is especially important in the behavioural sciences where the aim is to discover the underlying motives of human behaviour. Through such research we can analyse the various factors which motivate people to behave in a particular manner or which make people like or dislike a particular thing. It may be stated, however, that to apply qualitative research in practice is relatively a difficult job and therefore, while doing such research, one should seek guidance from experimental psychologists.

1.3.4 Conceptual vs. Empirical

- Conceptual research is that related to some abstract idea(s) or theory. It is generally used by philosophers and thinkers to develop new concepts or to reinterpret existing ones. On the other hand, empirical research relies on experience or observation alone, often without due regard for system and theory. It is data-based research, coming up with conclusions which are capable of being verified by observation or experiment.
- We can also call it as experimental type of research. In such a research, it is necessary to get at facts firsthand, at their source, and actively to go about doing certain things to stimulate the production of desired information. In such a research, the researcher must first provide himself with a working hypothesis or guess as to the probable results. The researcher then works to get enough facts (data) to prove or disprove his hypothesis. He then sets up experimental designs which he thinks will manipulate the persons or the materials concerned so as to bring forth the desired information.
- Such research is thus characterised by the experimenter's control over the variables under study and his deliberate manipulation of one of them to study its effects. Empirical research is appropriate when proof is sought that certain variables affect other variables in some way. Evidence gathered through experiments or empirical studies is today considered to be the most powerful support possible for a given hypothesis.

1.4 Research Approaches and Significance of Research

Research approaches and significance of research are explained below.

Research approaches

- The above description of the types of research brings to light the fact that there are two basic approaches to research, viz., quantitative approach and the qualitative approach. The former involves the generation of data in quantitative form which can be subjected to rigorous quantitative analysis in a formal and rigid fashion. This approach can be further sub-classified into inferential, experimental and simulation approaches to research.
- The purpose of inferential approach to research is to form a data base from which to infer characteristics or relationships of population. This usually means survey research where a sample of population is studied (questioned or observed) to determine its characteristics, and it is then inferred that the population has the same characteristics.
- Experimental approach is characterised by much greater control over the research environment and in this case some variables are manipulated to observe their effect on other variables. Simulation approach involves the construction of an artificial environment within which relevant information and data can be generated. This permits an observation of the dynamic behaviour of a system (or its sub-system) under controlled conditions.

- The term 'simulation' in the context of business and social sciences applications refers to "the operation of a numerical model that represents the structure of a dynamic process. Given the values of initial conditions, parameters and exogenous variables, a simulation is run to represent the behaviour of the process over time." 5 Simulation approach can also be useful in building models for understanding future conditions.
- Qualitative approach to research is concerned with subjective assessment of attitudes, opinions and behaviour. Research in such a situation is a function of researcher's insights and impressions. Such an approach to research generates results either in non-quantitative form or in the form which are not subjected to rigorous quantitative analysis.

Significance of research

- Research inspires scientific and inductive thinking and it promotes the development of logical habits of thinking and organisation. The role of research in several fields of applied economics, whether related to business or to the economy as a whole, has greatly increased in modern times. The increasingly complex nature of business and government has focused attention on the use of research in solving operational problems.
- Research, as an aid to economic policy, has gained added importance, both for government and business. Research provides the basis for nearly all government policies in our economic system. For instance, government's budgets rest in part on an analysis of the needs and desires of the people and on the availability of revenues to meet these needs.
- The cost of needs has to be equated to probable revenues and this is a field where research is most needed. Through research we can devise alternative policies and can as well examine the consequences of each of these alternatives.
- Decision-making may not be a part of research, but research certainly facilitates the decisions of the policy maker. Government has also to chalk out programmes for dealing with all facets of the country's existence and most of these will be related directly or indirectly to economic conditions.
- Research is considered necessary with regard to the allocation of nation's resources. Another area in government, where research is necessary, is collecting information on the economic and social structure of the nation. Such information indicates what is happening in the economy and what changes are taking place. Collecting such statistical information is by no means a routine task, but it involves a variety of research problems.
- These days, nearly all governments maintain large staff of research technicians or experts to carry on this work. Thus, in the context of government, research as a tool to economic policy has three distinct phases of operation, viz.,
  - investigation of economic structure through continual compilation of facts;
  - diagnosis of events that are taking place and the analysis of the forces underlying them; and
  - the prognosis, i.e., the prediction of future developments
- Research has its special significance in solving various operational and planning problems of business and industry. Operations research and market research, along with motivational research, are considered crucial and their results assist, in more than one way, in taking business decisions. Market research is the investigation of the structure and development of a market for the purpose of formulating efficient policies for purchasing, production and sales.
- Operations research refers to the application of mathematical, logical and analytical techniques to the solution of business problems of cost minimisation or of profit maximisation or what can be termed as optimisation problems. Motivational research of determining why people behave as they do is mainly concerned with market characteristics.
- In other words, it is concerned with the determination of motivations underlying the consumer (market) behaviour. All these are of great help to people in business and industry who are responsible for taking business decisions. Research with regard to demand and market factors has great utility in business. Given knowledge of future demand, it is generally not difficult for a firm, or for an industry to adjust its supply schedule within the limits of its projected capacity.

- Market analysis has become an integral tool of business policy these days. Business budgeting, which ultimately results in a projected profit and loss account, is based mainly on sales estimates which in turn depend on business research. Once sales forecasting is done, efficient production and investment programmes can be set up around which are grouped the purchasing and financing plans. Research, thus, replaces intuitive business decisions by more logical and scientific decisions.
- Research is equally important for social scientists in studying social relationships and in seeking answers to various social problems. It provides the intellectual satisfaction of knowing a few things just for the sake of knowledge and also has practical utility for the social scientist to know for the sake of being able to do something better or in a more efficient manner. Research in social sciences is concerned both with knowledge for its own sake and with knowledge for what it can contribute to practical concerns.
- "This double emphasis is perhaps especially appropriate in the case of social science. On the one hand, its responsibility as a science is to develop a body of principles that make possible the understanding and prediction of the whole range of human interactions. On the other hand, because of its social orientation, it is increasingly being looked to for practical guidance in solving immediate problems of human relations."
- In addition to what has been stated above, the significance of research can also be understood keeping in view the following points:
  - To those students who are to write a master's or Ph.D. thesis, research may mean a careerism or a way to attain a high position in the social structure;
  - To professionals in research methodology, research may mean a source of livelihood;
  - To philosophers and thinkers, research may mean the outlet for new ideas and insights;
  - To literary men and women, research may mean the development of new styles and creative work;
  - To analysts and intellectuals, research may mean the generalisations of new theories.
- Thus, research is the fountain of knowledge for the sake of knowledge and an important source of providing guidelines for solving different business, governmental and social problems. It is a sort of formal training which enables one to understand the new developments in one's field in a better way.

1.5 Research Methods versus Methodology

Difference between research methods and methodology is detailed below.

- Research methods may be understood as all those methods/techniques that are used for conduction of research.
- At times, a distinction is also made between research techniques and research methods. Research techniques refer to the behaviour and instruments we use in performing research operations such as making observations, recording data, techniques of processing data and the like. Research methods refer to the behaviour and instruments used in selecting and constructing research technique. For instance, the difference between methods and techniques of data collection can better be understood from the details given in the following chart:

Type	Methods	Techniques
Library Research	• Analysis of historical records • Analysis of documents	• Recording of notes, content analysis, tape and film listening and analysis. • Statistical complications and manipulations, reference and abstract guides, content analysis.
Field Research	• Non participant direct observation. • Participant observation. • Mass observation • Mail questionnaire • Opinionnaire • Personal interview • Focused interview • Group interview • Telephone survey • Case study and life history	• Observational behavioural scales, use of score cards etc. • Interactional recording, possible use of tape recorder, photo graphic techniques. • Recording mass behavior, interview using independent observers in public place. • Identification of social and economic background of respondents. • Use of attitude scales, projective techniques, use of sociometric scales. • Interviewer uses a detailed scheduled with open and closed questions. • Interviewer focuses attention upon a given experience and its effect. • Small groups of respondents are interviewed simultaneously. • Used as a survey technique for information and for discerning opinion; may also be used as a follow up of questionnaire. • Cross sectional collection of data for information for intensive analysis, longitudinal collection of data of intensive character.
Laboratory Research	Small group study of random behavior, play and role analysis.	Use of audio-visual recording devices, use of observers, etc.

Table 1.1 Methods and techniques of data collection

- From what has been stated above, we can say that methods are more general. It is the methods that generate techniques. However, in practice, the two terms are taken as interchangeable and when we talk of research methods we do, by implication, include research techniques within their compass use in performing research operations. In other words, all those methods which are used by the researcher during the course of studying his research problem are termed as research methods.

- Since the object of research, particularly the applied research, it to arrive at a solution for a given problem, the available data and the unknown aspects of the problem have to be related to each other to make a solution possible. Keeping this in view, research methods can be put into the following three groups:
  - In the first group we include those methods which are concerned with the collection of data. These methods will be used where the data already available are not sufficient to arrive at the required solution;
  - The second group consists of those statistical techniques which are used for establishing relationships between the data and the unknowns;
  - The third group consists of those methods which are used to evaluate the accuracy of the results obtained.
- Research methods falling in the above stated last two groups are generally taken as the analytical tools of research. Research methodology is a way to systematically solve the research problem. It may be understood as a science of studying how research is done scientifically. In it we study the various steps that are generally adopted by a researcher in studying his research problem along with the logic behind them. It is necessary for the researcher to know not only the research methods/techniques but also the methodology.
- Researchers not only need to know how to develop certain indices or tests, how to calculate the mean, the mode, the median or the standard deviation or chi-square, how to apply particular research techniques, but they also need to know which of these methods or techniques, are relevant and which are not, and what would they mean and indicate and why. Researchers also need to understand the assumptions underlying various techniques and they need to know the criteria by which they can decide that certain techniques and procedures will be applicable to certain problems and others will not.
- All this means that it is necessary for the researcher to design his methodology for his problem as the same may differ from problem to problem. For example, an architect, who designs a building, has to consciously evaluate the basis of his decisions, i.e., he has to evaluate why and on what basis he selects particular size, number and location of doors, windows and ventilators, uses particular materials and not others and the like.
- Similarly, in research, the scientist has to evaluate the research decisions before they are implemented. He has to specify very clearly and precisely what decisions he selects and why he selects them so that they can be evaluated by others also. From what has been stated above, we can say that research methodology has many dimensions and research methods do constitute a part of the research methodology.
- The scope of research methodology is wider than that of research methods. Thus, when we talk of research methodology we not only talk of the research methods but also consider the logic behind the methods we use in the context of our research study and explain why we are using a particular method or technique and why we are not using others so that research results are capable of being evaluated either by the researcher himself or by others.

1.6 Research and Scientific Method

Research and scientific methods are explained below.

- For a clear perception of the term 'research', one should know the meaning of scientific method. The two terms, research and scientific method, are closely related. Research, as we have already stated, can be termed as “an inquiry into the nature of, the reasons for, and the consequences of any particular set of circumstances, whether these circumstances are experimentally controlled or recorded just as they occur.
- Further, research implies the researcher is interested in more than particular results; he is interested in the repeatability of the results and in their extension to more complicated and general situations.” On the other hand, the philosophy common to all research methods and techniques, although they may vary considerably from one science to another, is usually given the name of scientific method.
- Scientific method is the pursuit of truth as determined by logical considerations. The ideal of science is to achieve a systematic interrelation of facts. Scientific method attempts to achieve “this ideal by experimentation, observation, logical arguments from accepted postulates and a combination of these three in varying proportions.” In scientific method, logic aids in formulating propositions explicitly and accurately so that their possible alternatives become clear.

- Further, logic develops the consequences of such alternatives, and when these are compared with observable phenomena, it becomes possible for the researcher or the scientist to state which alternative is most in harmony with the observed facts. All this is done through experimentation and survey investigations which constitute the integral parts of scientific method. Experimentation is done to test hypotheses and to discover new relationships if any, among variables.
- But the conclusions drawn on the basis of experimental data are generally criticised for faulty assumptions, poorly designed experiments, badly executed experiments or faulty interpretations. As such the researcher must pay all possible attention while developing the experimental design and must state only probable inferences.
- The purpose of survey investigations may also be to provide scientifically gathered information to work as a basis for the researchers for their conclusions. The scientific method is, thus, based on certain basic postulates which can be stated as under:
  - It relies on empirical evidence;
  - It utilises relevant concepts.
  - It is committed to only objective considerations.
  - It presupposes ethical neutrality, i.e., it aims at nothing but making only adequate and correct statements about population objects.
  - It results into probabilistic predictions.
  - Its methodology is made known to all concerned for critical scrutiny are for use in testing the conclusions through replication.
  - It aims at formulating most general axioms or what can be termed as scientific theories.

1.7 Research Process

Before embarking on the details of research methodology and techniques, it seems appropriate to present a brief overview of the research process. Research process consists of series of actions or steps necessary to effectively carry out research and the desired sequencing of these steps.

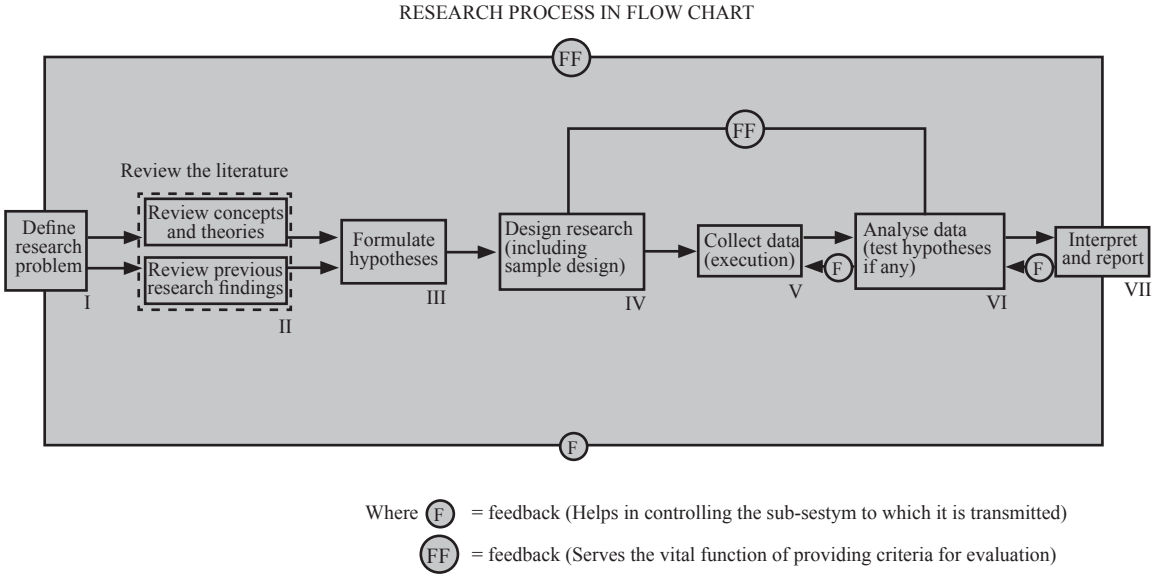


Fig. 1.2 Research process in flow chart

(Source: http://www.scribd.com/doc/19680689/Research-MethodologyMethods-and-Techniques-by-CR-Kothari)

Where F = Feed back (help in controlling the sub-system to which it is transmitted)
FF = Feed Forward (serves the vital function of providing criteria for evaluation)

- The chart indicates that the research process consists of a number of closely related activities. But such activities overlap continuously rather than following a strictly prescribed sequence. At times, the first step determines the nature of the last step to be undertaken. If subsequent procedures have not been taken into account in the early stages, serious difficulties may arise which may even prevent the completion of the study.
- One should remember that the various steps involved in a research process are not mutually exclusive; nor are they separate and distinct. They do not necessarily follow each other in any specific order and the researcher has to be constantly anticipating at each step in the research process the requirements of the subsequent steps.
- However, the following order concerning various steps provides a useful procedural guideline regarding the research process:
  - formulating the research problem
  - extensive literature survey
  - developing the hypothesis
  - preparing the research design
  - determining sample design
  - collecting the data
  - execution of the project
  - analysis of data
  - hypothesis testing
  - generalisations and interpretation, and
  - preparation of the report or presentation of the results, i.e., formal write-up of conclusions reached

Formulating the research problem

- There are two types of research problems, viz., those which relate to states of nature and those which relate to relationships between variables. At the very outset, the researcher must single out the problem he wants to study, i.e., he must decide the general area of interest or aspect of a subject-matter that he would like to inquire into.
- Initially, the problem may be stated in a broad general way and then the ambiguities, if any, relating to the problem be resolved. Then, the feasibility of a particular solution has to be considered before a working formulation of the problem can be set up. The formulation of a general topic into a specific research problem, thus, constitutes the first step in a scientific enquiry.
- Essentially two steps are involved in formulating the research problem, viz., understanding the problem thoroughly, and rephrasing the same into meaningful terms from an analytical point of view. The best way of understanding the problem is to discuss it with one's own colleagues or with those having some expertise in the matter. In an academic institution the researcher can seek the help from a guide who is usually an experienced man and has several research problems in mind.
- Often, the guide puts forth the problem in general terms and it is up to the researcher to narrow it down and phrase the problem in operational terms. In private business units or in governmental organisations, the problem is usually earmarked by the administrative agencies with which the researcher can discuss as to how the problem originally came about and what considerations are involved in its possible solutions.
- The researcher must at the same time examine all available literature to get himself acquainted with the selected problem. He may review two types of literature—the conceptual literature concerning the concepts and theories, and the empirical literature consisting of studies made earlier which are similar to the one proposed. The basic outcome of this review will be the knowledge as to what data and other materials are available for operational purposes which will enable the researcher to specify his own research problem in a meaningful context.
- After this the researcher rephrases the problem into analytical or operational terms, i.e., to put the problem in as specific terms as possible. This task of formulating, or defining, a research problem is a step of greatest importance in the entire research process. The problem to be investigated must be defined unambiguously for that will help discriminating relevant data from irrelevant ones. Care must; however, be taken to verify the objectivity and validity of the background facts concerning the problem.

Extensive literature survey

- Once the problem is formulated, a brief summary of it should be written down. It is compulsory for a research worker writing a thesis for a Ph.D. degree to write a synopsis of the topic and submit it to the necessary Committee or the Research Board for approval. At this juncture the researcher should undertake extensive literature survey connected with the problem.
- For this purpose, the abstracting and indexing journals and published or unpublished bibliographies are the first place to go to. Academic journals, conference proceedings, government reports, books etc., must be tapped depending on the nature of the problem. In this process, it should be remembered that one source will lead to another. The earlier studies, if any, which are similar to the study in hand, should be carefully studied. A good library will be a great help to the researcher at this stage.

Development of working hypotheses

- After extensive literature survey, researcher should state in clear terms the working hypothesis or hypotheses. Working hypothesis is tentative assumption made in order to draw out and test its logical or empirical consequences. As such the manner in which research hypotheses are developed is particularly important since they provide the focal point for research.
- They also affect the manner in which tests must be conducted in the analysis of data and indirectly the quality of data which is required for the analysis. In most types of research, the development of working hypothesis plays an important role. Hypothesis should be very specific and limited to the piece of research in hand because it has to be tested. The role of the hypothesis is to guide the researcher by delimiting the area of research and to keep him on the right track.
- It sharpens his thinking and focuses attention on the more important facets of the problem. It also indicates the type of data required and the type of methods of data analysis to be used. How does one go about developing working hypotheses? The answer is by using the following approach:
  - Discussions with colleagues and experts about the problem, its origin and the objectives in seeking a solution;
  - Examination of data and records, if available, concerning the problem for possible trends, peculiarities and other clues;
  - Review of similar studies in the area or of the studies on similar problems; and
  - Exploratory personal investigation which involves original field interviews on a limited scale with interested parties and individuals with a view to secure greater insight into the practical aspects of the problem.
- Thus, working hypotheses arise as a result of a-priori thinking about the subject, examination of the available data and material including related studies and the counsel of experts and interested parties. Working hypotheses are more useful when stated in precise and clearly defined terms. It may as well be remembered that occasionally we may encounter a problem where we do not need working hypotheses, especially in the case of exploratory or formulative researches which do not aim at testing the hypothesis. But as a general rule, specification of working hypotheses in another basic step of the research process in most research problems.

Preparing the research design

- The research problem having been formulated in clear cut terms, the researcher will be required to prepare a research design, i.e., he will have to state the conceptual structure within which research would be conducted. The preparation of such a design facilitates research to be as efficient as possible yielding maximal information. In other words, the function of research design is to provide for the collection of relevant evidence with minimal expenditure of effort, time and money. But how all these can be achieved depends mainly on the research purpose. Research purposes may be grouped into four categories, viz.,
  - Exploration
  - Description
  - Diagnosis
  - Experimentation

- A flexible research design which provides opportunity for considering many different aspects of a problem is considered appropriate if the purpose of the research study is that of exploration. But when the purpose happens to be an accurate description of a situation or of an association between variables, the suitable design will be one that minimises bias and maximises the reliability of the data collected and analysed.
- There are several research designs, such as, experimental and non-experimental hypothesis testing. Experimental designs can be either informal designs (such as before-and-after without control, after-only with control, before-and-after with control) or formal designs (such as completely randomised design, randomised block design, Latin square design, simple and complex factorial designs), out of which the researcher must select one for his own project.
- The preparation of the research design, appropriate for a particular research problem, involves usually the consideration of the following:
  - the means of obtaining the information
  - the availability and skills of the researcher and his staff (if any)
  - explanation of the way in which selected means of obtaining information will be organised and the reasoning leading to the selection
  - the time available for research
  - the cost factor relating to research, i.e., the finance available for the purpose

Determining sample design

- All the items under consideration in any field of inquiry constitute a 'universe' or 'population'. A complete enumeration of all the items in the 'population' is known as a census inquiry. It can be presumed that in such an inquiry when all the items are covered no element of chance is left and highest accuracy is obtained. But in practice this may not be true.
- Even the slightest element of bias in such an inquiry will get larger and larger as the number of observations increases. Moreover, there is no way of checking the element of bias or its extent except through a resurvey or use of sample checks. Besides, this type of inquiry involves a great deal of time, money and energy.
- The researcher must decide the way of selecting a sample or what is popularly known as the sample design. In other words, a sample design is a definite plan determined before any data are actually collected for obtaining a sample from a given population. With probability samples each element has a known probability of being included in the sample but the non-probability samples do not allow the researcher to determine this probability.
- Probability samples are those based on simple random sampling, systematic sampling, stratified sampling, cluster/area sampling whereas non-probability samples are those based on convenience sampling, judgement sampling and quota sampling techniques. A brief mention of the important sample designs is as follows:

Deliberate sampling

- It is also known as purposive or non-probability sampling. This sampling method involves purposive or deliberate selection of particular units of the universe for constituting a sample which represents the universe. When population elements are selected for inclusion in the sample based on the ease of access, it can be called convenience sampling.
- If a researcher wishes to secure data from, say, gasoline buyers, he may select a fixed number of petrol stations and may conduct interviews at these stations. This would be an example of convenience sample of gasoline buyers. At times such a procedure may give very biased results particularly when the population is not homogeneous. On the other hand, in judgement sampling the researcher's judgement is used for selecting items which he considers as representative of the population.

Simple random sampling

- This type of sampling is also known as chance sampling or probability sampling where each and every item in the population has an equal chance of inclusion in the sample and each one of the possible samples, in case of finite universe, has the same probability of being selected.

- This procedure gives each item an equal probability of being selected. In case of infinite population, the selection of each item in a random sample is controlled by the same probability and that successive selections are independent of one another.

Systematic sampling

- In some instances, the most practical way of sampling is to select every 15th name on a list, every 10th house on one side of a street and so on. Sampling of this type is known as systematic sampling. An element of randomness is usually introduced into this kind of sampling by using random numbers to pick up the unit with which to start.
- This procedure is useful when sampling frame is available in the form of a list. In such a design the selection process starts by picking some random point in the list and then every nth element is selected until the desired number is secured.

Stratified sampling

- If the population from which a sample is to be drawn does not constitute a homogeneous group, then stratified sampling technique is applied so as to obtain a representative sample. In this technique, the population is stratified into a number of non overlapping subpopulations or strata and sample items are selected from each stratum.
- If the items selected from each stratum is based on simple random sampling the entire procedure, first stratification and then simple random sampling, is known as stratified random sampling.

Quota sampling

- In stratified sampling, the cost of taking random samples from individual strata is often so expensive that interviewers are simply given quota to be filled from different strata, the actual selection of items for sample being left to the interviewer's judgement.
- This is called quota sampling. The size of the quota for each stratum is generally proportionate to the size of that stratum in the population. Quota sampling is thus an important form of non-probability sampling. Quota samples generally happen to be judgement samples rather than random samples.

Cluster sampling and area sampling

- Cluster sampling involves grouping the population and then selecting the groups or the clusters rather than individual elements for inclusion in the sample. The sample size must often be larger than the simple random sample to ensure the same level of accuracy because is cluster sampling procedural potential for order bias and other sources of error are usually accentuated. The clustering approach can, however, make the sampling procedure relatively easier and increase the efficiency of field work, especially in the case of personal interviews.
- Area sampling is quite close to cluster sampling and is often talked about when the total geographical area of interest happens to be big one. Under area sampling, we first divide the total area into a number of smaller non-overlapping areas, generally called geographical clusters, then a number of these smaller areas are randomly selected, and all units in these small areas are included in the sample. Area sampling is especially helpful where we do not have the list of the population concerned. It also makes the field interviewing more efficient since interviewer can do many interviews at each location.

Multi-stage sampling

- This is a further development of the idea of cluster sampling. This technique is meant for big inquiries extending to a considerably large geographical area like an entire country. Under multi-stage sampling the first stage may be to select large primary sampling units such as states, then districts, then towns and finally certain families within towns. If the technique of random-sampling is applied at all stages, the sampling procedure is described as multi-stage random sampling.

Sequential sampling

- This is somewhat a complex sample design where the ultimate size of the sample is not fixed in advance but is determined according to mathematical decisions on the basis of information yielded as survey progresses. This design is usually adopted under acceptance sampling plan in the context of statistical quality control.
- In practice, several of the methods of sampling described above may well be used in the same study in which case it can be called mixed sampling. It may be pointed out here that normally one should resort to random sampling so that bias can be eliminated and sampling error can be estimated.
- But purposive sampling is considered desirable when the universe happens to be small and a known characteristic of it is to be studied intensively. Also, there are conditions under which sample designs other than random sampling may be considered better for reasons like convenience and low costs. The sample design to be used must be decided by the researcher taking into consideration the nature of the inquiry and other related factors.

Data collection

- In dealing with any real life problem, it is often found that data at hand are inadequate, and hence, it becomes necessary to collect data that are appropriate. There are several ways of collecting the appropriate data which differ considerably in context of money costs, time and other resources at the disposal of the researcher. Primary data can be collected either through experiment or through survey. If the researcher conducts an experiment, he observes some quantitative measurements, or the data, with the help of which he examines the truth contained in his hypothesis. But in the case of a survey, data can be collected by any one or more of the following ways:
  - By observation: This method implies the collection of information by way of investigator's own observation, without interviewing the respondents. The information obtained relates to what is currently happening and is not complicated by either the past behaviour or future intentions or attitudes of respondents. This method is no doubt an expensive method and the information provided by this method is also very limited. As such this method is not suitable in inquiries where large samples are concerned.
  - Through personal interview: The investigator follows a rigid procedure and seeks answers to a set of pre-conceived questions through personal interviews. This method of collecting data is usually carried out in a structured way where output depends upon the ability of the interviewer to a large extent.
  - Through telephone interviews: This method of collecting information involves contacting the respondents on telephone itself. This is not a very widely used method but it plays an important role in industrial surveys in developed regions, particularly, when the survey has to be accomplished in a very limited time.
  - By mailing of questionnaires: The researcher and the respondents do come in contact with each other if this method of survey is adopted. Questionnaires are mailed to the respondents with a request to return after completing the same. It is the most extensively used method in various economic and business surveys. Before applying this method, usually a Pilot Study for testing the questionnaire is conduced which reveals the weaknesses, if any, of the questionnaire? Questionnaire to be used must be prepared very carefully so that it may prove to be effective in collecting the relevant information.
  - Through schedules: Under this method, the enumerators are appointed and given training. They are provided with schedules containing relevant questions. These enumerators go to respondents with these schedules. Data are collected by filling up the schedules by enumerators on the basis of replies given by respondents. Much depends upon the capability of enumerators so far as this method is concerned. Some occasional field checks on the work of the enumerators may ensure sincere work. The researcher should select one of these methods of collecting the data taking into consideration the nature of investigation, objective and scope of the inquiry, financial resources, available time and the desired degree of accuracy. Though he should pay attention to all these factors but much depends upon the ability and experience of the researcher.

Execution of the project

- Execution of the project is a very important step in the research process. If the execution of the project proceeds on correct lines, the data to be collected would be adequate and dependable. The researcher should see that the project is executed in a systematic manner and in time. If the survey is to be conducted by means of structured questionnaires, data can be readily machine-processed. In such a situation, questions as well as the possible answers may be coded. If the data are to be collected through interviewers, arrangements should be made for proper selection and training of the interviewers.
- The training may be given with the help of instruction manuals which explain clearly the job of the interviewers at each step. Occasional field checks should be made to ensure that the interviewers are doing their assigned job sincerely and efficiently. A careful watch should be kept for unanticipated factors in order to keep the survey as much realistic as possible. This, in other words, means that steps should be taken to ensure that the survey is under statistical control so that the collected information is in accordance with the pre-defined standard of accuracy.
- If some of the respondents do not cooperate, some suitable methods should be designed to tackle this problem. One method of dealing with the non-response problem is to make a list of the non-respondents and take a small sub-sample of them, and then with the help of experts vigorous efforts can be made for securing response.

Analysis of data

- After the data have been collected, the researcher turns to the task of analysing them. The analysis of data requires a number of closely related operations such as establishment of categories, the application of these categories to raw data through coding, tabulation and then drawing statistical inferences. The unwieldy data should necessarily be condensed into a few manageable groups and tables for further analysis.
- Thus, researcher should classify the raw data into some purposeful and usable categories. Coding operation is usually done at this stage through which the categories of data are transformed into symbols that may be tabulated and counted. Editing is the procedure that improves the quality of the data for coding. With coding the stage is ready for tabulation.
- Tabulation is a part of the technical procedure wherein the classified data are put in the form of tables. The mechanical devices can be made use of at this juncture. A great deal of data, especially in large inquiries, is tabulated by computers. Computers not only save time but also make it possible to study large number of variables affecting a problem simultaneously.
- Analysis work after tabulation is generally based on the computation of various percentages, coefficients, etc., by applying various well defined statistical formulae. In the process of analysis, relationships or differences supporting or conflicting with original or new hypotheses should be subjected to tests of significance to determine with what validity data can be said to indicate any conclusion(s).
- For instance, if there are two samples of weekly wages, each sample being drawn from factories in different parts of the same city, giving two different mean values, then our problem may be whether the two mean values are significantly different or the difference is just a matter of chance. Through the use of statistical tests we can establish whether such a difference is a real one or is the result of random fluctuations. If the difference happens to be real, the inference will be that the two samples come from different universes and if the difference is due to chance, the conclusion would be that the two samples belong to the same universe.
- Similarly, the technique of analysis of variance can help us in analysing whether three or more varieties of seeds grown on certain fields yield significantly different results or not. In brief, the researcher can analyse the collected data with the help of various statistical measures.

Hypothesis-testing

- After analysing the data as stated above, the researcher is in a position to test the hypotheses, if any, he had formulated earlier. Do the facts support the hypotheses or they happen to be contrary? This is the usual question which should be answered while testing hypotheses. Various tests, such as Chi square test, t-test, F-test, have been developed by statisticians for the purpose. The hypotheses may be tested through the use of one or more of such tests, depending upon the nature and object of research inquiry.

- Hypothesis-testing will result in either accepting the hypothesis or in rejecting it. If the researcher had no hypotheses to start with, generalisations established on the basis of data may be stated as hypotheses to be tested by subsequent researches in times to come.

Generalisations and interpretation

- If a hypothesis is tested and upheld several times, it may be possible for the researcher to arrive at generalisation, i.e., to build a theory. As a matter of fact, the real value of research lies in its ability to arrive at certain generalisations. If the researcher had no hypothesis to start with, he might seek to explain his findings on the basis of some theory. It is known as interpretation. The process of interpretation may quite often trigger off new questions which in turn may lead to further researches.

Preparation of the report or the thesis

Finally, the researcher has to prepare the report of what has been done by him. Writing of report must be done with great care keeping in view the following:

- The layout of the report should be as follows:
  - the preliminary pages
  - the main text
  - the end matter

In its preliminary pages, the report should carry title and date followed by acknowledgements and foreword. Then there should be a table of contents followed by a list of tables and list of graphs and charts, if any, given in the report.

- The main text of the report should have the following parts:
  - **Introduction**: It should contain a clear statement of the objective of the research and an explanation of the methodology adopted in accomplishing the research. The scope of the study along with various limitations should as well be stated in this part.
  - **Summary of findings**: After introduction there would appear a statement of findings and recommendations in non-technical language. If the findings are extensive, they should be summarised.
  - **Main report**: The main body of the report should be presented in logical sequence and broken-down into readily identifiable sections.
  - **Conclusion:** Towards the end of the main text, researcher should again put down the results of his research clearly and precisely. In fact, it is the final summing up.
- At the end of the report, appendices should be enlisted in respect of all technical data. Bibliography, i.e., list of books, journals, reports, etc., consulted, should also be given in the end. Index should also be given specially in a published research report. Report should be written in a concise and objective style in simple language avoiding vague expressions such as 'it seems,' 'there may be', and the like.
- Charts and illustrations in the main report should be used only if they present the information more clearly and forcibly.
- Calculated 'confidence limits' must be mentioned and the various constraints experienced in conducting research operations may as well be stated.

1.8 Criteria of Good Research

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

- The purpose of the research should be clearly defined and common concepts be used.
- The research procedure used should be described in sufficient detail to permit another researcher to repeat the research for further advancement, keeping the continuity of what has already been attained.
- The procedural design of the research should be carefully planned to yield results that are as objective as possible.

- The researcher should report with complete frankness, flaws in procedural design and estimate their effects upon the findings.
- The analysis of data should be sufficiently adequate to reveal its significance and the methods of analysis used should be appropriate. The validity and reliability of the data should be checked carefully.
- Conclusions should be confined to those justified by the data of the research and limited to those for which the data provide an adequate basis.
- Greater confidence in research is warranted if the researcher is experienced, has a good reputation in research and is a person of integrity.

The qualities of a good research

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

1.9 Problems Encountered by Researchers in India

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

- Lack of scientific training
  - The lack of a scientific training in the methodology of research is a great impediment for researchers in our country. There is paucity of competent researchers. Many researchers take a leap in the dark without knowing research methods. Most of the work, which goes in the name of research, is not methodologically sound. Research to many researchers and even to their guides, is mostly a scissor and paste job without any insight shed on the collated materials.
  - The consequence is obvious, viz., the research results, quite often, do not reflect the reality or realities. Thus, a systematic study of research methodology is an urgent necessity. Before undertaking research projects, researchers should be well equipped with all the methodological aspects. As such, efforts should be made to provide short duration intensive courses for meeting this requirement.
- Insufficient interaction
  - There is insufficient interaction between the university research departments on one side and business establishments, government departments and research institutions on the other side. A great deal of primary data of non-confidential nature remains untouched / untreated by the researchers for want of proper contacts.
  - Efforts should be made to develop satisfactory liaison among all concerned for better and realistic researches. There is need for developing some mechanisms of a university—industry interaction programme so that academics can get ideas from practitioners on what needs to be researched and practitioners can apply the research done by the academics.
- Lack of secrecy
  - Most of the business units in our country do not have the confidence that the material supplied by them to researchers will not be misused and as such they are often reluctant in supplying the needed information to researchers.

- The concept of secrecy seems to be sacrosanct to business organisations in the country so much so that it proves an impermeable barrier to researchers. Thus, there is the need for generating the confidence that the information/data obtained from a business unit will not be misused.

- Identification of research problems
  - Research studies overlapping one another are undertaken quite often for want of adequate information. This results in duplication and fritters away resources. This problem can be solved by proper compilation and revision, at regular intervals, of a list of subjects on which and the places where the research is going on.
  - Due attention should be given toward identification of research problems in various disciplines of applied science which are of immediate concern to the industries.

- Code of conduct
  - There is no existence of a code of conduct for researchers and inter-university and interdepartmental rivalries are also quite common. Hence, there is need for developing a code of conduct for researchers which, if adhered sincerely, can win over this problem.
  - Many researchers in our country also face the difficulty of adequate and timely secretarial assistance, including computerial assistance. This causes unnecessary delays in the completion of research studies. All possible efforts be made in this direction so that efficient secretarial assistance is made available to researchers and that too well in time. University Grants Commission must play a dynamic role in solving this difficulty.

- Library management
  - Library management and functioning is not satisfactory at many places and much of the time and energy of researchers are spent in tracing out the books, journals, reports, etc., rather than in tracing out relevant material from them.
  - There is also the problem that many of our libraries are not able to get copies of old and new Acts/Rules, reports and other government publications in time. This problem is felt more in libraries which are away in places from Delhi and/or the state capitals. Thus, efforts should be made for the regular and speedy supply of all governmental publications to reach our libraries.

- Availability of data
  - There is also the difficulty of timely availability of published data from various government and other agencies doing this job in our country. Researcher also faces the problem on account of the fact that the published data vary quite significantly because of differences in coverage by the concerning agencies.
  - There may, at times, take place the problem of conceptualisation and also problems relating to the process of data collection and related things.

Summary

- Research in common phrasing refers to a search for knowledge. It can be defined as a scientific and systematic search for applicable information on a specific topic.
- The systematic approach concerning generalisation and the formulation of a theory is also research. As such the term 'research' refers to the systematic method consisting of enunciating the problem, formulating a hypothesis, collecting the facts or data, analysing the facts and reaching certain conclusions either in the form of solutions(s) towards the concerned problem or in certain generalisations for some theoretical formulation.
- In social science, a lot of research is being carried on for both purposes and has resulted in the discovery of new facts as well as modification of old concepts. Constant verification of the old concepts is also needed, especially in case of dynamic sciences.
- Fundamental research may be entirely new discovery, the knowledge of which has not existed so far. The researcher is often born-genius, has a sharp intellect, is thirsty of knowledge and eventually has an ocean of knowledge in his possession and emerges gems of knowledge. This discovery may have nothing to do with an existing theory.
- Applied research has practical utility in the developing countries. Developing countries may benefit by the discoveries, scientific and otherwise, that have already been made, either by direct application or by making some modifications wherever necessary.
- Descriptive research as the name suggests is designed to describe something. For example, the characteristics of users of a given product, the degree to which product use varies with income etc. to be of maximum benefit, a descriptive study must collect data for a definite purpose. Descriptive studies vary in the degree to which a specific hypothesis is the guide.
- Action research is also a recent addition to the categories of research known to a modern social scientist. By its very definition, it is a research through launching of a direct action with the objective of obtaining workable solutions to the given problems. This also aims at collecting information from other sources that have direct or indirect bearing on the research programme.

References

- Kothari, C.R., 2004. *Research Methodology*. New Age International.
- Murthy, C., 2009.Research *methodology*, 1st ed., Vrinda Publications (P) Ltd.
- *Research Methodology,* [pdf] Available at: <http://www.ihmctan.edu/PDF/notes/Research_Methodology.pdf>. [Accessed 4 October 2011].
- *Research Methodology,* [Online] Available at: <https://www.tutorsindia.com/research-methodology-help.html>. [Accessed 4 October 2011].
- *Types of Research & Research Designs -- Rey Ty*., 2009. [Video Online] Available at: <http://www.youtube.com/watch?v=49_u-pXwx7g>. [Accessed 4 October 2011].
- *When Should I Use Qualitative Vs. Quantitative Research*?, 2009. [Video Online] Available at: <http://www.youtube.com/watch?v=638W_s5tRq8&feature=related>. [Accessed 4 October 2011].

Recommended Reading

- Khanzode, V. V., 2004. *Research Methodology*., APH Publishing.
- Jonker, J. and Pennink S., 2009. *The Essence of Research Methodology: A Concise Guide for Master and PhD Students in Management Science*., Springer Publication.
- Krishnaswamy, K. N., Sivakumar, I. A, and Mathirajan, M., 2009. *Management Research Methodology: Integration of Methods and Techniques.* Pearson Education India.

Self Assessment

1. _____ studies the behavior of human beings as members of society, and their feelings, response, attitudes under different circumstances.
   a. Social research
   b. Scientific research
   c. Scientific analysis
   d. Social analysis

2. In _____, the traders, farmers, and wage earners from whom the data were collected may not reveal their business secrets which also set certain limitations.
   a. personal interview
   b. social survey
   c. questionnaire
   d. research

3. _____ may be entirely new discovery, the knowledge of which has not existed so far.
   a. Social research
   b. Social survey
   c. Fundamental research
   d. Fundamental survey

4. Which of the following statements is true?
   a. The assumptions of a theory should always be well defined and implausible.
   b. The assumptions of a theory are always complicated and plausible.
   c. The assumptions of a theory should always be well defined and plausible.
   d. The assumptions of a theory should always be clear.

5. The _____ is conducted to test the empirical content or the basic assumptions or the very validity of theory under given conditions.
   a. social research
   b. fundamental research
   c. exploratory research
   d. applied research

6. A part of _____ is the investigation of relationships among variables without knowing why they are studied.
   a. exploratory research
   b. action research
   c. fundamental research
   d. qualitative research

7. The methods used for _____ are usually personal interviews method and the survey method.
   a. exploratory research
   b. action research
   c. qualitative research
   d. quantitative research

8. Experimentation will refer to that process of research in which _____.
   a. only one variable is manipulated under conditions, which permit the collection of data, which show the effects.
   b. maximum variables are manipulated under conditions, which permit the collection of data, which show the effects.
   c. one or more variables are manipulated under conditions, which permit the collection f data, which show the effects.
   d. evaluation takes place at different phases or stages of performance of the project.

9. _____ has become very popular these days as a scientific method for discovering relevant impact and inter-relationships of social psychological variables from given populations.
   a. Survey research
   b. Empirical research
   c. Action research
   d. Field investigation research

10. Match the following.

Type of research	Definition
1. Qualitative research	A. the scientist does not have direct control of independent variables because their manifestations have already occurred or because they are inherently not manipulative
2. Quantitative research	B. research through launching of a direct action with the objective of obtaining workable solutions to the given problems
3. Action research	C. based on the measurement of quantity or amount
4. Ex-post facto research	D. aims at discovering the underlying motives and desires, using in depth interviews for the purpose

   a. 1-D; 2-C; 3-B; 4-A
   b. 1-C; 2-D; 3-A; 4-B
   c. 1-A; 2-B; 3-D; 4-C
   d. 1-B; 2-A; 3-C; 4-D

Chapter II

Research Problem

Aim

The aim of this chapter is to:

- define the research problem
- highlight the necessity of defining the problem
- introduce the techniques involved in defining a problem

Objectives

The objectives of this chapter are to:

- elucidate research design
- discuss the need for research design
- elaborate features of a good design

Learning outcome

At the end of this chapter, you will be able to:

- discuss important concepts relating to research design
- enlist different research designs
- understand the basic principles of experimental designs

2.1 Defining the Research Problem

In research process, the first and foremost step is that of selecting and properly defining a research problem. A researcher must find the problem and formulate it, so that it becomes susceptible to research. Like a medical doctor, a researcher must examine all the symptoms (presented to him or observed by him) concerning a problem before he can diagnose correctly. To define a problem correctly, a researcher must know what a problem is.

2.1.1 What is a Research Problem?

A research problem, in general, refers to some difficulty which a researcher experiences in the context of either a theoretical or practical situation and wants to obtain a solution for the same. Usually we say that a research problem does exist if the following conditions are met with:

- There must be an individual (or a group or an organisation), let us call it 'I,' to whom the problem can be attributed. The individual or the organisation, as the case may be, occupies an environment, say 'N', which is defined by values of the uncontrolled variables, Yj.
- There must be at least two courses of action, say C1 and C2, to be pursued. A course of action is defined by one or more values of the controlled variables. For example, the number of items purchased at a specified time is said to be one course of action.
- There must be at least two possible outcomes, say O1 and O2, of the course of action, of which one should be preferable to the other. In other words, this means that there must be at least one outcome that the researcher wants, i.e., an objective.
- The courses of action available must provide some chance of obtaining the objective, but they cannot provide the same chance, otherwise the choice would not matter. Thus, if P (Oj | I, Cj, N) represents the probability that an outcome Oj will occur, if I select Cj in N, then PbO1| I , C1, Ng ¹ PbO1| I , C2 , Ng . In simple words, we can say that the choices must have unequal efficiencies for the desired outcomes.
- Over and above these conditions, the individual or the organisation can be said to have the problem only if 'I' does not know what course of action is best, i.e., 'I', must be in doubt about the solution.
- Thus, an individual or a group of persons can be said to have a problem which can be technically described as a research problem, if they (individual or the group), having one or more desired outcomes, are confronted with two or more courses of action that have some but not equal efficiency for the desired objective(s) and are in doubt about which course of action is best.
- We can, thus, state the components of a research problem as under:
  - There must be an individual or a group which has some difficulty or the problem.
  - There must be some objective(s) to be attained at. If one wants nothing, one cannot have a problem.
  - There must be alternative means (or the courses of action) for obtaining the objective(s) one wishes to attain. This means that there must be at least two means available to a researcher for if he has no choice of means, he cannot have a problem.
  - There must remain some doubt in the mind of a researcher with regard to the selection of alternatives. This means that research must answer the question concerning the relative efficiency of the possible alternatives.
  - There must be some environment(s) to which the difficulty pertains.
- Thus, a research problem is one which requires a researcher to find out the best solution for the given problem, i.e., to find out by which course of action the objective can be attained optimally in the context of a given environment. There are several factors which may result in making the problem complicated.
- For instance, the environment may change affecting the efficiencies of the courses of action or the values of the outcomes; the number of alternative courses of action may be very large; persons not involved in making the decision may be affected by it and react to it favourably or unfavourably, and similar other factors. All such elements (or at least the important ones) may be thought of in context of a research problem.

2.1.2 Selecting a Problem

The research problem undertaken for study must be carefully selected. The task is a difficult one, although it may not appear to be so. Help may be taken from a research guide in this connection. A research guide can at the most only help a researcher choose a subject. However, the following points may be observed by a researcher in selecting a research problem or a subject for research:

- Subject, which is overdone, should not be normally chosen, for it will be a difficult task to throw any new light in such a case.
- Controversial subject should not become the choice of an average researcher.
- Too narrow or too vague problems should be avoided.
- The subject selected for research should be familiar and feasible so that the related research material or sources of research are within one's reach. Even then it is quite difficult to supply definitive ideas concerning how a researcher should obtain ideas for his research. For this purpose, a researcher should contact an expert or a professor in the University who is already engaged in research. He may as well read articles published in current literature available on the subject and may think how the techniques and ideas discussed therein might be applied to the solution of other problems. He may discuss with others what he has in mind concerning a problem. In this way he should make all possible efforts in selecting a problem.
- The importance of the subject, the qualifications and the training of a researcher, the costs involved, and the time factor are few other criteria that must also be considered in selecting a problem. In other words, before the final selection of a problem is done, a researcher must ask himself the following questions:
  - Whether he is well equipped in terms of his background to carry out the research?
  - Whether the study falls within the budget he can afford?
  - Whether the necessary cooperation can be obtained from those who must participate in research as subjects? If the answers to all these questions are in the affirmative, one may become sure so far as the practicability of the study is concerned.
- The selection of a problem must be preceded by a preliminary study. This may not be necessary when the problem requires the conduct of a research closely similar to one that has already been done. But when the field of inquiry is relatively new and does not have available a set of well developed techniques, a brief feasibility study must always be undertaken.
- If the subject for research is selected properly by observing the above mentioned points, the research will not be a boring work. In fact, passion for work is a must. The subject or the problem selected must involve the researcher and must have an upper most place in his mind so that he may undertake all pains needed for the study.

2.1.3 Necessity of Defining a Problem

- The problem to be investigated must be defined decidedly for that will help to discriminate relevant data from the irrelevant ones. A proper definition of research problem will enable the researcher to be on the track whereas an ill-defined problem may create hurdles.
- Questions like: What data are to be collected? What characteristics of data are relevant and need to be studied? What relations are to be explored? What techniques are to be used for the purpose? And similar other questions crop up in the mind of the researcher who can well plan his strategy and find answers to all such questions only when the research problem has been well defined.
- Thus, defining a research problem properly is a condition for any study and is a step of the highest importance. In fact, formulation of a problem is often more essential than its solution. It is only on careful detailing the research problem that we can work out the research design and can smoothly carry on all the consequential steps involved while doing research.

2.1.4 Technique Involved in Defining a Problem

- Defining a problem involves the task of laying down boundaries within which a researcher shall study the problem with a pre-determined objective in view. How to define a research problem is undoubtedly a phenomenal task. However, it is a task that must be tackled intelligently to avoid the perplexity encountered in a research operation.
- The usual approach is that the researcher should himself pose a question (or in case someone else wants the researcher to carry on research, the concerned individual, organisation or an authority should pose the question to the researcher) and set-up techniques and procedures for throwing light on the question concerned for formulating or defining the research problem. But such an approach generally does not produce definitive results because the question phrased in such a fashion is usually in broad general terms and as such may not be in a form suitable for testing.
- Defining a research problem properly and clearly is a crucial part of a research study and must in no case be accomplished hurriedly. However, in practice this frequently overlooked which causes a lot of problems later on. Hence, the research problem should be defined in a systematic manner, giving due weightage to all relating points. The technique for the purpose involves the undertaking of the following steps generally one after the other:
  - statement of the problem in a general way
  - understanding the nature of the problem
  - surveying the available literature
  - developing the ideas through discussions
  - rephrasing the research problem into a working proposition

A brief description of all these points will be helpful.

Statement of the problem in a general way:

- First of all the problem should be stated in a broad general way, keeping in view either some practical concern or some scientific or intellectual interest. For this purpose, the researcher must immerse himself thoroughly in the subject matter concerning which he wishes to pose a problem. In case of social research, it is considered advisable to do some field observation and as such the researcher may undertake some sort of preliminary survey or what is often called pilot survey.
- Then the researcher can himself state the problem or he can seek the guidance of the guide or the subject expert in accomplishing this task. Often, the guide puts forth the problem in general terms, and it is then up to the researcher to narrow it down and phrase the problem in operational terms. In case there is some directive from an organisational authority, the problem then can be stated accordingly.
- The problem stated in a broad general way may contain various ambiguities which must be resolved by cool thinking and rethinking over the problem. At the same time the feasibility of a particular solution has to be considered and the same should be kept in view while stating the problem.

Understanding the nature of the problem:

- The next step in defining the problem is to understand its origin and nature clearly. The best way of understanding the problem is to discuss it with those who first raised it in order to find out how the problem originally came about and with what objectives in view. If the researcher has stated the problem himself, he should consider once again all those points that induced him to make a general statement concerning the problem.
- For a better understanding of the nature of the problem involved, he can enter into discussion with those who have a good knowledge of the problem concerned or similar other problems. The researcher should also keep in view the environment within which the problem is to be studied and understood.

Surveying the available literature:

- All available literature concerning the problem at hand must necessarily be surveyed and examined before a definition of the research problem is given. This means that the researcher must be well-conversant with relevant theories in the field, reports and records as also all other relevant literature. He must devote sufficient time in reviewing of research already undertaken on related problems. This is done to find out what data and other materials, if any, are available for operational purposes.
- "Knowing what data are available often serves to narrow the problem itself as well as the technique that might be used." This would also help a researcher to know if there are certain gaps in the theories, or whether the existing theories applicable to the problem under study are inconsistent with each other, or whether the findings of the different studies do not follow a pattern consistent with the theoretical expectations and so on.
- All this will enable a researcher to take new steps in the field for continuance of knowledge, i.e., he can move up starting from the existing premise. Studies on related problems are useful for indicating the type of difficulties that may be encountered in the present study as also the possible analytical shortcomings. At times such studies may also suggest useful and even new lines of approach to the present problem.

Developing the ideas through discussions:

- Discussion concerning a problem often produces useful information. Various new ideas can be developed through such an exercise. Hence, a researcher must discuss his problem with his colleagues and others who have enough experience in the same area or in working on similar problems. This is quite often known as an experience survey.
- People with big experience are in a position to enlighten the researcher on different aspects of his proposed study and their advice and comments are usually invaluable to the researcher. They help him sharpen his focus of attention on specific aspects within the field. Discussions with such persons should not only be confined to the formulation of the specific problem at hand, but should also be concerned with the general approach to the given problem, techniques that might be used, possible solutions, etc.
- Rephrasing the research problem: Finally, the researcher must sit to rephrase the research problem into a working proposition. Once the nature of the problem has been clearly understood, the environment (within which the problem has got to be studied) has been defined, discussions over the problem have taken place and the available literature has been surveyed and examined, rephrasing the problem into analytical or operational terms is not a difficult task. Through rephrasing, the researcher puts the research problem in as specific terms as possible so that it may become operationally viable and may help in the development of working hypotheses.
- In addition to what has been stated above, the following points must also be observed while defining a research problem:
  - Technical terms and words or phrases, with special meanings used in the statement of the problem, should be clearly defined.
  - Basic assumptions or postulates (if any) relating to the research problem should be clearly stated.
  - A straight forward statement of the value of the investigation (i.e., the criteria for the selection of the problem) should be provided.
  - The suitability of the time-period and the sources of data available must also be considered by the researcher in defining the problem.
  - The scope of the investigation or the limits within which the problem is to be studied must be mentioned explicitly in defining a research problem.

2.2 Research Design

The next step after stating the management problem, research purpose, and research hypotheses and questions, is to formulate a research design. The starting point for the research design is, in fact, the research questions and hypotheses that have been so carefully developed. In essence, the research design answers the question: How are we going to get answers to these research questions and test these hypotheses? The research design is a plan of action indicating the specific steps that are necessary to provide answers to those questions, test the hypotheses, and thereby achieve the research purpose that helps choose among the decision alternatives to solve the management problem or capitalise on the market opportunity.

2.2.1 Meaning of Research Design

- The difficult problem that follows the task of defining the research problem is the preparation of the design of the research project, popularly known as the "research design". Decisions regarding what, where, when, how much, by what means concerning an inquiry or a research study constitute a research design.
- "A research design is the arrangement of conditions for collection and analysis of data in a manner that aims to combine relevance to the research purpose with economy in procedure." In fact, the research design is the conceptual structure within which research is conducted; it constitutes the blueprint for the collection, measurement and analysis of data. As such the design includes an outline of what the researcher will do from writing the hypothesis and its operational implications to the final analysis of data. More explicitly, the design decisions happen to be in respect of:
  - What is the study about?
  - Why is the study being made?
  - Where will the study be carried out?
  - What type of data is required?
  - Where can the required data are found?
  - What periods of time will the study include?
  - What will be the sample design?
  - What techniques of data collection will be used?
  - How will the data be analysed?
  - In what style will the report be prepared?

Keeping in view the above stated design decisions; one may split the overall research design into the following parts:

- the sampling design which deals with the method of selecting items to be observed for the given study;
- the observational design which relates to the conditions under which the observations are to be made;
- the statistical design which concerns with the question of how many items are to be observed and how the information and data gathered are to be analysed; and
- the operational design which deals with the techniques by which the procedures specified in the sampling, statistical and observational designs can be carried out

2.2.2 Need for Research Design

- Research design is needed because it facilitates the smooth sailing of the various research operations, thereby making research as efficient as possible yielding maximal information with minimal expenditure of effort, time and money. Just as for better, economical and attractive construction of a house, we need a blueprint (or what is commonly called the map of the house) well thought out and prepared by an expert architect, similarly we need a research design or a plan in advance of data collection and analysis for our research project.
- Research design stands for advance planning of the methods to be adopted for collecting the relevant data and the techniques to be used in their analysis, keeping in view the objective of the research and the availability of staff, time and money. Preparation of the research design should be done with great care as any error in it may upset the entire project.

- Research design, in fact, has a great bearing on the reliability of the results arrived at and as such constitutes the firm foundation of the entire building of the research work. Even then the need for a well thought out research design is at times not realised by many. The importance which this problem deserves is not given to it. As a result, many researches do not serve the purpose for which they are undertaken. In fact, they may even give misleading conclusions.
- Thoughtlessness in designing the research project may result in rendering the research exercise is useless. It is, therefore, essential that an efficient and appropriate design must be prepared before starting research operations. The design helps the researcher to organise his ideas in a form whereby it will be possible for him to look for flaws and inadequacies. Such a design can even be given to others for their comments and critical evaluation. In the absence of such a course of action, it will be difficult for the critic to provide a comprehensive review of the proposed study.

2.2.3 Features of a Good Design

- A good design is often characterised by adjectives like flexible, appropriate, efficient, and economical and so on. Generally, the design which minimises bias and maximises the reliability of the data collected and analysed is considered a good design. The design which gives the smallest experimental error is supposed to be the best design in many investigations.
- Similarly, a design which yields maximal information and provides an opportunity for considering many different aspects of a problem is considered most appropriate and efficient design in respect of many research problems. Thus, the question of good design is related to the purpose or objective of the research problem and also with the nature of the problem to be studied. A design may be quite suitable in one case, but may be found wanting in one respect or the other in the context of some other research problem.
- One single design cannot serve the purpose of all types of research problems. A research design appropriate for a particular research problem, usually involves the consideration of the following factors:
  - the means of obtaining information
  - the availability and skills of the researcher and his staff, if any
  - the objective of the problem to be studied
  - the nature of the problem to be studied
  - the availability of time and money for the research work
- If the research study happens to be an exploratory or a formulative one, wherein the major emphasis is on discovery of ideas and insights, the research design most appropriate must be flexible enough to permit the consideration of many different aspects of a phenomenon. But when the purpose of a study is accurate description of a situation or of an association between variables (or in what are called the descriptive studies), accuracy becomes a major consideration and a research design which minimises bias and maximises the reliability of the evidence collected is considered a good design.
- Studies involving the testing of a hypothesis of a causal relationship between variables require a design which will permit inferences about causality in addition to the minimisation of bias and maximisation of reliability. But in practice it is the most difficult task to put a particular study in a particular group, for a given research may have in it elements of two or more of the functions of different studies. It is only on the basis of its primary function that a study can be categorised either as an exploratory or descriptive or hypothesis-testing study and accordingly the choice of a research design may be made in case of a particular study.
- Besides, the availability of time, money, skills of the research staff and the means of obtaining the information must be given due weightage while working out the relevant details of the research design such as experimental design, survey design, sample design and the like.

2.2.4 Important Concepts Related to Research Design

Before describing the different research designs, it will be appropriate to explain the various concepts relating to designs so that these may be better and easily understood.

Dependent and independent variables:

- A concept which can take on different quantitative values is called a variable. As such the concepts like weight, height, income are all examples of variables. Qualitative phenomena (or the attributes) are also quantified on the basis of the presence or absence of the concerning attribute(s). Phenomena which can take on quantitatively different values even in decimal points are called 'continuous variables'. But all variables are not continuous.
- If they can only be expressed in integer values, they are non-continuous variables or in statistical language 'discrete variables'. Age is an example of continuous variable, but the number of children is an example of non-continuous variable. If one variable depends upon or is a consequence of the other variable, it is termed as a dependent variable, and the variable that is antecedent to the dependent variable is termed as an independent variable.
- For instance, if we say that height depends upon age, then height is a dependent variable and age is an independent variable. Further, if in addition to being dependent upon age, height also depends upon the individual's sex, then height is a dependent variable and age and sex are independent variables. Similarly, readymade films and lectures are examples of independent variables, whereas behavioural changes, occurring as a result of the environmental manipulations, are examples of dependent variables.

Extraneous variable:

- Independent variables that are not related to the purpose of the study, but may affect the dependent variable are termed as extraneous variables. Suppose the researcher wants to test the hypothesis that there is a relationship between children's gains in social studies achievement and their self-concepts. In this case self-concept is an independent variable and social studies achievement is a dependent variable.
- Intelligence may as well affect the social studies achievement, but since it is not related to the purpose of the study undertaken by the researcher, it will be termed as an extraneous variable. Whatever effect is noticed on dependent variable as a result of extraneous variable(s) is technically described as an 'experimental error'. A study must always be so designed that the effect upon the dependent variable is attributed entirely to the independent variable(s), and not to some extraneous variable or variables.

Control:

- One important characteristic of a good research design is to minimise the influence or effect of extraneous variable(s). The technical term 'control' is used when we design the study minimising the effects of extraneous independent variables. In experimental researches, the term 'control' is used to refer to restrain experimental conditions.

Confounded relationship:

- When the dependent variable is not free from the influence of extraneous variable(s), the relationship between the dependent and independent variables is said to be confounded by an extraneous variable(s).

Research hypothesis:

- When a prediction or a hypothesised relationship is to be tested by scientific methods, it is termed as research hypothesis. The research hypothesis is a predictive statement that relates an independent variable to a dependent variable. Usually a research hypothesis must contain, at least, one independent and one dependent variable. Predictive statements which are not to be objectively verified or the relationships that are assumed but not to be tested are not termed research hypotheses.

Experimental and non-experimental hypothesis-testing research:

- When the purpose of research is to test a research hypothesis, it is termed as hypothesis-testing research. It can be of the experimental design or of the non-experimental design. Research in which the independent variable is manipulated is termed 'experimental hypothesis-testing research' and a research in which an independent variable is not manipulated is called 'non-experimental hypothesis-testing research'.

- For instance, suppose a researcher wants to study whether intelligence affects reading ability for a group of students and for this purpose he randomly selects 50 students and tests their intelligence and reading ability by calculating the coefficient of correlation between the two sets of scores. This is an example of non-experimental hypothesis-testing research because herein the independent variable, intelligence, is not manipulated.
- But now suppose that our researcher randomly selects 50 students from a group of students who are to take a course in statistics and then divides them into two groups by randomly assigning 25 to Group A, the usual studies programme, and 25 to Group B, the special studies programme.
- At the end of the course, he administers a test to each group in order to judge the effectiveness of the training programme on the student's performance-level. This is an example of experimental hypothesis-testing research because in this case the independent variable, viz., the type of training programme, is manipulated.

Experimental and control groups:

- In an experimental hypothesis-testing research when a group is exposed to usual conditions, it is termed a 'control group', but when the group is exposed to some novel or special condition, it is termed an 'experimental group'. In the above illustration, the Group A can be called a control group and the Group B an experimental group. If both groups A and B are exposed to special studies programmes, then both groups would be termed 'experimental groups.'
- It is possible to design studies which include only experimental groups or studies which include both experimental and control groups.

Treatments:

- The different conditions under which experimental and control groups are put are usually referred to as 'treatments'. In the illustration taken above, the two treatments are the usual studies programme and the special studies programme. Similarly, if we want to determine through an experiment the comparative impact of three varieties of fertilisers on the yield of wheat, in that case the three varieties of fertilisers will be treated as three treatments.

Experiment:

- The process of examining the truth of a statistical hypothesis, relating to some research problem, is known as an experiment. For example, we can conduct an experiment to examine the usefulness of a certain newly developed drug. Experiments can be of two types, viz., absolute experiment and comparative experiment.
- If we want to determine the impact of a fertiliser on the yield of a crop, it is a case of absolute experiment; but if we want to determine the impact of one fertiliser as compared to the impact of some other fertiliser, our experiment then will be termed as a comparative experiment. Often, we undertake comparative experiments when we talk of designs of experiments.

Experimental unit(s):

The pre-determined plots or the blocks, where different treatments are used, are known as experimental units. Such experimental units must be selected (defined) very carefully.

2.3 Different Research Designs

Different research designs can be conveniently described if they are categorised as:

- research design in case of exploratory research studies;
- research design in case of descriptive and diagnostic research studies, and
- research design in case of hypothesis-testing research studies

Each category is taken up separately.

Research design in case of exploratory research studies:

- Exploratory research studies are also termed as formulative research studies. The main purpose of such studies is that of formulating a problem for more precise investigation or of developing the working hypotheses from an operational point of view. The major emphasis in such studies is on the discovery of ideas and insights.
- As such the research design appropriate for such studies must be flexible enough to provide opportunity for considering different aspects of a problem under study. Inbuilt flexibility in research design is needed because the research problem, broadly defined initially, is transformed into one with more precise meaning in exploratory studies, which fact may necessitate changes in the research procedure for gathering relevant data.
- Generally, the following three methods in the context of research design for such studies are talked about:
  - the survey of concerning literature;
  - the experience survey and
  - the analysis of 'insight-stimulating' examples
- The survey of concerning literature happens to be the most simple and fruitful method of formulating precisely the research problem or developing hypothesis. Hypotheses stated by earlier workers may be reviewed and their usefulness be evaluated as a basis for further research. It may also be considered whether the already stated hypotheses suggest new hypothesis. In this way the researcher should review and build upon the work already done by others, but in cases where hypotheses have not yet been formulated, his task is to review the available material for deriving the relevant hypotheses from it.
- Besides, the bibliographical survey of studies, already made in one's area of interest may as well as made by the researcher for precisely formulating the problem. He should also make an attempt to apply concepts and theories developed in different research contexts to the area in which he is himself working. Sometimes the works of creative writers also provide a fertile ground for hypothesis formulation and as such may be looked into by the researcher.
- Experience survey means the survey of people who have had practical experience with the problem to be studied. The object of such a survey is to obtain insight into the relationships between variables and new ideas relating to the research problem. For such a survey people who are competent and can contribute new ideas may be carefully selected as respondents to ensure a representation of different types of experience. The respondents so selected may then be interviewed by the investigator.
- The researcher must prepare an interview schedule for the systematic questioning of informants. But the interview must ensure flexibility in the sense that the respondents should be allowed to raise issues and questions which the investigator has not previously considered. Generally, the experience collecting interview is likely to be long and may last for few hours.
- Hence, it is often considered desirable to send a copy of the questions to be discussed to the respondents well in advance. This will also give an opportunity to the respondents for doing some advance thinking over the various issues involved so that, at the time of interview, they may be able to contribute effectively. Thus, an experience survey may enable the researcher to define the problem more concisely and help in the formulation of the research hypothesis.
- This survey may as well provide information about the practical possibilities for doing different types of research. Analysis of 'insight-stimulating' examples is also a fruitful method for suggesting hypotheses for research. It is particularly suitable in areas where there is little experience to serve as a guide. This method consists of the intensive study of selected instances of the phenomenon in which one is interested. For this purpose the existing records, if any, may be examined, the unstructured interviewing may take place, or some other approach may be adopted.
- Attitude of the investigator, the intensity of the study and the ability of the researcher to draw together diverse information into a unified interpretation are the main features which make this method an appropriate procedure for evoking insights. Now, what examples are to be selected and studied? There is no clear cut answer to it. Experience indicates that for particular problems certain types of instances are more appropriate than others. One can mention few examples of 'insight-stimulating' cases such as the reactions of strangers, the reactions of marginal individuals, the study of individuals who are in transition from one stage to another, the reactions of individuals from different social strata and the like.

- In general, cases that provide sharp contrasts or have striking features are considered relatively more useful while adopting this method of hypotheses formulation. Thus, in an exploratory of formulative research study which merely leads to insights or hypotheses, whatever method or research design outlined above is adopted, the only thing essential is that it must continue to remain flexible so that many different facets of a problem may be considered as and when they arise and come to the notice of the researcher.

Research design in case of descriptive and diagnostic research studies:

- Descriptive research studies are those studies which are concerned with describing the characteristics of a particular individual, or of a group, whereas diagnostic research studies determine the frequency with which something occur or its association with something else. The studies concerning whether certain variables are associated are examples of diagnostic research studies.
- As against this, studies concerned with specific predictions, with narration of facts and characteristics concerning individual, group or situation are all examples of descriptive research studies. Most of the social research comes under this category. From the point of view of the research design, the descriptive as well as diagnostic studies share common requirements and as such we may group together these two types of research studies.
- In descriptive as well as in diagnostic studies, the researcher must be able to define clearly, what he wants to measure and must find adequate methods for measuring it along with a clear cut definition of 'population' he wants to study. Since the aim is to obtain complete and accurate information in the said studies, the procedure to be used must be carefully planned. The research design must make enough provision for protection against bias and must maximise reliability, with due concern for the economical completion of the research study.
- The design in such studies must be rigid and not flexible and must focus attention on the following:

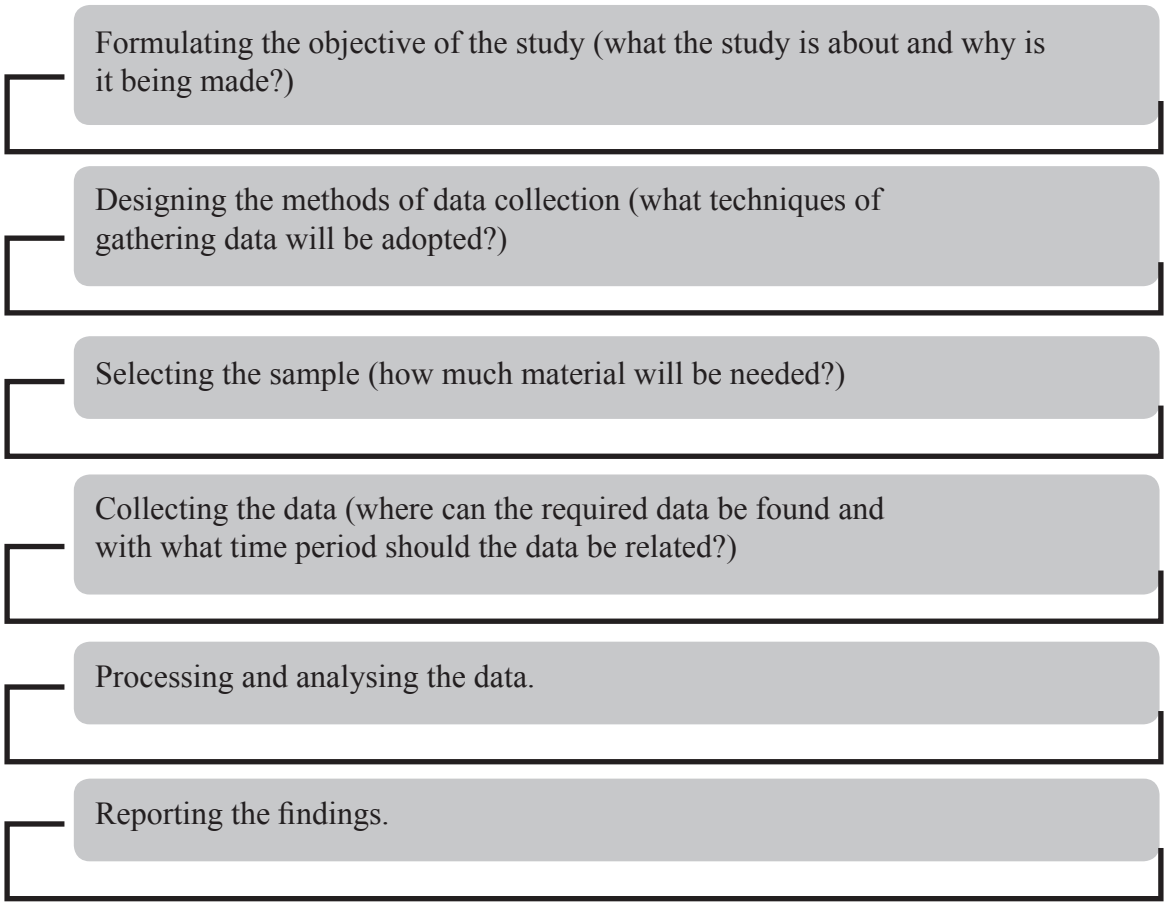


Fig. 2.1 Research design in case of descriptive and diagnostic research studies

- In a descriptive/diagnostic study, the first step is to specify the objectives with sufficient precision to ensure that the data collected are relevant. If this is not done carefully, the study may not provide the desired information. Then comes the question of selecting the methods by which the data are to be obtained. In other words, techniques for collecting the information must be devised.
- Several methods (viz., observation, questionnaires, interviewing, examination of records, etc.), with their merits and limitations, are available for the purpose and the researcher may user one or more of these methods which have been discussed in detail in later chapters. While designing data-collection procedure, adequate safeguards against bias and unreliability must be ensured.
- Whichever method is selected, questions must be well examined and be made unambiguous; interviewers must be instructed not to express their own opinion; observers must be trained so that they uniformly record a given item of behaviour. It is always desirable to pre-test the data collection instruments before they are finally used for the study purposes.
- In other words, we can say that "structured instruments" are used in such studies. In most of the descriptive/ diagnostic studies, the researcher takes out sample(s) and then wishes to make statements about the population on the basis of the sample analysis or analyses. More often than not, sample has to be designed. Different sample designs have been discussed in detail in a separate chapter in this book.
- Here, we may only mention that the problem of designing samples should be tackled in such a fashion that the samples may yield accurate information with a minimum amount of research effort. Usually, one or more forms of probability sampling, or what is often described as random sampling, are used.
- To obtain data free from errors introduced by those responsible for collecting them, it is necessary to supervise closely the staff of field workers as they collect and record information. Checks may be set up to ensure that the data collecting staffs performs their duty honestly and without prejudice. "As data are collected, they should be examined for completeness, comprehensibility, consistency and reliability."
- The data collected must be processed and analysed. This includes steps like coding the interview replies, observations, etc.; tabulating the data; and performing several statistical computations. To the extent possible, the processing and analysing procedure should be planned in detail before actual work is started. This will prove economical in the sense that the researcher may avoid unnecessary labour such as preparing tables for which he later finds he has no use or on the other hand, re-doing some tables because he failed to include relevant data.
- Coding should be done carefully to avoid error in coding and for this purpose the reliability of coders needs to be checked. Similarly, the accuracy of tabulation may be checked by having a sample of the tables re-done. In case of mechanical tabulation the material (i.e., the collected data or information) must be entered on appropriate cards which are usually done by punching holes corresponding to a given code.
- The accuracy of punching is to be checked and ensured. Finally, statistical computations are needed and as such averages, percentages and various coefficients must be worked out. Probability and sampling analysis may as well be used. The appropriate statistical operations, along with the use of appropriate tests of significance should be carried out to safeguard the drawing of conclusions concerning the study.
- Last of all comes the question of reporting the findings. This is the task of communicating the findings to others and the researcher must do it in an efficient manner. The layout of the report needs to be well planned so that all things relating to the research study may be well presented in simple and effective style.
- Thus, the research design in case of descriptive/diagnostic studies is a comparative design throwing light on all points narrated above and must be prepared keeping in view the objective(s) of the study and the resources available. However, it must ensure the minimisation of bias and maximisation of reliability of the evidence collected. The said design can be appropriately referred to as a survey design since it takes into account all the steps involved in a survey concerning a phenomenon to be studied.
- The difference between research designs in respect of the above two types of research studies can be conveniently summarised in tabular form as under:

Research design	Type of study	
	Exploratory of Formulate	**Descriptive / Diagnostic**
Overall design	Flexible design (design must provide opportunity for considering different aspects of the problem)	Rigid design (design must make enough provision for protection against bias and must maximise reliability)
Sampling design	Non probability sampling design (purposive or judgement sampling)	Probability sampling design (random sampling)
Statistical design	No pre-planned design for analysis	Pre-planned design for analysis
Observational design	Unstructured instruments for collection of data	Structured or well thought out instruments for collection of data
Operational design	No fixed decisions about the operational procedures	Advanced decisions about operational procedures

Table 2.1 Difference between research designs

Research design in case of hypothesis-testing research studies:

- Hypothesis-testing research studies (generally known as experimental studies) are those where the researcher tests the hypotheses of causal relationships between variables. Such studies require procedures that will not only reduce bias and increase reliability, but will permit drawing inferences about causality. Usually experiments meet this requirement. Hence, when we talk of research design in such studies, we often mean the design of experiments.
- Professor R.A. Fisher's name is associated with experimental designs. Beginning of such designs was made by him when he was working at Rothamsted Experimental Station (Centre for Agricultural Research in England). As such the study of experimental designs has its origin in agricultural research. Professor Fisher found that by dividing agricultural fields or plots into different blocks and then by conducting experiments in each of these blocks, whatever information is collected and inferences drawn from them, happens to be more reliable.
- This fact inspired him to develop certain experimental designs for testing hypotheses concerning scientific investigations. Today, the experimental designs are being used in researches relating to phenomena of several disciplines. Since experimental designs originated in the context of agricultural operations, we still use, though in a technical sense, several terms of agriculture (such as treatment, yield, plot, block etc.) in experimental designs.

2.4 Basic Principles of Experimental Designs

Professor Fisher has enumerated three principles of experimental designs:

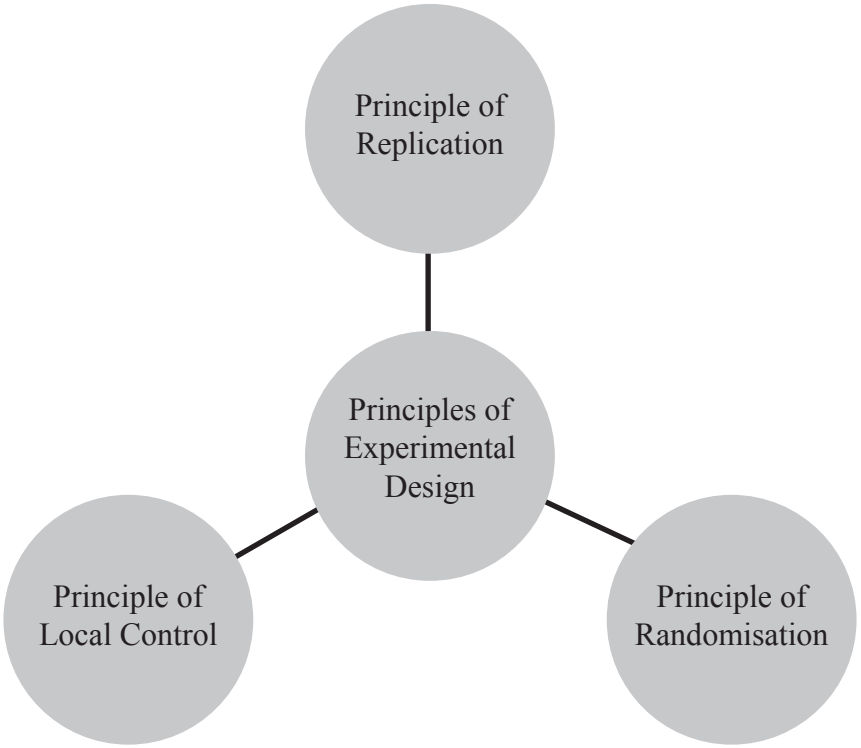


Fig. 2.2 Principles of experimental design

- According to the Principle of Replication, the experiment should be repeated more than once. Thus, each treatment is applied in many experimental units instead of one. By doing so, the statistical accuracy of the experiments is increased. For example, suppose we are to examine the effect of two varieties of rice. For this purpose, we may divide the field into two parts and grow one variety in one part and the other variety in the other part. We can then compare the yield of the two parts and draw conclusion on that basis.
- But if we are to apply the principle of replication to this experiment, then we first divide the field into several parts, grow one variety in half of these parts and the other variety in the remaining parts. We can then collect the data of yield of the two varieties and draw conclusion by comparing the same. The result so obtained will be more reliable in comparison to the conclusion we draw without applying the principle of replication.
- The entire experiment can even be repeated several times for better results. Conceptually replication does not present any difficulty, but computationally it does. For example, if an experiment requiring a two-way analysis of variance is replicated, it will then require a three-way analysis of variance since replication itself may be a source of variation in the data.
- However, it should be remembered that replication is introduced in order to increase the precision of a study; that is to say, to increase the accuracy with which the main effects and interactions can be estimated. The Principle of Randomisation provides protection, when we conduct an experiment, against the effect of extraneous factors by randomisation.
- In other words, this principle indicates that we should design or plan the experiment in such a way that the variations caused by extraneous factors can all be combined under the general heading of "chance." For instance, if we grow one variety of rice, say, in the first half of the parts of a field and the other variety is grown in the other half, then it is just possible that the soil fertility may be different in the first half in comparison to the other half.

- If this is so, our results would not be realistic. In such a situation, we may assign the variety of rice to be grown in different parts of the field on the basis of some random sampling technique, i.e., we may apply randomisation principle and protect ourselves against the effects of the extraneous factors (soil fertility differences in the given case).
- The Principle of Local Control is another important principle of experimental designs. Under it the extraneous factor, the known source of variability, is made to vary deliberately over as wide a range as necessary and this need to be done in such a way that the variability it causes can be measured and hence eliminated from the experimental error.
- This means that we should plan the experiment in a manner that we can perform a two-way analysis of variance, in which the total variability of the data is divided into three components attributed to treatments (varieties of rice in our case), the extraneous factor (soil fertility in our case) and experimental error.
- In other words, according to the principle of local control, we first divide the field into several homogeneous parts, known as blocks, and then each such block is divided into parts equal to the number of treatments. Then the treatments are randomly assigned to these parts of a block. Dividing the field into several homogenous parts is known as 'blocking'.
- In general, blocks are the levels at which we hold an extraneous factor fixed, so that we can measure its contribution to the total variability of the data by means of a two-way analysis of variance. In brief, through the principle of local control we can eliminate the variability due to extraneous factor(s) from the experimental error.

Summary

- A research problem, in general, refers to some difficulty which a researcher experiences in the context of either a theoretical or practical situation and wants to obtain a solution for the same.
- The subject selected for research should be familiar and feasible so that the related research material or sources of research are within one's reach. Even then it is quite difficult to supply definitive ideas concerning how a researcher should obtain ideas for his research. For this purpose, a researcher should contact an expert or a professor in the University who is already engaged in research.
- Defining a research problem properly is a prerequisite for any study and is a step of the highest importance. In fact, formulation of a problem is often more essential than its solution.
- The usual approach is that the researcher should himself pose a question and set-up techniques and procedures for throwing light on the question concerned for formulating or defining the research problem.
- The problem stated in a broad general way may contain various ambiguities which must be resolved by cool thinking and rethinking over the problem. At the same time the feasibility of a particular solution has to be considered and the same should be kept in view while stating the problem.
- All available literature concerning the problem at hand must necessarily be surveyed and examined before a definition of the research problem is given. This means that the researcher must be well-conversant with relevant theories in the field, reports and records as also all other relevant literature.
- People with big experience are in a position to enlighten the researcher on different aspects of his proposed study and their advice and comments are usually invaluable to the researcher.
- Research design is needed because it facilitates the smooth sailing of the various research operations, thereby making research as efficient as possible yielding maximal information with minimal expenditure of effort, time and money.

References

- Ross, S. M. and Morrison, G. R., *Experimental Research Methods* [PDF] Available at: <http://www.aect.org/edtech/ed1/38.pdf>. [Accessed 4 October 2011].
- Singh, Y. K. and Bajpai, R. B., 2008. *Research methodology: techniques and trends,* APH Publishing Corporation.
- Kothari, C.R., 2004. *Research Methodology*. New Age International.
- *Research Problem, Hypotheses and Variables* [Online] Available at: <http://www.scribd.com/doc/16739725/Research-Problem-Hypotheses-and-Variables>. [Accessed 4 October 2011].
- *The Research Problem* [Online] Available at: <http://www.public.asu.edu/~kroel/www500/The%20Research%20Problem.pdf>. [Accessed 4 October 2011].
- *Research Design*. 2011. [Video Online] Available at: <http://www.youtube.com/watch?v=H2v9Xh2iQV8&feature=related>. [Accessed 4 October 2011].
- *The Research Problem,* 2011. [Video Online] Available at: <http://www.youtube.com/watch?v=d3mwAK_lI1E>. [Accessed 4 October 2011].

Recommended Reading

- Marczyk, G. R., DeMatteo, D. and Festinger, D., 2010. *Essentials of Research Design and Methodology*. John Wiley and Sons.
- Dane, C. F., 2010. *Evaluating Research: Methodology for People Who Need to Read Research*. Sage Publications.
- William, W., 2009. *Research Methods in Education.* Pearson Education India.

Self Assessment

1. A _____ refers to some difficulty which a researcher experiences in the context of either a theoretical or practical situation and wants to obtain a solution for the same.
   a. research problem
   b. research
   c. research process
   d. problem

2. Subject which is overdone should not be normally chosen, _____.
   a. there is a need of variety of subjects.
   b. for it will be a difficult task to throw any new light in such a case.
   c. it will be easy to follow the problem
   d. else, there will be more resources of same subject

3. Which of the following statements is false?
   a. Formulation of a problem is often more essential than its solution.
   b. Defining a research problem properly is a condition for any study.
   c. A proper definition of research problem will enable the researcher to be on the track whereas an ill-defined problem may create hurdles.
   d. Subject which is overdone should be normally chosen

4. Discussion concerning a problem often produces useful information; this is known as _____.
   a. Experience survey
   b. Sharing of experiences
   c. Brain storming
   d. Sharing of problems

5. The next step after stating the management problem, research purpose, and research hypotheses and questions, is to formulate a _____.
   a. question
   b. survey
   c. research design
   d. interview

6. Generally, the design which minimises bias and maximises the reliability of the data collected and analysed is considered a_____.
   a. standard design
   b. milestone
   c. benchmark
   d. good design

7. A concept which can take on different quantitative values is called a _____.
   a. Variable
   b. Subject
   c. Factor
   d. Design

8. Phenomena which can take on quantitatively different values even in decimal points are called _____.
   a. variables
   b. continuous variables
   c. simple variables
   d. quantitative variables

9. Independent variables that are not related to the purpose of the study, but may affect the dependent variable are termed as _____.
   a. dependent variables
   b. independent variables
   c. extraneous variables
   d. continuous variables

10. Match the following.

Design	Definition
1. Sampling design	A. deals with the techniques by which the procedures specified in the sampling, statistical and observational designs can be carried out
2. Observational design	B. concerns with the question of how many items are to be observed and how the information and data gathered are to be analysed
3. statistical design	C. relates to the conditions under which the observations are to be made
4. operational design	D. deals with the method of selecting items to be observed for the given study

   a. 1-A; 2-B; 3-C; 4-D
   b. 1-B; 2-A; 3-D; 4-C
   c. 1-D; 2-C; 3-B; 4-A
   d. 1-C; 2-D; 3-A; 4-B

Chapter III

Sampling Fundamentals and Sampling Design

Aim

The aim of this chapter is to:

- define need for sampling
- explain important sampling distributions
- introduce the essentials of good sample

Objectives

The objectives of this chapter are to:

- elucidatee central limit theorem
- enlist different sampling methods
- explicate sampling theory

Learning outcome

At the end of this chapter, you will be able to:

- discuss Standler's A test
- understand the concept of standard error
- comprehend sample size and its discrimination

3.1 Introduction

Sampling may be defined as the selection of some part of an aggregate or totality on the basis of which a judgement or inference about the aggregate or totality is made. In other words, it is the process of obtaining information about an entire population by examining only a part of it. In most of the research work and surveys, the usual approach happens to be to make generalisations or to draw inferences based on samples about the parameters of population from which the samples are taking. The researcher quite often selects only a few items from the universe for his study purposes. All this is done on the assumption that the sample data will enable him to estimate the population parameters. The items so selected constitute what is technically called a sample, their selection process or technique is called sample design and the survey conducted on the basis of sample is described as sample survey. Sample should be truly representative of population characteristics without any bias so that it may result in valid and reliable conclusions.

3.1.1 Need for Sampling

Sampling is used in practice for a variety of reasons such as,

- Sampling can save time and money. A sample study is usually less expensive than a census study and produces results at a relatively faster speed.
- Sampling may enable more accurate measurements for a sample study is generally conducted by trained and experienced investigators.
- Sampling remains the only way when population contains infinitely many members.
- Sampling remains the only choice when a test involves the destruction of the item under study.
- Sampling usually enables to estimate the sampling errors and, thus, assists in obtaining information concerning some characteristic of the population.

3.1.2 Definition

Following are some fundamental definitions concerning sampling concepts and principles:

Universe/Population:

- From a statistical point of view, the term 'Universe' refers to the total of the items or units in any field of inquiry, whereas the term 'population' refers to the total of items about which information is desired.
- The attributes that are the object of study are referred to as characteristics and the units possessing them are called as elementary units. The aggregate of such units is generally described as population. Thus, all units in any field of inquiry form universe and all elementary units (on the basis of one characteristic or more) form population.
- Many a times it is difficult to find any difference between population and universe, and as such the two terms are taken as interchangeable. However, a researcher must necessarily define these terms precisely.
- The population or universe can be finite or infinite. The population is said to be finite if it consists of a fixed number of elements so that it is possible to enumerate it in its totality. For instance, the population of a city, the number of workers in a factory are examples of finite populations. The symbol 'N' is generally used to indicate how many elements (or items) are there in case of a finite population.
- An infinite population is that population in which it is theoretically impossible to observe all the elements. Thus, in an infinite population the number of items is infinite i.e., we cannot have any idea about the total number of items. The number of stars in a sky, possible rolls of a pair of dice are examples of infinite population.
- From a practical consideration, we then use the term infinite population for a population that cannot be enumerated in a reasonable period of time. This way we use the theoretical concept of infinite population as an approximation of a very large finite population.

Sampling frame:

- The elementary units or the group or cluster of such units may form the basis of sampling process in which case they are called as sampling units. A list containing all such sampling units is known as sampling frame. Thus, sampling frame consists of a list of items from which the sample is to be drawn.
- If the population is finite and the time frame is in the present or past, then it is possible for the frame to be identical with the population. In most cases they are not identical because it is often impossible to draw a sample directly from population.
- As such this frame is either constructed by a researcher for the purpose of his study or may consist of some existing list of the population. For instance, one can use telephone directory as a frame for conducting opinion survey in a city. Whatever the frame may be, it should be a good representative of the population.

Sampling design:

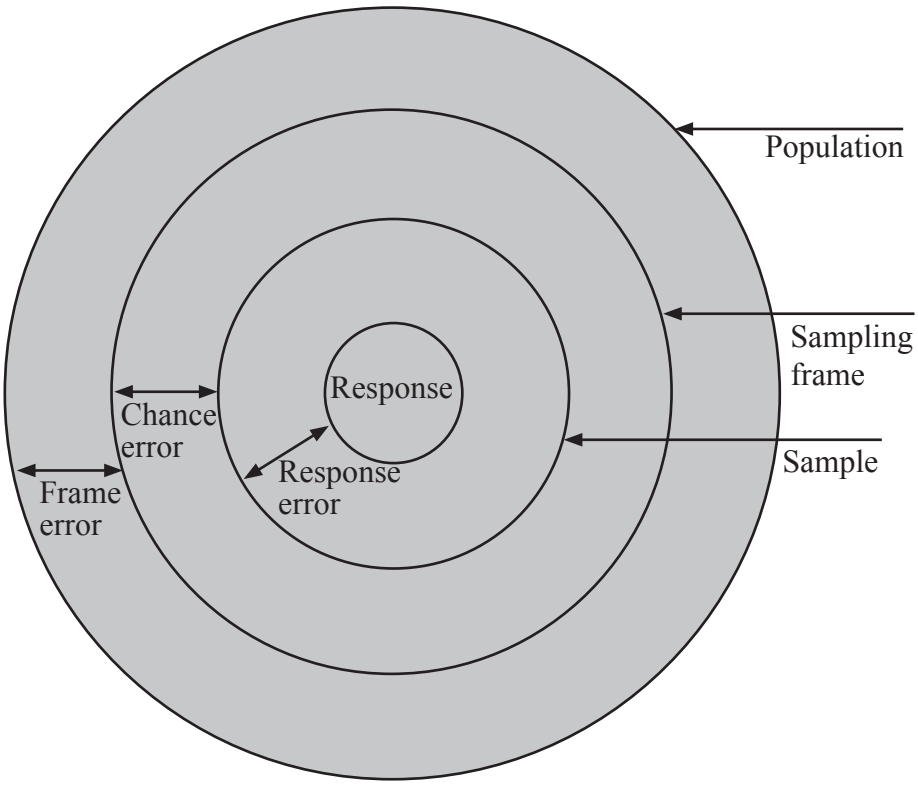
- A sample design is a definite plan for obtaining a sample from the sampling frame. It refers to the technique or the procedure the researcher would adopt in selecting some sampling units from which inferences about the population is drawn. Sampling design is determined before any data are collected.

Statistic(s) and parameter(s):

- A statistic is a characteristic of a sample, whereas a parameter is a characteristic of a population. Thus, when we work out certain measures such as mean, median, mode or the like ones from samples, then they are called statistic(s) for they describe the characteristics of a sample. But when such measures describe the characteristics of a population, they are known as parameter(s). For instance, the samples mean (X) is a statistic. To obtain the estimate of a parameter from a statistic constitutes the prime objective of sampling analysis.

Sampling error:

- Sample surveys do imply the study of a small portion of the population and as such there would naturally be a certain amount of inaccuracy in the information collected. This inaccuracy may be termed as sampling error or error variance.
- In other words, sampling errors are those errors, which arise on account of sampling and they generally happen to be random variations (in case of random sampling) in the sample estimates around the true population values. The meaning of sampling error can be easily understood from the following diagram:



Sampling error = Frame error
+ chance error + response error.
(If we add measurement error or the non-sampling error to sampling error, we get total error)

Fig. 3.1 Sampling error

(Source: Dr. Kothari, C. R., 2008. Research methodology: methods and techniques, 2nd ed., New Age International)

- Sampling error = Frame error + Chance error + Response error (If we add measurement error or the non-sampling error to sampling error, we get total error). Sampling errors occur randomly and are equally likely to be in either direction. The magnitude of the sampling error depends upon the nature of the universe; the more homogeneous the universe, the smaller the sampling error. Sampling error is inversely related to the size of the sample, i.e., sampling error decreases as the sample size increases and vice-versa.
- A measure of the random sampling error can be calculated for a given sample design and size and this measure is often called the precision of the sampling plan. Sampling error is usually worked out as the product of the critical value at a certain level of significance and the standard error.
- As opposed to sampling errors, we may have non-sampling errors which may creep in during the process of collecting actual information and such errors occur in all surveys whether census or sample. There is no way to measure non-sampling errors.

Precision:

- Precision is the range within which the population average (or other parameter) will lie in accordance with the reliability specified in the confidence level as a percentage of the estimate ± or as a numerical quantity. For instance, if the estimate is Rs 4000 and the precision desired is ± 4%, then the true value will be no less than Rs 3840 and no more than Rs 4160.
- This is the range (Rs 3840 to Rs 4160) within which the true answer should lie. But if we desire that the estimate should not deviate from the actual value by more than Rs 200 in either direction, in that case the range would be Rs 3800 to Rs 4200.

Confidence level and significance level:

- The confidence level or reliability is the expected percentage of times that the actual value will fall within the stated precision limits. Thus, if we take a confidence level of 95%, then we mean that there are 95 chances in 100 (or .95 in 1) that the sample results represent the true condition of the population within a specified precision range against 5 chances in 100 (or .05 in 1) that it does not.
- Precision is the range within which the answer may vary and still be acceptable; confidence level indicates the likelihood that the answer will fall within that range, and the significance level indicates the likelihood that the answer will fall outside that range. If the confidence level is 95%, then the significance level will be (100 – 95) i.e., 5%; if the confidence level is 99%, the significance level is (100 – 99) i.e., 1%, and so on.
- Even area of normal curve within precision limits for the specified confidence level constitute the acceptance region and the area of the curve outside these limits in either direction constitutes the rejection regions.*

Sampling distribution:

- We are often concerned with sampling distribution in sampling analysis. If we take certain number of samples and for each sample compute various statistical measures such as mean, standard deviation, etc., then we can find that each sample may give its own value for the statistic under consideration.
- All such values of a particular statistic, say mean, together with their relative frequencies will constitute the sampling distribution of the particular statistic, say mean.
- Accordingly, we can have sampling distribution of mean, or the sampling distribution of standard deviation or the sampling distribution of any other statistical measure. It may be noted that each item in a sampling distribution is a particular statistic of a sample.
- The sampling distribution tends quite closer to the normal distribution if the number of samples is large. The significance of sampling distribution follows from the fact that the mean of a sampling distribution is the same as the mean of the universe. Thus, the mean of the sampling distribution can be taken as the mean of the universe.

3.1.3 Important Sampling Distributions

Some important sampling distributions, which are commonly used, are:

- Sampling distribution of mean
- Sampling distribution of proportion
- Student's 't' distribution
- F distribution
- Chi-square distribution

Sampling distribution of mean:

- Sampling distribution of mean refers to the probability distribution of all the possible means of random samples of a given size that we take from a population. If samples are taken from a normal population, N (μ, σp) the sampling distribution of mean would also be normal with mean $\mu_{\bar{x}} = \mu$ and standard deviation = $\sigma p \sqrt{n}$, where μ is the mean of the population, σp is the standard deviation of the population and n means the number of items in a sample.
- But when sampling is from a population which is not normal (may be positively or negatively skewed), even then, as per the central limit theorem, the sampling distribution of mean tends quite closer to the normal distribution, provided the number of sample items is large i.e., more than 30.
- In case if we want to reduce the sampling distribution of mean to unit normal distribution, i.e., N (0,1), we can write the normal variate $z = \frac{\bar{x}-\mu}{\sigma_p/\sqrt{n}}$ for the sampling distribution of mean. This characteristic of the sampling distribution of mean is very useful in several decision situations for accepting or rejection of hypotheses.

Sampling distribution of proportion:

- Like sampling distribution of mean, we can as well have a sampling distribution of proportion. This happens in case of statistics of attributes. Assume that we have worked out the proportion of defective parts in large number of samples, each with say 100 items, that have been taken from an infinite population and plot a probability distribution of the said proportions, we obtain what is known as the sampling distribution of the said proportions, we obtain what is known as the sampling distribution of proportion.
- Usually the statistics of attributes correspond to the conditions of a binomial distribution that tends to become normal distribution as n becomes larger and larger. If p represents the proportion of defectives i.e., of successes and q the proportion of non defectives i.e., of failures (or q = 1 – p) and if p is treated as a random variable, then the sampling distribution of proportion of successes has a mean = p with standard deviation = $\sqrt{\frac{p.q}{n}}$, where n

is the sample size. Presuming the binomial distribution approximating the normal distribution for large n, the normal variant of the sampling distribution of proportion $z = \frac{\hat{p}-p}{\sqrt{(p.q)/n}}$

, where $\hat{p}$ (pronounced as p-hat) is the sample proportion of successes, can be used for testing of hypotheses.

Student's t-distribution:

- When population standard deviation ($\sigma_{p)}$ is not known and the sample is of a small size (i.e., n $\lesssim$ 30), we use t distribution for the sampling distribution of mean and workout t variable as:

$t = (\bar{X} - \mu)/(\sigma_s/\sqrt{n})$ where $\sigma_s = \sqrt{\frac{\sum(X_i - X)^2}{n}} - 1$

i.e. the sample standard deviation. t-distribution is also symmetrical and is very close to the distribution of standard normal variate, z, except for small values of n. The variable t differs from z in the sense that we use sample standard deviation (σ_s) in the calculation of t, whereas we use standard deviation of population (σ_s) in the calculation of z.

- There is a different t distribution for every possible sample size i.e., for different degrees of freedom. The degrees of freedom for a sample of size n is n – 1. As the sample size gets larger, the shape of the t distribution becomes approximately equal to the normal distribution.
- In fact for sample sizes of more than 30, the t distribution is so close to the normal distribution that we can use the normal to approximate the t-distribution. But when n is small, the t-distribution is far from normal but when n - α, t-distribution is identical with normal distribution.
- The t-distribution tables are available which give the critical values of t for different degrees of freedom at various levels of significance. The table value of t for given degrees of freedom at a certain level of significance is compared with the calculated value of t from the sample data, and if the latter is either equal to or exceeds, we infer that the null hypothesis cannot be accepted.

F distribution:

- If $(\sigma_{s1})^2$ and $(\sigma_{s2})^2$ of size n_1 and n_2 respectively taken from two independent normal populations, having the same variance, $(\sigma_{v1})^2 = (\sigma_{v2})^2$, the ratio $F = (\sigma_{s1})^2/(\sigma_{s2})^2$, where $(\sigma_{s1})^2 = \sum (\bar{X}_{1i} - \bar{X}_1)^2/n_1 - 1$ and $(\sigma_{s2})^2 = \sum (\bar{X}_{2i} - \bar{X}_2)^2/n_2 - 1$ has an F distribution with n_1 - 1 and n_2 – 1 degrees of freedom. F ratio is computed in a way that the larger variance is always in the numerator.
- Tables have been prepared for F distribution that gives critical values of F for various values of degrees of freedom for larger as well as smaller variance. The calculated value of F from the sample data is compared with the corresponding table value of F and if the former is equal to or exceeds the latter, then we infer that the null hypothesis of the variances being equal can not be accepted. We shall make use of the F ratio in the context of hypothesis testing and also in the context of ANOVA technique.

Chi-square (χ^2) distribution:

- Chi-square distribution is encountered when we deal with collections of values that involve adding up squares. Variances of samples require us to add a collection of squared quantities and thus have distributions that are related to chi-square distribution.
- If we take each one of a collection of sample variances, divide them by the known population variance and multiply these quotients by (n – 1), where n means the number of items in the sample, we shall obtain a chi-square distribution. Thus, $(\sigma_s^2 / \sigma_s^2)$ (n-1) would have the same distribution as chi-square distribution with (n – 1) degrees of freedom.
- Chi-square distribution is not symmetrical and all the values are positive. One must know the degrees of freedom for using chi-square distribution. This distribution may also be used for judging the significance of difference between observed and expected frequencies and also as a test of goodness of fit. The generalised shape of χ^2 distribution depends upon the d.f. and the χ^2 value is worked out as under:

$$\chi^2 = \sum_{i=1}^{k} \frac{(O_i - E_i)^2}{E_i}$$

3.1.4 Essentials of Good Samples

It is important that the sampling results must reflect the characteristics of the population. Therefore, while selecting the sample from the population under investigation it should be ensured that the sample has the following characteristics:

- A sample must represent a true picture of the population from which it is drawn.
- A sample must be unbiased by the sampling procedure.
- A sample must be taken at random so that every member of the population of data has an equal chance of selection.
- A sample must be sufficiently large but as economical as possible.
- A sample must be accurate and complete. It should not leave any information incomplete and should include all the respondents, units or items included in the sample.
- Adequate sample size must be taken considering the degree of precision required in the results of inquiry.

3.1.5 Central Limit Theorem

- When sampling is from a normal population, the means of samples drawn from such a population are themselves normally distributed. But when sampling is not from a normal population, the size of the sample plays a critical role. When n is small, the shape of the distribution will depend largely on the shape of the parent population, but as n gets large (n> 30), the thape of the sampling distribution will become more and more like a normal distribution, irrespective of the shape of the parent population.
- The theorem which explains this sort of relationship between the shape of the population distribution and the sampling distribution of the mean is known as the central limit theorem. This theorem is by far the most important theorem in statistical inference. It assures that the sampling distribution of the mean approaches normal distribution as the sample size increases.
- In formal terms, we may say that the central limit theorem states that “the distribution of means of random samples taken from a population having mean m and finite variance σ^2 approaches the normal distribution with mean m and variance σ^2 / n as n goes to infinity.”
- The significance of the central limit theorem lies in the fact that it permits us to use sample statistics to make inferences about population parameters without knowing anything about the shape of the frequency distribution of that population other than what we can get from the sample.”

3.2 Methods of Sampling

- If money, time, trained manpower and other resources were not a concern, the researcher could get most accurate data from surveying the entire population of interest. Since most often the resources are scarce, the researcher is forced to go for sampling. But the real purpose of the survey is to know the characteristics of the population. Then the question is with what level of confidence the researcher will be able to say that the characteristics of a sample represent the entire population.
- Using a combination of tasks of hypotheses and unbiased sampling methods, the researcher can collect data that actually represents the characteristics of the entire population from which the sample was taken. To ensure a high level of confidence that the sample represents the population it is necessary that the sample is unbiased and sufficiently large.
- It was scientifically proved that if we increase the sample size we shall be that much closer to the characteristics of the population. Ultimately, if we cover each and every unit of the population, the characteristics of the sample will be equal to the characteristics of the population. That is why in a census there is no sampling error. Larger the sample size, the less sampling error we have.
- The statistical meaning of bias is error. The sample must be error free to make it an unbiased sample. In practice, it is impossible to achieve an error free sample even using unbiased sampling methods. However, we can minimise the error by employing appropriate sampling methods. The various sampling methods can be classified into two categories. These are random sampling methods and non-random sampling methods.

3.2.1 Random Sampling Methods

- The random sampling method is also called probability sampling. In random sampling all units or items in the population have a chance of being chosen in the sample. In other words a random sample is a sample in which each element of the population has a known and non-zero chance of being selected.
- Random sampling always produces the smallest possible sampling error. In the real sense, the size of the sampling error in a random sample is affected only by a random chance. Because a random sample contains the least amount of sampling error, we may say that it is an unbiased sample.
- It is not like that a random sample contains no error, but rather the minimum possible amount of error. The major advantage of random sampling is that it is possible to quantify the magnitude of the likely error in the inference made and this will help in building confidence in drawing inferences.

The following are the important methods of random sampling:

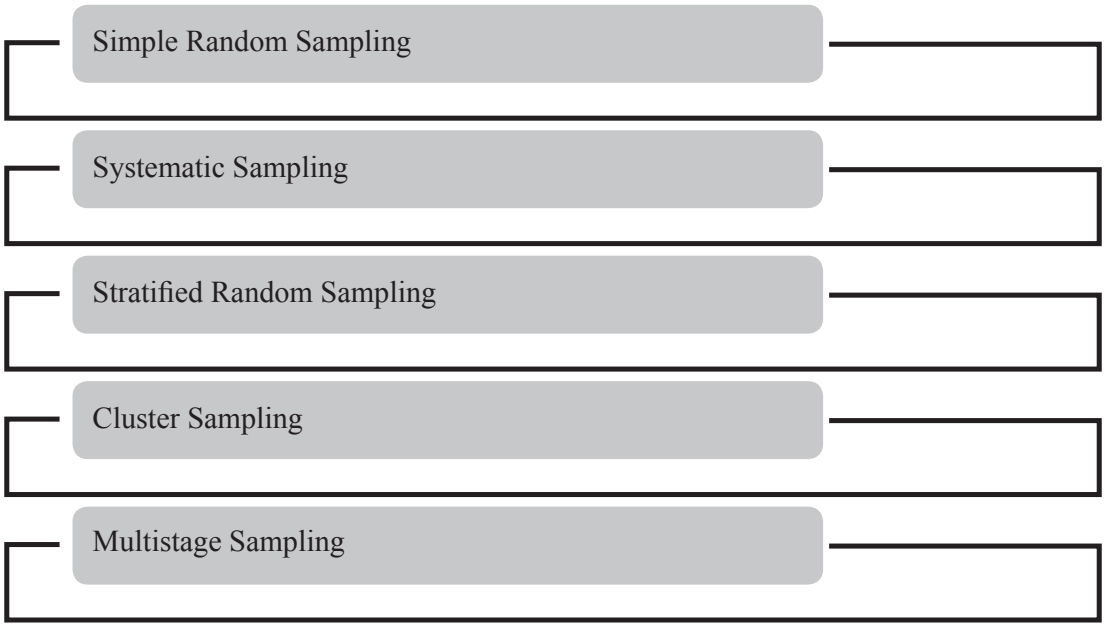


Fig. 3.2 Methods of random sampling

Simple random sampling:

- The most commonly used random sampling method is simple random sampling method. A simple random sample is one in which each item in the total population has an equal chance of being included in the sample. In addition, the selection of one item for inclusion in the sample should in no way influence the selection of another item.
- Simple random sampling should be used with a homogeneous population, that is, a population consisting of items that possess the same attributes that the researcher is interested in.
- The characteristics of homogeneity may include such as age, sex, income, social/religious/political affiliation, geographical region etc. The best way to choose a simple random sample is to use random number table. A random sampling method should meet the following criteria:
  - Every member of the population must have an equal chance of inclusion in the sample.
  - The selection of one member is not affected by the selection of previous members.
- The random numbers are a collection of digits generated through a probabilistic mechanism. The random numbers have the following properties:
  - The probability that each digit (0, 1, 2, 3, 4, 5, 6, 7, 8, or 9) will appear at any place is the same. That is 1/10.
  - The occurrence of any two digits in any two places is independent of each other.
- Each member of a population is assigned a unique number. The members of the population chosen for the sample will be those whose numbers are identical to the ones extracted from the random number table in succession until the desired sample size is reached. An example of a random number table is given below. To select a random sample using simple random sampling method we should follow the steps given below:
  - Determine the population size (N)
  - Determine the sample size (n)
  - Number each member of the population under investigation in serial order. Suppose there are 100 members number them from 00 to 99.
  - Determine the starting point of selecting sample by randomly picking up a page from random number tables and dropping your finger on the page blindly.
  - Choose the direction in which you want to read the numbers (from left to right, or right to left, or down or up).
  - Select the first 'n' numbers whose X digits are between 0 and N. If N = 100 then X would be 2, if N is a four digit number then X would be 3 and so on.
  - Once a number is chosen, do not use it again.
  - If you reach the end point of the table before obtaining 'n' numbers, pick another starting point and read in a different direction and then use the first X digit instead of the last X digits and continue until the desired sample is selected.

Advantages

- The simple random sample requires less knowledge about the characteristics of the population.
- Since sample is selected at random giving each member of the population equal chance of being selected the sample can be called as unbiased sample. Bias due to human preferences and influences is eliminated.
- Assessment of the accuracy of the results is possible by sample error estimation.
- It is a simple and practical sampling method provided population size is not large.

Limitations

- If the population size is large, a great deal of time must be spent listing and numbering the members of the population.

- A simple random sample will not adequately represent many population characteristics unless the sample is very large. That is, if the researcher is interested in choosing a sample on the basis of the distribution in the population of gender, age, social status, a simple random sample needs to be very large to ensure all these distributions are representative of the population.
- To obtain a representative sample across multiple population attributes we should use stratified random sampling.

Systematic sampling:

- In systematic sampling the sample units are selected from the population at equal intervals in terms of time, space or order. The selection of a sample using systematic sampling method is very simple. From a population of 'N' units, a sample of 'n' units may be selected by following the steps given below:
  - Arrange all the units in the population in an order by giving serial numbers from 1 to N.
  - Determine the sampling interval by dividing the population by the sample size. That is, K=N/n.
  - Select the first sample unit at random from the first sampling interval (1 to K).
  - Select the subsequent sample units at equal regular intervals.

Advantages

- The main advantage of using systematic sample is that it is more expeditious to collect a sample systematically since the time taken and work involved is less than in simple random sampling. For example, it is frequently used in exit polls and store consumers.
- This method can be used even when no formal list of the population units is available. For example, suppose if we are interested in knowing the opinion of consumers on improving the services offered by a store we may simply choose every kth (say 6th) consumer visiting a store provided that we know how many consumers are visiting the store daily (say 1000 consumers visit and we want to have 100 consumers as sample size).

Limitations

- If there is periodicity in the occurrence of elements of a population, the selection of sample using systematic sample could give a highly un-representative sample. For example, suppose the sales of a consumer store are arranged chronologically and using systematic sampling we select sample for 1st of every month. The 1st day of a month can not be a representative sample for the whole month. Thus in systematic sampling there is a danger of order bias.
- Every unit of the population does not have an equal chance of being selected and the selection of units for the sample depends on the initial unit selection. Regardless how we select the first unit of sample, subsequent units are automatically determined lacking complete randomness.

Stratified random sampling

- The stratified sampling method is used when the population is heterogeneous rather than homogeneous. A heterogeneous population is composed of unlike elements such as male/female, rural/urban, literate/illiterate, high income/low income groups, etc. In such cases, use of simple random sampling may not always provide a representative sample of the population.
- In stratified sampling, we divide the population into relatively homogenous groups called strata. Then we select a sample using simple random sampling from each stratum. There are two approaches to decide the sample size from each stratum, namely, proportional stratified sample and disproportional stratified sample. With either approach, the stratified sampling guarantees that every unit in the population has a chance of being selected. Following given are the two approaches of selecting samples.
- Proportional Stratified Sample: If the number of sampling units drawn from each stratum is in proportion to the corresponding stratum population size, it is said that the sample is proportional stratified sample. For example, let us say we want to draw a stratified random sample from a heterogeneous population (on some characteristics) consisting of rural/urban and male/female respondents. So we have to create 4 homogeneous sub groups called stratums as follows:

Urban		Rural	
Male	Female	Male	Female

To ensure each stratum in the sample will represent the corresponding stratum in the population we must ensure each stratum in the sample is represented in the same proportion to the stratums as they are in the population.

- Disproportional Stratified Sample: In a disproportional stratified sample, sample size for each stratum is not allocated on a proportional basis with the population size, but by analytical considerations of the researcher such as stratum variance, stratum population, time and financial constraints etc.

Advantages

- Since the sample are drawn from each of the stratums of the population, stratified sampling is more representative and thus more accurately reflects characteristics of the population from which they are chosen.
- It is more precise and to a great extent avoids bias.
- Since sample size can be less in this method, it saves a lot of time, money and other resources for data collection.

Limitations

- Stratified sampling requires a detailed knowledge of the distribution of attributes or characteristics of interest in the population to determine the homogeneous groups that lie within it. If we cannot accurately identify the homogeneous groups, it is better to use simple random sample since improper stratification can lead to serious errors.
- Preparing a stratified list is a difficult task as the lists may not be readily available.

Cluster samplings

- In cluster sampling, we divide the population into groups having heterogeneous characteristics called clusters and then select a sample of clusters using simple random sampling. It is assumed that each of the clusters is representative of the population as a whole. This sampling is widely used for geographical studies of many issues.
- The principles that are basic to the cluster sampling are as follows:
  - The differences or variability within a cluster should be as large as possible. As far as possible the variability within each cluster should be the same as that of the population.
  - The variability between clusters should be as small as possible. Once the clusters are selected, all the units in the selected clusters are covered for obtaining data.

Advantages

- The cluster sampling provides significant gains in data collection costs, since travelling costs are smaller.
- Since the researcher need not cover all the clusters and only a sample of clusters is covered, it becomes a more practical method which facilitates fieldwork.

Limitations

- The cluster sampling method is less precise than sampling of units from the whole population since the latter is expected to provide a better cross-section of the population than the former, due to the usual tendency of units in a cluster to be homogeneous.
- The sampling efficiency of cluster sampling is likely to decrease with the decrease in cluster size or increase in number of clusters.

The above advantages or limitations of cluster sampling suggest that, in practical situations where sampling efficiency is less important but the cost is of greater significance, the cluster sampling method is extensively used. If the division of clusters is based on the geographic sub-divisions, it is known as area sampling.

In cluster sampling instead of covering all the units in each cluster we can resort to sub-sampling as two-stage sampling. Here, the clusters are termed as primary units and the units within the selected clusters are taken as secondary units.

Multistage sampling

- Multistage sampling is a generalisation of two stage sampling. As the name suggests, multi stage sampling is carried out in different stages. In each stage, progressively smaller (population) geographic areas will be randomly selected.
- In this sampling method, it is possible to take as many stages as are necessary to achieve a representative sample. Each stage results in a reduction of sample size. In a multi stage sampling at each stage of sampling a suitable method of sampling is used. More number of stages is used to arrive at a sample of desired sampling units.

Advantages

- Multistage sampling provides cost gains by reducing the data collection on costs.
- Multistage sampling is more flexible and allows us to use different sampling procedures in different stages of sampling.
- If the population is spread over a very wide geographical area, multistage sampling is the only sampling method available in a number of practical situations.

Limitations

- If the sampling units selected at different stages are not representative multistage sampling becomes less precise and efficient.

3.2.2 Non-Random Sampling Methods

The non-random sampling methods are also often called non-probability sampling methods. In a non-random sampling method, the probability of any particular unit of the population being chosen is unknown. Here the method of selection of sampling units is quite arbitrary as the researchers rely heavily on personal judgment.

Non-random sampling methods usually do not produce samples that are representative of the general population from which they are drawn. The greatest error occurs when the researcher attempts to generalise the results on the basis of a sample to the entire population.

Such an error is insidious because it is not at all obvious from merely looking at the data, or even from looking at the sample. The easiest way to recognise whether a sample is representative or not is to determine whether the sample is selected randomly or not. Nevertheless, there are occasions where non-random samples are best suited for the researcher's purpose. The various non-random sampling methods commonly used are:

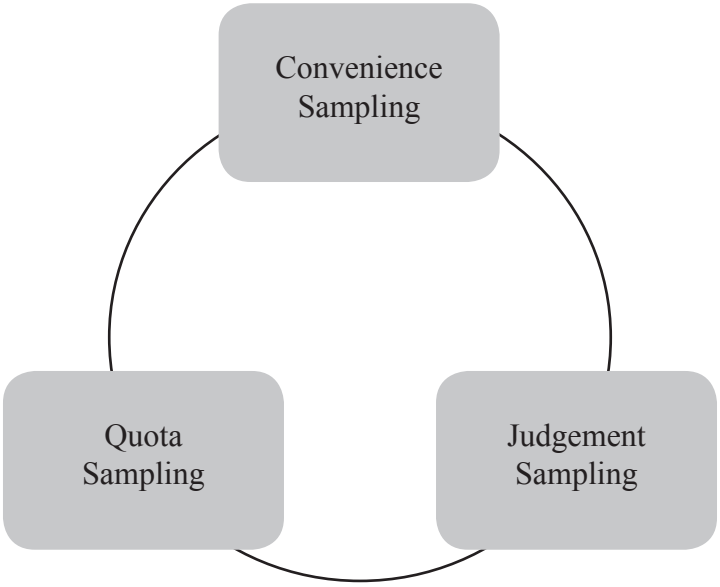


Fig. 3.3 Methods of non-random sampling

Convenience sampling:

- Convenience sampling refers to the method of obtaining a sample that is most conveniently available to the researcher. It may be cautioned that the generalisation of results based on convenience sampling beyond that particular sample may not be appropriate.
- Convenience samples are best used for exploratory research when additional research will be subsequently conducted with a random sample. Convenience sampling is also useful in testing the questionnaires designed on a pilot basis. Convenience sampling is extensively used in marketing studies.

Judgement sampling:

- Judgement sampling method is also known as purposive sampling. In this method of sampling the selection of sample is based on the researcher's judgment about some appropriate characteristic required of the sample units.
- In business judgment sampling is often used to measure the performance of salesmen/saleswomen. The salesmen/saleswomen are grouped into high, medium or low performers based on certain specified qualities. Then the sales manager may actually classify the salesmen/saleswomen working under him/her who in his/her opinion will fall in which group. Judgment sampling is also often used in forecasting election results.
- However, in the absence of any representative data, one may resort to this kind of non-random sampling.

Quota sampling:

- The quota sampling method is commonly used in marketing research studies. The samples are selected on the basis of some parameters such as age, gender, geographical region, education, income, occupation etc, in order to make them as representative samples. The investigators, then, are assigned fixed quotas of the sample meeting these population characteristics.
- The purpose of quota sampling is to ensure that various sub-groups of the population are represented on pertinent sample characteristics to the extent that the investigator desires. The stratified random sampling also has this objective but should not be confused with quota sampling.
- In the stratified sampling method, the researcher selects a random sample from each group of the population, where as, in quota sampling, the interviewer has a quota fixed for him/her to achieve.
- Quota sampling has the advantage that the sample confirms the selected characteristics of the population that the researcher desires. Also, the cost and time involved in collecting the data are also greatly reduced.

Limitations

- In quota sampling, the respondents are selected according to the convenience of the field investigator rather than on a random basis. This kind of selection of sample may be biased. Suppose in our example of soft drinks, after the sample is taken it was found that most of the respondents belong to the lower income group then the purpose of conducting the survey becomes useless and the results may not reflect the actual situation.
- If the numbers of parameters on which basis the quotas are fixed, are larger, then it becomes difficult for the researcher to fix the quota for each sub-group.
- The field workers have the tendency to cover the quota by going to those places where the respondents may be willing to provide information and avoid those with unwilling respondents. For example, the investigators may avoid places where high income group respondents stay and cover only low income group areas.

3.3 Sampling Theory

Sampling theory is a study of relationships existing between a population and samples drawn from the population. Sampling theory is applicable only to random samples. For this purpose, the population or a universe may be defined as an aggregate of items possessing a common trait or traits. In other words, a universe is the complete group of items about which knowledge is sought. The universe may be finite or infinite. Finite universe is one which has a definite and certain number of items, but when the number of items is uncertain and infinite, the universe is said to be an infinite universe. Similarly, the universe may be hypothetical or existent. In the former case the universe in fact does not exist and we can only imagine the items constituting it. Tossing of a coin or throwing a dice is examples of hypothetical universe. Existent universe is a universe of concrete objects i.e., the universe where the items constituting it really exist.

On the other hand, the term 'sample' refers to that part of the universe which is selected for the purpose of investigation. The theory of sampling studies the relationships that exist between the universe and the sample or samples drawn from it. The main problem of sampling theory is the problem of relationship between a parameter and a statistic. The theory of sampling is concerned with estimating the properties of the population from those of the sample and also with gauging the precision of the estimate. This sort of movement from particular (sample) towards general (universe) is what is known as statistical induction or statistical inference.

In order to be able to follow this inductive method, we first follow a deductive argument which is that we imagine a population or universe (finite or infinite) and investigate the behaviour of the samples drawn from this universe applying the laws of probability." The methodology dealing with all this is known as sampling theory. Sampling theory is designed to attain one or more of the following objectives:

- Statistical estimation: Sampling theory helps in estimating unknown population parameters from a knowledge of statistical measures based on sample studies. In other words, to obtain an estimate of parameter from statistic is the main objective of the sampling theory. The estimate can either be a point estimate or it may be an interval estimate. Point estimate is a single estimate expressed in the form of a single figure, but interval estimate has two limits, viz., the upper limit and the lower limit within which the parameter value may lie. Interval estimates are often used in statistical induction.
- Testing 01 hypothesis: The second objective of sampling theory is to enable us to decide whether to accept or reject hypothesis; the sampling theory helps in determining whether observed differences are actually due to chance or whether they are really significant.
- Statistical inference: Sampling theory helps in making generalisation about the population of universe from the studies based on samples drawn from it. It also helps in determining the accuracy of such generalisations.

3.3.1 Sandler's A Test

Joseph Sandler has developed an alternate approach based on a simplification of t-test. His approach is described as Sandler's A -test that serves the same purpose as is accomplished by t-test relating to paired data. Psychologists generally use this test in case of two groups that are matched with respect to some extraneous variable(s)

- While using A-test, we work out A-statistic that yields exactly the same results as Student's t-test*. A-statistic is found as follows: A= the sum of squares of the difference the squares of sum of the difference The number of degrees of freedom (dJ.) in A-test is the same as with Student's t-test i.e., d.f. = n - I, n being equal to the number of pairs.
- The critical value of A" at a given level of significance for given d.f., can be obtained from the table of A- statistic (given in appendix at the end of the book). One has to compare the computed value of A with its corresponding table value for drawing inference concerning acceptance or rejection of null hypothesis. If the calculated value of A is equal to or less than the table value, in that case A-statistic is considered significant where upon we reject Ho and accept H. But if the calculated value of A is more than its table value, then a statistic is taken as insignificant and accordingly we accept Ho. This is so because the two test statistics, viz., t and A are inversely related.
- We can write these two statistics in terms of one another in this way:
  - 'in terms of 'r can be expressed as n-l 1 A= - - + nxt 2 - n
  - 't' in terms of -:4' can be expressed as t- - ftE 1 Axn-l Computational work concerning A-statistic is relatively simple. As such the use of A-statistic result in considerable saving of time and labour, especially when matched groups are to be compared with respect to a large number of variables.
- Accordingly researchers may replace Student's t-test by Sandler's A-test whenever correlated sets of scores are employed. Sandler's A-statistic can as well be used "in the one sample case as a direct substitute for the Student t-ratio. This is so because Sandler's A is an algebraically equivalent to the Student's t. When we use A-test in one sample case, the following steps are involved:
  - Subtract the hypothesised mean of the population (μ_H) from each individual score (X_1) to obtain D_1 and then work out LD;
  - Square each D_i and then obtain the sum of such squares i.e., $\sum D_1^{\,2}$
  - Find A-statistic as under:
  - $A = \sum D_i^{\,2} / (\sum D_i)^2$
  - Read the table of A-statistic for (n -1) degrees of freedom at a given level of significance (using one tailed or two-tailed values depending upon H_a) to find the critical value of A.
  - Finally, draw the inference as under: When calculated value of A is equal to or less than the table value, then reject H_o(or accept H_a) but when computed A is greater than its table value, then accept H_o

3.3.2 Concept of Standard Error

- The standard deviation of sampling distribution of a statistic is known as its standard error (S.E) and is considered the key to sampling theory. The utility of the concept of standard error in statistical induction arises on account of the following reasons:
  - The standard error helps in testing whether the difference between observed and expected frequencies could arise due to chance.
  - The criterion usually adopted is that if a difference is less than times the S.E., the difference is supposed to exist as a matter of chance and if the difference is equal to or more than times the S.E., chance fails to account for it, and we conclude the difference as significant difference. This criterion is based on the fact that at $\bar{X} \pm 3$ (S.E.) the normal curve covers an area of 99.73 per cent. Sometimes the criterion of 2 S.E. is also used in place of 3 S.E.
  - Thus the standard error is an important measure in significance tests or in examining hypotheses. If the estimated parameter differs from the calculated statistic by more than 1.96 times the S.E., the difference is taken as significant at 5 per cent level of significance.
  - This, in other words, means that the difference is outside the limits i.e., it lies in the 5 per cent area (2.5 per cent on both sides) outside the 95 per cent area of the sampling distribution. Hence we can say with 95 per cent confidence that the said difference is not due to fluctuations of sampling. In such a situation our hypothesis that there is no difference is rejected at 5 per cent level of significance.

- But if the difference is less than 1.96 times the S.E., then it is considered not significant at 5 per cent level and we can say with 95 per cent confidence that it is because of the fluctuations of sampling. In such a situation our null hypothesis stands true. 1.96 is the critical value at 5 per cent level.
- The product of the critical value at a certain level of significance and the S.E. is often described as 'Sampling Error' at that particular level of significance. We can test the difference at certain other levels of significance as well depending upon our requirement.

- The standard error enables us to specify the limits within which the parameters of the population are expected to lie with a specified degree of confidence. Such an interval is usually known as confidence interval. Important formulae for computing the standard errors concerning various measures based on samples are as under:

In case of sampling of attributes:

- Standard error of number of success = $\sqrt{n.p.q}$

  Where n = number of events in each sample,

  p = probability of success in each event,

  q = probability of failure in each event
- Standard error of proportion of successes $\sqrt{\frac{(p.q)}{n}}$
- Standard error of the difference between proportions of two samples:

  $\sigma_{p_1-p_2} = \sqrt{p.q\ [\frac{1}{n_1}+\frac{1}{n_2}]}$

  Where p = best estimate of proportion in the population and is worked out as under:

  $P = \frac{n_1 p_1 + n_2 p_2}{n_1 + n_2}$

  Q = 1- p

  n_1 = number of events in sample one

  n_2 = number of events in sample two

  Note: Instead of the above formula, we use the following formula:

  $\sigma_{p_1-p_2} = \sqrt{\frac{p_1 q_1}{n_1} + \frac{p_2 q_2}{n_2}}$

  When samples are drawn from two heterogeneous populations where we cannot have the best estimate of proportion in the universe on the basis of given sample data. Such a situation often arises in study of occasion of attributes.

In case of sampling of variables (large samples):

- Standard error of mean when population standard deviation is known:

  $\sigma_{\bar{x}} = \frac{\sigma_p}{\sqrt{n}}$

  Where

  σ_p = standard deviation of population

  n = number of items in the sample

  Note: this formula is used when n is 30 or less.
- Standard error of mean when population standard deviation is unknown:

  $\sigma_{\bar{x}} = \frac{\sigma_s}{\sqrt{n}}$

  Where

  σ_s = standard deviation of the sample and is worked out as under

$$\sigma_s = \sqrt{\frac{\sum(X_i - \bar{X})^2}{n-1}}$$

n = number of items in the sample.

- Standard error of standard deviation when population standard deviation is known:

$$\sigma_{\sigma_s} = \frac{\sigma_p}{\sqrt{2n}}$$

- Standard error of standard deviation when population standard deviation is unknown:

$$\sigma_{\sigma_s} = \frac{\sigma_s}{\sqrt{2n}}$$

Where

$$\sigma_s = \sqrt{\frac{\sum(X_i - \bar{X})^2}{n-1}}$$

n = number of items in the sample.

- Standard error of the coefficient of simple correlation:

$$\sigma_r = \frac{1 - r^2}{\sqrt{n}}$$

r = coefficient of simple correlation
n = number of items in the sample

- Standard error of difference between means of two samples:
  - When two samples are drawn from the same population:

$$\sigma_{\bar{x}_i - \bar{x}_2} = \sqrt{\sigma_p{}^2 \ \frac{1}{n_1} + \frac{1}{n_2}}$$

(If σ_p is not known, sample standard deviation for combined samples ($\sigma_{s_{1.2}}$)* may be substituted).

  - When two samples are drawn from different populations:

$$\sigma_{\bar{x}_1 - \bar{x}_2} = \sqrt{\frac{(\sigma_{p1})^2}{n_1} + \frac{(\sigma_{\rho 2})^2}{n_2}}$$

(If $\sigma_{\rho 1}$ and $\sigma_{\rho 2}$ are not known, then in their places σ_{s1} and σ_{s2}respectively may be substituted).

  - In case of sampling of variables (small samples):

Standard error of mean when σ_ρ is unknown:

$$\sigma_{s_{1.2}} = \sqrt{\frac{n_1(\sigma_{s1})^2 + n_2(\sigma_{s2})^2 + n_1(\bar{X}_1 - \bar{X}_{1.2})^2 + n_2(\bar{X}_2 - \bar{X}_{1.2})^2}{n_1 + n_2}}$$

Where $\bar{X}_{1.2} = \frac{n_1(\bar{X}_1) + n_2(\bar{X}_2)}{n_1 + n_2} \bar{X}_{1.2} = \frac{n_1(\bar{X}_1) + n_2(\bar{X}_2)}{n_1 + n_2}$

Note: (1) All these formulae apply in case of infinite population. But in case of finite population where sampling is done without replacement and the sample is more than 5% of the population, we must as well use the finite population multiplier in our standard error formulae. For instance, $SE_{\bar{X}}$ in case of finite population will be as under:

$$SE_{\bar{X}} = \frac{\sigma_p}{\sqrt{n}} \cdot \sqrt{\frac{(N-n)}{(N-1)}}$$

It may be remembered that in cases in which the population is very large in relation to the size of the sample, the finite population multiplier is close to one and has little effect on the calculation of S.E. as such when sampling fraction is less than 0.5; the finite population multiplier is generally not used.

3.4 Sample Size and its Discrimination

In sampling analysis the most difficult question is: What should be the size of the sample or how large or small should be n? If the sample size (nj is too small, it may not serve to achieve the objectives and if it is too large, we may incur huge cost and waste resources.

As a general rule, one can say that the sample must be of an optimum size i.e., it should neither be excessively large nor too small. Technically, the sample size should be large enough to give a confidence interval of desired width and as such the size of the sample must be chosen by some logical process before sample is taken from the universe. Size of the sample should be determined by a researcher keeping in view the following points:

3.4.1 Nature of Universe

- Universe may be either homogenous or heterogeneous in nature. If the items of the universe are homogenous, a small sample can serve the purpose. But if the items are heterogeneous, a large sample would be required. Technically, this can be termed as the dispersion factor.
- Number of classes proposed: If many class-groups (groups and sub-groups) are to be formed, a large sample would be required because a small sample might not be able to give a reasonable number of items in each class-group.
- Nature of study: If items are to be intensively and continuously studied, the sample should be small. For a general survey the size of the sample should be large, but a small sample is considered appropriate in technical surveys.
- Type of sampling: Sampling technique plays an important part in determining the size of the sample. A small random sample is apt to be much superior to a larger but badly selected sample.
- Standard of accuracy and acceptable confidence level: If the standard of accuracy or the level of precision is to be kept high, we shall require relatively larger sample. For doubling the accuracy for a fixed significance level, the sample size has to be increased fourfold. .
- Availability of finance: In practice, size of the sample depends upon the amount of money available for the study purposes. This factor should be kept in view while determining the size of sample for large samples result in increasing the cost of sampling estimates.
- Other considerations: Nature of units, size of the population, size of questionnaire, availability of trained investigators, the conditions under which the sample is being conducted, the time available for completion of the study are a few other considerations to which a researcher must pay attention while selecting the size of the sample. There are two alternative approaches for determining the size of the sample.
  - The first approach is to specify the precision of estimation desired and then to determine the sample size necessary to insure it. It is capable of giving a mathematical solution, and as such is a frequently used technique of determining ‘n’. The limitation of this technique is that it does not analyse the cost of gathering information vis-a-vis the expected value of information.

- The second approach uses Bayesian statistics to weigh the cost of additional information against the expected value of the additional information. It is theoretically optimal, but it is seldom used because of the difficulty involved in measuring the value of information. Hence, we shall mainly concentrate here on the first approach.

3.5 Sampling Design

Following is the description of various types of sampling design.

3.5.1 Census and Sample Survey

- All items in any field of inquiry constitute a 'Universe' or 'Population.' A complete enumeration of all items in the 'population' is known as a census inquiry. It can be presumed that in such an inquiry, when all items are covered, no element of chance is left and highest accuracy is obtained. But in practice, this may not be true.
- Even the slightest element of bias in such an inquiry will get larger and larger as the number of observation increases. Moreover, there is no way of checking the element of bias or its extent except through a resurvey or use of sample checks. Besides, this type of inquiry involves a great deal of time, money and energy. Therefore, when the field of inquiry is large, this method becomes difficult to adopt because of the resources involved.
- At times, this method is practically beyond the reach of ordinary researchers. Perhaps, government is the only institution which can get the complete enumeration carried out. Even the government adopts this in very rare cases such as population census conducted once in a decade. Further, many a time it is not possible to examine every item in the population, and sometimes it is possible to obtain sufficiently accurate results by studying only a part of total population. In such cases there is no utility of census surveys.
- However, it needs to be emphasised that when the universe is a small one, it is no use resorting to a sample survey. When field studies are undertaken in practical life, considerations of time and cost almost invariably lead to a selection of respondents, i.e., selection of only a few items. The respondents selected should be as representative of the total population as possible in order to produce a miniature cross-section.
- The selected respondents constitute what is technically called a 'sample' and the selection process is called 'sampling technique.' The survey so conducted is known as 'sample survey'. Algebraically, let the population size be N and if a part of size n (which is < N) of this population is selected according to some rule for studying some characteristic of the population, the group consisting of these n units is known as 'sample'. Researcher must prepare a sample design for his study i.e., he must plan how a sample should be selected and of what size such a sample would be.

3.5.2 Implications of a Sample Design

- A sample design is a definite plan for obtaining a sample from a given population. It refers to the technique or the procedure the researcher would adopt in selecting items for the sample. Sample design may as well lay down the number of items to be included in the sample, i.e., the size of the sample.
- Sample design is determined before data are collected. There are many sample designs from which a researcher can choose. Some designs are relatively more precise and easier to apply than others. Researcher must select/ prepare a sample design which should be reliable and appropriate for his research study.

3.5.3 Steps in Sample Design

While developing a sampling design, the researcher must pay attention to the following points:

- Type of universe:
  - The first step in developing any sample design is to clearly define the set of objects, technically called the Universe, to be studied. The universe can be finite or infinite. In finite universe the number of items is certain, but in case of an infinite universe the number of items is infinite, i.e., we cannot have any idea about the total number of items.

- The population of a city, the number of workers in a factory and the like are examples of finite universes, whereas the number of stars in the sky, listeners of a specific radio programme, throwing of a dice etc. are examples of infinite universes.

- Sampling unit: A decision has to be taken concerning a sampling unit before selecting sample. Sampling unit may be a geographical one such as state, district, village, etc., or a construction unit such as house, flat, etc., or it may be a social unit such as family, club, school, etc., or it may be an individual. The researcher will have to decide one or more of such units that he has to select for his study.
- Source list: It is also known as 'sampling frame' from which sample is to be drawn. It contains the names of all items of a universe (in case of finite universe only). If source list is not available, researcher has to prepare it. Such a list should be comprehensive, correct, reliable and appropriate. It is extremely important for the source list to be as representative of the population as possible.
- Size of sample:
  - This refers to the number of items to be selected from the universe to constitute a sample. This is a major problem before a researcher. The size of sample should neither be excessively large, nor too small. It should be optimum. An optimum sample is one which fulfils the requirements of efficiency, representativeness, reliability and flexibility.
  - While deciding the size of sample, researcher must determine the desired precision as also an acceptable confidence level for the estimate. The size of population variance needs to be considered as in case of larger variance usually a bigger sample is needed.
  - The size of population must be kept in view for this also limits the sample size. The parameters of interest in a research study must be kept in view, while deciding the size of the sample.
  - Costs too dictate the size of sample that we can draw. As such, budgetary constraint must invariably be taken into consideration when we decide the sample size.
- Parameters of interest: In determining the sample design, one must consider the question of the specific population parameters which are of interest. For instance, we may be interested in estimating the proportion of persons with some characteristic in the population, or we may be interested in knowing some average or the other measure concerning the population. There may also be important sub-groups in the population about whom we would like to make estimates. All this has a strong impact upon the sample design we would accept.
- Budgetary constraint: Cost considerations, from practical point of view, have a major impact upon decisions relating to not only the size of the sample but also to the type of sample. This fact can even lead to the use of a non-probability sample.
- Sampling procedure: Finally, the researcher must decide the type of sample he will use i.e., he must decide about the technique to be used in selecting the items for the sample. In fact, this technique or procedure stands for the sample design itself. There are several sample designs (explained in the pages that follow) out of which the researcher must choose one for his study. Obviously, he must select that design which, for a given sample size and for a given cost, has a smaller sampling error.

3.5.4 Criteria of Selecting a Sampling Procedure

- In this context, one must remember that two costs are involved in a sampling analysis, viz., the cost of collecting the data and the cost of an incorrect inference resulting from the data.
- Researcher must keep in view the two causes of incorrect inferences, viz., systematic bias and sampling error. Systematic bias results from errors in the sampling procedures, and it cannot be reduced or eliminated by increasing the sample size. At best the causes responsible for these errors can be detected and corrected. Usually, a systematic bias is the result of one or more of the following factors:
  - Inappropriate sampling frame: If the sampling frame is inappropriate, i.e., a biased representation of the universe, it will result in a systematic bias.
  - Defective measuring device: If the measuring device is constantly in error, it will result in systematic bias. In survey work, systematic bias can result if the questionnaire or the interviewer is biased. Similarly, if the physical measuring device is defective there will be systematic bias in the data collected through such a measuring device.

- Non-respondents: If we are unable to sample all the individuals initially included in the sample, there may arise a systematic bias. The reason is that in such a situation the likelihood of establishing contact or receiving a response from an individual is often correlated with the measure of what is to be estimated.
- Indeterminacy principle: Sometimes we find that individuals act differently when kept under observation than what they do when kept in non-observed situations. For instance, if workers are aware that somebody is observing them in course of a work study on the basis of which the average length of time to complete a task will be determined and accordingly the quota will be set for piece work, they generally tend to work slowly in comparison to the speed with which they work if kept unobserved. Thus, the indeterminacy principle may also be a cause of a systematic bias.
- Natural bias in the reporting of data:
  - Natural bias of respondents in the reporting of data is often the cause of a systematic bias in many inquiries. There is usually a downward bias in the income data collected by government taxation department, whereas we find an upward bias in the income data collected by some social organisation.
  - People in general understate their incomes if asked about it for tax purposes, but they overstate the same if asked for social status or their affluence. Generally in psychological surveys, people tend to give what they think is the 'correct' answer rather than revealing their true feelings. Sampling errors are the random variations in the sample estimates around the true population parameters.
  - Since they occur randomly and are equally likely to be in either direction, their nature happens to be of compensatory type and the expected value of such errors happens to be equal to zero. Sampling error decreases with the increase in the size of the sample, and it happens to be of a smaller magnitude in case of homogeneous population.
  - Sampling error can be measured for a given sample design and size. The measurement of sampling error is usually called the 'precision of the sampling plan'. If we increase the sample size, the precision can be improved. But increasing the size of the sample has its own limitations, viz., a large sized sample increases the cost of collecting data and also enhances the systematic bias.
  - Thus, the effective way to increase precision is usually to select a better sampling design which has a smaller sampling error for a given sample size at a given cost. In practice, however, people prefer a less precise design because it is easier to adopt the same and also because of the fact that systematic bias can be controlled in a better way in such a design.

- In brief, while selecting a sampling procedure, researcher must ensure that the procedure causes a relatively small sampling error and helps to control the systematic bias in a better way.

3.5.5 Characteristics of a Good Sample Design

The characteristics of a good sample design as under:

- Sample design must result in a truly representative sample.
- Sample design must be such which results in a small sampling error.
- Sample design must be viable in the context of funds available for the research study.
- Sample design must be such so that systematic bias can be controlled in a better way.
- Sample should be such that the results of the sample study can be applied, in general, for the universe with a reasonable level of confidence

3.6 Different Types of Sample Designs

There are different types of sample designs based on two factors, viz., the representation basis and the element selection technique. On the representation basis, the sample may be probability sampling or it may be non-probability sampling. Probability sampling is based on the concept of random selection, whereas non-probability sampling is 'non-random' sampling.

- On element selection basis, the sample may be either unrestricted or restricted. When each sample element is drawn individually from the population at large, then the sample so drawn is known as 'unrestricted sample', whereas all other forms of sampling are covered under the term 'restricted sampling'. The following chart exhibits the sample designs as explained above.
- Thus, sample designs are basically of two types, viz., non-probability sampling and probability sampling. These two designs are taken up separately.

Element selection technique ↓	**Representation basis** ↓	
	Probability sampling	Non probability sampling
Unrestricted sampling	Simple random sampling	Haphazard sampling or convenience sampling
Restricted sampling	Complex random sampling (such as cluster sampling, systematic sampling, stratified sampling etc.)	Purposive sampling (such as quota sampling, judgement sampling)

Table 3.1 Types of sample designs

3.6.1 Non-Probability Sampling

- Non-probability sampling is that sampling procedure which does not afford any basis for estimating the probability that each item in the population has of being included in the sample. Non-probability sampling is also known by different names such as deliberate sampling, purposive sampling and judgement sampling.
- In this type of sampling, items for the sample are selected deliberately by the researcher; his choice concerning the items remains supreme. In other words, under non-probability sampling the organisers of the inquiry purposively choose the particular units of the universe for constituting a sample on the basis that the small mass that they so select out of a huge one will be typical or representative of the whole.
- For instance, if economic conditions of people living in a state are to be studied, a few towns and villages may be purposively selected for intensive study on the principle that they can be representative of the entire state. Thus, the judgement of the organisers of the study plays an important part in this sampling design. In such a design, personal element has a great chance of entering into the selection of the sample.
- The investigator may select a sample which shall yield results favourable to his point of view and if that happens, the entire inquiry may get vitiated. Thus, there is always the danger of bias entering into this type of sampling technique. But in the investigators are impartial, work without bias and have the necessary experience so as to take sound judgement, the results obtained from an analysis of deliberately selected sample may be tolerably reliable.
- However, in such a sampling, there is no assurance that every element has some specifiable chance of being included. Sampling error in this type of sampling cannot be estimated and the element of bias, great or small, is always there. As such this sampling design in rarely adopted in large inquires of importance.
- However, in small inquiries and researches by individuals, this design may be adopted because of the relative advantage of time and money inherent in this method of sampling. Quota sampling is also an example of non-probability sampling. Under quota sampling the interviewers are simply given quotas to be filled from the different strata, with some restrictions on how they are to be filled.
- In other words, the actual selection of the items for the sample is left to the interviewer's discretion. This type of sampling is very convenient and is relatively inexpensive. But the samples so selected certainly do not possess the characteristic of random samples. Quota samples are essentially judgement samples and inferences drawn on their basis are not amenable to statistical treatment in a formal way.

3.6.2 Probability Sampling

- Probability sampling is also known as 'random sampling' or 'chance sampling'. Under this sampling design, every item of the universe has an equal chance of inclusion in the sample. It is, so to say, a lottery method in which individual units are picked up from the whole group not deliberately but by some mechanical process.
- Here it is blind chance alone that determines whether one item or the other is selected. The results obtained from probability or random sampling can be assured in terms of probability i.e., we can measure the errors of estimation or the significance of results obtained from a random sample, and this fact brings out the superiority of random sampling design over the deliberate sampling design.
- Random sampling ensures the law of Statistical Regularity which states that if on an average the sample chosen is a random one, the sample will have the same composition and characteristics as the universe. This is the reason why random sampling is considered as the best technique of selecting a representative sample.
- Random sampling from a finite population refers to that method of sample selection which gives each possible sample combination an equal probability of being picked up and each item in the entire population to have an equal chance of being included in the sample. This applies to sampling without replacement i.e., once an item is selected for the sample, it cannot appear in the sample again (Sampling with replacement is used less frequently in which procedure the element selected for the sample is returned to the population before the next element is selected.
- In such a situation the same element could appear twice in the same sample before the second element is chosen). In brief, the implications of random sampling (or simple random sampling) are:
  - It gives each element in the population an equal probability of getting into the sample; and all choices are independent of one another.
  - It gives each possible sample combination an equal probability of being chosen.
- Keeping this in view we can define a simple random sample (or simply a random sample) from a finite population as a sample which is chosen in such a way that each of the N_{C_n} possible samples has the same probability, $1/N_{C_n}$, of being selected. To make it more clear we take a certain finite population consisting of six elements (say a, b, c, d, e, f) i.e., N = 6.
- Suppose that we want to take a sample of size n = 3 from it. Then there are 6_{C_3} = 20 possible distinct samples of the required size, and they consist of the elements abc, abd, abe, abf, acd, ace, acf, ade, adf, aef, bcd, bce, bcf, bde, bdf, bef, cde, cdf, cef, and def. If we choose one of these samples in such a way that each has the probability 1/20 of being chosen, we will then call this a random sample.

3.6.3 Determination of Appropriate Sampling Design

Determining an appropriate sampling design is a challenging issue and has greater implications on the application of the research findings. Apart from considering the theoretical components, sampling issues, advantages and drawbacks of different sampling techniques, the decision should take into consideration the following factors:

Research objectives:
A clear understanding of the statement of the problem and the objectives will provide the initial guidelines for determining the appropriate sampling design. If the research objectives include the need to generalise the findings of the research study, then a probability sampling method should be opted of research, viz., exploratory or descriptive will also influence the type of the sampling design.

Scope of the research:
The scope of the research project is local, regional, national or international has an implication on the choice of the sampling method. The geographical proximity of the defined target population elements will influence not only the researcher's ability to compile needed list of sampling units, but also the selection design. When the target population is equally distributed geographically a cluster sampling method may become more attractive than other available methods. If the geographical area to be covered is more extensive then complex sampling method should be adopted to ensure proper representation of the target population.

Availability of resources:
The researchers command over the financial and human resources should be considered in deciding the sampling method. If the financial and human resource availability is limited, some of the more time consuming, complex probability sampling methods cannot be selected for the study.

Time frame:
The researcher who has meet a short deadline will be more likely to select a simple, less time consuming sampling method rather than a more complex and accurate method.

Advanced knowledge of the target population:
If the complete lists of the entire population elements are not available to the researcher, the possibilities of the probability sampling method are ruled out. It may dictate that a preliminary study be conducted to generate information to build a sampling frame for the study. The researcher must gain a strong understanding of the key descriptor factors that make up the true members of any target population.

Degree of accuracy:
The degree of accuracy required or the level of tolerance for error may vary from one study to another. If the researcher wants to make predictions or inferences about 'true' position of all members of the defined target population, then some type of probability sampling method should be selected. If the researcher aims to solely identify and obtain preliminary insights into the defined target population, non-probability methods might prove to be more appropriate.

Perceived statistical analysis needs:
The need for statistical projections or estimates based on the sample results is to be considered. Only probability sampling techniques allow the researcher to adequately use statistical analysis for estimates beyond the sample respondents. Though the statistical method can be applied on the non-probability samples of people and objects, the researcher's ability to accurately generalise the results and findings to the larger defined target population is technically inappropriate and questionable. The researcher should also decide on the appropriateness of sample size as it has a direct impact on the data quality, statistical precision and generalisation of findings.

Summary

- Sample should be truly representative of population characteristics without any bias so that it may result in valid and reliable conclusions.
- The attributes that are the object of study are referred to as characteristics and the units possessing them are called as elementary units. The aggregate of such units is generally described as population. Thus, all units in any field of inquiry form universe and all elementary units (on the basis of one characteristic or more) form population.
- An infinite population is that population in which it is theoretically impossible to observe all the elements. Thus, in an infinite population the number of items is infinite.
- A sample design is a definite plan for obtaining a sample from the sampling frame. It refers to the technique or the procedure the researcher would adopt in selecting some sampling units from which inferences about the population is drawn.
- A measure of the random sampling error can be calculated for a given sample design and size and this measure is often called the precision of the sampling plan. Sampling error is usually worked out as the product of the critical value at a certain level of significance and the standard error.
- Chi-square distribution is encountered when we deal with collections of values that involve adding up squares. Variances of samples require us to add a collection of squared quantities and thus have distributions that are related to chi-square distribution.
- The central limit theorem states that "the distribution of means of random samples taken from a population having mean m and finite variance σ^2 approaches the normal distribution with mean m and variance σ^2 / n as n goes to infinity.
- Sampling theory is a study of relationships existing between a population and samples drawn from the population. Sampling theory is applicable only to random samples. For this purpose the population or a universe may be defined as an aggregate of items possessing a common trait or traits.

References

- Dhawan, S., 2010. *Research Methodology for Business and Management Studies.* Global Media Publisher.
- Murthy, C., 2009. Research *methodology*, 1st ed., Vrinda Publications (P) Ltd.
- Kothari, C.R., 2004. *Research Methodology*. New Age International.
- Sridhar, S. M. *Research Methodology* [Online] Available at: <http://www.scribd.com/doc/2034905/Research-Methodology-Part-5-Sampling-Sampling-Strategy-or-Plan>. [Accessed 4 October 2011].
- Agarwal, S. C. and Khurana, S. K. *Research Methodology and Statistical Analysis* [Online] Available at: <http://books.google.co.in/books?id=3OEf_n4sxgYC&pg=PA301&dq=sampling+plan+in+research+methodology&hl=en&ei=Hk1oTsX5EpCGrAeFvbz3Cg&sa=X&oi=book_result&ct=result&resnum=7&ved=0CFgQ6AEwBg#v=onepage&q=sampling%20plan%20in%20research%20methodology&f=false pg 301>. [Accessed 4 October 2011]
- *Sampling methods*, 2008. [Video Online] Available at: <http://www.youtube.com/watch?v=wHgI6TolySw&feature=related>. [Accessed 4 October 2011].
- *5.5 How sample size is determined*, 2010. [Video Online] Available at: <http://www.youtube.com/watch?v=53EI30_uTkQ&feature=related>. [Accessed 4 October 2011].

Recommended Reading

- Panneerselvam, R., 2004. *Research Methodology*. PHI Learning Pvt. Ltd.
- Krishnaswamy, K.N ., Sivakumar, I. A. and Mathirajan, M., 2009. *Management Research Methodology: Integration of Methods and Techniques* Pearson Education India.
- Kumar, R., 2010. *Research Methodology: A Step-by-Step Guide for Beginners*. Sage Publication Ltd.

Self Assessment

1. The selection process of the items selected composes what is technically called 'a sample', this technique is called _____.
   a. sample design
   b. ample survey
   c. random technique
   d. simple random

2. The survey conducted on the basis of sample is described as _____.
   a. Sample Design
   b. Sample Survey
   c. Random Sample
   d. Chi square Distribution

3. From a statistical point of view, the term _____ refers to the total of the items or units in any field of inquiry, whereas the term _____ refers to the total of items about which information is desired.
   a. Universe, Population
   b. Universal, Popular
   c. Global, Pollution
   d. Population, Universe

4. Which of the following statements is false?
   a. An infinite population is that population in which it is theoretically impossible to observe all the elements.
   b. Chi-square distribution is encountered when we deal with collections of values that involve adding up squares.
   c. Sampling distribution of mean refers to the probability distribution of all the possible means of random samples of a given size that we take from a population.
   d. A chi-square distribution is a definite plan for obtaining a sample from the sampling frame.

5. Which of the following statements is true?
   a. Frame error = Sampling error + Chance error + Response error
   b. Chance error = Frame error + Sampling error + Response error
   c. Sampling error = Frame error + Chance error + Response error
   d. Response error = Frame error + Chance error + Sampling error

6. _____ is the range within which the population average (or other parameter) will lie in accordance with the reliability specified in the confidence level as a percentage of the estimate ± or as a numerical quantity.
   a. Precision
   b. Chance error
   c. Frame error
   d. Sampling

7. The theorem which explains the relationship between the shape of the population distribution and the sampling distribution of the mean is known as the _____.
   a. F-distribution
   b. Simple random sampling
   c. Central limit theorem
   d. Systematic sampling

8. In _____ , all units or items in the population have a chance of being chosen in the sample.
   a. random sampling
   b. simple random sampling
   c. systematic sampling
   d. random sampling

9. A simple random sample is one in which each item in the total population has _____ chance of being included in the sample.
   a. no
   b. an equal
   c. maximum
   d. minimum

10. Match the following.

Type of sampling	Definition
1. Multi stage sampling	A. Used when the population is heterogeneous rather than homogeneous
2. Stratified random sampling	B. It is a generalisation of two stage sampling.
3. Non random sampling methods	C. Obtaining a sample that is most conveniently available to the researcher
4. convenience sampling	D. Usually do not produce samples that are representative of the general population from which they are drawn

   a. 1-A; 2-B; 3-C; 4-D
   b. 1-B; 2-A; 3-D; 4-C
   c. 1-C; 2-B; 3-D; 4-A
   d. 1-D; 2-C; 3-B; 4-A

Chapter IV

Measurement and Scaling Techniques

Aim

The aim of this chapter is to:

- define measurement in research
- explain meaning of attitude
- introduce selection of measurement of scale

Objectives

The objectives of this chapter are to:

- elucidate sources of error in management
- highlight the tests of sound measurement
- explain scaling

Learning outcome

At the end of this chapter, you will be able to:

- discuss scale classification base
- understand scale construction techniques
- comprehend technique of developing measurement tools

4.1 Introduction

The data consists of quantitative variables like price, income, sales etc., and qualitative variables like knowledge, performance, character etc. The qualitative information must be converted into numerical form for further analysis. This is possible through measurement and scaling techniques. A common feature of survey based research is to have respondent's feelings, attitudes, opinions, etc. in some measurable form. For example, a bank manager may be interested in knowing the opinion of the customers about the services provided by the bank. Similarly, a fast food company having a network in a city may be interested in assessing the quality and service provided by them. As a researcher you may be interested in knowing the attitude of the people towards the government announcement of a metro rail in Delhi.

4.1.1 Measurement in Research

- Measurement is a relatively complex and demanding task, especially so when it concerns qualitative or abstract phenomena. By measurement we mean the process of assigning numbers to objects or observations, the level of measurement being a function of the rules under which the numbers are assigned.
- It is easy to assign numbers in respect of properties of some objects, but it is relatively difficult in respect of others. For instance, measuring such things as social conformity, intelligence, or marital adjustment is much less obvious and requires much closer attention than measuring physical weight, biological age or a person's financial assets. In other words, properties like weight, height, etc., can be measured directly with some standard unit of measurement, but it is not that easy to measure properties like motivation to succeed, ability to stand stress and the like.
- We can expect high accuracy in measuring the length of pipe with a yard stick, but if the concept is abstract and the measurement tools are not standardised, we are less confident about the accuracy of the results of measurement. Technically speaking, measurement is a process of mapping aspects of a domain onto other aspects of a range according to some rule of correspondence.
- In measuring, we devise some form of scale in the range (in terms of set theory, range may refer to some set) and then transform or map the properties of objects from the domain (in terms of set theory, domain may refer to some other set) onto this scale. For example, in case we are to find the male to female attendance ratio while conducting a study of persons who attend some show, then we may tabulate those who come to the show according to sex.
- In terms of set theory, this process is one of mapping the observed physical properties of those coming to the show (the domain) on to a sex classification (the range). The rule of correspondence is: If the object in the domain appears to be male, assign to "0" and if female assign to "1". Similarly, we can record a person's marital status as 1, 2, 3 or 4, depending on whether the person is single, married, widowed or divorced. We can as well record "Yes or No" answers to a question as "0" and "1" (or as 1 and 2 or perhaps as 59 and 60).
- In this artificial or nominal way, categorical data (qualitative or descriptive) can be made into numerical data and if we thus code the various categories, we refer to the numbers we record as nominal data. Nominal data are numerical in name only, because they do not share any of the properties of the numbers we deal in ordinary arithmetic.
- In those situations when we cannot do anything except set up inequalities, we refer to the data as ordinal data.
- When in addition to setting up inequalities we can also form differences, we refer to the data as interval data.
- When in addition to setting up inequalities and forming differences we can also form quotients (i.e., when we can perform all the customary operations of mathematics), we refer to such data as ratio data. In this sense, ratio data includes all the usual measurement (or determinations) of length, height, money amounts, weight, volume, area, pressures etc.
- The above stated distinction between nominal, ordinal, interval and ratio data is important for the nature of a set of data may suggest the use of particular statistical techniques. A researcher has to be quite alert about this aspect while measuring properties of objects or of abstract concepts.

4.1.2 Meaning of Attitude

- Attitudes are affected by attributes and beliefs. The salient factors that go into the building of the overall attitude of the individual towards the object are:
  - his/her beliefs about the attributes possessed by the object,
  - his/her preference or otherwise for those attributes, and
  - the relative importance of each attribute to the individual's decision making process
- In attitude measurement, the researcher is primarily interested in measuring the "state of mind" of the respondent (s). It may include factors such as awareness, attitudes and decision processes. An interesting characteristic of these measures is that their verification is rather difficult.
- There is no way to determine whether the answer given by a respondent to the level of liking for a new product, such as ice-cream mix, represents the "truth" or not. The researcher, unless he is a "telepathist", cannot actually observe the states of mind like preference, likes and dislikes, etc. Such things can only be inferred.

4.2 Measurement Scales

- The level of measurement refers to the relationship among the values that are assigned to the attributes, feelings or opinions for a variable. For example, the variable 'whether the taste of fast food is good' has a number of attributes, namely, very good, good, neither good nor bad, bad and very bad. For the purpose of analysing the results of this variable, we may assign the values 1, 2, 3, 4 and 5 to the five attributes respectively.
- The level of measurement describes the relationship among these five values. Here, we are simply using the numbers as shorter placeholders for the lengthier text terms. We don't mean that higher values mean 'more' of something or lower values mean 'less' of something. We don't assume that 'good' which has a value of 2 is twice of 'very good' which has a value of 1. We don't even assume that 'very good' which is assigned the value '1' has more preference than 'good' which is assigned the value '2'. We simply use the values as a shorter name for the attributes, opinions, or feelings. The assigned values of attributes allow the researcher more scope for further processing of data and statistical analysis.

4.2.1 Selection of Measurement of Scale

Scaling technique should be used which will yield the highest level of information feasible in a given situation. Also, if possible the technique should permit you the use of a variety of statistical analysis. A number of issues decide the choice of scaling technique. Some significant issues are:

Problem Definition and Statistical Analysis: The Choice between ranking, sorting, or rating techniques is determined by the problem definition and the type of statistical analysis likely to be performed. For example, ranking provides only ordinal data that limits the use of statistical techniques.

The Choice between Comparative and Non-comparative Scales: Some times it is better to use a comparative scale rather than a non-comparative scale.

Type of Category Labels: Many researchers use verbal categories since they believe that these categories are understood well by the respondents. The maturity and the education level of the respondents influence this decision.

Number of Categories: While there is no single, optimal number of categories, traditional guidelines suggest that there should be between five and nine categories. Also, if a neutral or indifferent scale response is possible for at least some of the respondents, an odd number of categories should be used. However, the researcher must determine the number of meaningful positions that are best suited for a specific problem.

Balanced versus Unbalanced Scale: In general, the scale should be balanced to obtain objective data.

Forced versus No forced Categories: In situations where the respondents are expected to have no opinion, the accuracy of data may be improved by a non forced scale that provides a 'no opinion' category.

4.2.2 Measurement of Scales

From what has been stated above, we can write that scales of measurement can be considered in terms of their mathematical properties. The most widely used classification of measurement scales are:

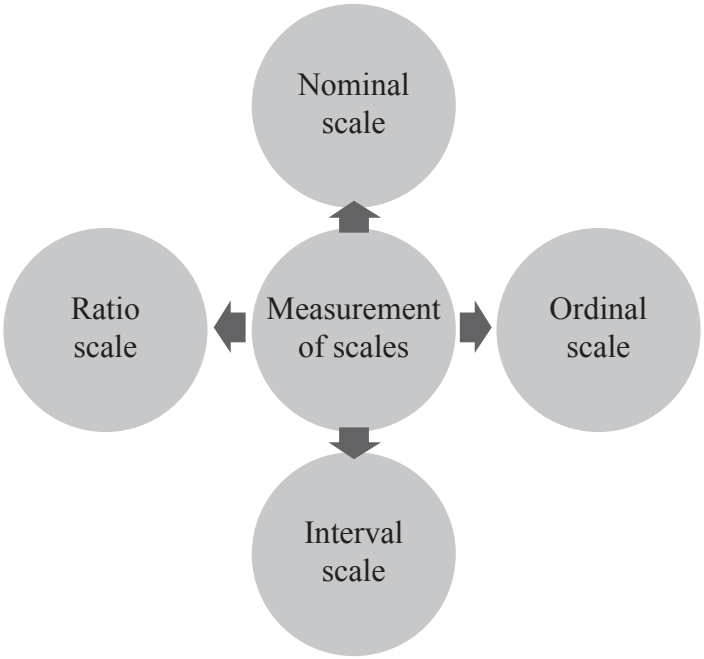


Fig. 4.1 Classification of measurement of scales

Nominal scale:

- Nominal scale is simply a system of assigning number symbols to events in order to label them. The usual example of this is the assignment of numbers of basketball players in order to identify them. Such numbers cannot be considered to be associated with an ordered scale for their order is of no consequence; the numbers are just convenient labels for the particular class of events and as such have no quantitative value.
- Nominal scales provide convenient ways of keeping track of people, objects and events. One cannot do much with the numbers involved. For example, one cannot usefully average the numbers on the back of a group of football players and come up with a meaningful value. Neither can one usefully compare the numbers assigned to one group with the numbers assigned to another.
- The counting of members in each group is the only possible arithmetic operation when a nominal scale is employed. Accordingly, we are restricted to use mode as the measure of central tendency. There is no generally used measure of dispersion for nominal scales.
- Chi-square test is the most common test of statistical significance that can be utilised, and for the measures of correlation, the contingency coefficient can be worked out.
- Nominal scale is the least powerful level of measurement. It indicates no order or distance relationship and has no arithmetic origin. A nominal scale simply describes differences between things by assigning them to categories. Nominal data are, thus, counted data. The scale wastes any information that we may have about varying degrees of attitude, skills, understandings, etc.
- In spite of all this, nominal scales are still very useful and are widely used in surveys and other ex-post-facto research when data are being classified by major sub-groups of the population.

Ordinal scale:

- The lowest level of the ordered scale that is commonly used is the ordinal scale. The ordinal scale places events in order, but there is no attempt to make the intervals of the scale equal in terms of some rule. Rank orders represent ordinal scales and are frequently used in research relating to qualitative phenomena.

- Ordinal scales only permit the ranking of items from highest to lowest. Ordinal measures have no absolute values, and the real differences between adjacent ranks may not be equal. All that can be said is that one person is higher or lower on the scale than another, but more precise comparisons cannot be made.
- Thus, the use of an ordinal scale implies a statement of 'greater than' or 'less than' (an equality statement is also acceptable) without our being able to state how much greater or less. The real difference between ranks 1 and 2 may be more or less than the difference between ranks 5 and 6.
- Since the numbers of this scale have only a rank meaning, the appropriate measure of central tendency is the median. A percentile or quartile measure is used for measuring dispersion. Correlations are restricted to various rank order methods. Measures of statistical significance are restricted to the non-parametric methods.

Interval scale:

- In the case of interval scale, the intervals are adjusted in terms of some rule that has been established as a basis for making the units equal. The units are equal only in so far as one accepts the assumptions on which the rule is based. Interval scales can have an arbitrary zero, but it is not possible to determine for them what may be called an absolute zero or the unique origin.
- The primary limitation of the interval scale is the lack of a true zero; it does not have the capacity to measure the complete absence of a trait or characteristic. Interval scales provide more powerful measurement than ordinal scales for interval scale also incorporates the concept of equality of interval.
- As such more powerful statistical measures can be used with interval scales. Mean is the appropriate measure of central tendency, while standard deviation is the most widely used measure of dispersion. Product moment correlation techniques are appropriate and the generally used tests for statistical significance are the't' test and 'F' test.

Ratio scale:

- Ratio scales have an absolute or true zero of measurement. The term 'absolute zero' is not as precise as it was once believed to be. We can conceive of an absolute zero of length and similarly we can conceive of an absolute zero of time.
- The ratio involved does have significance and facilitates a kind of comparison which is not possible in case of an interval scale. Ratio scale represents the actual amounts of variables. Measures of physical dimensions such as weight, height, distance, etc. are examples. Generally, all statistical techniques are usable with ratio scales and all manipulations that one can carry out with real numbers can also be carried out with ratio scale values.
- Multiplication and division can be used with this scale but not with other scales mentioned above. Geometric and harmonic means can be used as measures of central tendency and coefficients of variation may also be calculated. Thus, proceeding from the nominal scale (the least precise type of scale) to ratio scale (the most precise), relevant information is obtained increasingly.
- If the nature of the variables permits, the researcher should use the scale that provides the most precise description. Researchers in physical sciences have the advantage to describe variables in ratio scale form but the behavioural sciences are generally limited to describe variables in interval scale form, a less precise type of measurement.

4.3 Sources of Error in Measurement

Measurement should be precise and unambiguous in an ideal research study. This objective, however, is often not met with in entirety. As such the researcher must be aware about the sources of error in measurement. The following are the possible sources of error in measurement.

4.3.1 Respondent

At times the respondent may be reluctant to express strong negative feelings or it is just possible that he may have very little knowledge but may not admit his ignorance. All this reluctance is likely to result in an interview of 'guesses.' Transient factors like fatigue, boredom, anxiety, etc. may limit the ability of the respondent to respond accurately and fully.

4.3.2 Situation

Situational factors may also come in the way of correct measurement. Any condition which places a strain on interview can have serious effects on the interviewer-respondent rapport. For instance, if someone else is present, he can distort responses by joining in or merely by being present. If the respondent feels that anonymity is not assured, he may be reluctant to express certain feelings.

4.3.3 Measurer

The interviewer can distort responses by rewording or reordering questions. His behaviour, style and looks may encourage or discourage certain replies from respondents. Careless mechanical processing may distort the findings. Errors may also creep in because of incorrect coding, faulty tabulation and/or statistical calculations, particularly in the data-analysis stage.

4.3.4 Instrument

- Error may arise because of the defective measuring instrument. The use of complex words, beyond the comprehension of the respondent, ambiguous meanings, poor printing, inadequate space for replies, response choice omissions, etc. are a few things that make the measuring instrument defective and may result in measurement errors.
- Another type of instrument deficiency is the poor sampling of the universe of items of concern. Researcher must know that correct measurement depends on successfully meeting all of the problems listed above. He must, to the extent possible, try to eliminate, neutralise or otherwise deal with all the possible sources of error so that the final results may not be contaminated.

4.4 Tests of Sound Measurement

- Sound measurement must meet the tests of validity, reliability and practicality. In fact, these are the three major considerations one should use in evaluating a measurement tool.
  - Validity refers to the extent to which a test measures what we actually wish to measure.
  - Reliability has to do with the accuracy and precision of a measurement procedure.
  - Practicality is concerned with a wide range of factors of economy, convenience, and interpretability.

4.4.1 Test of Validity

- Validity is the most critical criterion and indicates the degree to which an instrument measures what it is supposed to measure. Validity can also be thought of as utility. In other words, validity is the extent to which differences found with a measuring instrument reflect true differences among those being tested.
- But the question arises: how can one determine validity without direct confirming knowledge? The answer may be that we seek other relevant evidence that confirms the answers we have found with our measuring tool. What is relevant, evidence often depends upon the nature of the research problem and the judgement of the researcher. But one can certainly consider three types of validity in this connection:
  - Content validity is the extent to which a measuring instrument provides adequate coverage of the topic under study. If the instrument contains a representative sample of the universe, the content validity is good. Its determination is primarily judgemental and intuitive. It can also be determined by using a panel of persons who shall judge how well the measuring instrument meets the standards, but there is no numerical way to express it.
  - Criterion-related validity relates to our ability to predict some outcome or estimate the existence of some current condition. This form of validity reflects the success of measures used for some empirical estimating purpose. The concerned criterion must possess the following qualities:
    - Relevance: (A criterion is relevant if it is defined in terms we judge to be the proper measure.)
    - Freedom from bias: (Freedom from bias is attained when the criterion gives each subject an equal opportunity to score well.)
- Reliability: (A reliable criterion is stable or reproducible.)

- Availability: (The information specified by the criterion must be available.) In fact, a Criterion-related validity is a broad term that actually refers to:
  - Predictive validity: the usefulness of a test in predicting some future performance
  - Concurrent validity: the usefulness of a test in closely relating to other measures of known validity.
  - Criterion-related validity is expressed as the coefficient of correlation between test scores and some measure of future performance or between test scores and scores on another measure of known validity.
  - Construct validity: It is the most complex and abstract. A measure is said to possess construct validity to the degree that it confirms to predicted correlations with other theoretical propositions. Construct validity is the degree to which scores on a test can be accounted for by the explanatory constructs of a sound theory. For determining construct validity, we associate a set of other propositions with the results received from using our measurement instrument. I measurements on our devised scale correlate in a predicted way with these other propositions, we can conclude that there is some construct validity. If the above stated criteria and tests are met with, we may state that our measuring instrument is valid and will result in correct measurement; otherwise we shall have to look for more information and/or resort to exercise of judgement.

4.4.2 Test of Reliability

- The test of reliability is another important test of sound measurement. A measuring instrument is reliable if it provides consistent results. Reliable measuring instrument does contribute to validity, but a reliable instrument need not be a valid instrument.
- If the quality of reliability is satisfied by an instrument, then while using it we can be confident that the transient and situational factors are not interfering. Two aspects of reliability, viz., stability and equivalence deserve special mention. The stability aspect is concerned with securing consistent results with repeated measurements of the same person and with the same instrument.
- We usually determine the degree of stability by comparing the results of repeated measurements. The equivalence aspect considers how much error may get introduced by different investigators or different samples of the items being studied. A good way to test for the equivalence of measurements by two investigators is to compare their observations of the same events. Reliability can be improved in the following two ways:
  - By standardising the conditions under which the measurement takes place i.e., we must ensure that external sources of variation such as boredom, fatigue, etc., are minimised to the extent possible. That will improve stability aspect.
  - By carefully designed directions for measurement with no variation from group to group, by using trained and motivated persons to conduct the research and also by broadening the sample of items used. This will improve equivalence aspect.

4.4.3 Test of Practicality

- The practicality characteristic of a measuring instrument can be judged in terms of economy, convenience and interpretability. From the operational point of view, the measuring instrument ought to be practical i.e., it should be economical, convenient and interpretable.
- Economy consideration suggests that some trade-off is needed between the ideal research project and that which the budget can afford. The length of measuring instrument is an important area where economic pressures are quickly felt. Although more items give greater reliability as stated earlier, but in the interest of limiting the interview or observation time, we have to take only few items for our study purpose.
- Similarly, data-collection methods to be used are also dependent at times upon economic factors. Convenience test suggests that the measuring instrument should be easy to administer. For this purpose one should give due attention to the proper layout of the measuring instrument. For instance, a questionnaire, with clear instructions (illustrated by examples), is certainly more effective and easier to complete than one which lacks these features.
- Interpretability consideration is especially important when persons other than the designers of the test are to interpret the results. The measuring instrument, in order to be interpretable, must be supplemented by

- detailed instructions for administering the test
- scoring keys
- evidence about the reliability
- guides for using the test and for interpreting results

4.4.4 Technique of Developing Measurement Tools

The technique of developing measurement tools involves a four-stage process, consisting of the following:

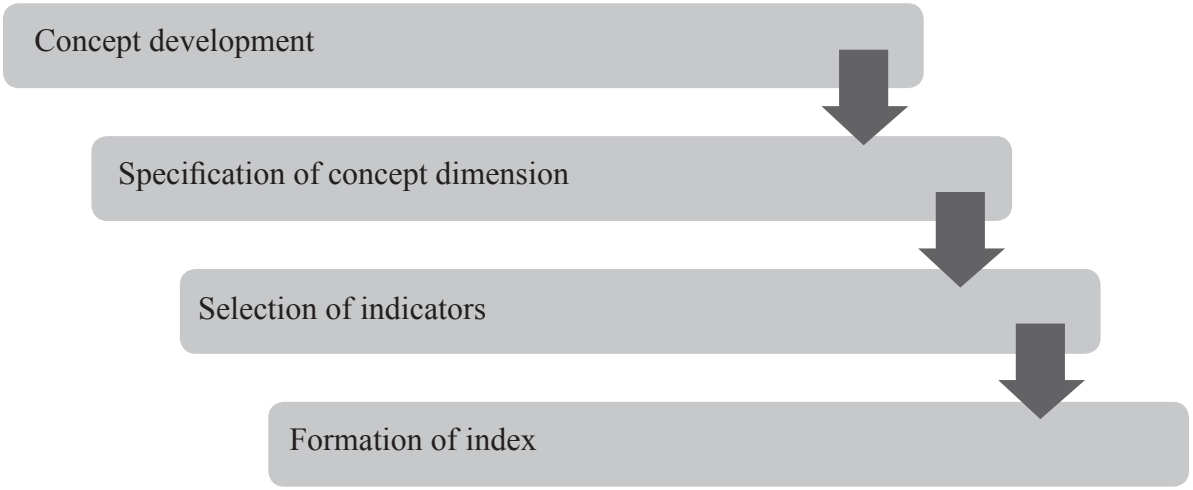


Fig. 4.2 Technique of developing measurement tools process

- The first and foremost step is that of concept development which means that the researcher should arrive at an understanding of the major concepts pertaining to his study. This step of concept development is more apparent in theoretical studies than in the more pragmatic research, where the fundamental concepts are often already established.
- The second step requires the researcher to specify the dimensions of the concepts that he developed in the first stage. This task may either be accomplished by deduction i.e., by adopting a more or less intuitive approach or by empirical correlation of the individual dimensions with the total concept and/or the other concepts.
- For instance, one may think of several dimensions such as product reputation, customer treatment, corporate leadership, concern for individuals, and sense of social responsibility and so forth when one is thinking about the image of a certain company. Once the dimensions of a concept have been specified, the researcher must develop indicators for measuring each concept element. Indicators are specific questions, scales, or other devices by which respondent's knowledge, opinion, expectation, etc., are measured.
- As there is seldom a perfect measure of a concept, the researcher should consider several alternatives for the purpose. The use of more than one indicator gives stability to the scores and it also improves their validity.
- The last step is that of combining the various indicators into an index, i.e., formation of an index. When we have several dimensions of a concept or different measurements of a dimension, we may need to combine them into a single index. One simple way for getting an overall index is to provide scale values to the responses and then sum up the corresponding scores.

4.5 Scaling

- In research we quite often face measurement problem (since we want a valid measurement but may not obtain it), especially when the concepts to be measured are complex and abstract and we do not possess the standardised measurement tools. Alternatively, we can say that while measuring attitudes and opinions, we face the problem of their valid measurement.
- Similar problem may be faced by a researcher, of course in a lesser degree, while measuring physical or institutional concepts. As such we should study some procedures which may enable us to measure abstract concepts more accurately. This brings us to the study of scaling techniques.

4.5.1 Meaning of Scaling

- Scaling describes the procedures of assigning numbers to various degrees of opinion, attitude and other concepts. This can be done in two ways, viz.,
  - making a judgement about some characteristic of an individual and then placing him directly on a scale that has been defined in terms of that characteristic and
  - Constructing questionnaires in such a way that the score of individual's responses assigns him a place on a scale. It may be stated here that a scale is a continuum, consisting of the highest point (in terms of some characteristic, example, preference, favourableness, etc.) and the lowest point along with several intermediate points between these two extreme points.
- These scale-point positions are so related to each other that when the first point happens to be the highest point, the second point indicates a higher degree in terms of a given characteristic as compared to the third point and the third point indicates a higher degree as compared to the fourth and so on. Numbers for measuring the distinctions of degree in the attitudes/opinions are, thus, assigned to individuals corresponding to their scale-positions.
- All this is better understood when we talk about scaling technique(s). Hence the term 'scaling' is applied to the procedures for attempting to determine quantitative measures of subjective abstract concepts. Scaling has been defined as a "procedure for the assignment of numbers (or other symbols) to a property of objects in order to impart some of the characteristics of numbers to the properties in question."

4.5.2 Scale Classification Bases

The number assigning procedures or the scaling procedures may be broadly classified on one or more of the following bases:

- **Subject orientation:** Under it, a scale may be designed to measure characteristics of the respondent who completes it or to judge the stimulus object which is presented to the respondent. In respect of the former, we presume that the stimuli presented are sufficiently homogeneous, so that the between stimuli variation is small as compared to the variation among respondents. In the latter approach, we ask the respondent to judge some specific object in terms of one or more dimensions and we presume that the between-respondent variation will be small as compared to the variation among the different stimuli presented to respondents for judging.
- **Response form:** Under this, we may classify the scales as categorical and comparative. Categorical scales are also known as rating scales. These scales are used when a respondent scores some object without direct reference to other objects. Under comparative scales, which are also known as ranking scales, the respondent is asked to compare two or more objects. In this sense, the respondent may state that one object is superior to the other or those three models of pen rank in order 1, 2 and 3. The essence of ranking is, in fact, a relative comparison of a certain property of two or more objects.
- **Degree of subjectivity:** With this basis, the scale data may be based on whether we measure subjective personal preferences or simply make non-preference judgements. In the former case, the respondent is asked to choose which person he favours or which solution he would like to see employed, whereas in the latter case he is simply asked to judge which person is more effective in some aspect or which solution will take fewer resources without reflecting any personal preference.
- **Scale properties:** Considering scale properties, one may classify the scales as nominal, ordinal, interval and ratio scales. Nominal scales merely classify without indicating order, distance or unique origin. Ordinal scales indicate magnitude relationships of 'more than' or 'less than', but indicate no distance or unique origin. Interval scales have both, order and distance values, but no unique origin. Ratio scales possess all these features.
- **Number of dimensions:** In respect of this basis, scales can be classified as 'unidimensional' and 'multidimensional' scales. Under the former we measure only one attribute of the respondent or object, whereas multidimensional scaling recognises that an object might be described better by using the concept of an attribute space of 'n' dimensions, rather than a single-dimension continuum.
- **Scale construction techniques:** Following are the five main techniques by which scales can be developed.
  - Arbitrary approach: It is an approach where scale is developed on ad hoc basis. This is the most widely used approach. It is presumed that such scales measure the concepts for which they have been designed, although there is little evidence to support such an assumption.

- Consensus approach: Here a panel of judges evaluate the items chosen for inclusion in the instrument in terms of whether they are relevant to the topic area and unambiguous in implication.
- Item analysis approach: Under it, a number of individual items are developed into a test which is given to a group of respondents. After administering the test, the total scores are calculated for every one. Individual items are then analysed to determine which items discriminate between persons or objects with high total scores and those with low scores.
- Cumulative scales are chosen on the basis of their conforming to some ranking of items with ascending and descending discriminating power. For instance, in such a scale the endorsement of an item representing an extreme position should also result in the endorsement of all items indicating a less extreme position.
- Factor scales may be constructed on the basis of inter correlations of items which indicate that a common factor accounts for the relationship between items. This relationship is typically measured through factor analysis method.

4.5.3 Important Scaling Techniques

Some of the important scaling techniques often used in the context of research especially in context of social or business research.

Rating scales:

- The rating scale involves qualitative description of a limited number of aspects of a thing or of traits of a person. When we use rating scales (or categorical scales), we judge an object in absolute terms against some specified criteria, i.e., we judge properties of objects without reference to other similar objects.
- These ratings may be in such forms as "like-dislike", "above average, average, below average", or other classifications with more categories such as "like very much—like some what—neutral—dislike somewhat—dislike very much"; "excellent—good—average—below average—poor", "always—often—occasionally—rarely—never", and so on. There is no specific rule whether to use a two-point scale, three-point scale or scale with still more points.
- In practice, three to seven points' scales are generally used for the simple reason that more points on a scale provide an opportunity for greater sensitivity of measurement. Rating scale may be either a graphic rating scale or an itemised rating scale.
- Ranking scales: Under ranking scales (or comparative scales) we make relative judgements against other similar objects. The respondents under this method directly compare two or more objects and make choices among them. There are two generally used approaches of ranking scales, viz.,
  - Method of paired comparisons: Under it the respondent can express his attitude by making a choice between two objects, say between a new flavour of soft drink and an established brand of drink. But when there are more than two stimuli to judge, the number of judgements required in a paired comparison is given by the formula:

$$N = \frac{n\,(n-1)}{2}$$

  where N = number of judgements

  n = number of stimuli or objects to be judged.
- For instance, if there are ten suggestions for bargaining proposals available to a workers union, there are 45 paired comparisons that can be made with them. When N happens to be a big figure, there is the risk of respondents giving ill considered answers or they may even refuse to answer.
- We can reduce the number of comparisons per respondent either by presenting to each one of them only a sample of stimuli or by choosing a few objects which cover the range of attractiveness at about equal intervals and then comparing all other stimuli to these few standard objects. Thus, paired-comparison data may be treated in several ways.

- If there is substantial consistency, we will find that if X is preferred to Y, and Y to Z, then X will consistently be preferred to Z. If this is true, we may take the total number of preferences among the comparisons as the score for that stimulus. It should be remembered that paired comparison provides ordinal data, but the same may be converted into an interval scale by the method of the Law of Comparative Judgement developed by L.L. Thurstone.
- This technique involves the conversion of frequencies of preferences into a table of proportions which are then transformed into Z matrix by referring to the table of area under the normal curve. J.P. Guilford in his book "Psychometric Methods" has given a procedure which is relatively easier. The method is known as the Composite Standard Method.
- By following the composite standard method, we can develop an interval scale from the paired comparison ordinal data given in the table for which purpose we have to adopt the following steps in order:
  - Using the data in the table, we work out the column mean with the help of the formula given below:

$$M_p = \frac{C + 5\,(N)}{nN}$$

Where,
Mp = the mean proportion of the columns
C = the total number of choices for a given suggestion
n = number of stimuli (proposals in the given problem)
N = number of items in the sample.

The column means have been shown in the Mp row in the above table.

- The Z values for the Mp are secured from the table giving the area under the normal curve. When the Mp value is less than .5, the Z value is negative and for all Mp values higher than .5, the Z values are positive.* These Z values are shown in Zj row in the above table.
- As the Zj values represent an interval scale, zero is an arbitrary value. Hence we can eliminate negative scale values by giving the value of zero to the lowest scale value (this being (–).11 in our example which we shall take equal to zero) and then adding the absolute value of this lowest scale value to all other scale items. This scale has been shown in Rj row in the above table.

Method of rank order:

- Under this method of comparative scaling, the respondents are asked to rank their choices. This method is easier and faster than the method of paired comparisons stated above. Moreover, a complete ranking at times is not needed in which case the respondents may be asked to rank only their first, say, four choices while the number of overall items involved may be more than four, say, it may be 15 or 20 or more. To secure a simple ranking of all items involved we simply total rank values received by each item.
- There are methods through which we can as well develop an interval scale of these data. But then there are limitations of this method. The first one is that data obtained through this method are ordinal data and hence rank ordering is an ordinal scale with all its limitations. Then there may be the problem of respondents becoming careless in assigning ranks particularly when there are many (usually more than 10) items.

4.6 Scale Construction Techniques

- In social science studies, while measuring attitudes of the people, we generally follow the technique of preparing the opinionnaire* (or attitude scale) in such a way that the score of the individual responses assigns him a place on a scale. Under this approach, the respondent expresses his agreement or disagreement with a number of statements relevant to the issue. While developing such statements, the researcher must note the following two points:
  - That the statements must elicit responses which are psychologically related to the attitude being measured;

- That the statements need be such that they discriminate not merely between extremes of attitude but also among individuals who differ slightly.

- Researchers must as well be aware that inferring attitude from what has been recorded in opinionnaires has several limitations. People may conceal their attitudes and express socially acceptable opinions. They may not really know how they feel about a social issue. People may be unaware of their attitude about an abstract situation; until confronted with a real situation, they may be unable to predict their reaction.
- Even behaviour itself is at times not a true indication of attitude. For instance, when politicians kiss babies, their behaviour may not be a true expression of affection toward infants. Thus, there is no sure method of measuring attitude; we only try to measure the expressed opinion and then draw inferences from it about people's real feelings or attitudes.
- With all these limitations in mind, psychologists and sociologists have developed several scale construction techniques for the purpose. The researcher should know these techniques so as to develop an appropriate scale for his own study. Some of the important approaches, along with the corresponding scales developed under each approach to measure attitude are as follows:

Name of the scale construction approach	Name of the scale developed
Arbitrary approach	Arbitrary scales
Consensus scale approach	Differential scales (For example, Thurstone Differentialscale)
Item analysis approach	Summated scales (For example, Likert scale)
Cumulative scale approach	Cumulative scales (For example, Guttman's Scalogram)
Factor analysis approach	Factor scales (For example, Osgood's Semantic Differential, Multi-dimensional Scaling etc.)

Table 4.1 Different scales for measuring attitudes of people

4.6.1 Arbitrary Scales

- Arbitrary scales are developed on ad hoc basis and are designed largely through the researcher's own subjective selection of items. The researcher first collects few statements or items which he believes are unambiguous and appropriate to a given topic. Some of these are selected for inclusion in the measuring instrument and then people are asked to check in a list the statements with which they agree.
- The chief merit of such scales is that they can be developed very easily, quickly and with relatively less expense. They can also be designed to be highly specific and adequate. Because of these benefits, such scales are widely used in practice.
- At the same time, there are some limitations of these scales. The most important one is that we do not have objective evidence that such scales measure the concepts for which they have been developed. We have simply to rely on researcher's insight and competence.

4.6.2 Differential Scales (or Thurstone-type Scales)

The name of L.L. Thurstone is associated with differential scales which have been developed using consensus scale approach. Under such an approach, the selection of items is made by a panel of judges who evaluate the items in terms of whether they are relevant to the topic area and unambiguous in implication. The detailed procedure is as under:

- The researcher gathers a large number of statements, usually twenty or more, that express various points of view toward a group, institution, idea, or practice (i.e., statements belonging to the topic area).
- These statements are then submitted to a panel of judges, each of whom arranges them in eleven groups or piles ranging from one extreme to another in position. Each of the judges is requested to place generally in the first pile the statements which he thinks are most unfavourable to the issue, in the second pile to place those statements which he thinks are next most unfavourable and he goes on doing so in this manner till in the eleventh pile he puts the statements which he considers to be the most favourable.

- This sorting by each judge yields a composite position for each of the items. In case of marked disagreement between the judges in assigning a position to an item, that item is discarded.
- For items that are retained, each is given its median scale value between one and eleven as established by the panel. In other words, the scale value of any one statement is computed as the 'median' position to which it is assigned by the group of judges.
- A final selection of statements is then made. For this purpose a sample of statements, whose median scores are spread evenly from one extreme to the other is taken. The statements so selected, constitute the final scale to be administered to respondents. The position of each statement on the scale is the same as determined by the judges.
- After developing the scale as stated above, the respondents are asked during the administration of the scale to check the statements with which they agree. The median value of the statements that they check is worked out and this establishes their score or quantifies their opinion. It may be noted that in the actual instrument the statements are arranged in random order of scale value.
- If the values are valid and if the opinionnaire deals with only one attitude dimension, the typical respondent will choose one or several contiguous items (in terms of scale values) to reflect his views. However, at times divergence may occur when a statement appears to tap a different attitude dimension.
- The Thurstone method has been widely used for developing differential scales which are utilised to measure attitudes towards varied issues like war, religion, etc. Such scales are considered most appropriate and reliable when used for measuring a single attitude. But an important deterrent to their use is the cost and effort required to develop them. Another weakness of such scales is that the values assigned to various statements by the judges may reflect their own attitudes. The method is not completely objective; it involves ultimately subjective decision process. Critics of this method also opine that some other scale designs give more information about the respondent's attitude in comparison to differential scales.

4.6.3 Summated Scales (or Likert-type Scales)

- Summated scales (or Likert-type scales) are developed by utilising the item analysis approach wherein a particular item is evaluated on the basis of how well it discriminates between those persons whose total score is high and those whose score is low. Those items or statements that best meet this sort of discrimination test are included in the final instrument.
- Thus, summated scales consist of a number of statements which express either a favourable or unfavourable attitude towards the given object to which the respondent is asked to react. The respondent indicates his agreement or disagreement with each statement in the instrument. Each response is given a numerical score, indicating its favourableness or unfavourableness, and the scores are totalled to measure the respondent's attitude.
- In other words, the overall score represents the respondent's position on the continuum of favourable-unfavourableness towards an issue. Most frequently used summated scales in the study of social attitudes follow the pattern devised by Likert. For this reason they are often referred to as Likert-type scales.
- In a Likert scale, the respondent is asked to respond to each of the statements in terms of several degrees, usually five degrees (but at times 3 or 7 may also be used) of agreement or disagreement. For example, when asked to express opinion whether one considers his job quite pleasant, the respondent may respond in any one of the following ways:
  - strongly agree
  - agree
  - undecided
  - disagree
  - strongly disagree
- We find that these five points constitute the scale. At one extreme of the scale there is strong agreement with the given statement and at the other, strong disagreement, and between them lie intermediate points. Each point on the scale carries a score. Response indicating the least favourable degree of job satisfaction is given the least score (say 1) and the most favourable is given the highest score (say 5). These score—values are normally not printed on the instrument but are shown here just to indicate the scoring pattern.

- The Likert scaling technique, thus, assigns a scale value to each of the five responses. The same thing is done in respect of each and every statement in the instrument. This way, the instrument yields a total score for each respondent, which would then measure the respondent's favourableness toward the given point of view. If the instrument consists of, say 30 statements, the following score values would be revealing.

  $30 \times 5 = 150$ Most favourable response possible

  $30 \times 3 = 90$ A neutral attitude

  $30 \times 1 = 30$ Most unfavourable attitude.

- The scores for any individual would fall between 30 and 150. If the score happens to be above 90, it shows favourable opinion to the given point of view, a score of below 90 would mean unfavourable opinion and a score of exactly 90 would be suggestive of a neutral attitude.
- Procedure: The procedure for developing a Likert-type scale is as follows:
  - As a first step, the researcher collects a large number of statements which are relevant to the attitude being studied and each of the statements expresses definite favourableness or unfavourableness to a particular point of view or the attitude and that the number of favourable and unfavourable statements is approximately equal.
  - After the statements have been gathered, a trial test should be administered to a number of subjects. In other words, a small group of people, from those who are going to be studied finally, are asked to indicate their response to each statement by checking one of the categories of agreement or disagreement using a five point scale as stated above.
  - The response to various statements are scored in such a way that a response indicative of the most favourable attitude is given the highest score of 5 and that with the most unfavourable attitude is given the lowest score, say, of 1.
  - Then, the total score of each respondent is obtained by adding his scores that he received for separate statements.
  - The next step is to array these total scores and find out those statements which have a high discriminatory power. For this purpose, the researcher may select some part of the highest and the lowest total scores say the top 25 per cent and the bottom 25 per cent. These two extreme groups are interpreted to represent the most favourable and the least favourable attitudes and are used as criterion groups by which to evaluate individual statements. This way we determine which statements consistently correlate with low favourability and which with high favourability.
  - Only those statements that correlate with the total test should be retained in the final instrument and all others must be discarded from it.
- **Advantages:** The Likert-type scale has several advantages. Mention may be made of the important ones.
  - It is relatively easy to construct the Likert-type scale in comparison t Thurstone-type scale because Likert-type scale can be performed without a panel of judges.
  - Likert-type scale is considered more reliable because under it respondents answer each statement included in the instrument. As such it also provides more information and data than does the Thurstone-type scale.
  - Each statement, included in the Likert-type scale, is given an empirical test for discriminating ability and as such, unlike Thurstone-type scale, the Likert-type scale permits the use of statements that are not manifestly related (to have a direct relationship) to the attitude being studied.
  - Likert-type scale can easily be used in respondent-centred and stimulus-centred studies i.e., through it we can study how responses differ between people and how responses differ between stimuli.
  - Likert-type scale takes much less time to construct, it is frequently used by the students of opinion research. Moreover, it has been reported in various research studies* that there is high degree of correlation between Likert-type scale and Thurstone-type scale.
- **Limitations:** There are several limitations of the Likert-type scale as well.
  - One important limitation is that, with this scale, we can simply examine whether respondents are more or less favourable to a topic, but we cannot tell how much more or less they are. There is no basis for belief that the five positions indicated on the scale are equally spaced. The interval between 'strongly agree' and

'agree', may not be equal to the interval between "agree" and "undecided". This means that Likert scale does not rise to a stature more than that of an ordinal scale, whereas the designers of Thurstone scale claim the Thurstone scale to be an interval scale.

- One further disadvantage is that often the total score of an individual respondent has little clear meaning since a given total score can be secured by a variety of answer patterns. It is unlikely that the respondent can validly react to a short statement on a printed form in the absence of real-life qualifying situations.
- Moreover, there "remains a possibility that people may answer according to what they think they should feel rather than how they do feel." This particular weakness of the Likert-type scale is met by using a cumulative scale.

- In spite of all the limitations, the Likert-type summated scales are regarded as the most useful in a situation wherein it is possible to compare the respondent's score with a distribution of scores from some well defined group. They are equally useful when we are concerned with a programme of change or improvement in which case we can use the scales to measure attitudes before and after the programme of change or improvement in order to assess whether our efforts have had the desired effects.
- We can as well correlate scores on the scale to other measures without any concern for the absolute value of what is favourable and what is unfavourable. All this accounts for the popularity of Likert-type scales in social studies relating to measuring of attitudes.

4.7 Multidimensional Scaling

- Multidimensional scaling (MDS) is relatively more complicated scaling device, but with this sort of scaling one can scale objects, individuals or both with a minimum of information. Multidimensional scaling (or MDS) can be characterised as a set of procedures for portraying perceptual or affective dimensions of substantive interest.
- It "provides useful methodology for portraying subjective judgements of diverse kinds." MDS is used when all the variables (whether metric or non-metric) in a study are to be analysed simultaneously and all such variables happen to be independent. The underlying assumption in MDS is that people (respondents) "perceive a set of objects as being more or less similar to one another on a number of dimensions (usually uncorrelated with one another) instead of only one."
- Through MDS techniques, one can represent geometrically the locations and interrelationships among a set of points. In fact, these techniques attempt to locate the points, given the information about a set of interpoint distances, in space of one or more dimensions such as to best summarise the information contained in the interpoint distances.
- The distances in the solution space then optimally reflect the distances contained in the input data. For instance, if objects, say X and Y, are thought of by the respondent as being most similar as compared to all other possible pairs of objects, MDS techniques will position objects X and Y in such a way that the distance between them in multidimensional space is shorter than that between any two other objects.
- Two approaches, viz., the metric approach and the non-metric approach, are usually talked about in the context of MDS, while attempting to construct a space containing m points such that $m(m-1)/2$ interpoint distances reflect the input data. The metric approach to MDS treats the input data as interval scale data and solves applying statistical methods for the additive constant* which minimises the dimensionality of the solution space.
- This approach utilises all the information in the data in obtaining a solution. The data (i.e., the metric similarities of the objects) are often obtained on a bipolar similarity scale on which pairs of objects are rated one at a time. If the data reflect exact distances between real objects in an r-dimensional space, their solution will reproduce the set of interpoint distances. But as the true and real data are rarely available, we require random and systematic procedures for obtaining a solution.
- Generally, the judged similarities among a set of objects are statistically transformed into distances by placing those objects in a multidimensional space of some dimensionality.
- The non-metric approach first gathers the non-metric similarities by asking respondents to rank order all possible pairs that can be obtained from a set of objects. Such non-metric data is then transformed into some arbitrary metric space and then the solution is obtained by reducing the dimensionality.

- In other words, this non-metric approach seeks "a representation of points in a space of minimum dimensionality such that the rank order of the interpoint distances in the solution space maximally corresponds to that of the data. This is achieved by requiring only that the distances in the solution be monotone with the input data." The non-metric approach has come into prominence during the sixties with the coming into existence of high speed computers to generate metric solutions for ordinal input data.
- The significance of MDS lies in the fact that it enables the researcher to study "the perceptual structure of a set of stimuli and the cognitive processes underlying the development of this structure. Psychologists, for example, employ multidimensional scaling techniques in an effort to scale psychophysical stimuli and to determine appropriate labels for the dimensions along which these stimuli vary."
- The MDS techniques, infact, do away with the need in the data collection process to specify the attribute(s) along which the several brands, say of a particular product, may be compared as ultimately the MDS analysis itself reveals such attribute(s) that presumably underlie the expressed relative similarities among objects. Thus, MDS is an important tool in attitude measurement and the techniques falling under MDS promise "a great advance from a series of unidimensional measurements (example, a distribution of intensities of feeling towards single attribute such as colour, taste or a preference ranking with indeterminate intervals), to a perceptual mapping in multidimensional space of objects, company images, advertisement brands, etc."
- In spite of all the merits stated above, the MDS is not widely used because of the computation complications involved under it. Many of its methods are quite laborious in terms of both the collection of data and the subsequent analyses. However, some progress has been achieved (due to the pioneering efforts of Paul Green and his associates) during the last few years in the use of non-metric MDS in the context of market research problems.
- The techniques have been specifically applied in "finding out the perceptual dimensions, and the spacing of stimuli along these dimensions, that people, use in making judgements about the relative similarity of pairs of Stimuli."12 But, "in the long run, the worth of MDS will be determined by the extent to which it advances the behavioural sciences."

Summary

- Measurement is a relatively complex and demanding task, especially so when it concerns qualitative or abstract phenomena. By measurement we mean the process of assigning numbers to objects or observations, the level of measurement being a function of the rules under which the numbers are assigned.
- Nominal scales provide convenient ways of keeping track of people, objects and events. One cannot do much with the numbers involved. For example, one cannot usefully average the numbers on the back of a group of football players and come up with a meaningful value. Neither can one usefully compare the numbers assigned to one group with the numbers assigned to another.
- The lowest level of the ordered scale that is commonly used is the ordinal scale. The ordinal scale places events in order, but there is no attempt to make the intervals of the scale equal in terms of some rule. Rank orders represent ordinal scales and are frequently used in research relating to qualitative phenomena.
- Validity is the most critical criterion and indicates the degree to which an instrument measures what it is supposed to measure. Validity can also be thought of as utility. In other words, validity is the extent to which differences found with a measuring instrument reflect true differences among those being tested.
- The rating scale involves qualitative description of a limited number of aspects of a thing or of traits of a person. When we use rating scales (or categorical scales), we judge an object in absolute terms against some specified criteria, i.e., we judge properties of objects without reference to other similar objects.
- The non-metric approach first gathers the non-metric similarities by asking respondents to rank order all possible pairs that can be obtained from a set of objects. Such non-metric data is then transformed into some arbitrary metric space and then the solution is obtained by reducing the dimensionality.
- In spite of all the limitations, the Likert-type summated scales are regarded as the most useful in a situation wherein it is possible to compare the respondent's score with a distribution of scores from some well defined group.

References

- *Levels of Measurement and Scaling* [Online] Available at: <http://www.fao.org/docrep/W3241E/w3241e04.htm>. [Accessed 4 October 2011].
- Kothari, C. R., 2004. *Research Methodology.* New Age International.
- *Measurement and scaling techniques* [Online] Available at: <http://www.egyankosh.ac.in/bitstream/123456789/2404/1/Unit%205.pdf>.
- [Accessed 4 October 2011].
- *Attitude measurement and scales* [Online] Available at: <http://www.egyankosh.ac.in/bitstream/123456789/7661/1/Unit-6.pdf>. [Accessed 4 October 2011].
- *Scales.,* 2009. [Video Online] Available at: <http://www.youtube.com/watch?v=GMIpnzLQtTQ >. [Accessed 4 October 2011].
- *Introduction to Statistics: Levels of Measurement,* 2011. [Video Online] Available at: <http://www.youtube.com/watch?v=B0ABvLa_u88&feature=related>. [Accessed 4 October 2011].

Recommended Reading

- *Phophalia, A.K., 2010.* Modern Research Methodology: New Trends and Techniques*. Global Media.*
- *Khan., Research Methodology. APH Publishing.*
- *Panneerselva*m, R., 2004. *Research Methodology*. PHI Learning Pvt. Ltd.

Self Assessment

1. In those situations when we cannot do anything except set up inequalities, we refer to the data as _____.
   a. Ordinal data
   b. Interval data
   c. Retrieval data
   d. Data

2. When in addition to setting up inequalities we can also form differences, we refer to the data as _____.
   a. Ordinal data
   b. Retrieval data
   c. interval data
   d. process data

3. In attitude measurement, the researcher is primarily interested in measuring the _____ of the respondent (s).
   a. Attitude
   b. Personality
   c. IQ
   d. state of mind

4. _____ is simply a system of assigning number symbols to events in order to label them.
   a. Nominal scale
   b. Ordinal scale
   c. Interval scale
   d. Ratio scale

5. The use of an _____ implies a statement of 'greater than' or 'less than' (an equality statement is also acceptable) without our being able to state how much greater or less.
   a. Nominal scale
   b. Ordinal scale
   c. Interval scale
   d. Ratio scale

6. _____ is the appropriate measure of central tendency, while _____ is the most widely used measure of dispersion.
   a. Mean, mode
   b. Mode, standard deviation
   c. Mean, standard deviation
   d. Standard deviation, mean

7. _____ is the extent to which differences found with a measuring instrument reflect true differences among those being tested.
   a. Validity
   b. Reliability
   c. Availability
   d. Elasticity

8. Which statement is false?
   a. Construct validity is the degree to which scores on a test can be accounted for by the explanatory constructs of a sound theory.
   b. Predictive validity is the usefulness of a test in predicting some future performance
   c. Criterion related validity is the usefulness of a test in closely relating to other measures of known validity.
   d. Content validity is the extent to which a measuring instrument provides adequate coverage of the topic under study.

9. _____ describes the procedures of assigning numbers to various degrees of opinion, attitude and other concepts.
   a. Scaling
   b. Validity
   c. Mapping
   d. Ratio scale

10. The researcher must develop indicators for measuring each_____.
   a. concept element
   b. index
   c. variant
   d. scale

Chapter V

Methods of Data Collection

Aim

The aim of this chapter is to:

- introduce the meaning and need for data
- explain collection of data through questionnaires
- define sources of data

Objectives

The objectives of this chapter are to:

- elucidate modern methods of data collection
- explicate problems in response questionnaire
- enlist different types of collection of data through schedule

Learning outcome

At the end of this chapter, you will be able to:

- discuss collection of secondary data
- comprehend appropriate method for data collection
- understand guidelines for constructing questionnaire / schedule

5.1 Introduction

A research design is a blue print which directs the plan of action to complete the research work. The quality and credibility of the results derived from the application of research methodology depends upon the relevant, accurate and adequate data. In this unit, we shall study about the various sources of data and methods of collecting primary and secondary data with their merits and limitations and also the choice of suitable method for data collection.

5.2 Meaning and Need for Data

- Data is required to make a decision in any business situation. The researcher is faced with one of the most difficult problems of obtaining suitable, accurate and adequate data. Utmost care must be exercised while collecting data because the quality of the research results depends upon the reliability of the data.
- Suppose you are the Director of your company. Your Board of Directors has asked you to find out why the profit of the company has decreased since the last two years. Your Board wants you to present facts and figures. What are you going to do?
- The first and foremost task is to collect the relevant information to make an analysis for the above mentioned problem. It is, therefore, the information collected from various sources, which can be expressed in quantitative form, for a specific purpose, which is called data. The rational decision maker seeks to evaluate information in order to select the course of action that maximises objectives.
- For decision making, the input data must be appropriate. This depends on the appropriateness of the method chosen for data collection. The application of a statistical technique is possible when the questions are answerable in quantitative nature, for instance; the cost of production and profit of the company measured in rupees, age of the workers in the company measured in years. Therefore, the first step in statistical activities is to gather data. The data may be classified as primary and secondary data. Let us now discuss these two kinds of data in detail.

5.2.1 Source of Data

Data sources can be broadly categorised into three types.

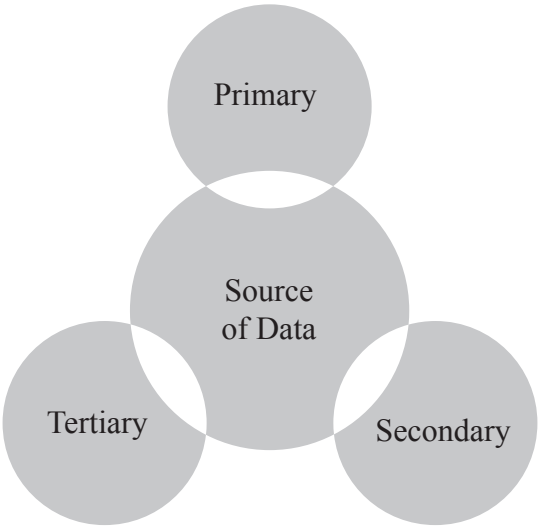


Fig. 5.1 Types of data sources

- **Primary data sources:** Primary data refers to information gathered first hand by the researcher for the specific purpose of the study. It is raw data without interpretation and represents the personal or official opinion or position. Primary sources are most authorative since the information is not filtered or tampered. Some examples of the sources of primary data are individuals, focus groups, panel of respondents, internet etc. Data collection from individuals can be made through interviews, observation etc.

- **Secondary data sources:** There is hardly same secondary data available on a spatial basis and on a micro level. Some of it is available in an aggregate level which does not cater to the needs of this study. There is very little data of important environmental factors like air, water and noise pollution and traffic congestion on special level. Data on all diseases and child mortality have been collected from the patient's registers of various hospitals onward basis, these are: density of population, density of houses.
- **Tertiary sources:** Tertiary sources are an interpretation of a secondary source. It is generally represented by index, bibliographies, dictionaries, encyclopaedias, handbooks, directories and other finding aids like the internet search engines.

5.2.2 Primary and Secondary Data

- The Primary data are original data which are collected for the first time for a specific purpose. Such data are published by authorities who themselves are responsible for their collection.
- The Secondary data on the other hand, are those which have already been collected by some other agency and which have already been processed. Secondary data may be available in the form of published or unpublished sources.
- For instance, population census data collected by the Government in a country is primary data for that Government. But the same data becomes secondary for those researchers who use it later. In case you have decided to collect primary data for your investigation, you have to identify the sources from where you can collect that data.
- For example, if you wish to study the problems of the workers of X Company Ltd., then the workers who are working in that company are the source. On the other hand, if you have decided to use secondary data, you have to identify the secondary source that has already collected the related data for their study purpose.
- With the above discussion, we can understand that the difference between primary and secondary data is only in terms of degree. That is that the data which is primary in the hands of one becomes secondary in the hands of another.

5.2.3 Methods and Techniques of Data Collection

- The task of data collection begins after a research problem has been defined and research design/ plan chalked out. While deciding about the method of data collection to be used for the study, the researcher should keep in mind two types of data, viz., primary and secondary.
- The primary data are those which are collected afresh and for the first time, and thus happen to be original in character. The secondary data, on the other hand, are those which have already been collected by someone else and which have already been passed through the statistical process.
- The researcher would have to decide which sort of data he would be using (thus collecting) for his study and accordingly he will have to select one or the other method of data collection. The methods of collecting primary and secondary data differ since primary data are to be originally collected, while in case of secondary data the nature of data collection work is merely that of compilation.

5.2.4 Collection of Primary Data

- Primary data is collected during the course of doing experiments in an experimental research but in case research done of the descriptive type and perform surveys, whether sample surveys or census surveys, then we can obtain primary data either through observation or through direct communication with respondents in one form or another or through personal interviews.*
- This, in other words, means that there are several methods of collecting primary data, particularly in surveys and descriptive researches. Important ones are:
  - observation method
  - interview method
  - through questionnaires
  - through schedules
  - other methods which include

- warranty cards
- distributor
- audits
- pantry audits
- consumer panels
- using mechanical devices
- through projective techniques
- depth interviews
- content analysis

5.2.5 Observation Method

- The observation method is the most commonly used method especially in studies relating to behavioural sciences. In a way we all observe things around us, but this sort of observation is not scientific observation. Observation becomes a scientific tool and the method of data collection for the researcher, when it serves a formulated research purpose, is systematically planned and recorded and is subjected to checks and controls on validity and reliability.
- Under the observation method, the information is sought by way of investigator's own direct observation without asking from the respondent. For instance, in a study relating to consumer behaviour, the investigator instead of asking the brand of wrist watch used by the respondent, may himself look at the watch. The main advantage of this method is that subjective bias is eliminated, if observation is done accurately.
- Secondly, the information obtained under this method relates to what is currently happening; it is not complicated by either the past behaviour or future intentions or attitudes.
- Thirdly, this method is independent of respondents' willingness to respond and as such is relatively less demanding of active cooperation on the part of respondents as happens to be the case in the interview or the questionnaire method. This method is particularly suitable in studies which deal with subjects (i.e., respondents) who are not capable of giving verbal reports of their feelings for one reason or the other.
- However, observation method has various limitations. Firstly, it is an expensive method. Secondly, the information provided by this method is very limited. Thirdly, sometimes unforeseen factors may interfere with the observational task.
- At times, the fact that some people are rarely accessible to direct observation creates obstacle for this method to collect data effectively. While using this method, the researcher should keep in mind things like:
  - What should be observed?
  - How the observations should be recorded? Or
  - How the accuracy of observation can be ensured?
- In case the observation is characterised by a careful definition of the units to be observed, the style of recording the observed information, standardised conditions of observation and the selection of pertinent data of observation, then the observation is called as structured observation. But when observation is to take place without these characteristics to be thought of in advance, the same is termed as unstructured observation.
- Structured observation is considered appropriate in descriptive studies, whereas in an exploratory study the observational procedure is most likely to be relatively unstructured. We often talk about participant and non-participant types of observation in the context of studies, particularly of social sciences.
- This distinction depends upon the observer's sharing or not sharing the life of the group he is observing. If the observer observes by making himself, more or less, a member of the group he is observing so that he can experience what the members of the group experience, the observation is called as the participant observation. But when the observer observes as a detached emissary without any attempt on his part to experience through participation what others feel, the observation of this type is often termed as non-participant observation. (When the observer is observing in such a manner that his presence may be unknown to the people he is observing, such an observation is described as disguised observation.)

- There are several merits of the participant type of observation:
  - The researcher is enabled to record the natural behaviour of the group.
  - The researcher can even gather information which could not easily be obtained if he observes in a disinterested fashion.
  - The researcher can even verify the truth of statements made by informants in the context of a questionnaire or a schedule.
- But there are also certain demerits of this type of observation, viz., the observer may lose the objectivity to the extent he participates emotionally; the problem of observation-control is not solved; and it may narrow-down the researcher's range of experience.
- Sometimes, we talk of controlled and uncontrolled observation. If the observation takes place in the natural setting, it may be termed as uncontrolled observation, but when observation takes place according to definite pre-arranged plans, involving experimental procedure, the same is then termed controlled observation.
- In non-controlled observation, no attempt is made to use precision instruments. The major aim of this type of observation is to get a spontaneous picture of life and persons. It has a tendency to supply naturalness and completeness of behaviour, allowing sufficient time for observing it.
- But in controlled observation, we use mechanical (or precision) instruments as aids to accuracy and standardisation. Such observation has a tendency to supply formalised data upon which generalisations can be built with some degree of assurance.
- The main pitfall of non-controlled observation is that of subjective interpretation. There is also the danger of having the feeling that we know more about the observed phenomena than we actually do. Generally, controlled observation takes place in various experiments that are carried out in a laboratory or under controlled conditions, whereas uncontrolled observation is resorted to in case of exploratory researches.

5.2.6 Interview Method

The interview method of collecting data involves presentation of oral-verbal stimuli and reply in terms of oral-verbal responses. This method can be used through personal interviews and, if possible, through telephone interviews.

Personal interviews

- Personal interview method requires a person known as the interviewer asking questions generally in a face-to-face contact to the other person or persons. (At times the interviewee may also ask certain questions and the interviewer responds to these, but usually the interviewer initiates the interview and collects the information.)
- This sort of interview may be in the form of direct personal investigation or it may be indirect oral investigation. In the case of direct personal investigation the interviewer has to collect the information personally from the sources concerned. He has to be on the spot and has to meet people from whom data have to be collected.
- This method is particularly suitable for intensive investigations. But in certain cases it may not be possible or worthwhile to contact directly the persons concerned or on account of the extensive scope of enquiry, the direct personal investigation technique may not be used. In such cases an indirect oral examination can be conducted under which the interviewer has to cross-examine other persons who are supposed to have knowledge about the problem under investigation and the information, obtained is recorded.
- Most of the commissions and committees appointed by government to carry on investigations make use of this method. The method of collecting information through personal interviews is usually carried out in a structured way. As such we call the interviews as structured interviews. Such interviews involve the use of a set of predetermined questions and of highly standardised techniques of recording.
- Thus, the interviewer in a structured interview follows a rigid procedure laid down, asking questions in a form and order prescribed. As against it, the unstructured interviews are characterised by a flexibility of approach to questioning. Unstructured interviews do not follow a system of pre-determined questions and standardised techniques of recording information.

- In a non-structured interview, the interviewer is allowed much greater freedom to ask, in case of need, supplementary questions or at times he may omit certain questions if the situation so requires. He may even change the sequence of questions. He has relatively greater freedom while recording the responses to include some aspects and exclude others.
- But this sort of flexibility results in lack of comparability of one interview with another and the analysis of unstructured responses becomes much more difficult and time-consuming than that of the structured responses obtained in case of structured interviews.
- Unstructured interviews also demand deep knowledge and greater skill on the part of the interviewer. Unstructured interview, however, happens to be the central technique of collecting information in case of exploratory or formulative research studies. But in case of descriptive studies, we quite often use the technique of structured interview because of its being more economical, providing a safe basis for generalisation and requiring relatively lesser skill on the part of the interviewer.
- We may as well talk about focussed interview, clinical interview and the non-directive interview. Focussed interview is meant to focus attention on the given experience of the respondent and its effects. Under it the interviewer has the freedom to decide the manner and sequence in which the questions would be asked and has also the freedom to explore reasons and motives.
- The main task of the interviewer in case of a focussed interview is to confine the respondent to a discussion of issues with which he seeks conversance. Such interviews are used generally in the development of hypotheses and constitute a major type of unstructured interviews. The clinical interview is concerned with broad underlying feelings or motivations or with the course of individual's life experience.
- The method of eliciting information under it is generally left to the interviewer's discretion. In case of non-directive interview, the interviewer's function is simply to encourage the respondent to talk about the given topic with a bare minimum of direct questioning. The interviewer often acts as a catalyst to a comprehensive expression of the respondents' feelings and beliefs and of the frame of reference within which such feelings and beliefs take on personal significance.
- Despite the variations in interview-techniques, the major advantages and weaknesses of personal interviews can be enumerated in a general way. The chief merits of the interview method are as follows:
  - More information and that too in greater depth can be obtained.
  - Interviewer by his own skill can overcome the resistance, if any, of the respondents; the interview method can be made to yield an almost perfect sample of the general population.
  - There is greater flexibility under this method as the opportunity to restructure questions is always there, especially in case of unstructured interviews.
  - Observation method can as well be applied to recording verbal answers to various questions.
  - Personal information can as well be obtained easily under this method.
  - Samples can be controlled more effectively as there arises no difficulty of the missing returns; non-response generally remains very low.
  - The interviewer can usually control which person(s) will answer the questions. This is not possible in mailed questionnaire approach. If so desired, group discussions may also be held.
  - The interviewer may catch the informant off-guard and thus may secure the most spontaneous reactions than would be the case if mailed questionnaire is used.
  - The language of the interview can be adapted to the ability or educational level of the person interviewed and as such misinterpretations concerning questions can be avoided.
  - The interviewer can collect supplementary information about the respondent's personal characteristics and environment which is often of great value in interpreting results. But there are also certain weaknesses of the interview method. Among the important weaknesses, mention may be made of the following:
    - It is a very expensive method, especially when large and widely spread geographical sample is taken.
    - There remains the possibility of the bias of interviewer as well as that of the respondent; there also remains the headache of supervision and control of interviewers.

- Certain types of respondents such as important officials or executives or people in high income groups may not be easily approachable under this method and to that extent the data may prove inadequate.
- This method is relatively more-time-consuming, specially when the sample is large and recalls upon the respondents are necessary.
- The presence of the interviewer on the spot may over-stimulate the respondent, sometimes even to the extent that he may give imaginary information just to make the interview interesting.
- Under the interview method the organisation required for selecting, training and supervising the field-staff is more complex with formidable problems.
- Interviewing at times may also introduce systematic errors.
- Effective interview presupposes proper rapport with respondents that would facilitate free and frank responses. This is often a very difficult requirement.

Pre-requisites and basic tenets of interviewing:

- For successful implementation of the interview method, interviewers should be carefully selected, trained and briefed. They should be honest, sincere, hardworking, and impartial and must possess the technical competence and necessary practical experience.
- Occasional field checks should be made to ensure that interviewers are neither cheating, nor deviating from instructions given to them for performing their job efficiently. In addition, some provision should also be made in advance so that appropriate action may be taken if some of the selected respondents refuse to cooperate or are not available when an interviewer calls upon them.
- In fact, interviewing is an art governed by certain scientific principles. Every effort should be made to create friendly atmosphere of trust and confidence, so that respondents may feel at ease while talking to and discussing with the interviewer. The interviewer must ask questions properly and intelligently and must record the responses accurately and completely. At the same time, the interviewer must answer legitimate question(s), if any, asked by the respondent and must clear any doubt that the latter has.
- The interviewers approach must be friendly, courteous, conversational and unbiased. The interviewer should not show surprise or disapproval of a respondent's answer but he must keep the direction of interview in his own hand, discouraging irrelevant conversation and must make all possible effort to keep the respondent on the track.

Telephone Interviews

This method of collecting information consists in contacting respondents on telephone itself. It is not a very widely used method, but plays important part in industrial surveys, particularly in developed regions. The chief merits of such a system are:

- It is more flexible in comparison to mailing method.
- It is faster than other methods i.e., a quick way of obtaining information.
- It is cheaper than personal interviewing method; here the cost per response is relatively low.
- Recall is easy; call backs are simple and economical.
- There is a higher rate of response than what we have in mailing method; the non-response is generally very low.
- Replies can be recorded without causing embarrassment to respondents.
- Interviewer can explain requirements more easily.
- At times, access can be gained to respondents who otherwise cannot be contacted for one reason or the other.
- No field staffs are required.
- Representative and wider distribution of sample is possible.

But this system of collecting information is not free from demerits. Some of these may be highlighted.

- Little time is given to respondents for considered answers; interview period is not likely to exceed five minutes in most cases.
- Surveys are restricted to respondents who have telephone facilities.
- Extensive geographical coverage may get restricted by cost considerations.
- It is not suitable for intensive surveys where comprehensive answers are required to various questions.
- Possibility of the bias of the interviewer is relatively more.
- Questions have to be short and to the point; probes are difficult to handle.

5.3 Collection of Data through Questionnaires

This method of data collection is quite popular, particularly in case of big enquiries. It is being adopted by private individuals, research workers, private and public organisations and even by governments. In this method a questionnaire is sent (usually by post) to the persons concerned with a request to answer the questions and return the questionnaire. A questionnaire consists of a number of questions printed or typed in a definite order on a form or set of forms. The questionnaire is mailed to respondents who are expected to read and understand the questions and write down the reply in the space meant for the purpose in the questionnaire itself. The respondents have to answer the questions on their own.

5.3.1 Merits and Demerits of a Questionnaire

The method of collecting data by mailing the questionnaires to respondents is most extensively employed in various economic and business surveys. The merits claimed on behalf of this method are as follows:

- There is low cost even when the universe is large and is widely spread geographically.
- It is free from the bias of the interviewer; answers are in respondents' own words.
- Respondents have adequate time to give well thought out answers.
- Respondents, who are not easily approachable, can also be reached conveniently.
- Large samples can be made use of and thus the results can be made more dependable and reliable.

The main demerits of this system can also be listed here:

- Low rate of return of the duly filled in questionnaires; bias due to no-response is often indeterminate.
- It can be used only when respondents are educated and cooperating.
- The control over questionnaire may be lost once it is sent.
- There is inbuilt inflexibility because of the difficulty of amending the approach once questionnaires have been despatched.
- There is also the possibility of ambiguous replies or omission of replies altogether to certain questions; interpretation of omissions is difficult.
- It is difficult to know whether willing respondents are truly representative.
- This method is likely to be the slowest of all.

Before using this method, it is always advisable to conduct 'pilot study' (Pilot Survey) for testing the questionnaires. In a big enquiry the significance of pilot survey is felt very much. Pilot survey is in fact the replica and rehearsal of the main survey. Such a survey, being conducted by experts, brings to the light the weaknesses (if any) of the questionnaires and also of the survey techniques. From the experience gained in this way, improvement can be effected.

5.3.2 Main Aspects of a Questionnaire

Quite often questionnaire is considered as the heart of a survey operation. Hence it should be very carefully constructed. If it is not properly set up, then the survey is bound to fail. This fact requires us to study the main aspects of a questionnaire, viz., the general form, question sequence and question formulation and wording. Researcher should note the following with regard to these three main aspects of a questionnaire:

General form:

- So far as the general form of a questionnaire is concerned, it can either be structured or unstructured questionnaire. Structured questionnaires are those questionnaires in which there are definite, concrete and pre-determined questions. The questions are presented with exactly the same wording and in the same order to all respondents.
- Resort is taken to this sort of standardisation to ensure that all respondents reply to the same set of questions. The form of the question may be either closed (i.e., of the type 'yes' or 'no') or open (i.e., inviting free response) but should be stated in advance and not constructed during questioning. Structured questionnaires may also have fixed alternative questions in which responses of the informants are limited to the stated alternatives.
- Thus a highly structured questionnaire is one in which all questions and answers are specified and comments in the respondent's own words are held to the minimum. When these characteristics are not present in a questionnaire, it can be termed as unstructured or non-structured questionnaire.
- More specifically, we can say that in an unstructured questionnaire, the interviewer is provided with a general guide on the type of information to be obtained, but the exact question formulation is largely his own responsibility and the replies are to be taken down in the respondent's own words to the extent possible; in some situations tape recorders may be used to achieve this goal.
- Structured questionnaires are simple to administer and relatively inexpensive to analyse. The provision of alternative replies, at times, helps to understand the meaning of the question clearly. But such questionnaires have limitations too.
- For instance, wide range of data and that too in respondent's own words cannot be obtained with structured questionnaires. They are usually considered inappropriate in investigations where the aim happens to be to probe for attitudes and reasons for certain actions or feelings. They are equally not suitable when a problem is being first explored and working hypotheses sought. In such situations, unstructured questionnaires may be used effectively.
- Then on the basis of the results obtained in pre-test (testing before final use) operations from the use of unstructured questionnaires, one can construct a structured questionnaire for use in the main study.

Question sequence:

- In order to make the questionnaire effective and to ensure quality to the replies received, a researcher should pay attention to the question-sequence in preparing the questionnaire. A proper sequence of questions reduces considerably the chances of individual questions being misunderstood. The question-sequence must be clear and smoothly-moving, meaning thereby that the relation of one question to another should be readily apparent to the respondent, with questions that are easiest to answer being put in the beginning.
- The first few questions are particularly important because they are likely to influence the attitude of the respondent and in seeking his desired cooperation. The opening questions should be such as to arouse human interest. The following type of questions should generally be avoided as opening questions in a questionnaire:
  - questions that put too great a strain on the memory or intellect of the respondent;
  - questions of a personal character;
  - questions related to personal wealth, etc.
- Following the opening questions, we should have questions that are really vital to the research problem and a connecting thread should run through successive questions. Ideally, the question sequence should conform to the respondent's way of thinking. Knowing what information is desired, the researcher can rearrange the order of the questions (this is possible in case of unstructured questionnaire) to fit the discussion in each particular case.

- But in a structured questionnaire the best that can be done is to determine the question-sequence with the help of a Pilot Survey which is likely to produce good rapport with most respondents. Relatively difficult questions must be relegated towards the end so that even if the respondent decides not to answer such questions, considerable information would have already been obtained.
- Thus, question-sequence should usually go from the general to the more specific and the researcher must always remember that the answer to a given question is a function not only of the question itself, but of all previous questions as well. For instance, if one question deals with the price usually paid for coffee and the next with reason for preferring that particular brand, the answer to this latter question may be couched largely in terms of price differences.

Question formulation and wording:

- With regard to this aspect of questionnaire, the researcher should note that each question must be very clear for any sort of misunderstanding can do irreparable harm to a survey. Question should also be impartial in order not to give a biased picture of the true state of affairs. Questions should be constructed with a view to their forming a logical part of a well thought out tabulation plan. In general, all questions should meet the following standards:
  - Should be easily understood
  - should be simple, i.e., should convey only one thought at a time
  - should be concrete and should conform as much as possible to the respondent's way of thinking. (For instance, instead of asking. "How many razor blades do you use annually?" The more realistic question would be to ask, "How many razor blades did you use last week?"
- Concerning the form of questions, we can talk about two principal forms, viz., multiple choice questions and the open-end question. In the former the respondent selects one of the alternative possible answers put to him, whereas in the latter he has to supply the answer in his own words. The question with only two possible answers (usually 'Yes' or 'No') can be taken as a special case of the multiple choice question, or can be named as a 'closed question.'
- There are some advantages and disadvantages of each possible form of question. Multiple choice or closed questions have the advantages of easy handling, simple to answer, quick and relatively inexpensive to analyse. They are most amenable to statistical analysis. Sometimes, the provision of alternative replies helps to make clear the meaning of the question.
- But the main drawback of fixed alternative questions is that of "putting answers in people's mouths" i.e., they may force a statement of opinion on an issue about which the respondent does not in fact have any opinion. They are not appropriate when the issue under consideration happens to be a complex one and also when the interest of the researcher is in the exploration of a process.
- In such situations, open-ended questions which are designed to permit a free response from the respondent rather than one limited to certain stated alternatives are considered appropriate. Such questions give the respondent considerable latitude in phrasing a reply. Getting the replies in respondent's own words is, thus, the major advantage of open-ended questions.
- But one should not forget that, from an analytical point of view, open-ended questions are more difficult to handle, raising problems of interpretation, comparability and interviewer bias.*
- In practice, one rarely comes across a case when one questionnaire relies on one form of questions alone. The various forms complement each other. As such questions of different forms are included in one single questionnaire. For instance, multiple-choice questions constitute the basis of a structured questionnaire, particularly in a mail survey. But even there, various open-ended questions are generally inserted to provide a more complete picture of the respondent's feelings and attitudes.

- Researcher must pay proper attention to the wordings of questions since reliable and meaningful returns depend on it to a large extent. Since words are likely to affect responses, they should be properly chosen. Simple words, which are familiar to all respondents, should be employed. Words with ambiguous meanings must be avoided. Similarly, danger words, catch-words or words with emotional connotations should be avoided. Caution must also be exercised in the use of phrases which reflect upon the prestige of the respondent. Question wording, in no case, should bias the answer. In fact, question wording and formulation is an art and can only be learnt by practice.

5.3.3 Essentials of a Good Questionnaire

- To be successful, questionnaire should be comparatively short and simple i.e., the size of the questionnaire should be kept to the minimum. Questions should proceed in logical sequence moving from easy to more difficult questions. Personal and intimate questions should be left to the end.
- Technical terms and vague expressions capable of different interpretations should be avoided in a questionnaire. Questions may be dichotomous (yes or no answers), multiple choice (alternative answers listed) or open-ended.
- The latter type of questions is often difficult to analyse and hence should be avoided in a questionnaire to the extent possible. There should be some control questions in the questionnaire which indicate the reliability of the respondent.
- For instance, a question designed to determine the consumption of particular material may be asked first in terms of financial expenditure and later in terms of weight. The control questions, thus, introduce a cross-check to see whether the information collected is correct or not. Questions affecting the sentiments of respondents should be avoided. Adequate space for answers should be provided in the questionnaire to help editing and tabulation. There should always be provision for indications of uncertainty, example, "do not know," "no preference" and so on. Brief directions with regard to filling up the questionnaire should invariably be given in the questionnaire itself.
- Finally, the physical appearance of the questionnaire affects the cooperation the researcher receives from the recipients and as such an attractive looking questionnaire, particularly in mail surveys, is a plus point for enlisting cooperation. The quality of the paper, along with its colour, must be good so that it may attract the attention of recipients.

5.3.4 Functions of Questionnaire

A questionnaire serves five functions:

- Give the respondent clear comprehension of the question.
- Induce the respondent to cooperate and to trust that answers will be treated confidentially.
- Stimulate response through greater introspection, plumbing of memory or reference to records.
- Give instructions on what is wanted and the manner of responding.
- Identify what needs to be known to classify and verify and interview.

5.3.5 Problem of Response Questionnaires

In case of questionnaire, the greatest problem is that of response. In an interview the field worker is personally present and very few people can refuse to grant interview at the face of the interviewer. In case of questionnaire, this facility is not present and the researcher has to rely on other means to get a good response.

Factors affecting response: A number of factors are responsible for varying degree of response. The chief of them are as follows:

- Special Characteristics of the Groups: the response generally differs according to special characteristics of the respondent group. These special characteristics are sex, age, economic status, educational level etc.

- Prestige Sponsoring Groups: if the institution carrying on research has a good reputation and credit, the percentage of responses is very high. People generally do not care to reply if the sponsoring authority is unknown, inconspicuous or a doubtful integrity, and the response is likely to be low.
- Importance of the problem under study: If the problem under study is important and significant a higher response is expected. If it is only of a general nature the response is likely to be low. A higher response is expected specially, if the problem concerns the respondents directly.
- Nature of questionnaire: If the questionnaire is short and simple its form is attractive and layout good, a higher response may be expected. If on the other hand, the questionnaire is clumsy. Long or difficult the respondents are, most likely to be indifferent and apathetic. Therefore, in order to get a higher response, proper attention should be paid towards form and layout of the questionnaire.
- Nature of reaction: according to Katz and Cantril, a very high percentage of respondents are those who have a strong feeling either in favour or against it. If the respondents are strongly in favour or against the problem, they are most likely to respond. If on the other hand, they do not possess any strong feelings and are more or less indifferent about the problem, they are expected not to reply.
- Size of questionnaire: The size of the questionnaire also is an important factor in determining reason. Small size questionnaires generally have a better response than large size ones. But this advantage is available only when the size of the question s sufficiently small. After that small size limit is crossed, any further increase in pages has no effect upon the response.
- Sequence of the questions: According to Frazen, "the length of the questionnaire does not affect the response so much as the number of unrelated subjects dealt with." Thus, if the questions are not in one sequence and appear as broken subjects the response may be much lower than if they are in sequential order. Sletto's experiments also prove this fact. When complete questionnaire of 35 pages was sent it had a better response than 25 pages incomplete questionnaire.

5.3.6 Pilot testing or Pre-testing the Questionnaire

Before finalising the questionnaire, it is desirable to carry out a preliminary experiment on a sample basis. The investigator should examine each question to ensure that the question is not confusing, leading to biased responses etc. The real test of a questionnaire is how it performs under actual conditions of data collection. This test can be carried out among small groups of subjects in order to provide an estimate of the time needed for responding to the survey. The questionnaire pre-test serves the same role in questionnaire design as testing a new product in the market. As test marketing provides the real test of customer reactions to the product and the accompanying marketing programmes, in the same way, the pre-test provides the real test of the questionnaire. Therefore this work must be done with utmost care and caution to yield good results.

5.3.7 Specimen Questionnaire

The following specimen questionnaire incorporates most of the qualities which we have discussed above. It relates to 'Computer User Survey'.

Computer User Survey

1. What brand of Computer do you primarily use?
   a. IBM
   b. Compaq
   c. HCL
   d. Dell
   e. Siemens
   f. Any other
   g. _____ (please specify)

2. Where was the computer purchased?
   a. Computer store
   b. Mail order
   c. Manufacturer
   d. Company Dealer
   e. Any other _____

3. How long have you been using computers?
   _____years _____months.

4. In a week about how many hours do you spend on the computer? _____

5. Which database management package do you use most often?
   a. Dbase-II
   b. Dbase-III
   c. Lotus 1,2,3
   d. MS-Excel
   e. Oracle
   f. Any other _____ (please specify)

6. Does the computer, that you primarily use, have a hard disk?
   a. Yes
   b. No

7. Where did you obtain the software that you use?
   a. Computer user group
   b. Regular dealer
   c. Mail order
   d. Directly from Software
   e. Any other _____ dealer

8. 8. On the following 9-point scale, rate the degree of difficulty that you have encountered in using the computer.
   Extremely difficult 1 2 3 4 5 6 7 8 9 Not difficult

9. If you have to purchase a personal computer today, which one would, you be most likely to purchase?
   a. IBM
   b. Compaq
   c. HCL
   d. Dell
   e. Siemens
   f. Any other _____ (please specify)

10. What is your gender?
    a. Male
    b. Female

11. Please state your date of birth _____. _____ _____

12. Your Qualifications
    a. Secondary
    b. Sr. Secondary
    c. Graduate
    d. Post-graduate
    e. Doctorate
    f. Any other _____ (please specify)

13. Which of the following best describes your primary field of employment?
    a. Medical
    b. Education
    c. Business
    d. Government
    e. Technical
    f. Any other _____ (please specify)

14. What is your current salary?

5.4 Collection of Data through Schedules

The schedule is the form containing some questions or blank which are to be filled by the workers after getting information from the informants. The schedule may, thus, contain two types of questions firstly those that are in form of direct questions and secondly those that are in form of a table. There are some kinds of information that can be procured only by putting a question, for example, questions for eliciting the informant's opinion, attitude preferences of his suggestion about some matter. There are some kinds of information that can be procured only by putting a question, for example, questions for eliciting the informant's opinion, attitude preferences of his suggestions about some matter. There are others which may better be put in form of tables. Generally, most of them are inter changeable.

5.4.1 Purpose of Schedule

To attain objectivity: The purpose of schedule is to provide a standardised tool for observation or interview in order to attain objectivity. By schedule every informant has to reply the same questions put in the same language and the field worker ha no choice o get the desired reply by putting the different question or changing the language of the same question. The order of the question is also the same and thus the whole interview takes place under standardised conditions, and the data received is perfectly comparable. It has not been proved that responses of the people regarding same matter differ, if the language of the question is even slightly different or evens the place of question is changed. A perfectly standardised form is, therefore, needed for any objective study.

To Act as Memory: The other purpose of schedule is to act as memory tickler. In the absence of any schedule the field worker may put different number of questions to different people. He may forget to inquire about some important aspects and then he would be required to go over the whole process again to collect that missing information too, the schedule keeps his memory refreshed, and keeps him reminded of the different aspects that are to be particularly observed.

5.4.2 Essentials of a Good Schedule

According to P.V.Young, there are two essential conditions of a good schedule.

- **Accurate communication**: Accurate communication is achieved when respondent's communication is achieved, when respondents understand the questions in the same sense which they are expected to convey. Thus the basis of accurate communications is proper wording of questions. The questions should be so worded that they may clearly carry the desired sense without any ambiguity. Various tools for achieving accuracy of communication have been discussed in the succeeding paragraphs.
- **Accurate response:** Accurate response is said to have been achieved when replies contain the information sought for. Thus information should be unbiased and true. The respondents may also co-operate in filling the schedule, by giving out correct information.

5.4.3 Difference between Questionnaires and Schedules

Both questionnaire and schedule are popularly used methods of collecting data in research surveys. There is much resemblance in the nature of these two methods and this fact has made many people to remark that from a practical point of view, the two methods can be taken to be the same. But from the technical point of view there is difference between the two. The important points of difference are as under:

- The questionnaire is generally sent through mail to informants to be answered as specified in a covering letter, but otherwise without further assistance from the sender. The schedule is generally filled out by the research worker or the enumerator, who can interpret questions when necessary.
- To collect data through questionnaire is relatively cheap and economical since we have to spend money only in preparing the questionnaire and in mailing the same to respondents. Here no field staff required. To collect data through schedules is relatively more expensive since considerable amount of money has to be spent in appointing enumerators and in importing training to them. Money is also spent in preparing schedules.
- Non-response is usually high in case of questionnaire as many people do not respond and many return the questionnaire without answering all questions. Bias due to non-response often remains indeterminate. As against this, non-response is generally very low in case of schedules because these are filled by enumerators who are able to get answers to all questions. But there remains the danger of interviewer bias and cheating.
- In case of questionnaire, it is not always clear as to who replies, but in case of schedule the identity of respondent is known.
- The questionnaire method is likely to be very slow since many respondents do not return the questionnaire in time despite several reminders, but in case of schedules the information is collected well in time as they are filled in by enumerators.
- Personal contact is generally not possible in case of the questionnaire method as questionnaires are sent to respondents by post who also in turn returns the same by post. But in case of schedules direct personal contact is established with respondents.
- Questionnaire method can be used only when respondents are literate and cooperative, but in case of schedules the information can be gathered even when the respondents happen to be illiterate.
- Wider and more representative distribution of sample is possible under the questionnaire method, but in respect of schedules there usually remains the difficulty in sending enumerators over a relatively wider area.
- Risk of collecting incomplete and wrong information is relatively more under the questionnaire method, particularly when people are unable to understand questions properly. But in case of schedules, the information collected is generally complete and accurate as enumerators can remove the difficulties, if any, faced by respondents in correctly understanding the questions. As a result, the information collected through schedules is relatively more accurate than that obtained through questionnaires.
- The success of questionnaire method lies more on the quality of the questionnaire itself, but in the case of schedules much depends upon the honesty and competence of enumerators.
- In order to attract the attention of respondents, the physical appearance of questionnaire must be quite attractive, but this may not be so in case of schedules as they are to be filled in by enumerators and not by respondents.
- Along with schedules, observation method can also be used but such a thing is not possible while collecting data through questionnaires.

5.5 Collection of Data through Interview

Interview may be regarded as a systematic a systematic method by which a person enters more or less imaginatively into the life of a comparative stranger. It is like a tape recorder in which the past incidences, feelings and reactions of the subject are played back to the living present to be listened by the interviewer with a scientific approach, always ready to find sequences of fundamental trials of human behavior, underlying universal laws guiding and motivating human actions and reaction. The researcher tries to penetrate deeply in his imagination, into the circumstances being narrated by the subject and realise the full significance of the feelings being expressed by him.

5.5.1 Purpose of Interview

The purpose of interview is two fold:

- Securing certain information from the subject which is, known only to him self and cannot be gathered from any source. Here, purpose the interviewer tells the topic to the subject and he narrates the incidents of his life, his feelings and reactions pertaining to it. The interviewer listens to these narratives, with cautions and alert, trying to find how far they fit in the hypothesis framed by him.
- Laboratory study of verbal behavior under given circumstances. In this, the interest of the researcher is more centerd in the attitude and expressions of the subject than the actual facts. Here he acts more as a social psychologist than as a sociologist. His attention is more focused at the reactions and changes in the moods and gestures of the subject while he is relating to a particular incident.

5.5.2 Kinds of Interviews

Interviews are generally of the following types:

Structured Interview: It is also known as controlled, guided or direct interview. In this kind of interview a complete schedule is used. The interviewer is asked to get the answers to those questions only. He generally does not add any thing from his own side. The language too is not changed. He can only interpret or amplify the statement wherever necessary.

Unstructured Interview: It is known as uncontrolled, unguided or undirected interview. No direct or predetermined questions used are in this type of interview. The field worker may be told certain broad topics upon which the information is to be collected. It is generally held in the form of free discussion or story type narrative. The subject is asked to narrate the incidents of his life, his own feeling and reactions and the researcher has to draw his own conclusions from it.

Focused Interview: As the name suggests, its main focus is on social and psychological effects of mass communication, for example, the reactions of a film show or radio programmes. Merton has given the following characteristics of a focussed interview. Interviewee is known to have been involved in a particular concrete situation. The situation under study is the one that has already been analysed prior to interviewer. The field worker tries to focus his attention to the particulars aspect of the problem, and tries to know his experiences, attitudes and emotional response regarding the concrete situation under study.

Repetitive Interview: Interview is repetitive in nature when it is desired to note the gradual influence of some social or psychological process. There are some social changes that have a far reaching influence upon the people and it is sometimes desired to know the effect of such factors in time sequence. Thus, for example, suppose a village is linked with some road connecting it with the city. Naturally, it will have its own influence upon the life of the people. The influence of course would not be sudden. There will be a gradual change in economic status, standard of living, attitudes and opinions, inter-relationships of the people. In order to study this influence in time sequence a study has to be conducted at regular intervals to mark the gradual change taking place.

5.5.3 Technique of Interview

- The general qualities that an interviewer should possess can be identified as extroversion, impartiality to different points of view, friendliness without getting personal, ability to remain calm under many different circumstances, conscientiousness, perseverance, accuracy of reporting and adaptability.
- In addition, the interviewer must plan in advance and should fully know the problem under consideration. He must choose a suitable time and place so that the interviewee may feel at ease during the interview. In order to establish proper rapport especially in the rural areas if a significant part of the respondents are women, female interviewer' must be employed to obtain the responses.
- If the language of the questionnaire is English, translation in the regional languages must be available so that the interviewer has no ambiguity about the questions. While collecting data regarding various industrial problems, the interviewer should proceed very cautiously and systematically as the success of his project would depend entirely on the co-operation of the others.
- The interviewer is also expected to be familiar with all the material already published by the organisation. In an administrative hierarchy, it is preferable to begin interviews by contacting the officials at the middle and lower levels and to collect all relevant material available with them. The selection, training and supervision of interviewers are very important administrative functions for collection of reliable data.

5.6 Collection of Data through Observation

Observation is probably the oldest method used by man in scientific investigation. In the physical sciences, the observation is the only method for new experiment, although guided and controlled observation has replaced the ancient form of unaided visual observation.

5.6.1 Kinds of Observation

Observation can be classified into various types according to the method used and the type of control exercised. Following are the chief types of observation.

Participant observation: The observation may be participant or non participant. When the observer participates with the activities of the group under study, it is known as participant observation. Thus, a participant observer makes himself part of the group under study. He need not necessarily carry out all the activities as carried out by other members of the group, but, his presence as an active member of the group is necessary.

Non – participant observation: When the observer does not actually participate in the activities of the group, but simply observes them from a distance, it is, known as a non participant observation. Purely non participant observation is extremely difficult.

Non controlled observation: When the observation is made in the natural surroundings and the activities are performed in their usual course without being influenced or guided by any external force it is known as non controlled observation. Non controlled observation is generally not very reliable. The observation itself may be biased and coloured by the views of the observer, because there is no check upon him. Various observers may observe the same thing differently and draw different conclusions. The greatest difficulty is that the observer may be so overpowered by uncontrolled and stray events that he may regard them to be absolutely true while they are far from being so.

Controlled observation: Controlled observation afford grater precision and objectivity and can be repeatedly observed under identical conditions. The main purpose of a controlled observation is, thus, to check any bias due to faulty perception, inaccurate data and influence of outside factors on the particular incident. The control exercised in the observation is of two types:

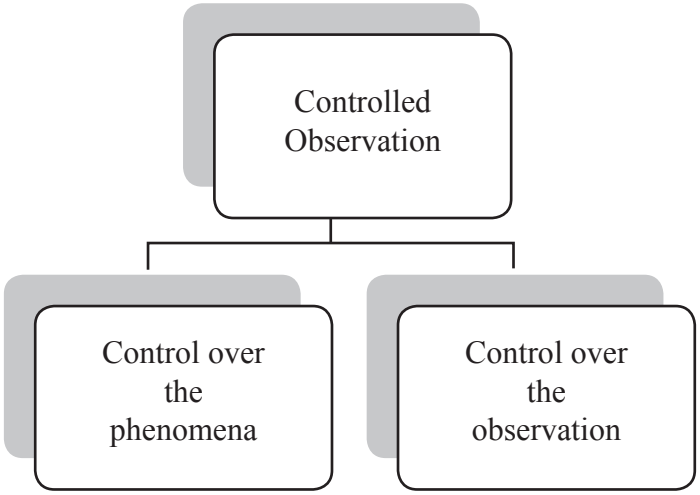


Fig. 5.2 Types of control exercised in the observation

5.6.2 Importance of Observation in Social Research

Observation as method of social research has been used quite extensively everywhere. Following are the chief advantages of the method.

Simplest method: Observation is the simplest and the most non technical method. Observation is also perfected with the theoretical knowledge of the phenomena which can be gained more easily.

Useful in framing hypothesis: The social researcher observes various activities of the people around him. Continuous observation even if causal and unplanned may reveal certain sequences and trends. These trends and sequences generally form the basis of new hypothesis. He must read and at the same time observe the activities of the people around him, see how far they agree with the theoretical principals and in what way do they differ.

Grater accuracy: Observation at times affords grater accuracy than other methods. In case of interview or case history the researcher has to depend on the information supplied by other people. In case of observation the researcher need not depend upon information supplied by other people blind folded. He can see if for himself. The data collected through observation is thus generally more valid and reliable, than data collected through observation, than data collected either through interview or case history.

More convincing results: When the information has been supplied by others, there is always some doubt in the mind of the people about its validity. Such data may, therefore, lack the force of conviction. In case of observation the researcher has observed thing for himself and no doubt about their correctness.

Test of validity: Even if the researcher himself is fully convinced of the hypothesis, he has sometimes to prove its validity to others. This can be done through observation. He can let people observe the same things and this would convince them of the validity of the statement. Researcher himself also has to observe phenomena from time to see whether his generalisations hold good.

A common method for all sciences: Observation is a common method for all sciences. Interview is the method that is used specially in case of social sciences, but observation is common to all the sciences, and naturally has greater universality of practice. The rules of observation are the same for all sciences; and accepted physical or social. A common method is most commonly followed and accepted as a tool.

5.6.3 Limitations of Observation as a Method of Social Research

Though observation has been used widely in social research; it has its own limitations. Not all occurrences are open to observation, can be observed when the observer is at hand, not all occurrences lend themselves to study by observational techniques. The first draw back of observation is its uncontrolled nature. A systematic and pre planned observation is very difficult. There are three types of difficulties in the use of observation.

Some of the occurrences may not be open to observation: Some of the people concerned do not allow the observer to note the activities. Most of the personal activities are thus not open to actual observation by an outsider.

Not all occurrences open to observation can be observed when observer is at hand: This type of incapacity is caused by uncertainty of the event. Many social events are most uncertain, so that the time and place of their occurrences can not be pre planned. The event may take place when the observer is not present, on the other hand even when the observer is constantly present, the event may take place.

Not all occurrences lend themselves to study observational techniques: Most of the social phenomena to be studied are abstract in nature, example, feelings, reactions etc. They are not open to our senses and hence can not be subjected to study by observational techniques. For the study of such phenomena we have to adopt other techniques like case study, interview etc. the other limitation of the use of observation is caused by the lack of validity of results. Even when the observational study has been made and data thus collected one problem remains to be solved. How far the data collected is valid to the problem under study. How far the information thus gathered is reliable.

5.7 Modern Methods of Data Collection

Radio Buttons: These are small circles that are placed next to response options of close ended questions. By default, only one radio button within any given group of radio buttons can be selected at a time. It can be used in case of mutually exclusive options. They closely resemble the paper based questionnaire answer formats. It demands a relatively high degree of mouse precision and users with limited computer exposure may find it frustrating to click the options or to change the options.

Check Boxes: These are typically small squares that contain a tick mark when checked and allow multiple options rather than exclusive options. They also require high degree of mouse precision. The advantage of using check boxes and radio buttons within the same questionnaire is that their appearance is visibly different and so, respondents are given visual clues as to how to answer any question using either of the two response formats.

5.8 Other Methods of Data Collection

Let us consider some other methods of data collection:

Warranty cards:

- Warranty cards are usually postal sized cards which are used by dealers of consumer durables to collect information regarding their products. The information sought is printed in the form of questions on the 'warranty cards' which is placed inside the package along with the product with a request to the consumer to fill in the card and post it back to the dealer.

Distributor or store audits:

- Distributor or store audits are performed by distributors as well as manufactures through their salesmen at regular intervals. Distributors get the retail stores audited through salesmen and use such information to estimate market size, market share, and seasonal purchasing pattern and so on.
- The data are obtained in such audits not by questioning but by observation. The principal advantage of this method is that it offers the most efficient way of evaluating the effect on sales of variations of different techniques of in-store promotion.

Pantry audits:

- In this type of audit, the investigator collects an inventory of types, quantities and prices of commodities consumed. Thus in pantry audit data are recorded from the examination of consumer's pantry. The usual objective in a pantry audit is to find out what types of consumers buy certain products and certain brands, the assumption being that the contents of the pantry accurately portray consumer's preferences.
- Quite often, pantry audits are supplemented by direct questioning relating to reasons and circumstances under which particular products were purchased in an attempt to relate these factors to purchasing habits. An important limitation of pantry audit approach is that, at times, it may not be possible to identify consumers' preferences from the audit data alone, particularly when promotion devices produce a marked rise in sales.

Consumer panels:

- An extension of the pantry audit approach on a regular basis is known as 'consumer panel', where a set of consumers are arranged to come to an understanding to maintain detailed daily records of their consumption and the same is made available to investigator on demands. Mostly consume panels are of two types, viz., the transitory consumer panel and the continuing consumer panel.
  - A transitory consumer panel is set up to measure the effect of a particular phenomenon. Usually such a panel is conducted on a before-and-after-basis. Initial interviews are conducted before the phenomenon takes place to record the attitude of the consumer.
  - A second set of interviews is carried out after the phenomenon has taken place to find out the consequent changes that might have occurred in the consumer's attitude. It is a favourite tool of advertising and of social research.
- A continuing consumer panel is often set up for an indefinite period with a view to collect data on a particular aspect of consumer behaviour over time, generally at periodic intervals or may be meant to serve as a general purpose panel for researchers on a variety of subjects. Such panels have been used in the area of consumer expenditure, public opinion and radio and TV listenership among others.
- Most of these panels operate by mail. The representativeness of the panel relative to the population and the effect of panel membership on the information obtained after the two major problems associated with the use of this method of data collection.

Use of mechanical devices: The use of mechanical devices has been widely made to collect information by way of indirect means. Eye camera, Pupil metric camera, Psycho galvanometer, Motion picture camera and Audiometer are the principal devices so far developed and commonly used by modern big business houses, mostly in the developed world for the purpose of collecting the required information.

- Eye cameras are designed to record the focus of eyes of a respondent on a specific portion of a sketch or diagram or written material. Such information is useful in designing advertising material.
- Pupil metric cameras record dilation of the pupil as a result of a visual stimulus. The extent of dilation shows the degree of interest aroused by the stimulus.
- Psycho galvanometer is used for measuring the extent of body excitement as a result of the visual stimulus.
- Motion picture cameras can be used to record movement of body of a buyer while deciding to buy a consumer good from a shop or big store.
- Audiometers are used by some TV concerns to find out the type of programmes as well as stations preferred by people. A device is fitted in the television instrument itself to record these changes. Such data may be used to find out the market share of competing television stations.

Projective techniques:

- Projective techniques (or what are sometimes called as indirect interviewing techniques) for the collection of data have been developed by psychologists to use projections of respondents for inferring about underlying motives, urges, or intentions which are such that the respondent either resists to reveal them or is unable to figure out himself.

- In projective techniques the respondent in supplying information tends unconsciously to project his own attitudes or feelings on the subject under study. Projective techniques play an important role in motivational researches or in attitude surveys.
- The use of these techniques requires intensive specialised training. In such techniques, the individual's responses to the stimulus-situation are not taken at their face value. The stimuli may arouse many different kinds of reactions. The nature of the stimuli and the way in which they are presented under these techniques do not clearly indicate the way in which the response is to be interpreted. The stimulus may be a photograph, a picture, an inkblot and so on.
- Responses to these stimuli are interpreted as indicating the individual's own view, his personality structure, his needs, tensions, etc. in the context of some pre-established psychological conceptualisation of what the individual's responses to the stimulus mean.

Depth interviews:

- Depth interviews are those interviews that are designed to discover underlying motives and desires and are often used in motivational research. Such interviews are held to explore needs, desires and feelings of respondents. Unless the researcher has specialised training, depth interviewing should not be attempted. Depth interview may be projective in nature or it may be a non-projective interview.
- The difference lies in the nature of the questions asked. Indirect questions on seemingly irrelevant subjects provide information that can be related to the informant's behaviour or attitude towards the subject under study.

Content-analysis:

- Content-analysis consists of analysing the contents of documentary materials such as books, magazines, newspapers and the contents of all other verbal materials which can be either spoken or printed. The analysis of content is a central activity whenever one is concerned with the study of the nature of the verbal materials.
- A review of research in any area, for instance, involves the analysis of the contents of research articles that have been published. The analysis may be at a relatively simple level or may be a subtle one. It is at a simple level when we pursue it on the basis of certain characteristics of the document or verbal materials that can be identified and counted (such as on the basis of major scientific concepts in a book). It is at a subtle level when researcher makes a study of the attitude, say of the press towards education by feature writers.

5.9 Collection of Secondary Data

Secondary data means data that are already available i.e., they refer to the data which have already been collected and analysed by someone else. When the researcher utilises secondary data, then he has to look into various sources from where he can obtain them. In this case he is certainly not confronted with the problems that are usually associated with the collection of original data. Secondary data may either be published data or unpublished data. Usually published data are available in:

- various publications of the central, state are local governments
- various publications of foreign governments or of international bodies and their subsidiary organisations
- technical and trade journals
- books, magazines and newspapers
- reports and publications of various associations connected with business and industry, banks, stock exchanges, etc.
- reports prepared by research scholars, universities, economists, etc. in different fields
- Public records and statistics, historical documents, and other sources of published information

Researcher must be very careful in using secondary data. He must make a minute scrutiny because it is just possible that the secondary data may be unsuitable or may be inadequate in the context of the problem which the researcher wants to study. By way of caution, the researcher, before using secondary data, must see that they possess following characteristics:

5.9.1 Reliability of Data

The reliability can be tested by finding out such things about the said data:

- Who collected the data?
- What were the sources of data?
- Were proper methods adopted to collect the data?
- At what time were they collected?
- Was there any bias involved?
- What level of accuracy was desired? Was it achieved?

5.9.2 Suitability of Data

- The data that are suitable for one enquiry may not necessarily be found suitable in another enquiry. Hence, if the available data are found to be unsuitable, they should not be used by the researcher. In this context, the researcher must very carefully scrutinise the definition of various terms and units of collection used at the time of collecting the data from the primary source originally.
- Similarly, the object, scope and nature of the original enquiry must also be studied. If the researcher finds differences in these, the data will remain unsuitable for the present enquiry and should not be used.

5.9.3 Adequacy of Data

- If the level of accuracy achieved in data is found inadequate for the purpose of the present enquiry, they will be considered as inadequate and should not be used by the researcher. The data will also be considered inadequate, if they are related to an area which may be either narrower or wider than the area of the present enquiry.
- From all this we can say that it is very risky to use the already available data. The already available data should be used by the researcher only when he finds them reliable, suitable and adequate. But he should not blindly discard the use of such data if they are readily available from authentic sources and are also suitable and adequate for in that case it will not be economical to spend time and energy in field surveys for collecting information. At times, there may be wealth of usable information in the already available data which must be used by an intelligent researcher but with due precaution.

5.10 Selection of Appropriate Method for Data Collection

There are various methods of data collection. As such the researcher must judiciously select the method/methods for his own study, keeping in view the following factors:

- **Nature, scope and object of enquiry:** This constitutes the most important factor affecting the choice of a particular method. The method selected should be such that it suits the type of enquiry that is to be conducted by the researcher. This factor is also important in deciding whether the data already available (secondary data) are to be used or the data not yet available (primary data) are to be collected.
- **Availability of funds:** Availability of funds for the research project determines to a large extent the method to be used for the collection of data. When funds at the disposal of the researcher are very limited, he will have to select a comparatively cheaper method which may not be as efficient and effective as some other costly method. Finance, in fact, is a big constraint in practice and the researcher has to act within this limitation.
- **Time factor:** Availability of time has also to be taken into account in deciding a particular method of data collection. Some methods take relatively more time, whereas with others the data can be collected in a comparatively shorter duration. The time at the disposal of the researcher, thus, affects the selection of the method by which the data are to be collected.
- **Precision required:** Precision required is yet another important factor to be considered at the time of selecting the method of collection of data. But one must always remember that each method of data collection has its uses and none is superior in all situations. The secondary data may be used in case the researcher finds them reliable, adequate and appropriate for his research. While studying motivating influences in market researches or studying people's attitudes in psychological/social surveys, we can resort to the use of one or more of the projective techniques stated earlier. Such techniques are of immense value in case the reason is obtainable from

the respondent who knows the reason but does not want to admit it or the reason relates to some underlying psychological attitude and the respondent is not aware of it. But when the respondent knows the reason and can tell the same if asked, than a non-projective questionnaire, using direct questions, may yield satisfactory results even in case of attitude surveys.

Thus, the most desirable approach with regard to the selection of the method depends on the nature of the particular problem and on the time and resources (money and personnel) available along with the desired degree of accuracy. But, over and above all this, much depends upon the ability and experience of the researcher.

5.11 Case Study Method

Meaning:

- The case study method is a very popular form of qualitative analysis and involves a careful and complete observation of a social unit, be that unit a person, a family, an institution, a cultural group or even the entire community. The case study places more emphasis on the full analysis of a limited number of events or conditions and their interrelations.
- The case study deals with the processes that take place and their interrelationship. Thus, case study is essentially an intensive investigation of the particular unit under consideration. The object of the case study method is to locate the factors that account for the behaviour-patterns of the given unit as an integrated totality.
- According to H. Odum, "The case study method is a technique by which individual factor whether it is an institution or just an episode in the life of an individual or a group is analysed in its relationship to any other in the group." Thus, a fairly exhaustive study of a person (as to what he does and has done, what he thinks he does and had done and what he expects to do and says he ought to do) or group is called a life or case history.
- Burgess has used the words "the social microscope" for the case study method." Pauline V. Young describes case study as "a comprehensive study of a social unit be that unit a person, a group, a social institution, a district or a community."
- In brief, we can say that case study method is a form of qualitative analysis where in careful and complete observation of an individual or a situation or an institution is done; efforts are made to study each and every aspect of the concerning unit in minute details and then from case data generalisations and inferences are drawn.

Characteristics:

The important characteristics of the case study method are as under:

- Under this method the researcher can take one single social unit or more of such units for his study purpose; he may even take a situation to study the same comprehensively.
- Here, the selected unit is studied intensively, i.e., it is studied in minute details. Generally, the study extends over a long period of time to ascertain the natural history of the unit so as to obtain enough information for drawing correct inferences.
- In the context of this method, we make complete study of the social unit covering all facets. Through this method we try to understand the complex of factors that are operative within a social unit as an integrated totality.
- Under this method the approach happens to be qualitative and not quantitative. Mere quantitative information is not collected. Every possible effort is made to collect information concerning all aspects of life. As such, case study deepens our perception and gives us a clear insight into life. For instance, under this method we not only study how many crimes a man has done but shall peep into the factors that forced him to commit crimes when we are making a case study of a man as a criminal. The objective of the study may be to suggest ways to reform the criminal.
- In respect of the case study method an effort is made to know the mutual inter-relationship of causal factors.
- Under case study method the behaviour pattern of the concerning unit is studied directly and not by an indirect and abstract approach.
- Case study method results in fruitful hypotheses along with the data which may be helpful in testing them, and thus it enables the generalised knowledge to get richer and richer. In its absence, generalised social science may get handicapped.

Evolution and scope:
The case study method is a widely used systematic field research technique in sociology these days. The credit for introducing this method to the field of social investigation goes to Frederic Le Play who used it as a hand-maiden to statistics in his studies of family budgets. Herbert Spencer was the first to use case material in his comparative study of different cultures. Dr. William Healy resorted to this method in his study of juvenile delinquency, and considered it as a better method over and above the mere use of statistical data. Similarly, anthropologists, historians, novelists and dramatists have used this method concerning problems pertaining to their areas of interests. Even management experts use case study methods for getting clues to several management problems. In brief, case study method is being used in several disciplines. Not only this, its use is increasing day by day.

Assumptions: The case study method is based on several assumptions. The important assumptions may be listed as follows:

- The assumption of uniformity in the basic human nature in spite of the fact that human behaviour may vary according to situations.
- The assumption of studying the natural history of the unit concerned.
- The assumption of comprehensive study of the unit concerned.

Major phases involved:
Major phases involved in case study are as follows:

- Recognition and determination of the status of the phenomenon to be investigated or the unit of attention.
- Collection of data, examination and history of the given phenomenon.
- Diagnosis and identification of causal factors as a basis for remedial or developmental treatment.
- Application of remedial measures i.e., treatment and therapy (this phase is often characterised as case work).
- Follow-up programme to determine effectiveness of the treatment applied.

Advantages:
There are several advantages of the case study method that follow from the various characteristics outlined above. Mention may be made here of the important advantages.

- Being an exhaustive study of a social unit, the case study method enables us to understand fully the behaviour pattern of the concerned unit.
- Through case study a researcher can obtain a real and enlightened record of personal experiences which would reveal man's inner strivings, tensions and motivations that drive him to action along with the forces that direct him to adopt a certain pattern of behaviour.
- This method enables the researcher to trace out the natural history of the social unit and its relationship with the social factors and the forces involved in its surrounding environment.
- It helps in formulating relevant hypotheses along with the data which may be helpful in testing them. Case studies, thus, enable the generalised knowledge to get richer and richer.
- The method facilitates intensive study of social units which is generally not possible if we use either the observation method or the method of collecting information through schedules. This is the reason why case study method is being frequently used, particularly in social researches.
- Information collected under the case study method helps a lot to the researcher in the task of constructing the appropriate questionnaire or schedule for the said task requires thorough knowledge of the concerning universe.
- The researcher can use one or more of the several research methods under the case study method depending upon the prevalent circumstances. In other words, the use of different methods such as depth interviews, questionnaires, documents, study reports of individuals, letters, and the like is possible under case study method.
- Case study method has proved beneficial in determining the nature of units to be studied along with the nature of the universe. This is the reason why at times the case study method is alternatively known as "mode of organising data".

- This method is a means to well understand the past of a social unit because of its emphasis of historical analysis. Besides, it is also a technique to suggest measures for improvement in the context of the present environment of the concerned social units.
- Case studies constitute the perfect type of sociological material as they represent a real record of personal experiences which very often escape the attention of most of the skilled researchers using other techniques.
- Case study method enhances the experience of the researcher and this in turn increases his analysing ability and skill.
- This method makes possible the study of social changes. On account of the minute study of the different facets of a social unit, the researcher can well understand the social change then and now. This also facilitates the drawing of inferences and helps in maintaining the continuity of the research process. In fact, it may be considered the gateway to and at the same time the final destination of abstract knowledge.
- Case study techniques are indispensable for therapeutic and administrative purposes. They are also of immense value in taking decisions regarding several management problems. Case data are quite useful for diagnosis, therapy and other practical case problems.

Limitations:

Important limitations of the case study method may as well be highlighted.

- Case situations are seldom comparable and as such the information gathered in case studies is often not comparable. Since the subject under case study tells history in his own words, logical concepts and units of scientific classification have to be read into it or out of it by the investigator.
- Read Bain does not consider the case data as significant scientific data since they do not provide knowledge of the "impersonal, universal, non-ethical, non-practical, repetitive aspects of phenomena." Real information is often not collected because the subjectivity of the researcher does enter in the collection of information in a case study.
- The danger of false generalisation is always there in view of the fact that no set rules are followed in collection of the information and only few units are studied.
- It consumes more time and requires lot of expenditure. More time is needed under case study method since one study the natural history cycles of social units and that too minutely.
- The case data are often vitiated because the subject, according to Read Bain, may write what he thinks the investigator wants; and the greater the rapport, the more subjective the whole process is.
- Case study method is based on several assumptions which may not be very realistic at times, and as such the usefulness of case data is always subject to doubt.
- Case study method can be used only in a limited sphere; it is not possible to use it in case of a big society. Sampling is also not possible under a case study method.
- Response of the investigator is an important limitation of the case study method. He often thinks that he has full knowledge of the unit and can he answer about it. In case the same is not true, then consequences follow. In fact, this is more the fault of the researcher rather than that of the case method.

Conclusion

Despite the above stated limitations, we find that case studies are being undertaken in several disciplines, particularly in sociology, as a tool of scientific research in view of the several advantages indicated earlier. Most of the limitations can be removed if researchers are always conscious of these and are well trained in the modern methods of collecting case data and in the scientific techniques of assembling, classifying and processing the same. Besides, case studies, in modern times, can be conducted in such a manner that the data are amenable to quantification and statistical treatment. Possibly, this is also the reason why case studies are becoming popular day by day.

5.12 Guidelines for Constructing Questionnaire/Schedule

The researcher must pay attention to the following points in constructing an appropriate and effective questionnaire or a schedule:

- The researcher must keep in view the problem he is to study for it provides the starting point for developing the Questionnaire/Schedule. He must be clear about the various aspects of his research problem to be dealt with in the course of his research project.
- Appropriate form of questions depends on the nature of information sought, the sampled respondents and the kind of analysis intended. The researcher must decide whether to use closed or open-ended question. Questions should be simple and must be constructed with a view to their forming a logical part of a well thought out tabulation plan. The units of enumeration should also be defined precisely so that they can ensure accurate and full information.
- Rough draft of the Questionnaire/Schedule is prepared, giving due thought to the appropriate sequence of putting questions. Questionnaires or schedules previously drafted (if available) may as well be looked into at this stage.
- Researcher must invariably re-examine, and in case of need may revise the rough draft for a better one. Technical defects must be minutely scrutinised and removed.
- Pilot study should be undertaken for pre-testing the questionnaire. The questionnaire may be edited in the light of the results of the pilot study.
- Questionnaire must contain simple but straight forward directions for the respondents so that they may not feel any difficulty in answering the questions.

5.13 Choice of Suitable Method

- You have noticed that there are various methods and techniques for the collection of primary data. You should be careful while selecting the method which should be appropriate and effective. The selection of the methods depends upon various factors like scope and objectives of the inquiry, time, availability of funds, subject matter of the research, the kind of information required, degree of accuracy etc.
- As appraised, every method has its own merits and demerits. For example, the observation method is suitable for field surveys when the incident is really happening; the interview method is suitable where direct observation is not possible. Local reporter/correspondent method is suitable when information is required at regular intervals.
- The questionnaire method is appropriate in extensive enquiries where sample is large and scattered over large geographical areas and the respondents are able to express their responses in writing. The Schedule method is suitable in case respondents are illiterate.

Summary

- Data is required to make a decision in any business situation. The researcher is faced with one of the most difficult problems of obtaining suitable, accurate and adequate data. Utmost care must be exercised while collecting data because the quality of the research results depends upon the reliability of the data.
- Primary data refers to information gathered first hand by the researcher for the specific purpose of the study. It is raw data without interpretation and represents the personal or official opinion or position. Primary sources are most authorative since the information is not filtered or tampered.
- The task of data collection begins after a research problem has been defined and research design/ plan chalked out. While deciding about the method of data collection to be used for the study, the researcher should keep in mind two types of data, viz., primary and secondary.
- Under the observation method, the information is sought by way of investigator's own direct observation without asking from the respondent.
- Structured observation is considered appropriate in descriptive studies, whereas in an exploratory study the observational procedure is most likely to be relatively unstructured.
- If the observer observes by making himself, more or less, a member of the group he is observing so that he can experience what the members of the group experience, the observation is called as the participant observation.
- In non-controlled observation, no attempt is made to use precision instruments. The major aim of this type of observation is to get a spontaneous picture of life and persons.
- Controlled observation takes place in various experiments that are carried out in a laboratory or under controlled conditions, whereas uncontrolled observation is resorted to in case of exploratory researches.
- The interview method of collecting data involves presentation of oral-verbal stimuli and reply in terms of oral-verbal responses. This method can be used through personal interviews and, if possible, through telephone interviews.
- The main task of the interviewer in case of a focussed interview is to confine the respondent to a discussion of issues with which he seeks conversance. Such interviews are used generally in the development of hypotheses and constitute a major type of unstructured interviews.
- In order to make the questionnaire effective and to ensure quality to the replies received, a researcher should pay attention to the question-sequence in preparing the questionnaire.
- To be successful, questionnaire should be comparatively short and simple i.e., the size of the questionnaire should be kept to the minimum. Questions should proceed in logical sequence moving from easy to more difficult questions. Personal and intimate questions should be left to the end.

References

- Fowler, F., 2002. *Survey research methods*, Sage Publications.
- Murthy, C., 2009.*Research methodology*, 1st ed., Vrinda Publications (P) Ltd.
- *Collection of data.* [PDF] Available at: <http://www.egyankosh.ac.in/bitstream/123456789/25801/1/Unit2.pdf>. [Accessed 4 October 2011].
- *Methods and techniques of data collection*. [PDF] Available at: <http://www.egyankosh.ac.in/bitstream/123456789/35367/1/Unit-4.pdf>. [Accessed 4 October 2011].
- *Data Collection*, 2009. [Video Online] Available at: <http://www.youtube.com/watch?v=Z4frC2h8SzQ&feature=related>. [Accessed 4 October 2011].
- *4.3 Different data capture methods*, 2010. [Video Online] Available at: <http://www.youtube.com/watch?v=Hyh91AC_tAM>. [Accessed 4 October 2011].

Recommended Reading

- Singh, Y. K., 2006. *Fundamental of Research Methodology and Statistics*, New Age International.
- Goddard , W. and Melville, S., 2004. *Research Methodology:, An Introduction*. Juta and Company Ltd.
- Saunders., 2003. *Research Methods for Business Students,* Pearson Education India.

Self Assessment

1. A research design is a _____ which directs the plan of action to complete the research work.
   a. blue print
   b. red print
   c. basic print
   d. design

2. _____ is raw data without interpretation and represents the personal or official opinion or position.
   a. Secondary data
   b. Primary data
   c. Research design
   d. Tertiary data

3. Data on all diseases and child mortality have been collected from the patient's registers of various hospitals is an example of _____.
   a. Primary data
   b. Tertiary data
   c. Research data
   d. Secondary data

4. _____ is generally represented by index, bibliographies, dictionaries, encyclopaedias, handbooks, directories and other finding aids like the internet search engines.
   a. Primary data
   b. Tertiary data
   c. Research data
   d. Secondary data

5. Which statement is true?
   a. The research method is the most commonly used method especially in studies relating to behavioural sciences.
   b. The questionnaire method is the most commonly used method especially in studies relating to behavioural sciences.
   c. The observation method is the most commonly used method especially in studies relating to behavioural sciences.
   d. The interview method is the most commonly used method especially in studies relating to behavioural sciences.

6. By _____ every informant has to reply the same questions put in the same language and the field worker has no choice o get the desired reply by putting the different question or changing the language of the same question.
   a. Schedule
   b. Interview
   c. Observation
   d. Questionnaire

7. When the observation is made in the natural surroundings and the activities are performed in their usual course without being influenced or guided by any external force it is known as _____.
   a. Participant observation
   b. Non participant observation
   c. Controlled observation
   d. Non controlled observation

8. _____ are used by dealers of consumer durables to collect information regarding their products.
   a. Warranty cards
   b. Radio buttons
   c. Check boxes
   d. Panels

9. In _____ techniques the respondent in supplying information tends unconsciously to project his own attitudes or feelings on the subject under study.
   a. Check boxes
   b. Interview
   c. Observation
   d. Projective

10. Match the following.

A	B
1. Pantry audits	A. Designed to discover underlying motives and desires and are often used in motivational research.
2. Depth interviews	B. The usual objective is to find out what types of consumers buy certain products and certain brands, the assumption being that the contents of the pantry accurately portray consumer's preferences.
3. Controlled observation	C. No direct or predetermined questions used are in this type of interview.
4. Unstructured interview	D. Such observation has a tendency to supply formalised data upon which generalisations can be built with some degree of assurance.

   a. 1-A; 2-B; 3-C; 4-D
   b. 1-B; 2-A; 3-D; 4-C
   c. 1-C; 2-B; 3-A; 4-D
   d. 1-D; 2-C; 3-B; 4-A

Chapter VI

Processing and Analysing of Data

Aim

The aim of this chapter is to:

- define processing operations
- explain coding of data
- classify the data

Objectives

Objectives of this chapter are to:

- explain problems in processing
- explicate statistics in research
- elucidate the measures of research

Learning outcome

At the end of this chapter, you will be able to:

- discuss regression analysis
- comprehend partial correlation
- understand the measures of dispersion

6.1 Introduction

Once the collection of data is over, the next step is to organise data so that meaningful conclusions may be drawn. The information content of the observations has to be reduced to a relatively few concepts and aggregates. The data collected from the field has to be processed as laid down in the research plan. This is possible only through systematic processing of data. Data processing involves editing, coding, classification and tabulation of the data collected so that they are amenable to analysis. This is an intermediary stage between the collection of data and their analysis and interpretation. In this unit, therefore, we will learn about different stages of processing of data in detail.

6.1.1 Processing Operations

Following are the important steps in the operations process.

Editing of data

- Editing is the first stage in data processing. Editing may be broadly defined to be a procedure, which uses available information and assumptions to substitute inconsistent values in a data set. In other words, editing is the process of examining the data collected through various methods to detect errors and omissions and correct them for further analysis. While editing, care has to be taken to see that the data are as accurate and complete as possible, units of observations and number of decimal places are the same for the same variable. The following practical guidelines may be handy while editing the data:
  - The editor should have a copy of the instructions given to the interviewers.
  - The editor should not destroy or erase the original entry. Original entry should be crossed out in such a manner that they are still legible.
  - All answers, which are modified or filled in afresh by the editor, have to be indicated.
  - All completed schedules should have the signature of the editor and the date.
- For checking the quality of data collected, it is advisable to take a small sample of the questionnaire and examine them thoroughly. This helps in understanding the following types of problems:
  - whether all the questions are answered
  - whether the answers are properly recorded
  - whether there is any bias
  - whether there is any interviewer dishonesty
  - whether there are inconsistencies
- At times, it may be worthwhile to group the same set of questionnaires according to the investigators (whether any particular investigator has specific problems) or according to geographical regions (whether any particular region has specific problems) or according to the sex or background of the investigators, and corrective actions may be taken if any problem is observed.
- Before tabulation of data it may be good to prepare an operation manual to decide the process for identifying inconsistencies and errors and also the methods to edit and correct them. The following broad rules may be helpful.

Incorrect answers: It is quite common to get incorrect answers to many of the questions. A person with a thorough knowledge will be able to notice them.

For example, against the question "Which brand of biscuits do you purchase?" the answer may be "We purchase biscuits from ABC Stores". Now, this questionnaire can be corrected if ABC Stores stocks only one type of biscuits, otherwise not. Answer to the question "How many days did you go for shopping in the last week?" would be a number between 0 and 7. A number beyond this range indicates a mistake, and such a mistake cannot be corrected. The general rule is that changes may be made if one is absolutely sure, otherwise this question should not be used. Usually a schedule has a number of questions and although answers to a few questions are incorrect, it is advisable to use the other correct information from the schedule rather than discarding the schedule entirely.

Inconsistent answers: When there are inconsistencies in the answers or when there are incomplete or missing answers, the questionnaire should not be used. Suppose that in a survey, per capita expenditure on various items is reported as follows: Food – Rs. 700, Clothing – Rs.300, Fuel and Light – Rs. 200, other items – Rs. 550 and Total – Rs. 1600. The answers are obviously inconsistent as the total of individual items of expenditure is exceeding the total expenditure.

Modified answers: Sometimes it may be necessary to modify or qualify the answers. They have to be indicated for reference and checking. Numerical answers to be converted to same units: Against the question "What is the plinth area of your house?" answers could be either in square feet or in square metres. It will be convenient to convert all the answers to these questions in the same unit, square metre for example.

Coding of data

- Coding refers to the process by which data are categorised into groups and numerals or other symbols or both are assigned to each item depending on the class it falls in. Hence, coding involves:
  - deciding the categories to be used,
  - assigning individual codes to them
- In general, coding reduces the huge amount of information collected into a form that is amenable to analysis. A careful study of the answers is the starting point of coding. Next, a coding frame is to be developed by listing the answers and by assigning the codes to them. A coding manual is to be prepared with the details of variable names, codes and instructions. Normally, the coding manual should be prepared before collection of data, but for open-ended and partially coded questions. These two categories are to be taken care of after the data collection. The following are the broad general rules for coding:
  - Each respondent should be given a code number (an identification number).
  - Each qualitative question should have codes. Quantitative variables may or may not be coded depending on the purpose. Monthly income should not be coded if one of the objectives is to compute average monthly income. But if it is used as a classificatory variable it may be coded to indicate poor, middle or upper income group.
  - All responses including "don't know", "no opinion" "no response" etc., are to be coded.
- Sometimes it is not possible to anticipate all the responses and some questions are not coded before collection of data. Responses of all the questions are to be studied carefully and codes are to be decided by examining the essence of the answers. In partially coded questions, usually there is an option "Any other (specify)". Depending on the purpose, responses to this question may be examined and additional codes may be assigned.

Classification of data

Once the data is collected and edited, the next step towards further processing the data is classification. In most research studies, voluminous data collected through various methods needs to be reduced into homogeneous groups for meaningful analysis. This necessitates classification of data, which in simple terms is the process of dividing data into different groups or classes according to their similarities and dissimilarities.

The groups should be homogeneous within and heterogeneous between themselves. Classification condenses huge amount of data and helps in understanding the important underlying features. It enables us to make comparison, draw inferences, locate facts and also helps in bringing out relationships, so as to draw meaningful conclusions. In fact classification of data provides a basis for tabulation and analysis of data.

Types of classification

Data may be classified according to one or more external characteristics or one or more internal characteristics or both. Let us study these kinds with the help of illustrations.

Classification according to external characteristics

In this classification, data may be classified according to area or region (Geographical) and according to occurrences (Chronological).

- **Geographical:** In this type of classification, data are organised in terms of geographical area or region. State-wise production of manufactured goods is an example of this type. Data collected from an all India market survey may be classified geographically. Usually the regions are arranged alphabetically or according to the size to indicate the importance.
- **Chronological:** When data is arranged according to time of occurrence, it is called chronological classification. Profit of engineering industries over the last few years is an example. We may note that it is possible to have chronological classification within geographical classification and vice versa. For example, a large scale all India market survey spread over a number of years.

Classification according to internal characteristics

Data may be classified according to attributes (Qualitative characteristics which are not capable of being described numerically) and according to the magnitude of variables (Quantitative characteristics which are numerically described).

- **Classification according to attributes:** In this classification, data are classified by descriptive characteristic like sex, caste, occupation, place of residence etc. This is done in two ways – simple classification and manifold classification. In simple classification (also called classification according to dichotomy), data is simply grouped according to presence or absence of a single characteristics – male or female, employed or unemployed etc. In manifold classification (also known as multiple classification), data is classified according to more than one characteristic. First, the data may be divided into two groups according to one attribute (employee and unemployed, say). Then using the remaining attributes, data is sub-grouped again (male and female based on sex). This may go on based on other attributes, like married and unmarried, rural and urban so on… The following table is an example of manifold classification.
- **Classification according to magnitude of the variable:**
  - This classification refers to the classification of data according to some characteristics that can be measured. In this classification, there are two aspects: one is variables (age, weight, income etc) another is frequency (number of observations which can be put into a class).
  - Quantitative variables may be, generally, divided into two groups - discrete and continuous. A discrete variable is one which can take only isolated (exact) values, it does not carry any fractional value. The examples are number of children in a household, number of departments in an organisation, number of workers in a factory etc.
  - The variables that take any numerical value within a specified range are called continuous variables. The examples of continuous variables are the height of a person, profit/loss of companies etc. One point may be noted. In practice, even the continuous variables are measured up to some degree of precision and they also essentially become discrete variables.

Preparation of frequency distribution

When raw data is arranged in conveniently organised groups, it is called a frequency distribution. The number of data points in a particular group is called frequency. When a discrete variable takes a small number of values (not more than 8 or 10, say), each of the observed value is counted to form the discrete frequency distribution. In order to facilitate counting, prepare a column of "tallies". The following example illustrates it.

- **Construction of a continuous frequency distribution:** In continuous frequency distribution, the data is grouped into a small number of intervals instead of individual values of the variables. These groups are called classes. There are two different ways in which limits of classes may be arranged - exclusive and inclusive method. In the exclusive method, the class intervals are so arranged that the upper limit of one class is the lower limit of the next class, whereas in the inclusive method, the upper limit of a class is included in the class itself.

Tabulation of data

- Presentation of collected data in the tabular form is one of the techniques of data presentation. The two other techniques are diagrammatic and graphic presentation. Arranging the data in an orderly manner in rows and columns is called tabulation of data.

- Sometimes data collected by survey or even from publications of official bodies are so numerous that it is difficult to understand the important features of the data. Therefore it becomes necessary to summarise data through tabulation to an easily intelligible form. It may be noted that there may be loss of some minor information in certain cases, but the essential underlying features come out more clearly.
- Quite frequently, data presented in tabular form is much easier to read and understand than the data presented in the text. In classification, as discussed in the previous section, the data is divided on the basis of similarity and resemblance, whereas tabulation is the process of recording the classified facts in rows and columns. Therefore, after classifying the data into various classes, they should be shown in the tabular form.

Types of tables

Tables may be classified, depending upon the use and objectives of the data to be presented, into simple tables and complex tables. Let us discuss them along with illustrations.

Simple table: In this case data are presented only for one variable or characteristics. Therefore, this type of table is also known as one way table. The table showing the data relating to the sales of a company in different years will be an example of a single table.

Parts of a statistical table

A table should have the following four essential parts:

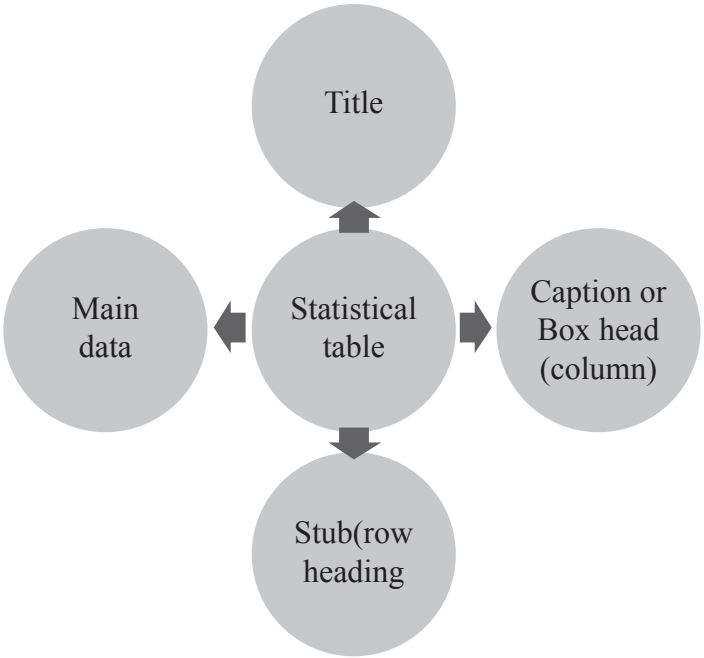


Fig. 6.1 Essentials of statistical table

- At times it may also contain an end note and source note below the table. The table should have a title, which is usually placed above the statistical table. The title should be clearly worded to give some idea of the table's contents. Usually a report has many tables. Hence the tables should be numbered to facilitate reference.
- Caption refers to the total of the columns. It is also termed as "box head".
- There may be sub-captions under the main caption. Stub refers to the titles given to the rows. Caption and stub should also be unambiguous. To the extent possible abbreviations should not be used in either caption or stub. But if they are used, the expansion must be given in the end note below. Notes pertaining to stub entries or box headings may be numerals. But, to avoid confusion, it is better to use some symbols (like *, **, @ etc) or alphabets for notes referring to the entries in the main body. If the table is based on outside information, it should be mentioned in the source note below.

This note should be complete with author, title, year of publication etc to enable the reader to go to the original source for crosschecking or for obtaining additional information. Columns and rows may be numbered for easy reference.

Arrangement of items in stub and box-head

There is no hard and fast rule about the arrangement of column and row headings in a table. It depends on the nature of data and type of analysis. A number of different methods are used - alphabetical, geographical, chronological / historical, magnitude-based and customary or conventional.

- Alphabetical: This method is suitable for general tables as it is easy to locate an item if it is arranged alphabetically. For example, population census data of India may be arranged in the alphabetical order of states/ union territories.
- **Geographical:** It can be used when the reader is familiar with the usual geographical classification.
- **Chronological:** A table containing data over a period of time may be presented in the chronological order. Population data (1961 to 2001) presented earlier are in chronological order. One may either start from the most recent year or the earliest year. However, there is a convention to start with the month of January whenever year and month data are presented.
- **Based on Magnitude:** At times, items in a table are arranged according to the value of the characteristic. Usually the largest item is placed first and other items follow in decreasing order. But this may be reversed also. Suppose that state-wise population data is arranged in order of decreasing magnitude. This will highlight the most populous state and the least populous state.
- **Customary or Conventional:** Traditionally some order is followed in certain cases. While presenting population census data, usually 'rural' comes before 'urban' and 'male' first and 'female' next. At times, conventional geographical order is also followed.

6.1.2 Some Problems in Processing

Following are the two problems of processing the data for analytical purposes:

The problem concerning "Don't know" (or DK) responses:

- While processing the data, the researcher often comes across some responses that are difficult to handle. One category of such responses may be 'Don't Know Response' or simply DK response. When the DK response group is small, it is of little significance. But when it is relatively big, it becomes a matter of major concern in which case the question arises: Is the question which elicited DK response useless? The answer depends on two point's, viz., the respondent actually may not know the answer or the researcher may fail in obtaining the appropriate information.
- In the first case the concerned question is said to be alright and DK response is taken as legitimate DK response. But in the second case, DK response is more likely to be a failure of the questioning process. How DK responses are to be dealt with by researchers? The best way is to design better type of questions. Good rapport of interviewers with respondents will result in minimising DK responses. But what about the DK responses that have already taken place?
- One way to tackle this issue is to estimate the allocation of DK answers from other data in the questionnaire. The other way is to keep DK responses as a separate category in tabulation where we can consider it as a separate reply category if DK responses happen to be legitimate, otherwise we should let the reader make his own decision.
- Yet another way is to assume that DK responses occur more or less randomly and as such we may distribute them among the other answers in the ratio in which the latter have occurred. Similar results will be achieved if all DK replies are excluded from tabulation and that too without inflating the actual number of other responses.

Use or percentages:

Percentages are often used in data presentation for they simplify numbers, reducing all of them to a 0 to 100 range. Through the use of percentages, the data are reduced in the standard form with base equal to 100 which fact facilitates relative comparisons. While using percentages, the following rules should be kept in view by researchers:

- Two or more percentages must not be averaged unless each is weighted by the group size from which it has been derived.
- Use of too large percentages should be avoided, since a large percentage is difficult to understand and tends to confuse, defeating the very purpose for which percentages are used.
- Percentages hide the base from which they have been computed. If this is not kept in view, the real differences may not be correctly read.
- Percentage decreases can never exceed 100 per cent and as such for calculating the percentage of decrease, the higher figure should invariably be taken as the base.
- Percentages should generally be worked out in the direction of the causal-factor in case of two-dimension tables and for this purpose we must select the more significant factor out of the two given factors as the causal factor.

6.1.3 Elements / Types of Analysis

- As stated earlier, by analysis we mean the computation of certain indices or measures along with searching for patterns of relationship that exist among the data groups. Analysis, particularly in case of survey or experimental data, involves estimating the values of unknown parameters of the population and testing of hypotheses for drawing inferences.
- Analysis may, therefore, be categorised as descriptive analysis and inferential analysis (Inferential analysis is often known as statistical analysis). "Descriptive analysis is largely the study of distributions of one variable. This study provides us with profiles of companies, work groups, persons and other subjects on any of a multiple of characteristics such as size, composition, efficiency, preferences, etc." This sort of analysis may be in respect of one variable (described as one-dimensional analysis), or in respect of two variables (described as bi variate analysis) or in respect of more than two variables (described as multivariate analysis).
- In this context we work out various measures that show the size and shape of a distribution(s) along with the study of measuring relationships between two or more variables. We may as well talk of correlation analysis and causal analysis. Correlation analysis studies the joint variation of two or more variables for determining the amount of correlation between two or more variables.
- Causal analysis is concerned with the study of how one or more variables affect changes in another variable. It is thus a study of functional relationships existing between two or more variables. This analysis can be termed as regression analysis. Causal analysis is considered relatively more important in experimental researches, whereas in most social and business researches our interest lies in understanding and controlling relationships between variables then with determining causes per se and as such we consider correlation analysis as relatively more important.
- In modern times, with the availability of computer facilities, there has been a rapid development of multivariate analysis which may be defined as "all statistical methods which simultaneously analyse more than two variables on a sample of observations". Usually the following analyses* are involved when we make a reference of multivariate analysis:
  - Multiple regression analysis: This analysis is adopted when the researcher has one dependent variable which is presumed to be a function of two or more independent variables. The objective of this analysis is to make a prediction about the dependent variable based on its covariance with all the concerned independent variables.
  - Multiple discriminate analysis: This analysis is appropriate when the researcher has a single dependent variable that cannot be measured, but can be classified into two or more groups on the basis of some attribute. The object of this analysis happens to be to predict an entity's possibility of belonging to a particular group based on several predictor variables.

- Multivariate analysis of variance (or multi-ANOVA): This analysis is an extension of two ways ANOVA, wherein the ratio of among group variance to within group variance is worked out on a set of variables.
- Canonical analysis: This analysis can be used in case of both measurable and non-measurable variables for the purpose of simultaneously predicting a set of dependent variables from their joint covariance with a set of independent variables.
- Inferential analysis: This is concerned with the various tests of significance for testing hypotheses in order to determine with what validity data can be said to indicate some conclusion or conclusions. It is also concerned with the estimation of population values. It is mainly on the basis of inferential analysis that the task of interpretation (i.e., the task of drawing inferences and conclusions) is performed.

6.1.4 Statistics in Research

- The role of statistics in research is to function as a tool in designing research, analysing its data and drawing conclusions there from. Most research studies result in a large volume of raw data which must be suitably reduced so that the same can be read easily and can be used for further analysis.
- Clearly the science of statistics cannot be ignored by any research worker, even though he may not have occasion to use statistical methods in all their details and ramifications. Classification and tabulation, as stated earlier, achieve this objective to some extent, but we have to go a step further and develop certain indices or measures to summarise the collected/classified data.
- Only after this we can adopt the process of generalisation from small groups (i.e., samples) to population. If fact, there are two major areas of statistics, viz., descriptive statistics and inferential statistics. Descriptive statistics concern the development of certain indices from the raw data, whereas inferential statistics concern with the process of generalisation. Inferential statistics are also known as sampling statistics and are mainly concerned with two major type of problems:
  - the estimation of population parameters, and
  - the testing of statistical hypotheses
- The important statistical measures that are used to summarise the survey/research data are:
  - measures of central tendency or statistical averages;
  - measures of dispersion;
  - measures of asymmetry;
  - measures of relationship; and
  - other measures
- Amongst the measures of central tendency, the three most important ones are the arithmetic average or mean, median and mode. Geometric mean and harmonic mean are also sometimes used. From among the measures of dispersion, variance, and its square root—the standard deviation are the most often used measures. Other measures such as mean deviation, range, etc. are also used. For comparison purpose, we use mostly the coefficient of standard deviation or the coefficient of variation.
- With respect to the measures of skewness and kurtosis, we mostly use the first measure of skewness based on mean and mode or on mean and median. Other measures of skewness, based on quartiles or on the methods of moments, are also used sometimes. Kurtosis is also used to measure the peakedness of the curve of the frequency distribution.
- Amongst the measures of relationship, Karl Pearson's coefficient of correlation is the frequently used measure in case of statistics of variables, whereas Yule's coefficient of association is used in case of statistics of attributes. Multiple correlation coefficients, partial correlation coefficient, regression analysis, etc., are other important measures often used by a researcher.
- Index numbers, analysis of time series, coefficient of contingency, etc., are other measures that may as well be used by a researcher, depending upon the nature of the problem under study. We give below a brief outline of some important measures (our of the above listed measures) often used in the context of research studies.

6.2 Data Measuring

Data measuring is explained below.

6.2.1 Measures of Central Tendency

- Measures of central tendency (or statistical averages) tell us the point about which items have a tendency to cluster. Such a measure is considered as the most representative figure for the entire mass of data. Measure of central tendency is also known as statistical average.
- Mean, median and mode are the most popular averages. Mean, also known as arithmetic average, is the most common measure of central tendency and may be defined as the value which we get by dividing the total of the values of various given items in a series by the total number of items we can work it out as under:

  Mean (or $\bar{X}$) = $\frac{\sum X_i}{n} = \frac{X_1 + X_2 + \ldots + X_n}{n}$

Where $\bar{X}$ = The symbol we use for mean (pronounced as X bar)
$\sum$ = Symbol for summation
X_i = Value of the ith item X, i = 1,2, ... , n
n = total number of items

- In case of a frequency distribution, we can work out mean in this way:

$$\bar{X} = \frac{\sum f_f X_i}{\sum f_i} = \frac{f_1 X_1 + f_2 X_2 + \ldots + f_n X_n}{f_1 + f_2 + \ldots + f_n} = n$$

- Sometimes, instead of calculating the simple mean, as stated above, we may workout the weighted mean for a realistic average. The weighted mean can be worked out as follows:

$$\overline{X_w} = \frac{\sum w_i X_i}{\sum w_i}$$

  Where X_w = weighted item
  w_i = weight of ith item X
  X_i = value of the ith item X

- Mean is the simplest measurement of central tendency and is a widely used measure. Its chief use consists in summarising the essential features of a series and in enabling data to be compared. It is amenable to algebraic treatment and is used in further statistical calculations. It is a relatively stable measure of central tendency. But it suffers from some limitations, viz., it is unduly affected by extreme items; it may not coincide with the actual value of an item in a series, and it may lead to wrong impressions, particularly when the item values are not given with the average. However, mean is better than other averages, especially in economic and social studies where direct quantitative measurements are possible.
- Median is the value of the middle item of series when it is arranged in ascending or descending order of magnitude. It divides the series into two halves; in one half all items are less than median, whereas in the other half all items have values higher than median. If the values of the items arranged in the ascending order are: 60, 74, 80, 90, 95, 100, and then the value of the 4th item, viz., 88 is the value of median. We can also write thus:

  If we assumed average A, then mean would be worked out as under:

$\bar{X} = A + \frac{\Sigma(X_i - A)}{n}$ or $\bar{X} = \frac{\Sigma f_i X_{i-A)}}{\Sigma f_i}$, in case of frequency distribution. This is also known as short cut method of finding $\bar{X}$

Median (M) = Value of $\left(\frac{n+1}{2}\right)^{th}$ item

- Median is a positional average and is used only in the context of qualitative phenomena, for example, in estimating intelligence, etc., which are often encountered in sociological fields. Median is not useful where items need to be assigned relative importance and weights. It is not frequently used in sampling statistics.
- Mode is the most commonly or frequently occurring value in a series. The mode in a distribution is that item around which there is maximum concentration. In general, mode is the size of the item which has the maximum frequency, but at items such an item may not be mode on account of the effect of the frequencies of the neighbouring items. Like median, mode is a positional average and is not affected by the values of extreme items. It is, therefore, useful in all situations where we want to eliminate the effect of extreme variations. Mode is particularly useful in the study of popular sizes.
- For example, a manufacturer of shoes is usually interested in finding out the size most in demand so that he may manufacture a larger quantity of that size. In other words, he wants a modal size to be determined for median or mean size would not serve his purpose. But there are certain limitations of mode as well. Geometric mean is also useful under certain conditions. It is defined as the nth root of the product of the values of n times in a given series. Symbolically, we can put it thus:
- Geometric mean (G.M.) = $\sqrt[n]{\pi X_i}$

= $\sqrt[n]{X_1.X_2.X_{3...}X_n}$

Where G.M. = Geometric Mean

n = number of items

X_i = ith value of the variable X

π = conventional product notation

For instance, the geometric mean of the numbers, 4, 6 and 9 is worked out as

G.M. = $\sqrt[3]{4.6.9}$

= 6

- The most frequently used application of this average is in the determination of average per cent of change, i.e., it is often used in the preparation of index numbers or when we deal in ratios. Harmonic mean is defined as the reciprocal of the average of reciprocals of the values of items of a series. Symbolically, we can express it as under:

Harmonic mean (H.M.) = Rec. $\frac{\sum Rec\ X_i}{n}$

= Rec. $\frac{Rec.X_1 + Rec.X_2 + \cdots + Rec.X_n}{n}$

Where H.M. = Harmonic mean

Rec. = Reciprocal

X_i = ith value of the variable X

n = number of items

For instance, the harmonic mean of the numbers 4, 5 and 10 is worked out as

H.M. = Rec $\frac{\frac{1}{4}+\frac{1}{5}+1/10}{3}$ = Rec $\frac{15+12+6/60}{3}$

$$\text{Rec}\left(\frac{33}{60} X \frac{1}{3}\right) = \frac{60}{11} = 5.45$$

- Harmonic mean is of limited application, particularly in cases where time and rate are involved. The harmonic mean gives largest weight to the smallest item and smallest weight to the largest item. As such it is used in cases like time and motion study where time is variable and distance constant. From what has been stated above, we can say that there are several types of statistical averages.
- Researcher has to make a choice for some average. There are no hard and fast rules for the selection of a particular average in statistical analysis for the selection of an average mostly depends on the nature, type of objectives of the research study. One particular type of average cannot be taken as appropriate for all types of studies. The chief characteristics and the limitations of the various averages must be kept in view; discriminate use of average is very essential for sound statistical analysis.

6.2.2 Measures of Dispersion

An average can represent a series only as best as a single figure can, but it certainly cannot reveal the entire story of any phenomenon under study. Especially it fails to give any idea about the scatter of the values of items of a variable in the series around the true value of average. In order to measure this scatter, statistical devices called measures of dispersion are calculated. Important measures of dispersion are:

Range

- It is the simplest possible measure of dispersion and is defined as the difference between the values of the extreme items of a series. Thus,
- Range = (Highest value of an item in a series) – (Lowest value of an item in a series)
- The utility of range is that it gives an idea of the variability very quickly, but the drawback is that range is affected very greatly by fluctuations of sampling. Its value is never stable, being based on only two values of the variable. As such, range is mostly used as a rough measure of variability and is not considered as an appropriate measure in serious research studies.

Mean deviation

- It is the average of difference of the values of items from some average of the series. Such a difference is technically described as deviation. In calculating mean deviation we ignore the minus sign of deviations while taking their total for obtaining the mean deviation. Mean deviation is, thus, obtained as under:

  Mean deviation from mean$(\delta_{\bar{x}}) = \frac{\sum|x_i - \bar{x}|}{n}$,

  if deviations, $|X_i - \bar{X}|$, are obtained from arithmetic average.

  Mean deviation from median (δ_m) =

- When mean deviation is divided by the average used in finding out the mean deviation itself, the resulting quantity is described as the coefficient of mean deviation. Coefficient of mean deviation is a relative measure of dispersion and is comparable to similar measure of other series. Mean deviation and its coefficient are used in statistical studies for judging the variability, and thereby render the study of central tendency of a series more precise by throwing light on the typicalness of an average. It is a better measure of variability than range as it takes into consideration the values of all items of a series. Even then it is not a frequently used measure as it is not amenable to algebraic process.

Standard deviation

- It is most widely used measure of dispersion of a series and is commonly denoted by the symbol 'σ' (pronounced as sigma). Standard deviation is defined as the square-root of the average of squares of deviations, when such deviations for the values of individual items in a series are obtained from the arithmetic average. It is worked out as under:

Standard deviation* $(\sigma) = \sqrt{\frac{\sum(X_i - \bar{X})^2}{n}}$

- * If we use assumed average, A, in place of $\bar{X}$ while finding deviations, then standard deviation would be worked out as under:

$$\sigma = \sqrt{\frac{\sum(X_i - A)^2}{n} - \left(\frac{\sum(X_i - A)^2}{n}\right)}$$

$$\sigma = \sqrt{\frac{\sum f_i(X_i - A)^2}{\sum f_i} - \left(\frac{\sum f_i(X_i - A)^2}{\sum f_i}\right)}$$

This is also known as the short cut method of finding σ.

Or

Standard deviation $(\sigma) = \sqrt{\frac{\sum f_i(X_i - \bar{X})^2}{\sum f_i}}$; in case of frequency distribution where, f_i means the frequency of the i^{th} item.

- When we divide the standard deviation by the arithmetic average of the series, the resulting quantity is known as coefficient of standard deviation which happens to be a relative measure and is often used for comparing with similar measure of other series. When this coefficient of standard deviation is multiplied by 100, the resulting figure is known as coefficient of variation. Sometimes, we work out the square of standard deviation, known as variance, which is frequently used in the context of analysis of variation.
- The standard deviation (along with several related measures like variance, coefficient of variation, etc.) is used mostly in research studies and is regarded as a very satisfactory measure of dispersion in a series. It is amenable to mathematical manipulation because the algebraic signs are not ignored in its calculation (as we ignore in case of mean deviation). It is less affected by fluctuations of sampling. These advantages make standard deviation and its coefficient a very popular measure of the scatteredness of a series. It is popularly used in the context of estimation and testing of hypotheses.

6.2.3 Measures of Relationship

- So far we have dealt with those statistical measures that we use in context of univariate population, i.e., the population consisting of measurement of only one variable. But if we have the data on two variables, we are said to have a bivariate population and if the data happen to be on more than two variables, the population is known as multivariate population. If for every measurement of a variable, X, we have corresponding value of a second variable, Y, the resulting pairs of values are called a bivariate population.
- In addition, we may also have a corresponding value of the third variable, Z, or the forth variable, W, and so on, the resulting pairs of values are called a multivariate population. In case of bivariate or multivariate populations, we often wish to know the relation of the two and/or more variables in the data to one another. We may like to know, for example, whether the number of hours students devote for studies is somehow related to their family income, to age, to sex or to similar other factor. There are several methods of determining the relationship between variables, but no method can tell us for certain that a correlation is indicative of causal relationship.
- Thus we have to answer two types of questions in bivariate or multivariate populations, viz.,
  - Does there exist association or correlation between the two (or more) variables? If yes, of what degree?
  - Is there any cause and effect relationship between the two variables in case of the bivariate population or between one variable on one side and two or more variables on the other side in case of multivariate population? If yes, of what degree and in which direction?

- The first question is answered by the use of correlation technique and the second question by the technique of regression. There are several methods of applying the two techniques, but the important ones are as under:
- In case of bivariate population: Correlation can be studied through
  - cross tabulation;
  - Charles Spearman's coefficient of correlation;
  - Karl Pearson's coefficient of correlation;

whereas cause and effect relationship can be studied through simple regression equations.

In case of multivariate population: Correlation can be studied through
- coefficient of multiple correlation;
- coefficient of partial correlation; whereas cause and effect relationship can be studied through multiple regression equations.

Following given is the brief explanation of each method:

Cross tabulation approach

- This is especially useful when the data are in nominal form. Under it we classify each variable into two or more categories and then cross classify the variables in these subcategories. Then we look for interactions between them which may be symmetrical, reciprocal or asymmetrical. A symmetrical relationship is one in which the two variables vary together, but we assume that neither variable is due to the other.
- A reciprocal relationship exists when the two variables mutually influence or reinforce each other. Asymmetrical relationship is said to exist if one variable (the independent variable) is responsible for another variable (the dependent variable). The cross classification procedure begins with a two-way table which indicates whether there is or there is not an interrelationship between the variables. This sort of analysis can be further elaborated in which case a third factor is introduced into the association through cross-classifying the three variables.
- By doing so we find conditional relationship in which factor X appears to affect factor Y only when factor Z is held constant. The correlation, if any, found through this approach is not considered a very powerful form of statistical correlation and accordingly we use some other methods when data happen to be either ordinal or interval or ratio data.

Charles Spearman's coefficient of correlation (or rank correlation)

- This is the technique of determining the degree of correlation between two variables in case of ordinal data where ranks are given to the different values of the variables. The main objective of this coefficient is to determine the extent to which the two sets of ranking are similar or dissimilar. This coefficient is determined as under:

  Spearman's coefficient of correlation or $(r_s) = 1 - \left[\frac{6\sum d_i^2}{n(n^2-1)}\right]$

  Where d_i = difference between ranks of i^{th} pair of the two variables;

  n = number of pairs of observations

  as rank correlation is a non-parametric technique for measuring relationship between paired observations of two variables when data are in the ranked form.
- Karl Pearson's coefficient of correlation (or simple correlation) is the most widely used method of measuring the degree of relationship between two variables. This coefficient assumes the following:
  - that there is linear relationship between the two variables;
  - that the two variables are casually related which means that one of the variables is independent and the other one is dependent; and
  - a large number of independent causes are operating in both variables so as to produce a normal distribution.
- Karl Pearson's coefficient of correlation can be worked out thus.

Karl Pearson's coefficient of correlation or (r)* =

$$\frac{\sum(X_i - \overline{X})\ (Y_i - \overline{Y}}{n.\ \sigma_x.\ \sigma_Y}$$

- * Alternatively, the formula can be written as:

$$r = \frac{\sum(X_i - \overline{X})\,(Y_i - \overline{Y}}{\sqrt{\sum(X_i - \overline{X})^2 . \sum(Y_i - Y^2}}$$

Or

$$r = \frac{Covariance\ between\ X\ and\ Y}{\sigma_x . \sigma_y} = \frac{\sum(X_i - \bar{X})(Y_i - \bar{Y})/n}{\sigma_x . \sigma_y}$$

$$r = \frac{\sum X_i Y_i - n\bar{X}.\bar{Y}}{\sqrt{\sum X_i^2 - n\bar{X}^2}\ \sqrt{\sum Y_i^2 - n\bar{Y}^2}}$$

(This applies when we take zero as the assumed mean for both variables, X and Y).

Where X_i = i[th] value of X variable

$\bar{X}$ = mean of X

Y_i = i[th] value of Y variable

$\bar{Y}$ = Mean of Y

n = number of pairs of observations of X and Y

σ_x = Standard deviation of X

σ_Y = Standard deviation of YIn case we use assumed means (Ax and Ay for variables X and Y respectively) in place of true means, then Karl Person's formula is reduced to:

$$\frac{\frac{\sum dx_i . dy_i}{n} - \left(\frac{\sum dx_i . dy_i}{n}\right)}{\sqrt{\frac{\sum dx_i^2}{n} - \left(\frac{\sum dx_i}{n}\right)^2}\ \sqrt{\frac{\sum dy_i^2}{n} - \left(\frac{\sum dy_i}{n}\right)^2}}$$

Where $\sum dx_i = \sum (X_i - A_x)$

$\sum dy_i = \sum (Y_i - A_y)$

$\sum dx_i^2 = \sum (X_i - A_x)^2$

$\sum dy_i^2 = \sum (Y_i - A_y)^2$

$\sum dx_i . dy_i = \sum (X_i - A_x)(Y_i - A_y)$

n = number of pairs of observations of X and Y

This is the short cut approach for finding 'r' in case of ungrouped data. If the data happen to be grouped data (i.e., the case of bivariate frequency distribution), we shall have to write Karl Pearson's coefficient of correlation as under:

$$\frac{\frac{\sum f_{ij}.\ dx_i . dy_j}{n} - \left(\frac{\sum f_i dx_i}{n} . \frac{\sum f_j\ dy_j}{n}\right)}{\sqrt{\frac{\sum f_i dx_i^2}{n} - \left(\frac{\sum f_i dx_i}{n}\right)}\ \sqrt{\frac{\sum f_i dy_j^2}{n} - \left(\frac{\sum f_j dy_i}{n}\right)^2}}$$

Where f_{ij} is the frequency of a particular cell in the correlation table and all other values are defined as earlier.

- Karl Pearson's coefficient of correlation is also known as the product moment correlation coefficient. The value of 'r' lies between ± 1. Positive values of r indicate positive correlation between the two variables (i.e., changes in both variables take place in the statement direction), whereas negative values of 'r' indicate negative correlation, i.e., changes in the two variables taking place in the opposite directions.
- A zero value of 'r' indicates that there is no association between the two variables. When r = (+) 1, it indicates perfect positive correlation and when it is (–) 1, it indicates perfect negative correlation, meaning thereby that variations in independent variable (X) explain 100% of the variations in the dependent variable (Y). We can also say that for a unit change in independent variable, if there happens to be a constant change in the dependent variable in the same direction, then correlation will be termed as perfect positive.
- But if such change occurs in the opposite direction, the correlation will be termed as perfect negative. The value of 'r' nearer to +1 or –1 indicates high degree of correlation between the two variables.

6.3 Regression Analysis

Following is the description of different types of regression analysis.

6.3.1 Simple Regression Analysis

Regression is the determination of a statistical relationship between two or more variables. In simple regression, we have only two variables, one variable (defined as independent) is the cause of the behaviour of another one (defined as dependent variable). Regression can only interpret what exists physically, i.e., there must be a physical way in which independent variable X can affect dependent variable Y. The basic relationship between X and Y is given by:

$$\hat{Y} = a + bX$$

Where the symbol $\hat{Y}$ denotes the estimated value of Y for a given value of X. This equation is known as the regression equation of Y on X (also represents the regression line of Y on X when drawn on a graph) which means that each unit change in X produces a change of b in Y, which is positive for direct and negative for inverse relationships.

Then generally used method to find the best fit that a straight line of this kind can give is the least-square method. To use it efficiently, we first determine

$$\sum x_i^2 = \sum X_i^2 - n\overline{X^2}$$

$$\sum y_i^2 = \sum Y_i^2 - n\overline{Y^2}$$

$$\sum x_i y_i = \sum X_i Y_i - n\bar{X}.\bar{Y}$$

Then $b = \frac{\sum x_i y_i}{\sum x_i^2}$, $a = \bar{Y} - b\bar{X}$

These measures define a and b which will give the best possible fit through the original X and Y points and the value of r can then be worked out as under:

$$r = \frac{b\sqrt{\sum x_i^2}}{\sqrt{\sum y_i^2}}$$

Thus, the regression analysis is a statistical method to deal with the formulation of mathematical model depicting relationship amongst variables which can be used for the purpose of prediction of the values of dependent variable, given the values of the independent variable.

(Alternatively, for fitting a regression equation of the type $\hat{Y}$ = a + bX to the given values of X and Y variables, we can find the values of the two constants, viz., a and b by using the following two normal equations:

$$\sum Y_i = na + b \sum X_i$$
$$\sum X_i Y_i = a \sum X_i + b \sum X_i^2$$

And then solving these equations for finding a and b values. Once these values are obtained and have been put in the equation $\hat{Y}$ = a + bX, we say that we have fitted the regression equation of Y on X to the given data. In a similar fashion, we can develop the regression equation of X and Y, viz., $\hat{X}$ = a + bX, presuming Y as an independent variable and X as dependent variable).

6.3.2 Multiple Correlation and Regression

When there are two or more than two independent variables, the analysis concerning relationship is known as multiple correlations and the equation describing such relationship as the multiple regression equation. We here explain multiple correlation and regression taking only two independent variables and one dependent variable (Convenient computer programs exist for dealing with a great number of variables). In this situation the results are interpreted as shown below:

Multiple regression equation assumes the form

$$\hat{Y} = a + b_1 X_1 + b_2 X_2$$

Where X1 and X2 are two independent variables and Y being the dependent variable, and the constants a, b_1 and b_2 can be solved by solving the following the following three normal equations:

$$\sum Y_i = na + b_1 \sum X_{1i} + b_2 \sum X_{2i}$$
$$\sum X_{1i} Y_i = a \sum X_{1i} + b_1 \sum X_{1i}^2 + b_2 \sum X_{1i} X_{2i}$$
$$\sum X_{2i} Y_i = a \sum X_{2i} + b1 \sum X_{1i} X_{2i} + b_2 \sum X_{2i}^2$$

(It may be noted that the number of normal equations would depend upon the number of independent variables. If there are 2 independent variables, then 3 equations, if there are 3 independent variables then 4 equations and so on, are used).

In multiple regression analysis, the regression coefficients (viz., b_1 b_2) become less reliable as the degree of correlation between the independent variables (viz., X_1 X_2) increases. If there is a high degree of correlation between independent variables, we have a problem of what is commonly described as the *problem of multi-co linearity*. In such a situation we should use only one set of the independent variable to make our estimate. In fact, adding a second variable, say X_2, that is correlated with the first variable, say X1, distorts the values of the regression coefficients. Nevertheless, the prediction for the dependent variable can be made even when multi-co linearity is present, but in such a situation enough care should be taken in selecting the independent variables to estimate a dependent variable so as to ensure that multi-co linearity is reduced to the minimum.

With more than one independent variable, we may make a difference between the collective effect of the two independent variables and the individual effect of each of them taken separately. The collective effect is given by the coefficient of multiple correlations.

$R_{y.x1x2}$ defined as under:

$$R_{y.x1x2} = \sqrt{\frac{b_1 \sum Y_i X_{1i} - n\bar{Y}\bar{X}_1 + b_2 \sum Y_i X_{2i} - n\bar{Y}\bar{X}_2}{\sum Y_i^2 - n\bar{Y}^2}}$$

Alternatively, we can write

$$R_{y.x1x2} = \sqrt{\frac{b_1 \sum x_{1i} y_i + b_2 \sum x_{2i} y_i}{\sum Y_i^2}}$$

Where $x_{1i} = (X_{1i} - \bar{X}_1)$

$x_{2i} = (X_{2i} - \bar{X}_2)$

$y_i = (Y_i - \bar{Y})$

and b_1 and b_2 are the regression coefficients.

6.3.3 Partial Correlation

Partial correlation measures separately the relationship between two variables in such a way that the effects of other related variables are eliminated. In other words, in partial correlation analysis, we aim at measuring the relation between a dependent variable and a particular independent variable by holding all other variables constant. Thus, each partial coefficient of correlation measures the effect of its independent variable on the dependent variable. To obtain it, it is first necessary to compute the simple coefficients of correlation between each set of pairs of variables as stated earlier. In the case of two independent variables, we shall have two partial correlation coefficients denoted $r_{yx1}.\ x_2$ and $r_{yx2\ .\ x1}$ which are worked out as under:

$$r_{yx1}.\ x_2 = \frac{R^2_{y.x1.x2} - r^2_{yx2}}{1 - r^2_{yx2}}$$

This measures the effort of X_1 on Y, more precisely, that proportion of the variation of Y not explained by X, which is explained by X_1. Also,

$$r_{yx2}.\ _{X1} = \frac{R^2_{y.x1.x2} - r^2_{yx1}}{1 - r^2_{yx1}}$$

in which X_1 and X_2 are simply interchanged, given the added effect of X_2 on Y.
Alternatively, we can work out the partial correlation coefficients thus:

$$r_{yx1.x2} = \frac{r_{yx1} - r_{yx2} \cdot r_{x1x2}}{\sqrt{1 - r^2_{yx2}}\sqrt{1 - r^2_{x1x2}}}$$

And

$$r_{yx2.x1} = \frac{r_{yx2} - r_{yx2} \cdot r_{x1x2}}{\sqrt{1 - r^2_{yx1}}\sqrt{1 - r^2_{x1x2}}}$$

These formulae of the alternative approach are based on simple coefficients of correlation (also known as zero order coefficients since no variable is held constant when simple correlation coefficients are worked out). The partial correlation coefficients are called first order coefficients when one variable is held constant as shown above; they are known as second order coefficients when two variables are held constant and so on.

6.4 Other Measures

Other data measures are explained below.

6.4.1 Index Numbers

- When series are expressed in same units, we can use averages for the purpose of comparison, but when the units in which two or more series are expressed happen to be different, statistical averages cannot be used to compare them. In such situations we have to rely upon some relative measurement which consists in reducing the figures to a common base.

- One such method is to convert the series into a series of index numbers. This is done when we express the given figures as percentages of some specific figure on a certain data. We can, thus, define an index number as a number which is used to measure the level of a given phenomenon as compared to the level of the same phenomenon at some standard date.
- The use of index number weights more as a special type of average, meant to study the changes in the effect of such factors which are incapable of being measured directly. But one must always remember that index numbers measure only the relative changes. Changes in various economic and social phenomena can be measured and compared through index numbers.
- Different indices serve different purposes. Specific commodity indices are to serve as a measure of changes in the phenomenon of that commodity only. Index numbers may measure cost of living of different classes of people. In economic sphere, index numbers are often termed as 'economic barometers measuring the economic phenomenon in all its aspects either directly by measuring the same phenomenon or indirectly by measuring something else which reflects upon the main phenomenon.
- But index numbers have their own limitations with which researcher must always keep him aware. For instance, index numbers are only approximate indicators and as such give only a fair idea of changes but cannot give an accurate idea. Chances of error also remain at one point or the other while constructing an index number but this does not diminish the utility of index numbers for they still can indicate the trend of the phenomenon being measured.
- However, to avoid fallacious conclusions, index numbers prepared for one purpose should not be used for other purposes or for the same purpose at other places.

6.4.2 Time Series Analysis

- In the context of economic and business researches, we may obtain quite often data relating to some time period concerning a given phenomenon. Such data is labelled as 'Time Series'. More clearly it can be stated that series of successive observations of the given phenomenon over a period of time are referred to as time series. Such series are usually the result of the effects of one or more of the following factors:
  - Secular trend or long term trend that shows the direction of the series in a long period of time. The effect of trend (whether it happens to be a growth factor or a decline factor) is gradual, but extends more or less consistently throughout the entire period of time under consideration. Sometimes, secular trend is simply stated as trend (or T).
  - Short time oscillations, i.e., changes taking place in the short period of time only and such changes can be the effect of the following factors:
    - Cyclical fluctuations (or C) are the fluctuations as a result of business cycles and are generally referred to as long term movements that represent consistently recurring rises and declines in an activity.
    - Seasonal fluctuations (or S) are of short duration occurring in a regular sequence at specific intervals of time. Such fluctuations are the result of changing seasons. Usually these fluctuations involve patterns of change within a year that tends to be repeated from year to year. Cyclical fluctuations and seasonal fluctuations taken together constitute short-period regular fluctuations.
    - Irregular fluctuations (or I), also known as Random fluctuations, are variations which take place in a completely unpredictable fashion.
- All these factors stated above are termed as components of time series and when we try to analyse time series, we try to isolate and measure the effects of various types of these factors on a series. To study the effect of one type of factor, the other type of factor is eliminated from the series. The given series is, thus, left with the effects of one type of factor only. For analysing time series, we usually have two models;
  - multiplicative model; and
  - additive model
- Multiplicative model assumes that the various components interact in a multiplicative manner to produce the given values of the overall time series and can be stated as under:

$$Y = T \times C \times S \times I$$

Where, Y = observed values of time series,
T = Trend
C = Cyclical fluctuations
S = Seasonal fluctuations
I = Irregular fluctuations

- Additive model considers the total of various components resulting in the given values of the overall time series and can be stated as:

$$Y = T + C + S + I$$

- There are various methods of isolating trend from the given series, viz., the free hand method, semi average method, method of moving averages, method of least squares and similarly there are methods of measuring cyclical and seasonal variations and whatever variations are left over are considered as random or irregular fluctuations.
- The analysis of time series is done to understand the dynamic conditions for achieving the short term and long-term goals of business firm(s). The past trends can be used to evaluate the success or failure of management policy or policies practiced hitherto. On the basis of past trends, the future patterns can be predicted and policy or policies may accordingly be formulated.
- We can as well study properly the effects of factors causing changes in the short period of time only, once we have eliminated the effects of trend. By studying cyclical variations, we can keep in view the impact of cyclical changes while formulating various policies to make them as realistic as possible. The knowledge of seasonal variations will be of great help to us in taking decisions regarding inventory, production, purchases and sales policies so as to optimise working results.
- Thus, analysis of time series is important in context of long term as well as short term forecasting and is considered a very powerful tool in the hands of business analysts and researchers.

Summary

- Data processing involves editing, coding, classification and tabulation of the data collected so that they are amenable to analysis. This is an intermediary stage between the collection of data and their analysis and interpretation.
- Coding refers to the process by which data are categorised into groups and numerals or other symbols or both are assigned to each item depending on the class it falls in.
- The groups should be homogeneous within and heterogeneous between themselves. Classification condenses huge amount of data and helps in understanding the important underlying features. It enables us to make comparison, draw inferences, locate facts and also helps in bringing out relationships, so as to draw meaningful conclusions. In fact classification of data provides a basis for tabulation and analysis of data.
- Causal analysis is concerned with the study of how one or more variables affect changes in another variable. It is thus a study of functional relationships existing between two or more variables.
- Multiple discriminate analyses are appropriate when the researcher has a single dependent variable that cannot be measured, but can be classified into two or more groups on the basis of some attribute.
- Median is the value of the middle item of series when it is arranged in ascending or descending order of magnitude. It divides the series into two halves; in one half all items are less than median, whereas in the other half all items have values higher than median.
- Mode is the most commonly or frequently occurring value in a series. The mode in a distribution is that item around which there is maximum concentration. In general, mode is the size of the item which has the maximum frequency, but at items such an item may not be mode on account of the effect of the frequencies of the neighbouring items.
- The use of index number weights more as a special type of average, meant to study the changes in the effect of such factors which are incapable of being measured directly. But one must always remember that index numbers measure only the relative changes.

References

- Kothari, C. R., 2004. *Research Methodology*. New Age International.
- Murthy, C., 2009. Research *methodology*, 1st ed., Vrinda Publications (P) Ltd.
- *Processing of data*. [Online] Available at: <http://www.egyankosh.ac.in/bitstream/123456789/2405/4/Unit%206.pdf>. [Accessed 4 October 2011].
- *Data Analysis*. [Online] Available at: <http://www.drtomoconnor.com/3760/3760lect07.htm>. [Accessed 4 October 2011].
- *Data analysis - example from a research project*, 2011. [Video Online] Available at: <http://vimeo.com/28346996>. [Accessed 4 October 2011].
- *2nd Part Quantitative Analysis,* 2009. [Video Online] Available at: <http://www.youtube.com/watch?v=oxkRAIn-bwg&feature=related>. [Accessed 4 October 2011].

Recommended Reading

- Singh, Y. K., 2006. *Fundamental of Research Methodology and Statistics*. New Age International.
- Marczyk, G. R., DeMatteo. D. and Festinger, D., 2010. *Essentials of Research Design and Methodology*. John Wiley and Sons.
- Kumar, R., 2010. *Research Methodology: A Step-by-Step Guide for Beginners*. Sage Publication Ltd.

Self Assessment

1. _____ is the process of examining the data collected through various methods to detect errors and omissions and correct them for further analysis.
   a. Editing
   b. Interview
   c. Questionnaire
   d. Observation

2. Once the data is collected and edited, the next step towards further processing the data is _____.
   a. Organising.
   b. Classification.
   c. Scheduling
   d. Interview

3. State-wise production of manufactured goods is an example of this type.
   a. Chronological
   b. Classified
   c. Geographical
   d. Social

4. A _____ is one which can take only isolated (exact) values, it does not carry any fractional value.
   a. Variable
   b. Discrete data
   c. Data
   d. Discrete variable

5. The variables that take any numerical value within a specified range are called _____.
   a. Continuous variable
   b. Discrete variable
   c. Variable
   d. Numerical variable

6. Which of the statements is true?
   a. When raw data is arranged in random groups, it is called a frequency distribution.
   b. When final data is arranged in conveniently organised groups, it is called a frequency distribution.
   c. When raw data is arranged in conveniently organised groups, it is called a frequency distribution.
   d. When raw data is arranged in conveniently organised groups, it is called a raw distribution.

7. Which statement is false?
   a. The number of data points in a particular group is called frequency.
   b. Presentation of collected data in the tabular form is one of the techniques of data presentation.
   c. Arranging the data in an orderly manner in rows and columns is called tabulation of data.
   d. Data presented in tabular form is much difficult to read and understand than the data presented in the text.

8. The table showing the data relating to the sales of a company in different years will be an example of a _____.
   a. Single table
   b. Statistical table
   c. Tabular table
   d. Alphabetical table

9. _____ are often used in data presentation for they simplify numbers, reducing all of them to a 0 to 100 range.
   a. Numbers
   b. Numeric
   c. Percentage
   d. Alphabets

10. Match the following.

1. Chronological classification	A. When data is arranged according to time of occurrence
2. Coding	B. It refers to the process by which data are categorised into groups and numerals or other symbols or both are assigned to each item depending on the class it falls in.
3. Magnitude table	C. It takes only isolated (exact) values, it does not carry any fractional value.
4. Discrete variable	D. Usually the largest item is placed first and other items follow in decreasing order.

   a. 1-A; 2-B; 3-C; 4-D
   b. 1-A; 2-B; 3-D; 4-C
   c. 1-D; 2-C; 3-B; 4-A
   d. 1-C; 2-B; 3-A; 4-D

Chapter VII

Interpretation and Report Writing

Aim

The aim of this chapter is to:

- define interpretation
- explain technique of interpretation
- introduce the essentials of interpretation

Objectives

The objectives of this chapter are to:

- enumerate report writing
- elucidate the purpose of a report
- enlist different steps in writing a report

Learning outcome

At the end of this chapter, you will be able to:

- understand the mechanics of writing a research report
- explain the characteristics of a good report
- discuss evaluation of the research report

7.1 Introduction

A researcher/statistician besides the collection and analysis of data has to draw inferences and explain their significance. Through interpretation, the meanings and implications of the study become clear. Analysis is not complete without interpretation, and interpretation can not proceed without analysis. Both are, thus, inter-dependent. In this unit, therefore, we will discuss the interpretation of analysed data, summarising the interpretation and statistical fallacies.

7.2 Meaning of Interpretation

Interpretation refers to the task of drawing inferences from the collected facts after an analytical and/or experimental study. In fact, it is a search for broader meaning of research findings. The task of interpretation has two major aspects, viz.,.

- the effort to establish continuity in research through linking the results of a given study with those of another, and
- the establishment of some explanatory concepts

In one sense, interpretation is concerned with relationships within the collected data, partially overlapping analysis. Interpretation also extends beyond the data of the study to include the results of other research, theory and hypotheses." Thus, interpretation is the device through which the factors that seem to explain what has been observed by researcher in the course of the study can be better understood and it also provides a theoretical conception which can serve as a guide for further researches.

7.2.1 Essentials of Interpretation

Certain points should be kept in mind before proceeding to draw conclusions from statistics. It is essential that:

- **The data are homogeneous:** It is necessary to ascertain that the data are strictly comparable. We must be careful to compare the like with the like and not with the unlike.
- **The data are adequate:** Sometimes it happens that the data are incomplete or insufficient and it is neither possible to analyse them scientifically nor is it possible to draw any inference from them. Such data must be completed first.
- **The data are suitable:** Before considering the data for interpretation, the researcher must confirm the required degree of suitability of the data. Inappropriate data are like no data. Hence, no conclusion is possible with unsuitable data.
- **The data are properly classified and tabulated:** Every care is to be taken as a pre-requisite, to base all types of interpretations on systematically classified and properly tabulated data and information.
- **The data are scientifically analysed:** Before drawing conclusions, it is necessary to analyse the data by applying scientific methods. Wrong analysis can play havoc with even the most carefully collected data. If interpretation is based on uniform, accurate, adequate, suitable and scientifically analysed data, there is every possibility of attaining a better and representative result.

Thus, from the above considerations we may conclude that it is essential to have all the pre-requisites/pre-conditions of interpretation satisfied to arrive at better conclusions.

7.2.2 Technique of Interpretation

The task of interpretation is not an easy job; rather it requires a great skill and dexterity on the part of researcher. Interpretation is an art that one learns through practice and experience. The researcher may, at times, seek the guidance from experts for accomplishing the task of interpretation. The technique of interpretation often involves the following steps:

- Researcher must give reasonable explanations of the relations which he has found and he must interpret the lines of relationship in terms of the underlying processes and must try to find out the thread of uniformity that lies under the surface layer of his diversified research findings. In fact, this is the technique of how generalisation should be done and concepts be formulated.

- Extraneous information, if collected during the study, must be considered while interpreting the final results of research study, for it may prove to be a key factor in understanding the problem under consideration.
- It is advisable, before embarking upon final interpretation, to consult someone having insight into the study and who is frank and honest and will not hesitate to point out omissions and errors in logical argumentation. Such a consultation will result in correct interpretation and, thus, will enhance the utility of research results.
- Researcher must accomplish the task of interpretation only after considering all relevant factors affecting the problem to avoid false generalisation. He must be in no hurry while interpreting results, for quite often the conclusions, which appear to be all right at the beginning, may not at all be accurate.

7.2.3 Precautions in Interpretation

One should always remember that even if the data are properly collected and analysed, wrong interpretation would lead to inaccurate conclusions. It is, therefore, absolutely essential that the task of interpretation be accomplished with patience in an impartial manner and also in correct perspective. Researcher must pay attention to the following points for correct interpretation:

- At the outset, researcher must invariably satisfy himself that
  - the data are appropriate, trustworthy and adequate for drawing inferences
  - the data reflect good homogeneity
  - proper analysis has been done through statistical methods
- The researcher must remain cautious about the errors that can possibly arise in the process of interpreting results. Errors can arise due to false generalisation and/or due to wrong interpretation of statistical measures, such as the application of findings beyond the range of observations, identification of correlation with causation and the like. Another major pitfall is the tendency to affirm that definite relationships exist on the basis of confirmation of particular hypotheses. In fact, the positive test results accepting the hypothesis must be interpreted as "being in accord" with the hypothesis, rather than as "confirming the validity of the hypothesis". The researcher must remain vigilant about all such things so that false generalisation may not take place. He should be well equipped with and must know the correct use of statistical measures for drawing inferences concerning his study.
- He must always keep in view that the task of interpretation is very much intertwined with analysis and cannot be distinctly separated. As such he must take the task of interpretation as a special aspect of analysis and accordingly must take all those precautions that one usually observes while going through the process of analysis, viz., precautions concerning the reliability of data, computational checks, validation and comparison of results.
- He must never lose sight of the fact that his task is not only to make sensitive observations of relevant occurrences, but also to identify and disengage the factors that are initially hidden to the eye. This will enable him to do his job of interpretation on proper lines. Broad generalisation should be avoided as most research is not amenable to it because the coverage may be restricted to a particular time, a particular area and particular conditions. Such restrictions, if any, must invariably be specified and the results must be framed within their limits.
- The researcher must remember that "ideally in the course of a research study, there should be constant interaction between initial hypothesis, empirical observation and theoretical conceptions. It is exactly in this area of interaction between theoretical orientation and empirical observation that opportunities for originality and creativity lie." He must pay special attention to this aspect while engaged in the task of interpretation.

7.3 Report Writing

The last and final phase of the journey in research is writing of the report. After the collected data has been analysed and interpreted and generalisations have been drawn the report has to be prepared. The task of research is incomplete till the report is presented. Writing of a report is the last step in a research study and requires a set of skills some what different from those called for in respect of the earlier stages of research. This task should be accomplished by the researcher with utmost care.

7.3.1 Purpose of Report Writing

- The report may be meant for the people in general, when the investigation has not been carried out at the instance of any third party. Research is essentially a cooperative venture and it is essential that every investigator should know what others have found about the phenomena under study. The purpose of a report is thus the dissipation of knowledge, broadcasting of generalisations so as to ensure their widest use.
- A report of research has only one function, "it must inform". It has to propagate knowledge. Thus, the purpose of a report is to convey to the interested persons the results and findings of the study in sufficient detail, and so arranged as to enable each reader to comprehend the data, and to determine for himself the validity of conclusions. Research results must invariably enter the general store of knowledge. A research report is always an addition to knowledge. All this explains the significance of writing a report.
- In a broader sense, report writing is common to both academics and organisations. However, the purpose may be different. In academics, reports are used for comprehensive and application-oriented learning. Whereas in organisations, reports form the basis for decision making.

7.3.2 Meaning of Report

- Reporting simply means communicating or informing through reports. The researcher has collected some facts and figures, analysed the same and arrived at certain conclusions. He has to inform or report the same to the parties interested. Therefore "reporting is communicating the facts, data and information through reports to the persons for whom such facts and data are collected and compiled".
- A report is not a complete description of what has been done during the period of survey/research. It is only a statement of the most significant facts that are necessary for understanding the conclusions drawn by the investigator. Thus, "a report by definition is simply an account". The report thus is an account describing the procedure adopted, the findings arrived at and the conclusions drawn by the investigator of a problem.

7.3.3 Importance of Social Research Report

The social research reports are prepared for the following purposes.

- Report enables the social problems in the society.
- The written report acts as guidelines for future course of action.
- Organise and regulate the social problems can be obtained through the reports.
- The information regarding specific problems or issues can be obtained by way of report.
- Information provided in the reports enables the decision making and policies of the planners and academicians etc.
- Report may also be prepared to convince the reader or sell an idea.
- Reports may be prepared to provide an insight into the problem and may also provide a final solution to the same.

7.3.4 Significance of Report Writing

Research report is considered a major component of the research study, as the research task remains incomplete till the report has been presented and/or written. As a matter of fact, even the most brilliant hypothesis, highly well designed and conducted research study, and the most striking generalisations and findings are of little value unless they are effectively communicated to others. The purpose of research is not well served unless the findings are made known to others. Research results must invariably enter the general store of knowledge. All this explains the significance of writing research report. There are people who do not consider writing of report as an integral part of the research process. But the general opinion is in favour of treating the presentation of research results or the writing of report as part and parcel of the research project. Writing of report is the last step in a research study and requires a set of skills somewhat different from those called for in respect of the earlier stages of research. This task should be accomplished by the researcher with utmost care; he may seek the assistance and guidance of experts for the purpose.

7.3.5 Different Steps in Writing Report

Research reports are the product of slow, painstaking, accurate inductive work. The usual steps involved in writing report are:

- logical analysis of the subject-matter;
- preparation of the final outline;
- preparation of the rough draft;
- rewriting and polishing;
- preparation of the final bibliography; and
- writing the final draft

Though all these steps are self explanatory, yet a brief mention of each one of these will be appropriate for better understanding.

Logical analysis of the subject matter: It is the first step which is primarily concerned with the development of a subject. There are two ways in which to develop a subject

- Logical: The logical development is made on the basis of mental connections and associations between the one thing and another by means of analysis. Logical treatment often consists in developing the material from the simple possible to the most complex structures.
- Chronological: This development is based on a connection or sequence in time or occurrence. The directions for doing or making something usually follow the chronological order.

Preparation of the final outline: It is the next step in writing the research report "Outlines are the framework upon which long written works are constructed. They are an aid to the logical organisation of the material and a reminder of the points to be stressed in the report."

Preparation of the rough draft: This follows the logical analysis of the subject and the preparation of the final outline. Such a step is of utmost importance for the researcher now sits to write down what he has done in the context of his research study. He will write down the procedure adopted by him in collecting the material for his study along with various limitations faced by him, the technique of analysis adopted by him, the broad findings and generalisations and the various suggestions he wants to offer regarding the problem concerned.

Rewriting and polishing of the rough draft: This step happens to be most difficult part of all formal writing. Usually this step requires more time than the writing of the rough draft. The careful revision makes the difference between a mediocre and a good piece of writing. While rewriting and polishing, one should check the report for weaknesses in logical development or presentation. The researcher should also "see whether or not the material, as it is presented, has unity and cohesion; does the report stand upright and firm and exhibit a definite pattern, like a marble arch? Or does it resemble an old wall of mouldering cement and loose brick." In addition the researcher should give due attention to the fact that in his rough draft he has been consistent or not. He should check the mechanics of writing - grammar, spelling and usage.

Preparation of the final bibliography:

- Next in order comes the task of the preparation of the final bibliography. The bibliography, which is generally appended to the research report, is a list of books in some way relevant to the research which has been done. It should contain all those works which the researcher has consulted. The bibliography should be arranged alphabetically and may be divided into two parts; the first part may contain the names of books and pamphlets, and the second part may contain the names of magazine and newspaper articles.
- Generally, this pattern of bibliography is considered convenient and satisfactory from the point of view of reader, though it is not the only way of presenting bibliography. The entries in bibliography should be made adopting the following order: For books and pamphlets the order may be as under:

- Name of author, last name first.
- Title, underlined to indicate italics.
- Place, publisher, and date of publication.
- Number of volumes.

Example
Kothari, C.R., Quantitative Techniques, New Delhi, Vikas Publishing House Pvt. Ltd., 1978.

For magazines and newspapers the order may be as under:
1. Name of the author, last name first.
2. Title of article, in quotation marks.
3. Name of periodical, underlined to indicate italics.
4. The volume or volume and number.
5. The date of the issue.
6. The pagination.

Example
Robert V. Roosa, "Coping with Short-term International Money Flows", The Banker, London, September, 1971, p. 995.

The above examples are just the samples for bibliography entries and may be used, but one should also remember that they are not the only acceptable forms. The only thing important is that, whatever method one selects, it must remain consistent.

Writing the final draft: This constitutes the last step. The final draft should be written in a concise and objective style and in simple language, avoiding vague expressions such as "it seems", "there may be", and the like ones. While writing the final draft, the researcher must avoid abstract terminology and technical jargon. Illustrations and examples based on common experiences must be incorporated in the final draft as they happen to be most effective in communicating the research findings to others. A research report should not be dull, but must stimulate people and maintain interest and must show originality. It must be remembered that every report should be an attempt to solve some intellectual problem and must contribute to the solution of a problem and must add to the knowledge of both the researcher and the reader.

7.3.6 Layout of the Research Report

Anybody, who is reading the research report, must necessarily be conveyed enough about the study so that he can place it in its general scientific context, judge the adequacy of its methods and thus form an opinion of how seriously the findings are to be taken. For this purpose there is the need of proper layout of the report. The layout of the report means as to what the research report should contain. A comprehensive layout of the research report should comprise:

Preliminary pages
In its preliminary pages the report should carry a title and date, followed by acknowledgements in the form of 'Preface' or 'Foreword'. Then there should be a table of contents followed by list of tables and illustrations so that the decision-maker or anybody interested in reading the report can easily locate the required information in the report.

Main text
The main text provides the complete outline of the research report along with all details. Title of the research study is repeated at the top of the first page of the main text and then follows the other details on pages numbered consecutively, beginning with the second page. Each main section of the report should begin on a new page. The main text of the report should have the following sections:

- Introduction: The purpose of introduction is to introduce the research project to the readers. It should contain a clear statement of the objectives of research. A brief summary of other relevant research may also be stated so that the present study can be seen in that context. The hypotheses of study, if any, and the definitions of the major concepts employed in the study should be explicitly stated in the introduction of the report. The methodology adopted in conducting the study must be fully explained. The statistical analysis adopted must also be clearly stated. In addition to all this, the scope of the study should be stated and the boundary lines be demarcated. The various limitations, under which the research project was completed, must also be narrated.
- Statement of findings and recommendations: After introduction, the research report must contain a statement of findings and recommendations in non-technical language so that it can be easily understood by all concerned. If the findings happen to be extensive, at this point they should be put in the summarised form.
- Results: A detailed presentation of the findings of the study, with supporting data in the form of tables and charts together with a validation of results, is the next step in writing the main text of the report. The result section of the report should contain statistical summaries and reductions of the data rather than the raw data. All the results should be presented in logical sequence and splitted into readily identifiable sections.
- Implications of the results: Toward the end of the main text, the researcher should again put down the results of his research clearly and precisely. He should, state the implications that flow from the results of the study, for the general reader is interested in the implications for understanding the human behaviour. Such implications may have three aspects as stated below:
  - A statement of the inferences drawn from the present study which may be expected to apply in similar circumstances.
  - The conditions of the present study which may limit the extent of legitimate generalisations of the inferences drawn from the study.
  - The relevant questions that still remain unanswered or new questions raised by the study along with suggestions for the kind of research that would provide answers for them.
- It is considered a good practice to finish the report with a short conclusion which summarises and recapitulates the main points of the study. The conclusion drawn from the study should be clearly related to the hypotheses that were stated in the introductory section. At the same time, a forecast of the probable future of the subject and an indication of the kind of research which needs to be done in that particular field is useful and desirable.
- Summary: It has become customary to conclude the research report with a very brief summary, resting in brief the research problem, the methodology, the major findings and the major conclusions drawn from the research results.

End matter

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

7.4 Types of Reports

Research reports vary greatly in length and type. In each individual case, both the length and the form are largely dictated by the problems at hand. The results of a research investigation can be presented in a number of ways, viz., a technical report, a popular report, an article, a monograph or at times even in the form of oral presentation. Method(s) of presentation to be used in a particular study depends on the circumstances under which the study arose and the nature of the results. A technical report is used whenever a full written report of the study is required whether for recordkeeping or for public dissemination. A popular report is used if the research results have policy implications. We give below a few details about the said two types of reports:

Technical report

In the technical report the main emphasis is on

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

A general outline of a technical report can be as follows:

- Summary of results: A brief review of the main findings just in two or three pages.
- Nature of the study: Description of the general objectives of study, formulation of the problem in operational terms, the working hypothesis, the type of analysis and data required, etc.
- Methods employed: Specific methods used in the study and their limitations. For instance, in sampling studies we should give details of sample design, viz., sample size, sample selection, etc.
- Data: Discussion of data collected their sources, characteristics and limitations. If secondary data are used, their suitability to the problem at hand be fully assessed. In case of a survey, the manner in which data were collected should be fully described.
- Analysis of data and presentation of findings: The analysis of data and presentation of the findings of the study with supporting data in the form of tables and charts be fully narrated. This, in fact, happens to be the main body of the report usually extending over several chapters.
- Conclusions: A detailed summary of the findings and the policy implications drawn from the results be explained.
- Bibliography: Bibliography of various sources consulted be prepared and attached.
- Technical appendices: Appendices be given for all technical matters relating to questionnaire, mathematical derivations, elaboration on particular technique of analysis and the like ones.
- Index: Index must be prepared and be given invariably in the report at the end.

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

Popular report

The popular report is one which gives emphasis on simplicity and attractiveness. The simplification should be sought through clear writing, minimisation of technical, particularly mathematical, details and liberal use of charts and diagrams. Attractive layout along with large print, many subheadings, even an occasional cartoon now and then is another characteristic feature of the popular report. Given below is a general outline of a popular report:

- The findings and their implications: Emphasis in the report is given on the findings of most practical interest and on the implications of these findings.
- Recommendations for action: Recommendations for action on the basis of the findings of the study is made in this section of the report.
- Objective of the study: A general review of how the problem arises is presented along with the specific objectives of the project under study.
- Methods employed: A brief and non-technical description of the methods and techniques used, including a short review of the data on which the study is based, is given in this part of the report.
- Results: This section constitutes the main body of the report wherein the results of the study are presented in clear and non-technical terms with liberal use of all sorts of illustrations such as charts, diagrams and the like ones.
- Technical appendices: More detailed information on methods used, forms, etc. is presented in the form of appendices. But the appendices are often not detailed if the report is entirely meant for general public.

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

Oral presentation

At times oral presentation of the results of the study is considered effective, particularly in cases where policy recommendations are indicated by project results. The merit of this approach lies in the fact that it provides an opportunity for give-and-take decisions which generally lead to a better understanding of the findings and their implications. But the main demerit of this sort of presentation is the lack of any permanent record concerning the research details and it may be just possible that the findings may fade away from people's memory even before an action is taken. In order to overcome this difficulty, a written report may be circulated before the oral presentation and referred to frequently during the discussion.

7.5 Mechanics of Writing A Research Report

There are very definite and set rules which should be followed in the actual preparation of the research report or paper. Once the techniques are finally decided, they should be carefully adhered to, and no deviation permitted. The criteria of format should be decided as soon as the materials for the research paper have been assembled. The following points deserve mention so far as the mechanics of writing a report are concerned:

- **Size and physical design:** The manuscript should be written on unruled paper 8 1 $2^2 \times 11^2$ in size. If it is to be written by hand, then black or blue-black ink should be used. A margin of at least one and one-half inches should be allowed at the left hand and of at least half an inch at the right hand of the paper. There should also be one-inch margins, top and bottom. The paper should be neat and legible. If the manuscript is to be typed, then all typing should be double-spaced on one side of the page only except for the insertion of the long quotations.
- **Procedure:** Various steps in writing the report should be strictly adhered (All such steps have already been explained earlier in this chapter).
- **Layout:** Keeping in view the objective and nature of the problem, the layout of the report should be thought of and decided and accordingly adopted (The layout of the research report and various types of reports have been described in this chapter earlier which should be taken as a guide for report-writing in case of a particular problem).
- **Treatment of quotations:** Quotations should be placed in quotation marks and double spaced, forming an immediate part of the text. But if a quotation is of a considerable length (more than four or five type written lines) then it should be single-spaced and indented at least half an inch to the right of the normal text margin.
- **The footnotes:** Regarding footnotes one should keep in view the followings:
  - The footnotes serve two purposes, viz., the identification of materials used in quotations in the report and the notice of materials not immediately necessary to the body of the research text but still of supplemental value. In other words, footnotes are meant for cross references, citation of authorities and sources, acknowledgement and elucidation or explanation of a point of view. It should always be kept in view that footnote is neither an end nor a means of the display of scholarship. The modern tendency is to make the minimum use of footnotes for scholarship does not need to be displayed.
  - Footnotes are placed at the bottom of the page on which the reference or quotation which they identify or supplement ends. Footnotes are customarily separated from the textual material by a space of half an inch and a line about one and a half inches long.
  - Footnotes should be numbered consecutively, usually beginning with 1 in each chapter separately. The number should be put slightly above the line, say at the end of a quotation. At the foot of the page, again, the footnote number should be indented and typed a little above the line. Thus, consecutive numbers must be used to correlate the reference in the text with its corresponding note at the bottom of the page, except in case of statistical tables and other numerical material, where symbols such as the asterisk (*) or the like one may be used to prevent confusion.
  - Footnotes are always typed in single space though they are divided from one another by double space.

- **Documentation style:** Regarding documentation, the first footnote reference to any given work should be complete in its documentation, giving all the essential facts about the edition used. Such documentary footnotes follow a general sequence. The common order may be described as under:
  - Regarding the single-volume reference
    - Author's name in normal order (and not beginning with the last name as in a bibliography) followed by a comma;
    - Title of work, underlined to indicate italics;
    - Place and date of publication;
    - Pagination references (The page number).

Example
John Gassner, Masters of the Drama, New York: Dover Publications, Inc. 1954, p. 315.

  - Regarding multi volumed reference
    - Author's name in the normal order;
    - Title of work, underlined to indicate italics;
    - Place and date of publication;
    - Number of volume;
    - Pagination references (The page number).
  - Regarding works arranged alphabetically
    - For works arranged alphabetically such as encyclopaedias and dictionaries, no pagination reference is usually needed. In such cases the order is illustrated as under:

Example 1
"Salamanca," Encyclopaedia Britannica, 14th Edition.

Example 2
"Mary Wollstonecraft Godwin," Dictionary of national biography.

But if there should be a detailed reference to a long encyclopaedia article, volume and pagination reference may be found necessary.

  - Regarding periodicals reference
    - Name of the author in normal order;
    - Title of article, in quotation marks;
    - Name of periodical, underlined to indicate italics;
    - Volume number;
    - Date of issuance;
    - Pagination.
  - Regarding anthologies and collections reference
    - Quotations from anthologies or collections of literary works must be acknowledged not only by author, but also by the name of the collector.
  - Regarding second-hand quotations reference

  In such cases the documentation should be handled as follows:
    - Original author and title
    - "quoted or cited in,"
    - Second author and work

Case of multiple authorship:

If there are more than two authors or editors, then in the documentation the name of only the first is given and the multiple authorship is indicated by "et al." or "and others". Subsequent references to the same work need not be as detailed as stated above. If the work is cited again without any other work intervening, it may be indicated as ibid, followed by a comma and the page number. A single page should be referred to as p., but more than one page be referred to as pp. If there are several pages referred to at a stretch, the practice is to use often the page number, for example, pp. 190ff, which means page number 190 and the following pages; but only for page 190 and the following page '190f'. Roman numerical is generally used to indicate the number of the volume of a book. Op. cit. (opera citato, in the work cited) or Loc. cit. (loco citato, in the place cited) is two of the very convenient abbreviations used in the footnotes. Op. cit. or Loc. cit. after the writer's name would suggest that the reference is to work by the writer which has been cited in detail in an earlier footnote but intervened by some other references.

- Punctuation and abbreviations in footnotes:
  - The first item after the number in the footnote is the author's name, given in the normal signature order. This is followed by a comma.
  - After the comma, the title of the book is given: the article (such as "A", "An", "The" etc.) is omitted and only the first word and proper nouns and adjectives are capitalised.
  - Information concerning the edition is given next. This entry is followed by a comma.
  - The place of publication is then stated; it may be mentioned in an abbreviated form, if the place happens to be a famous one such as Lond. for London, N.Y. for New York, N.D. for New Delhi and so on. This entry is followed by a comma.
  - Then the name of the publisher is mentioned and this entry is closed by a comma.
  - It is followed by the date of publication if the date is given on the title page. If the date appears in the copyright notice on the reverse side of the title page or elsewhere in the volume, the comma should be omitted and the date enclosed in square brackets [c 1978], [1978].
  - The entry is followed by a comma. Then follow the volume and page references and are separated by a comma if both are given. A period closes the complete documentary reference.
  - But one should remember that the documentation regarding acknowledgements from magazine articles and periodical literature follow a different form as stated earlier while explaining the entries in the bibliography.
  - Certain English and Latin abbreviations are quite often used in bibliographies and footnotes to eliminate tedious repetition. The following is a partial list of the most common abbreviations frequently used in report-writing (the researcher should learn to recognise them as well as he should learn to use them):

Abbreviation	Long form
anon	anonymous
ante	before
art	article
aug	augmented
bk	book
bull	bulletin
cf	compare
ch	chapter
col	column
diss	dissertation
ed	editor, edition, edited
ed. cit.	edition cited

e.g.	exempli gratia: for example
eng.	enlarged
et. al.	and others
et seq., et sequens	and the following
ex	example
f., ff.	and the following
fig (s)	figure (s)
fn	footnote
id, idem	the same
ill, illus, illust (s)	illustrated, illustrations
intro	introduction
l, ll	line (s)
MS., MSS.	Manuscript or manuscripts
N.B.	Nota bene: note well
n.d.	no date
n.p.	no place
no pub.	no publisher
no (s)	number (s)
o.p.	out of print
o.p. cit.	in the work cited
rev	revised
vid or vide	see, refer to
viz	namely
vol (s)	volume (s)

Table 7.1 List of abbreviations

- **Use of statistics, charts and graphs:** A judicious use of statistics in research reports is often considered a virtue for it contributes a great deal towards the clarification and simplification of the material and research results. One may well remember that a good picture is often worth more than a thousand words. Statistics are usually presented in the form of tables, charts, bars and line-graphs and pictograms. Such presentation should be self explanatory and complete in itself. It should be suitable and appropriate looking to the problem at hand. Finally, statistical presentation should be neat and attractive.
- **The final draft:** Revising and rewriting the rough draft of the report should be done with great care before writing the final draft. Sentences that seem crystal-clear to the writer may prove quite confusing to other people; a connection that had seemed self evident may strike others as a non-sequitur. A friendly critic, by pointing out passages that seem unclear or illogical, and perhaps suggesting ways of remedying the difficulties, can be an invaluable aid in achieving the goal of adequate communication."
- **Bibliography:** Bibliography should be prepared and appended to the research report as discussed earlier. Preparation of the index: At the end of the report, an index should invariably be given, the value of which lies in the fact that it acts as a good guide, to the reader. Index may be prepared both as subject index and as author index. The former gives the names of the subject-topics or concepts along with the number of pages on which they have appeared or discussed in the report, whereas the latter gives the similar information regarding the names of authors. The index should always be arranged alphabetically. Some people prefer to prepare only one index common for names of authors, subject-topics, concepts and the like ones.

7.6 Characteristics of a Good Research Report

Research report is a channel of communicating the research findings to the readers of the report. A good report is one which does this task efficiently and effectively. As such it should have the following characteristics/qualities.

- It must be clear in informing the what, why, who, whom, when, where and how of the research study.
- It should be neither too short nor too long. One should keep in mind the fact that it should be long enough to cover the subject matter but short enough to sustain the reader's interest.
- It should be written in an objective style and simple language, correctness, precision and clarity should be the watchwords of the scholar. Wordiness, indirection and pompous language are barriers to communication.
- A good report must combine clear thinking, logical organisation and sound interpretation.
- It should not be dull. It should be such as to sustain the reader's interest.
- It must be accurate. Accuracy is one of the requirements of a report. It should be factual with objective presentation. Exaggerations and superlatives should be avoided.
- Clarity is another requirement of presentation. It is achieved by using familiar words and unambiguous statements, explicitly defining new concepts and unusual terms.
- Coherence is an essential part of clarity. There should be logical flow of ideas (i.e. continuity of thought), sequence of sentences. Each sentence must be so linked with other sentences so as to move the thoughts smoothly.
- Readability is an important requirement of good communication. Even a technical report should be easily understandable. Technicalities should be translated into language understandable by the readers.
- A research report should be prepared according to the best composition practices. Ensure readability through proper paragraphing, short sentences, illustrations, examples, and section headings, use of charts, graphs and diagrams.
- Draw sound inferences/conclusions from the statistical tables. But don't repeat the tables in text (verbal) form.
- Footnote references should be in proper form. The bibliography should be reasonably complete and in proper form.
- The report must be attractive in appearance, neat and clean whether typed or printed.
- The report should be free from mistakes of all types, viz., language mistakes, factual mistakes, spelling mistakes, calculation mistakes etc.,

The researcher should try to achieve these qualities in his report as far as possible.

7.7 Precautions for Writing Research Reports

Research report is a channel of communicating the research findings to the readers of the report. A good research report is one which does this task efficiently and effectively. As such it must be prepared keeping the following precautions in view:

- While determining the length of the report (since research reports vary greatly in length), one should keep in view the fact that it should be long enough to cover the subject but short enough to maintain interest. In fact, report-writing should not be a means to learning more and more about less and less.
- A research report should not, if this can be avoided, be dull; it should be such as to sustain reader's interest.
- Abstract terminology and technical jargon should be avoided in a research report. The report should be able to convey the matter as simply as possible. This, in other words, means that report should be written in an objective style in simple language, avoiding expressions such as "it seems," "there may be" and the like.
- Readers are often interested in acquiring a quick knowledge of the main findings and as such the report must provide a ready availability of the findings. For this purpose, charts, graphs and the statistical tables may be used for the various results in the main report in addition to the summary of important findings.
- The layout of the report should be well thought out and must be appropriate and in accordance with the objective of the research problem.

- The reports should be free from grammatical mistakes and must be prepared strictly in accordance with the techniques of composition of report-writing such as the use of quotations, footnotes, documentation, proper punctuation and use of abbreviations in footnotes and the like.
- The report must present the logical analysis of the subject matter. It must reflect a structure wherein the different pieces of analysis relating to the research problem fit well.
- A research report should show originality and should necessarily be an attempt to solve some intellectual problem. It must contribute to the solution of a problem and must add to the store of knowledge.
- Towards the end, the report must also state the policy implications relating to the problem under consideration. It is usually considered desirable if the report makes a forecast of the probable future of the subject concerned and indicates the kinds of research still needs to be done in that particular field.
- Appendices should be enlisted in respect of all the technical data in the report.
- Bibliography of sources consulted is a must for a good report and must necessarily be given.
- Index is also considered an essential part of a good report and as such must be prepared and appended at the end.
- Report must be attractive in appearance, neat and clean, whether typed or printed.
- Calculated confidence limits must be mentioned and the various constraints experienced in conducting the research study may also be stated in the report.
- Objective of the study, the nature of the problem, the methods employed and the analysis techniques adopted must all be clearly stated in the beginning of the report in the form of introduction.

7.8 Presentation of Research Report

Presentation has become an important communication medium in organisations because a research report is properly understood if it is accompanied by a presentation.

Presentation skill: Research report presentation skills include the ability to mix in the right proportion various elements of:

- Communications
- Presentation package
- Use of audio – visual aids to achieve proper presentation.

The researcher needs to acquire the public conversation skills.

Communication dimension

The major elements of communication dimension, which are relevant to a presentation, are:

- **Purpose:** Researcher has to think the purpose of the presentation and focus the light on research analysis sharply. Researcher can try to achieve a variety of purpose such as informing, selling, exploring, and decision making persuading and changing attitude or behavior.
- **Audience:** research report presentation, multiple audiences interest at the same time. The researcher and the receiver of the message keep changing roles through clarification queries, question answer, dialogue and discussion; researcher should make live and dynamic shape of presentation of report.
- **Media:** Researcher report presentation, sound, sight and body language come into play. Therefore, the co-ordination of all three at one short becomes an important aspect of presentation.
- **Message:** Researcher has to think of focus of the message. The presentation situation is built on interaction between the presenter and the audience, the emotional content of the message and the audience should be considered.
- **Time:** The element of time in a presentation of research report situation depends on various factors like availability of the room, audience and right time if he has the choice. Time is the major aspect is how much time is given to the researcher to make the presentation.

- **Place:** Researcher may not have much choice in selecting the place. But the make the best of use of the place and the facilities available will depend on the researcher. On the room arrangement depend the kind of audio-visual tools that can be used and the type of interaction that the researcher can have with the audience.
- **Cost:** The preparation of a good presentation is time consuming and expensive. Researcher could use chapter production methods and aids than the ones he has chosen to put across the message.

Audio – visual aids
These can be classified as follows:

- Audio: Tape recorder and compact disc
- Visual:
  - Non projected: black board
    - Film charts
    - Models
  - Projected:
    - epidiascope
    - Overhead
    - Slide
    - Film strip
    - Slide projector with a timer
  - Audio - Visual:
    - film and video cassettes

Usefulness of audio visual aids: Visual aids are sound, greater credibility and clarity can be achieved in presentation. Since both sound and sight service are activated at the same time along with the body language, concentration, retention and recall can be obtained in presentation of research report.

Researcher's poise: Researcher himself is an essential part of the presentation. Researcher's posture and movement on the dias or at the speaking place and his hand gestures indicate the level of confidence of the researcher to present the report. Researcher's ability to maintain eye contact with the audience and keep his facial expressions suited to the subject become also important. The fluency, pace of delivery, level of the voice and command of the language shows the level of confidence and preparation of the researcher presenting the reports.

7.9 Evaluation of the Research Report

After the report has been submitted by the researcher, he should try to get the feedback on the same. Feedback will enable him to know the deficiencies of his report, both in regard to the subject matter and the method of write-up. A detailed evaluation of the research project should be undertaken. This evaluation and review should do by panel of experts related to the subject. With respect to each stage of the researcher process, researcher may have to keep in mind some important aspects are:

- Conformity with the actual requirement of the study
- Tools used in collection of data
- Collection instrument properly designed
- Appropriate sample survey
- Adequate sample size
- Sufficient control over survey, appropriate analysis of data
- Interpretation of data
- Suitable research findings

Research report should be free from different types of experimental errors. Thus, proper method of evaluation research report can lead to an improvement in the quality of research. A scientific researcher will find considerable improvement in the quality of research undertaken by him over a period of time.

Summary

- Interpretation refers to the task of drawing inferences from the collected facts after an analytical and/or experimental study. In fact, it is a search for broader meaning of research findings.
- Researcher must give reasonable explanations of the relations which he has found and he must interpret the lines of relationship in terms of the underlying processes and must try to find out the thread of uniformity that lies under the surface layer of his diversified research findings.
- Research is essentially a cooperative venture and it is essential that every investigator should know what others have found about the phenomena under study.
- The bibliography, which is generally appended to the research report, is a list of books in some way pertinent to the research which has been done. It should contain all those works which the researcher has consulted.
- Attractive layout along with large print, many subheadings, even an occasional cartoon now and then is another characteristic feature of the popular report.
- While determining the length of the report (since research reports vary greatly in length), one should keep in view the fact that it should be long enough to cover the subject but short enough to maintain interest.
- Presentation has become an important communication medium in organisations because a research report is properly understood if it is accompanied by a presentation.
- Researcher has to think the purpose of the presentation and focus the light on research analysis sharply. Researcher can try to achieve a variety of purpose such as informing, selling, exploring, and decision making persuading and changing attitude or behavior.
- Visual aids are sound, greater credibility and clarity can be achieved in presentation. Since both sound and sight service are activated at the same time along with the body language, concentration, retention and recall can be obtained in presentation of research report.

References

- Kothari, C.R., 2004. *Research Methodology*. New Age International.
- Murthy, C., 2009.Research *methodology*, 1st ed., Vrinda Publications (P) Ltd
- *Presentation of Report* [PDF] Available at: <http://www.egyankosh.ac.in/bitstream/123456789/35380/1/Unit-13.pdf>. [Accessed 4 October 2011].
- *Interpretation of Statistical data.* [PDF] Available at: <http://www.egyankosh.ac.in/bitstream/123456789/12397/1/Unit%2b18.pdf>. [Accessed 4 October 2011].
- Report Writing, 2009. [Video Online] Available at: <http://www.youtube.com/watch?v=AFGNKJruxdg>. [Accessed 4 October 2011].
- Organising your research paper, 2008. [Video Online] Available at: <http://www.youtube.com/watch?v=97HWWy_Y1Gs&feature=related>. [Accessed 4 October 2011].

Recommended Reading

- Singh, Y. K., 2006. *Fundamental of Research Methodology and Statistics.* New Age International.
- Misra, P. R., *Research Methodology: A Handbook.* Concept Publishing Company,
- Taylor, et. al., 2006 *Research Methodology: A Guide for Researchers in Management and Social Sciences*. PHI Learning Pvt. Ltd.

Self Assessment

1. _____ refers to the task of drawing inferences from the collected facts after an analytical and/or experimental study.
   a. Interpretation
   b. Report writing
   c. Research report
   d. Research

2. _____ is communicating the facts, data and information through reports to the persons for whom such facts and data are collected and compiled.
   a. Interpretation
   b. Reporting
   c. Researching
   d. Report writing

3. _____ is the first step which is primarily concerned with the development of a subject.
   a. Preparation of the final outline
   b. Preparation of the rough draft
   c. Logical analysis of the subject matter
   d. Rewriting and polishing of the rough draft

4. Which of the following is false?
   a. Logical treatment often consists in developing the material from the simple possible to the most complex structures.
   b. Chronological development is based on a connection or sequence in time or occurrence.
   c. Research reports are the product of slow, painstaking, accurate inductive work.
   d. Rewriting and polishing of the rough draft is the first step in writing the research report.

5. The _____ should be written in a concise and objective style and in simple language, avoiding vague expressions.
   a. Final draft
   b. Research report
   c. Bibliography
   d. References

6. If research report is to be written by hand, then _____ ink should be used.
   a. Red
   b. black or blue-black
   c. black
   d. blue
   e. red and green

7. _____ are placed at the bottom of the page on which the reference or quotation which they identify or supplement ends.
   a. Page numbers
   b. Foot message
   c. End notes
   d. Footnotes

8. The first item after the number in the footnote is the _____.
   a. author's name
   b. publication
   c. edition
   d. year of publication

9. Which of the following is false?
   a. Charts, graphs and the statistical tables should be omitted in the main report.
   b. The layout of the report should be well thought out and must be appropriate and in accordance with the objective of the research problem.
   c. Abstract terminology and technical jargon should be avoided in a research report.
   d. The layout of the report should be well thought out and must be appropriate and in accordance with the objective of the research problem.

10. Match the following.

Abbreviation	Long form
1. ante	A. enlarged
2. bull	B. column
3. col	C. bulletin
4. eng	D. before

   a. 1-A; 2-B; 3-C; 4-D
   b. 1-D; 2-C; 3-B; 4-A
   c. 1-B; 2-A; 3-D; 4-C
   d. 1-C; 2-D; 3-A; 4-B

Chapter VIII

Computer – Its Role in Research

Aim

The aim of this chapter is to:

- highlight the significance of reports
- introduce evolution of computers
- define computer system

Objectives

The objectives of this chapter are to:

- explain the characteristics of computers
- elucidate binary number system
- explicate binary fractions

Learning outcome

At the end of this chapter, you will be able to:

- understand computer applications
- discuss the phases of computer
- comprehend the use of internet in research

8.1 Introduction

The development of electronic devices, specially the computers, has given added force to problem solving activity. Problems which could not be solved earlier due to sheer amount of computations involved can now be tackled with the aid of computers accurately and rapidly. Computer is certainly one of the most versatile and ingenious developments of the modern technological age. To the researcher, the use of computer is to analyse complex data that has made complicated research designs practical. Electronic computers have become an indispensable part of research students in the physical and behavioural sciences as well as in the humanities. The research student, in this age of computer technology, must be exposed to the methods and use of computers. A basic understanding of the manner in which a computer works helps a person to appreciate the utility of this powerful tool.

8.2 Significance of Computers

Computer helps in various fields of research promotion and development. It is useful in research programme.

Speed and accuracy: A computer is a high speed device, capable of taking logical decisions, performing arithmetic and non-arithmetic operations for research. The high speed may be visualised by the fact that modern computer can perform a lot of calculations relating to research. The result of research is 100% accurate.

Logical decisions for research: A computer has the capabilities to take decision relating to research. It is capable of comparing data of research project and the results are compared to take appropriate action. Computer is handling both, numerical and non-numerical data in research.

Perfect memory relating to research: The most important characteristics of a computer in which it is very large and perfect memory relating to research work. A computer is capable of recalling the information stored in its memory at the rate of very faster in a second.

Versatility: A modern computer is a very versatile machine. Computer can be used to solve the problems relating various different research projects. It may be solving a complete scientific research problem. A computer machine is capable of doing any task which can be split into a number of logical steps.

Diligence: Computer is superior to human brain in respect of memory. Moreover, even after working for long hours, there is no loss of accuracy in results.

Automation: A computer machine is automatic in operation. Computer is automatic; it means that once the research data and research instructions are fed to a computer, human interventions are not required.

Achievement of modern knowledge: The discovery of computers which is considered as the greatest achievement of modern knowledge. Without computer knowledge, no nations can exist in development.

8.2.1 Evolution of Computers

- **Different engine:** In 1823, Babbage put together a model of the first mechanical computer. It is called the difference engine. His aim was to create a machine that would compute and check tables. Towards this purpose, he combined the two check tables. He combined these two operations of sequence control and automatic control, using punched cards.
- **Analytical engine:** Babbage improved on the difference engine and designed a more sophisticated and large calculating machine in 1834. This machine was capable of working on some form of charge and calculating up to 20 decimals and about 60 additions per unit. He initiated the first notions of imputing data via a device and the storage of data and information prior to the process. It is for this pioneering effort that he is known as the Father of computer.
- **First programming:** Lord Byron's daughter is famous as the first programmer, for having devised a suitable use of the Binary Number systems for programmes and data to be fed in to the Analytical computer.

- **Boolean algebra:** George Boole actually applied the solution of alternative choices and its link with computer development. His theoretical development for representing and manipulating logical evaluations were in terms of true or false.
- **Punctual cards:** A big step forward was taken in the development of information processing in 1890 census of U.S.A. The U.S.A. Census Department was booking for a quick way of processing the census information is put on punch cards.
- **Magnetic media:** Major developments were taking place in the areas of recording devices by Mr. Valdemer Poulson. He devised the method of tapes and drums coated with thin films of magnetic materials.
- **Thermionic value:** In 1906 Lee De Forest invented the Thermionic value. This invention facilitated amplification and switching of electrical signals or pulses, without the movement of any mechanical or electrical part.
- **First automatic electro-mechanical computer:** In 1930, at Harvard University, Howard Aiken and Grace Hopper got together and arranged for the IBM to build them a general purpose computer. This machine was very reliable and is regarded by some as finished engine.
- **Generation:** The term generation refers to major developments in computer science. From 1946 onwards, the development of computers occurred at a very fast pace and the development of Electronic Digital Computers which would be much faster and have grater storage facilities. At present we are in the middle of the 5th generation. Each generation was an improvement over the former in terms of technology, speed, storage, size, reliability and system cost. Computers are classified into three categories, namely, super computers, large computers and mini computers. Super computer is primarily used for highly scientific and research purposes. Computer is playing a vital role to social research as well as scientific research. Computers are used for a variety of tasks. The most common applications are:
  - Word Processing
  - Financial analysis
  - Data base access
  - Tabulation and classification of data
  - Graphic representation of research findings and communication network for scientific research.

8.2.2 The Computer System

In general, all computer system^{s} can be described as containing some kind of input devices, the CPU and some kind of output devices. The figure below depicts the components of a computer system and their inter-relationship.

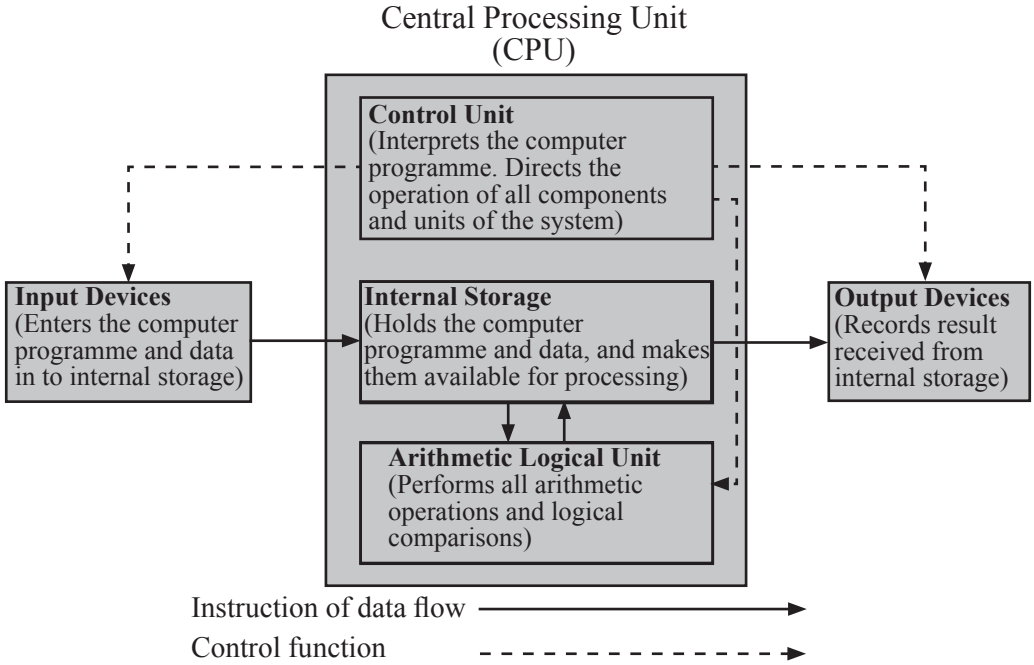


Fig. 8.1 The computer system

- The function of the input-output devices is to get information into, and out of, the CPU. The input devices translate the characters into binary, understandable by the CPU, and the output devices retranslate them back into the familiar character, i.e., in a human readable form. In other words, the purpose of the input-output devices is to act as translating devices between our external world and the internal world of the CPU, i.e., they act as an interface between man and the machine. So far as CPU is concerned, it has three segments, viz.,
  - internal storage
  - control unit and
  - arithmetic logical unit
- When a computer program or data is input into the CPU, it is in fact input into the internal storage of the CPU. The control unit serves to direct the sequence of computer system operation. Its function extends to the input and output devices as well and does not just remain confined to the sequence of operation within the CPU. The arithmetic logical unit is concerned with performing the arithmetic operations and logical comparisons designated in the computer program.
- In terms of overall sequence of events, a computer program is input into the internal storage and then transmitted to the control unit, where it becomes the basis for overall sequencing and control of computer system operations. Data that is input into the internal storage of the CPU is available for processing by the arithmetic logical unit, which conveys the result of the calculations and comparisons back to the internal storage. After the designated calculations and comparisons have been completed, output is obtained from the internal storage of the CPU.
- It would be appropriate to become familiar with the following terms as well in context of computers:
  - **Hardware:** All the physical components (such as CPU, Input-output devices, storage devices, etc.) of computer are collectively called hardware.
  - **Software:** It consists of computer programs written by the user which allow the computer to execute instructions.
  - **Firmware:** It is that software which is incorporated by the manufacturer into the electronic circuitry of computer.
  - System software: It is that program which tells the computer how to function. It is also known as operating software and is normally supplied by the computer manufacturer.
  - **Application software:** It is that program which tells the computer how to perform specific tasks such as preparation of company pay roll or inventory management. This software is either written by the user himself or supplied by 'software houses', the companies whose business is to produce and sell software.
  - **Integrated circuit (IC):** It is a complete electronic circuit fabricated on a single piece of pure silicon. Silicon is the most commonly used semiconductor—a material which is neither a good conductor of electricity nor a bad one. An IC may be small-scale, medium-scale or a large-scale depending upon the number of electronic components fabricated on the chip.
  - **Memory chips:** These ICs form the secondary memory or storage of the computer. They hold data and instructions not needed immediately by the main memory contained in the CPU.
  - **Two-state devices:** The transistors on an IC Chip take only two states—they are either on or off, conducting or non-conducting. The on-state is represented by 1 and the off-state by zero. These two binary digits are called bits. A string of eight bits is termed byte and a group of bits constitute a word. A chip is called 8-bit, 16-bit, 32-bit and so on, depending on the number of bits contained in its standard word.

8.2.3 Important Characteristics of Computers

The following characteristics of computers are note worthy:
Speed: Computers can perform calculations in just a few seconds that human beings would need weeks to do by hand. This has led to many scientific projects which were previously impossible.

Diligence: Being a machine, a computer does not suffer from the human traits of tiredness and lack of concentration. If two million calculations have to be performed, it will perform the two millionths with exactly the same accuracy and speed as the first.

Storage: Although the storage capacity of the present day computer is much more than its earlier counterpart, the internal memory of the CPU is only large enough to retain a certain amount of information just as the human brain selects and retains what it feels to be important and relegates unimportant details to the back of the mind or just forgets them. Hence, it is impossible to store all types of information inside the computer records. If needed, all unimportant information/data can be stored in auxiliary storage devices and the same may be brought into the main internal memory of the computer, as and when required for processing.

Accuracy: The computer's accuracy is consistently high. Errors in the machinery can occur but, due to increased efficiency in error-detecting techniques, these seldom lead to false results. Almost without exception, the errors in computing are due to human rather than to technological weaknesses, i.e., due to imprecise thinking by the programmer or due to inaccurate data or due to poorly designed systems.

Automation: Once a program is in the computer's memory, all that is needed is the individual instructions to it which are transferred one after the other, to the control unit for execution. The CPU follows these instructions until it meets a last instruction which says 'stop program execution'.

Binary digits: Computers use only the binary number system (a system in which all the numbers are represented by a combination of two digits—one and zero) and thus operates to the base of two, compared to the ordinary decimal arithmetic which operates on a base of ten. Computers use binary system because the electrical devices can understand only 'on' (1) or 'off' (0).

8.3 Introduction to Binary Number System

An arithmetic concept which uses two levels, instead of ten, but operates on the same logic is called the binary system. The binary system uses two symbols '0' and '1', known as bits, to form numbers. The base of this number system is 2. The system is called binary because it allows only two symbols for the formation of numbers. Binary numbers can be constructed just like decimal numbers except that the base is 2 instead of 10.

For example,

$523 \text{ (decimal)} = 5 \times 10^2 + 2 \times 10^1 + 3 \times 10^0$

Similarly,

$111 \text{ (binary)} = 1 \times 2^2 + 1 \times 2^1 + 1 \times 2^0 = 7 \text{ (decimal)}$

Thus, in the example, we see that in the decimal system, the first place is for 1s, 2nd place is for 10s and the 3rd place is for 100. On the other hand, in the binary system, the factor being 2 instead of 10, the first place is still for 1s but the 2nd place is for 2s, the 3rd for 4s, the 4th for 8s and so on.

- Decimal to Binary Conversion: A positive decimal integer can be easily converted to equivalent binary form by repeated division by 2. The method works as follows:
- Start by dividing the given decimal integer by 2. Let R_1 be the remainder and q_1 the quotient.
- Next, divide q_1 by 2 and let R_2 and q_2 be the remainder and quotient respectively. Continue this process of division by 2 until a 0 is obtained as quotient. The equivalent binary number can be formed by arranging the remainders as

$$Rk\ R_{k-1} \ldots R_1$$

where R_k and R_1 are the last and the first remainders respectively, obtained by the division process.

Illustration:
Find the binary equivalents of 26 and 45.
Solution: Table for conversion of 26 in to its Binary equivalent:

Number to be divided by 2	Quotient	Reminder
26	13	0
13	6	1
6	3	0
3	1	1
1	0	1

Table 8.1 Table for conversion

Collecting the reminders obtained in the above table we find that

26(decimal) =11010 (binary)

Or $(26)_{10} = (11010)_2$

8.3.1 Computations in Binary System

Binary addition: Binary addition is just like decimal addition except that the rules are much simpler. The binary addition rules are as shown below:

0	0	1	1
+ 0	+ 1	+ 0	+ 1
0	1	1	10

Note that sum of 1 and 1 is written as '10' (a zero sum with a 1 carries) which is the equivalent of decimal digit '2'. We can now look at two examples of binary additions which make use of the above rules. The computer performs all the other arithmetical operations (viz., ×, –, +) by a form of addition. This is easily seen in the case of multiplication, example, 6 × 8 may be thought of as essentially being determined by evaluating, with necessary carry over, 8 + 8 + 8 + 8 + 8 + 8. This idea of repeated addition may seem to be a longer way of doing things, but remember that computer is well suited to carry out the operation at great speed. Subtraction and division are handled essentially by addition using the principle of complementing.

Complementary subtraction: Three steps are involved in this method:

- Step 1: Find the ones complement of the number you are subtracting;
- Step 2: Add this to number from which you are taking away;
- Step 3: If there is a carry of 1 add it to obtain the result; if there is no carry, add 0, recomplement and attach a negative sign to obtain the result.

8.3.2 Binary Fractions

Just as we use a decimal point to separate the whole and decimal fraction parts of a decimal number, we can use a binary point in binary numbers to separate the whole and fractional parts. The binary fraction can be converted into decimal fraction as shown below:

0.101 (binary) = (1 × 2–1) + (0 × 2–2) + (1 × 2–3)

= 0.5 + 0.0 + 0.125

= 0.625 (decimal)

To convert the decimal fraction to binary fraction, the following rules are applied:

- Multiply the decimal fraction repeatedly by 2. The whole number part of the first multiplication gives the first 1 or 0 of the binary fraction.
- The fractional part of the result is carried over and multiplied by 2.
- The whole number part of the result gives the second 1 or 0 and so on.

Illustration
Convert 0.625 into its equivalent binary fraction.
Solution:
Applying the above rules, this can be done as under:
0.625 × 2 = 1.250 ® 1
0.250 × 2 = 0.500 ® 0
0.500 × 2 = 1.000 ® 1
Hence, 0.101 is the required binary equivalent.

8.4 Computer Applications

At present, computers are widely used for varied purposes. Educational, commercial, industrial, administrative, transport, medical, social financial and several other organisations are increasingly depending upon the help of computers to some degree or the other. Even if our work does not involve the use of computers in our everyday work, as individuals, we are affected by them. "The motorists, the air passenger, hospital patients and those working in large departmental stores, are some of the people for whom computers process information. Everyone who pays for electricity or telephone has their bills processed by computers. Many people who are working in major organisations and receive monthly salary have their salary slips prepared by computers. Thus, it is difficult to find anyone who in some way or the other does not have some information concerning them processed by computer".

"Computers can be used by just about anyone: doctors, policemen, pilots, scientists, engineers and recently even house-wives. Computers are used not only in numeric applications but also in nonnumeric applications such as proving theorems, playing chess, preparing menu, matrimonial matchmaking and so on. Without computers we might not have achieved a number of things. For example, man could not have landed on the moon nor could he have launched satellites. We might not have built 100 storied buildings or high speed trains and planes."

The following table depicts some of the important applications and uses of computers:

Applications	Some of the various uses
Education	
	Provide a large data bank of information
	Aid to time-tabling
	Carry out lengthy or complex calculations
	Assist teaching and learning processes
	Provide student's profiles
	Assist in career guidance
Commerce	
	Assist the production of text material such as reports, letters, circulars etc.
	Handle payroll of personal, office accounts, invoicing, records keeping, sales analysis, stock control and financial forecasting.

Banks and Financial institutions	Cheque handling
	Updating of accounts
	Printing of customer statements
	Interest calculations
Management	
	Planning of new enterprises
	Finding the best solution from several options
	Helpful in inventory management, sales forecasting and production planning.
	Useful in scheduling of projects
Industry	
	In process control
	In production control
	Used for load control by electricity authorities
	Computer aided designs to develop new products.
Communications and Transportation	
	Helpful in electronic mail
	Useful in aviation: Training of pilots,

Table 8.2 Important applications and uses of computers

8.5 Computers in Research

The computers are indispensable throughout the research process. The role of computer becomes more important when the research is on a large sample. Data can be stored in computers for immediate use or can be stored in auxiliary memories like floppy discs, compact discs, pen drives or memory cards, so that the same can be retrieved later. The computer assists the researcher throughout different phases of research process.

8.5.1 Phases of Research Process

There are five major phases of the research process. They are:

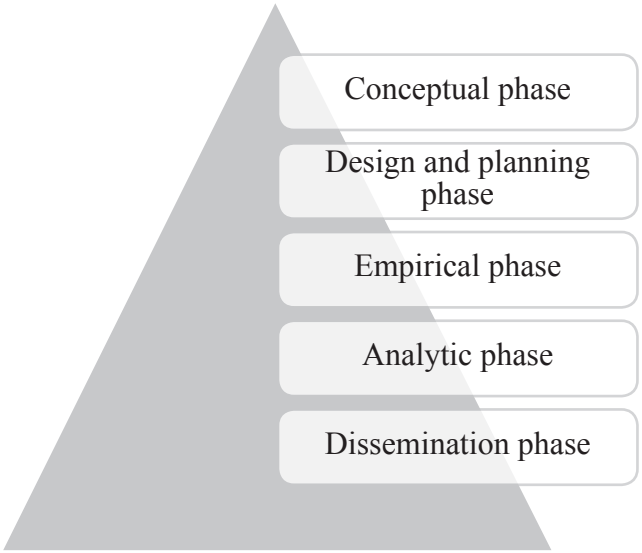


Fig. 8.2 Phases of research process

Role of computer in conceptual phase: The conceptual phase consists of formulation of research problem, review of literature, theoretical frame work and formulation of hypothesis.

Role of computers in design and planning phase: Design and planning phase consist of research design, population, research variables, sampling plan, reviewing research plan and pilot study.

Role of computers in empirical phase: Empirical phase consist of collecting and preparing the data for analysis.

Data Storage: the data obtained from the subjects are stored in computers as word files or excel spread sheets. This has the advantage of making necessary corrections or editing the whole layout of the tables if needed, which is impossible or time consuming in case of writing in papers. Thus, computers help in data entry, data editing, and data management including follow up actions etc. computers also allow for greater flexibility in recording the data while they are collected as well as greater ease during the analysis of these data.

In research studies, the preparation and inputting data is the most labor intensive and time consuming aspect of the work. Typically the data will be initially recorded on a questionnaire or record form suitable for its acceptance by the computer. To do this the researcher in conjunction with the statistician and the programmer, will convert the data into Microsoft word file or excel spreadsheet. These spreadsheets can be directly opened with statistical software for analysis.

Role of computers in data analysis: This phase consist of statistical analysis of the data and interpretation of results. Data analysis: Many software are now available to perform the mathematical part of the research process, i.e., the calculations using various statistical methods. Software like SPSS, NCSS-PASS, STATA and Sysat are some of the widely used. They can be like calculating the power of the study. Familiarity with any one package will suffice to carry out the most intricate statistical analyses. Computers are useful not only for statistical analyses, but also to monitor the accuracy and completeness of the data as they are collected.

Role of computers in research dissemination: This phase is the publication of the research study. Research publishing: The research article is typed in word format and converted to portable data format (PDF) and stored and / or published in the World Wide Web.

8.5.2 Further Uses of Computers

- Medical record linkage in longitudinal studies: In a more comprehensive way, medical record linkage provides continuous record of individual patients from birth to death, including illnesses, hospitalisation, operation, allergies and so on. Prior to the use of computers, the linking together of different events in an individual's medical case history was extremely difficult. Individual events (treatment by family doctor, stay in hospital, etc.) were, and indeed usually still are, recorded separately at different locations with no systematic procedures to bring together these elements of a case history into a single record.
- With computer based recording systems, such linking becomes more feasible, medical records can be put to better use and information stored efficiently; moreover, this information can be retrieved more rapidly than before. Cumulative files for individuals can be compiled and assembled into family groups, socio economic categories, and so on, for purposes of analysis.
- Longitudinal studies, which involve following up a given group of cohort over a period of time, are a well established application of the concept of medical record linkage, and the large and comprehensive computerised databases established through record linkage open up a wide range of potential research studies.
- For performing Meta-analysis and systematic reviews: the results from systematic reviews are graded as the level I evidence. These are obtained by the meta-analysis of the data from various studies. This is made easy by the use of specialised software for the purpose. This will be a mammoth task without a computer.
- For performing Rasch analysis: Many sophisticated software are available to analyses the psychometric properties of various outcome measures and performing a rasch analysis with the help of computers.

A note of caution

- The above description indicates clearly the usefulness of computer throughout the research process. Researchers using computers make their work faster with more accuracy and grater reliability. The developments taking place in the technology will further enhance and facilitate the use of computers for researchers.
- In spite of all these sophistications it is wise to remember that a computer is just a tool and a resource. It can only calculate or obey commands and cannot think. If the methods of handling the data are to be applied efficiently, adequate planning and suitable organisation is necessary. No facility can replace this aspect of planning. Further, it would be a disaster to replace the statistician analyses is built on sound principals of design, implementation and handling of exigencies in data collection, all of which require the expertise of a qualified statistician. The human brain remains supreme and will continue to be so for ever.

8.6 Use of Internet in Research

The internet, also known as the Information Super Highway, is possibly the greatest communication medium since the telephone. The World Wide Web (www or web) is the component of the Web and internet synonymously. The Internet and computers can be employed to aid the process of formulating a research design. The internet, in its capacity as a source of information can be useful in uncovering secondary data and also collecting primary data needed in conclusive research, descriptive research, and problem solving research, applied research, quantitative and qualitative research, exploratory research and laboratory research.

8.7 Importance of Internet

Importance of Internet is explained below.

- Source of Research Provider: There are many ways in which the Internet can be useful to researchers. It is a source of secondary data, a source of research software, and as a source for data gathering, survey etc.
- Helps in Project Management: The internet is also very useful for project management. Internet can be used to help find a job in marketing research.
- Information Supplier: Many sites include information on company history, products, clients and employees.
- Statistical Packages: The use of mainframe and Microcomputer versions of some popular statistical packages, SPSS, SAS, BMDP, Minitab and Excel.

- Computer Maps: Computer mapping which combines geographic, demographic and company information can be a valuable tool for marketing and business research.
- Sources of Secondary Data: The World Wide Web is an important source of secondary data for the researcher. Information on the web is of high quality because it comes from original sources and it is current.
- Sources of Government Data: One of the major sources of secondary data is the Central Government and State Government. The researcher can visit the Government Information Location Service. Extensive business statistics can be obtained from website. Thus the internet and computers can be used to access, analyse and store information available from secondary sources.
- E-mail Interview: The survey is written within the body of the e-mail message and sent to respondents. Respondents type the answers to either closed – ended or open – ended questions at designated places, and click or reply, responses are entered and tabulated. Data entry is typically required before any statistical analysis can be conducted.
- Consistency Check: Computer packages such as SPSS, SAS, BMDP, Minitab and Excel can be programmed to identify out-of-range values for each variable and print out the respondent code, variable code, variable name, record number, column number and out-of-range value. The correct response can be determined by going back to the edited and coded questionnaire.

Summary

- A computer is a high speed device, capable of taking logical decisions, performing arithmetic and non-arithmetic operations for research.
- The most important characteristics of a computer in which it is very large and perfect memory relating to research work. A computer is capable of recalling the information stored in its memory at the rate of very faster in a second.
- The input devices translate the characters into binary, understandable by the CPU, and the output devices retranslate them back into the familiar character, i.e., in a human readable form.
- Computers use only the binary number system (a system in which all the numbers are represented by a combination of two digits—one and zero) and thus operate to the base of two, compared to the ordinary decimal arithmetic which operates on a base of ten.
- Data can be stored in computers for immediate use or can be stored in auxiliary memories like floppy discs, compact discs, pen drives or memory cards, so that the same can be retrieved later.
- Computer mapping which combines geographic, demographic and company information can be a valuable tool for marketing and business research.
- Computer packages such as SPSS, SAS, BMDP, Minitab and Excel can be programmed to identify out-of-range values for each variable and print out the respondent code, variable code, variable name, record number, column number and out-of-range value.
- Many sophisticated software are available to analyses the psychometric properties of various outcome measures and performing a research analysis with the help of computers.
- It would be a disaster to replace the statistician analyses is built on sound principals of design, implementation and handling of exigencies in data collection, all of which require the expertise of a qualified statistician.

References

- Dhawan, S., 2010. *Research Methodology for Business and Management Studies.* Global Media Publisher.
- Kothari, C.R., 2004. *Research Methodology*. New Age International.
- *Role of Computer in Research.* [Online] Available at: <http://www.scribd.com/doc/11515555/Role-of-Computers-in-Research>. [Accessed 4 October 2011].
- *Binary Number System* [Online] Available at: <http://ed-thelen.org/comp-hist/lgp-21-prog-man4.pdf>. [Accessed 4 October 2011].
- *Binary Number*, 2007. [Video Online] Available at: <http://www.youtube.com/watch?v=zZBLD-2KgHM&feature=related>. [Accessed 4 October 2011].
- *Use Social Networking for Market Research*, 2008. [Video Online] Available at: <http://www.youtube.com/watch?v=cyGSmGFgyX4>. [Accessed 4 October 2011].

Recommended Reading

- Sachdeva, J.K., 2009. *Business Research Methodology*, Global Media.
- Singh, Y. K., 2006. *Fundamental of Research Methodology and Statistics.* New Age International.
- Geoffrey. R. and David. M., 2005. *Essentials of Research Design and Methodology*. Wiley Publishing.

Self Assessment

1. Which of the following statement is false?
   a. Major developments were taking place in the areas of recording devices by Mr. Valdemer Poulson.
   b. In 1823, Grace Hopper put together a model of the first mechanical computer.
   c. George Boole actually applied the solution of alternative choices and its link with computer development.
   d. In 1906 Lee De Forest invented the Thermionic value.

2. At present we are in the middle of the ___ generation of the computer.
   a. 5^{th}
   b. 4^{th}
   c. 7^{th}
   d. 3^{rd}

3. _____ is primarily used for highly scientific and research purposes.
   a. Computer
   b. Fastest computer
   c. Super computer
   d. Power computer

4. The function of the input-output devices is to get information into, and out of, the _____.
   a. Monitor
   b. CPU
   c. Screen
   d. Mother board

5. Data that is input into the internal storage of the CPU is available for processing by the_____, which conveys the result of the calculations and comparisons back to the internal storage.
   a. Numeric unit
   b. Logical unit
   c. Language unit
   d. arithmetic logical unit

6. _____ is the most commonly used semiconductor—a material which is neither a good conductor of electricity nor a bad one.
   a. Aluminium
   b. Glass
   c. Fibre
   d. Silicon

7. Computers use binary system because the electrical devices can understand only _____.
   a. 'on' (1) or 'off' (0).
   b. 'on' (0) or 'off' (1).
   c. 'on' (1) or 'off' (2).
   d. 'in' (1) or 'out' (0).

8. An arithmetic concept which uses two levels, instead of ten, but operates on the same logic is called the_____.
   a. Arithmetic concept
   b. Logical concept
   c. binary system
   d. secondary system

9. The _____ consists of formulation of research problem, review of literature, theoretical frame work and formulation of hypothesis.
   a. Design phase
   b. Planning phase
   c. Empirical phase
   d. conceptual phase

10. Match the following.

A	B
1. Firmware	a. It is that program which tells the computer how to perform specific tasks such as preparation of company pay roll or inventory management.
2. System software	b. It is that program which tells the computer how to function.
3. Application software	c. It is that software which is incorporated by the manufacturer into the electronic circuitry of computer.
4. Integrated circuit	d. It is a complete electronic circuit fabricated on a single piece of pure silicon.

   a. 1-a; 2-b; 3-c; 4-d
   b. 1-c; 2-b; 3-a; 4-d
   c. 1-d; 2-c; 3-b; 4-a
   d. 1-b; 2-a; 3-d; 4-c

Case Study I

(Source: Srivastava, 1958. *Business Research Methodology*, Tata McGraw-Hill Education.)

Q. The Jodhpur management programme is interested in collecting the information about the influence of textile advertisements on consumers. Draft a suitable questionnaire for the purpose.
Ans.: Given below a specimen questionnaire issued by the Jodhpur Management Programme to collect information about the influence of textile advertisements on consumers. The questionnaire by a polite letter to the respondents as given below:

Jodhpur Management Programme

Faculty of Commerce, University of Jodhpur,

Jodhpur

Mr. / Mrs. / Miss..... Date .....
.....

Dear Sir / Madam,
We are collecting information about the effect of textile advertisements on the consumers of textile goods. We hope you are conscious of the fact that millions of rupees are spent by textile companies on advertisements of their products. This money forms a part of their cost of production and is ultimately borne by you. We wish to find out whether it is worthwhile to spend so much money on advertisements which the consumer has to pay.

We would be extremely obliged to you if fill up the enclosed questionnaire and let us have you valued opinion on the various issues involved in the problem. The enquiry will help both the consumers as well as the producers of the textile goods.

Kindly fill up the enclosed questionnaire and return the same to us at your earliest. A self addressed stamped envelope is enclosed here with for your convenience.

The information supplied by you will be treated as confidential and copy of the findings will be sent to you, if you so desire.

Thanking you,

Yours truly

.....

Enclosed:

- Questionnaire
- Envelope

Jodhpur

Market Survey

Textile Advertisements and Consumers

Profile of the respondent:
Name: .....

Age: .....

Address: .....

Education: .....

Annual income: .....

Number of members in family: .....

Occupation: .....

Patterns of purchase decisions:
Do you have any predetermined idea about the clothes to be purchased when you enter the shop, in terms of:

	Mostly	Sometimes	Never
Price range			
Quality			
Textile of a particular company			
Textile exposed to you through advertisement			

If yes, list the factors that helped you most in forming this idea:

Suggestions offered by friends, relatives etc.	
Textile advertisement seen by you	
Window display in the shop	
Seeing someone else wearing it	

When you enter the shop, your enquiry begins with:

	Mostly	Sometimes	Never
Textile of a particular company			
Textile in a certain price range			
Textile which has been seen ads.			
Textile based on quality alone			

Factors influencing purchase decisions:

General Influence

What factors influence you most while taking a decision about the goods purchased.

(Rank in order of preference)

- Price
- Colour and design
- Quality and texture
- Impact of ads. Company image
- Seller's advice
- Friends' and relatives' advice
- Displays in the shop bearing company's name

Influence of the advertisements

Which type of textiles do you buy on the basis of advertisement only:

Types of Textiles	Textile for
Cotton textiles	Domestic purpose
Quality fibre textiles	Children's wear
Woollen clothes	For wife
Any textile offering incentives	For self
	Giving as gifts

Do you think the textiles the ads, have an impact on you:

	Mostly	Sometimes	Never
Only while looking at it			
Some time after having seen it			
When future purchases are planned			
It is retained in your memory when you enter the shop			
While making the purchase in the shop			
Only when incentives are offered through the advertisements			

Do you think textile ads seen by you are really helpful in your decision making:

Types of advertisement	Very helpful	Helpful	Not at all
Printed ads (newspaper etc.)			
Radio jingles			
Visual ads (cinema, T.V.) If yes, how?			

How do you learn about the companies' name and their products:

	Mostly	Sometimes	Never
By visiting the shop			
Through associates who patronise a particular company			
Through advertisement			
Old established companies, the image of which is retained			

Among the textile ads, to which you are exposed, mention those which you remember most (the company's name).

_____

_____

_____

You remember those ads, because

- They appear very often
- The slogan is very catchy
- Ads are of the most established companies
- Ads are very attractive
- They furnish the desired details

Your suggestions

What factors do you think should be incorporated in the ads so as to make it really helpful in decision making (rank the suggestion in order to preference).

1. They must furnish details in terms of

- Price range
- Product range
- Quality
- Colour range
- Place of availability
- Uses of the textile

2. Should appear often
3. Advertisement clothes should be made available i local markets
4. Should be made more attractive
5. Should be technicolor
6. Should be only in black and white
7. Should be in good taste
8. Any other:

Case Study II

Mattel: Toy Manufacturer

The world's largest toy maker Mattel Inc. recalled 1.5 million fisher-price infant toys world wide which were made in China, because of possible lead paint hazard for the children soon after in another incidence 9 million more Chinese made toys including popular 'Barbie', 'Poly pocket' and 'cars' were recalled because of lead paint and small magnets that could be swallowed by young ones. India has more than 130 million children below the age or 6 yrs. When children do chew and even swallow substances Indian toy manufac¬turer association wants to gauge the awareness among the Indian parents about the hazard of lead paint & small parts or the toys.

Source: Pune University, 2008. *Marketing Research Question paper* [Online] Available at: <http://www.careerbricks.com/marketing-research-question-papers-pune-university>. [Accessed 18 October 2011].

Questions:

1. Prepare a Research proposal to be submitted to the toy manufacturers association, highlighting the Research Method used by you.
2. Design a suitable tool for data collection, in order to collect data regarding price, availability of toys in the Indian Market.
3. Mention the sampling techniques and sample size to be included in the research

Case Study III

Neckties have been around in India for past several decades and now form an inevitable of men's formal dressing. Before 1960's, ties were mostly imported from France, England, Italy and a couple of other countries. It was 'Zodiac' which gave the impetus to domestic tie industry in the mid sixties. This was followed by a Delhi-based small company-Zorex Tie Manufacturing (P) Ltd., the ties were sold under the brand Zorex, which later took the name Bentlay. Zorex specialises only in neckwear, whereas other companies treat this segment as an additional item in Menswear market.

Today, the tie market in flooded with a number of brands. Bentlay is facing a stiff competition from the companies of Korea, Japan and China. There is a shift from domestic to foreign brands. This trend fuelled by liberalisation of imports is expected to gain further momentum in the years to come. Apart from this the domestic producers have also been hit by the onslaught of denim wear and casual dressing, thus reducing usage of ties. Bentlay is thinking seriously sponsor a comprehensive study on necktie market to study its customer profile, its image vis-a-vis other domestic and foreign brands and the segmentation pattern of tie market. Assume that you have been approached to conduct the market research.

Source: Pune University, 2008. *Marketing Research Question paper* [Online] Available at: <http://www.careerbricks.com/marketing-research-question-papers-pune-university>. [Accessed 18 October 2011].

Questions

1. Make a proposal for the market research to be submitted to the client.
2. Give detailed Research Design for the proposed study.
3. What tools of data collection will be used by you?

Bibliography

References

- *2nd Part Quantitative Analysis,* 2009. [Video Online] Available at: <http://www.youtube.com/watch?v=oxkRAInbwg&feature=related>. [Accessed 4 October 2011].
- *4.3 Different data capture methods, 2010. [Video Online] Available at:* <http://www.youtube.com/watch?v=Hyh91AC_tAM>. [Accessed 4 October 2011].
- *5.5 How sample size is determined, 2010.* [Video Online] Available at: <http://www.youtube.com/watch?v=53EI30_uTkQ&feature=related>. [Accessed 4 October 2011].
- Agarwal, S. C. and Khurana, S. K. *Research Methodology and Statistical Analysis* [Online] Available at: <http://books.google.co.in/books?id=3OEf_n4sxgYC&pg=PA301&dq=sampling+plan+in+research+methodology&hl=en&ei=Hk1oTsX5EpCGrAeFvbz3Cg&sa=X&oi=book_result&ct=result&resnum=7&ved=0CFgQ6AEwBg#v=onepage&q=sampling%20plan%20in%20research%20methodology&f=false pg 301 >. [Accessed 4 October 2011]
- *Attitude measurement and scales* [Online] Available at: <http://www.egyankosh.ac.in/bitstream/123456789/7661/1/Unit-6.pdf>. [Accessed 4 October 2011].
- *Binary Number System* [Online] Available at: <http://ed-thelen.org/comp-hist/lgp-21-prog-man4.pdf>. [Accessed 4 October 2011].
- *Binary Number,* 2007. [Video Online] Available at: <http://www.youtube.com/watch?v=zZBLD-2KgHM&feature=related>. [Accessed 4 October 2011].
- *Collection of data.* [PDF] Available at: <http://www.egyankosh.ac.in/bitstream/123456789/25801/1/Unit2.pdf>. [Accessed 4 October 2011].
- *Data analysis - example from a research project, 2011.* [Video Online] Available at: <http://vimeo.com/28346996>. [Accessed 4 October 2011].
- *Data Analysis. [Online] Available at: <http://www.drtomoconnor.com/3760/3760lect07.htm>. [Accessed 4 October 2011].*
- *Data Collection,* 2009. [Video Online] Available at: <http://www.youtube.com/watch?v=Z4frC2h8SzQ&feature=related>. [Accessed 4 October 2011].
- Dhawan, S., 2010. *Research Methodology for Business and Management Studies.* Global Media *Publisher.*
- Fowler, F., 2002. *Survey research methods,* Sage Publications.
- *Interpretation of Statistical data.* [PDF] Available at: <http://www.egyankosh.ac.in/bitstream/123456789/12397/1/Unit%2b18.pdf>. [Accessed 4 October 2011].
- *2011. Introduction to Statistics: Levels of Measurement,* [Video Online] Available at: <http://www.youtube.com/watch?v=B0ABvLa_u88&feature=related >. [Accessed 4 October 2011].
- Kothari, C.R., 2004. *Research Methodology.* New Age International.
- *Levels of Measurement and Scaling* [Online] Available at: <http://www.fao.org/docrep/W3241E/w3241e04.htm>. [Accessed 4 Oct 2011].
- *Measurement and scaling techniques* [Online] Available at: <http://www.egyankosh.ac.in/bitstream/123456789/2404/1/Unit%205.pdf>. [Accessed 4 Oct 2011]
- *Methods and techniques of data collection.* [PDF] Available at: <http://www.egyankosh.ac.in/bitstream/123456789/35367/1/Unit-4.pdf>. [Accessed 4 October 2011].
- Murthy, C., 2009. *Research methodology,* 1st ed., Vrinda Publications (P) Ltd.
- *Organising your research paper,* 2008. [Video Online] Available at: <http://www.youtube.com/watch?v=97HWWy_Y1Gs&feature=related>. [Accessed 4 October 2011].
- *Presentation of Report* [PDF] Available at: <http://www.egyankosh.ac.in/bitstream/123456789/35380/1/Unit-13.pdf>. [Accessed 4 October 2011].

- *Processing of data.* [Online] Available at: <http://www.egyankosh.ac.in/bitstream/123456789/2405/4/Unit%206.pdf>. [Accessed 4 October 2011].
- *Report Writing,* 2009. [Video Online] Available at: <http://www.youtube.com/watch?v=AFGNKJruxdg>. [Accessed 4 October 2011].
- *Research Design.* 2011. [Video Online] Available at: <http://www.youtube.com/watch?v=H2v9Xh2iQV8&feature=related>. [Accessed 4 October 2011].
- *Research Methodology,* [Online] Available at: <http://www.ihmctan.edu/PDF/notes/Research_Methodology.pdf>. [Accessed 4 October 2011].
- *Research Methodology,* [Online] Available at: <https://www.tutorsindia.com/research-methodology-help.html>. [Accessed 4 October 2011].
- *Research Problem, Hypotheses and Variables* [Online] Available at: <http://www.scribd.com/doc/16739725/Research-Problem-Hypotheses-and-Variables>. [Accessed 4 October 2011].
- *Role of Computer in Research.* [Online] Available at: <http://www.scribd.com/doc/11515555/Role-of-Computers-in-Research>. [Accessed 4 October 2011].
- Ross, S. M. and Morrison, G. R., *Experimental Research Methods* [PDF] Available at: <http://www.aect.org/edtech/ed1/38.pdf>. [Accessed 4 October 2011].
- *Sampling methods, 2008.* [Video Online] Available at: <http://www.youtube.com/watch?v=wHgI6TolySw&feature=related>. [Accessed 4 October 2011].
- *Scales.,* 2009. [Video Online] Available at: <http://www.youtube.com/watch?v=GMIpnzLQtTQ >. [Accessed 4 October 2011].
- Singh, Y. K. and Bajpai, R. B., 2008. Research methodology: techniques and trends, APH Publishing Corporation.
- Sridhar, S. M. *Research Methodology* [Online] Available at: <http://www.scribd.com/doc/2034905/Research-Methodology-Part-5-Sampling-Sampling-Strategy-or-Plan>. [Accessed 4 October 2011].
- *The Research Problem* [Online] Available at: <http://www.public.asu.edu/~kroel/www500/The%20Research%20Problem.pdf>. [Accessed 4 October 2011].
- *The Research Problem,* 2011. [Video Online] Available at: <http://www.youtube.com/watch?v=d3mwAK_lI1E>. [Accessed 4 October 2011].
- *Types of Research & Research Designs -- Rey Ty.*, 2009. [Video Online] Available at: <http://www.youtube.com/watch?v=49_u-pXwx7g>. [Accessed 4 October 2011].
- *Use Social Networking for Market Research,* 2008. [Video Online] Available at: <http://www.youtube.com/watch?v=cyGSmGFgyX4>. [Accessed 4 October 2011].
- *When Should I Use Qualitative Vs. Quantitative Research?,* 2009. [Video Online] Available at: <http://www.youtube.com/watch?v=638W_s5tRq8&feature=related>. [Accessed 4 October 2011].

Recommended Readings

- Dane, C. F., 2010. *Evaluating Research: Methodology for People Who Need to Read Research.* Sage Publications.
- Geoffrey. R. and David. M., 2005. *Essentials of Research Design and Methodology.* Wiley Publishing.
- Goddard , W., and Melville, S., 2004. *Research Methodology:, An Introduction.* Juta and Company Ltd.
- Jonker, J, and Pennink S., 2009. *The Essence of Research Methodology: A Concise Guide for Master and PhD* Students in Management Science*.,* Springer Publication.
- Khan., *Research Methodology.* APH Publishing.
- Khanzode, V.V., 2004. *Research Methodology.,* APH Publishing.
- Krishnaswamy , K.N ., Sivakumar, I. A, and Mathirajan, M., 2009. *Management Research Methodology: Integration of Methods and Techniques.* Pearson Education India.
- *Methodology: Integration of Methods and Techniques,* Pearson Education India.

- Kumar, R., 2010. *Research Methodology: A Step-by-Step Guide for Beginners.* Sage Publication Ltd.
- Marczyk, G. R., DeMatteo. D and Festinger, D., 2010. *Essentials of Research Design and Methodology.* John Wiley and Sons.
- Misra, P. R., *Research Methodology: A Handbook.* Concept Publishing Company,
- Panneerselvam, R., 2004. *Research Methodology.* PHI Learning Pvt. Ltd.
- Phophalia, A.K., 2010. *Modern Research Methodology: New Trends and Techniques.* Global Media.
- Sachdeva, J.K., 2009. *Business Research Methodology,* Global Media.
- Saunders.,2003. *Research Methods for Business Students,* Pearson Education India.
- Singh, Y. K., 2006. *Fundamental of Research Methodology and Statistics.* New Age International.
- Taylor, et. al., 2006 *Research Methodology: A Guide for Researchers in Management and Social Sciences.* PHI Learning Pvt. Ltd.
- William, W., 2009. *Research Methods in Education.* Pearson Education India.

Self Assessment Answers

Chapter I

1. a
2. b
3. c
4. c
5. d
6. a
7. b
8. c
9. a
10. a

Chapter II

1. a
2. b
3. d
4. a
5. c
6. d
7. a
8. b
9. c
10. c

Chapter III

1. a
2. b
3. a
4. d
5. c
6. a
7. c
8. a
9. b
10. b

Chapter IV

1. a
2. c
3. d
4. a
5. b
6. c
7. a
8. c
9. a
10. a

Chapter V

1. a
2. b
3. d
4. b
5. c
6. a
7. d
8. a
9. d
10. b

Chapter VI

1. a
2. b
3. c
4. d
5. a
6. c
7. d
8. a
9. c
10. b

Chapter VII

1. a
2. b
3. c
4. d
5. a
6. b
7. d
8. a
9. a
10. b

Chapter VIII

1. b
2. a
3. c
4. b
5. d
6. d
7. a
8. c
9. b
10. d