SYLLABI-BOOK MAPPING TABLE

Research Methodology

Syllabi	Mapping in Book
Unit I **Meaning of Research:** Objectives, types, approaches, significance of research, methods Vs methodology, research process, flow chart, criteria of good research, problems encountered by researches in India. **Measurement Scales:** Nominal scales, ordinal scales, interval scale, ratio scale, source of errors in measurement, tests of sound measurement, test of validity, test of reliability, test of practicability.	**Unit 1:** Meaning of Research **(Pages 3-24)**
Unit II **Meaning and Definition of a Research Problems:** Selection, sampling, steps, types, sample size, testing hypothesis-I, (parametric or standard test of hypothesis), testing hypothesis-II (non parametric or distribution), free tests.	**Unit 2:** Meaning and Definition of a Research Problem **(Pages 25-68)**
Unit III **Sampling Design:** Census and sample survey, implications of sample design, steps in sampling survey, types of universe, sampling unit, source list, size of sample, parameters interest, budgetary constraint, sampling procedure, criteria of selecting a sampling procedure, characteristics of a good sample design, different types of sample design, how to select a random sample, random sample from an infinitive universe complex random sample design. **Sampling Fundamentals:** Need for sampling, some fundamental definitions, important sampling distributions, central limit theorem, sampling theory, Sandier's A test, concept of standard error, estimation, estimating the population, mean, estimating population proportion, sample size through the approach based on precision rate and confidence level, deterministic sample size through the approach based on Bayesian statistics.	**Unit 3:** Sampling Design **(Pages 69-114)**
Unit IV **Research Design:** Meaning of research design, need for research design, features of a good design, important concepts relating to research design, different research designs, basic principles of experimental design, conclusions, developing research plan. Experimental Designs: Between group designs, within group designs, mixed designs, latin square designs. **Non experimental Design and Correlational Methods:** Non and quasi-experimental designs, correlational designs, newer social methods, advanced correlational methods, discriminant function analysis. **Qualitative Methods:** Definition and aim of qualitative research, construction of reality, subjectivity, dynamic research, process of documentation of qualitative research.	**Unit 4:** Research Design **(Pages 115-165)**
Unit V **Computer and Its Role in Research:** Computer applications, the computer system, important characteristics, the binary number system, computers and researchers. Interpretation and Report Writing: Meaning, reasons, techniques, precautions, steps in report writing, layout, types, oral presentation and precautions in report writing.	**Unit 5:** Role of Computers in Research, Interpretation and Report Writing **(Pages 167-189)**

CONTENTS

INTRODUCTION

NOTES

Doing research on the patterns and epidemiology of psychotic disorders forms an important part in the study of psychology. The book *Research Methodology* discusses different research techniques starting from sampling and research design to different statistical techniques.

Unit 1 discusses the meaning, definition and scope of research. The unit also explains the basics of research process, different types of research and the difference between research methods and methodology.

Unit 2 throws light on how to define a research problem and formulate a hypothesis. The unit further explains parametric and non-parametric tests and guides how to choose different kinds of tests to test a hypothesis.

Unit 3 explains the concept of sampling design and different types of sampling procedures. The unit also explains probability and non-probability sampling techniques, different types of sampling errors and how to minimize them.

Unit 4 explores the concept of research design and different types of researh. It also discusses how to differentiate between the qualitative and quantitative research.

Unit 5, examines how computers have revolutionized the research work. The unit also explains how to write a research report, interpret the results and draw conclusions.

Each unit, in this book, is supplemented with Summary, Key Terms, Answers to 'Check Your Progress', Questions and Exercises and Further Reading sections, to aid the reader.

UNIT 1 MEANING OF RESEARCH

NOTES

Structure

1.0 INTRODUCTION

In layman's terms research means the search for knowledge. Scientific research is a systematic and objective way of seeking answers to certain questions that require inquiry and insight or that have been raised on a particular topic. The purpose of research, therefore, is to discover and develop an organized body of knowledge in any discipline.

Research is a journey of discovery. It is a solution-oriented inquiry that must be objective and repeatable. It should inspire and guide further studies and should foster applications. Research will provide practical benefits if it can provide advanced understanding of a discipline or suggest ways to handle some situations that we confront.

Scientific research involves controlled observations, analysis of empirical data and interpretation of findings. This can further lead to the development of concepts, generalizations, etc., on the basis of which theories could be formulated. Such an investigation could help in determining cause and effect relationship. The

NOTES

ultimate aim of social science research is the control and prediction of behaviour.

Knowing how to do research: Every researcher should have the necessary training in gathering data, organizing materials suitably and engaging in field or laboratory work, as required. He should also have the competence in using statistics for treating the data and the ability to interpret the data collected meaningfully.

Training of the mind: Research needs discipline, right mental makeup, the ability to manage time effectively, objectivity, logical thinking, the capacity to evaluate the results of the research and ability to carefully assess the findings that are found by the research.

Research data allows people to make informed decisions by extrapolating the findings from the field or laboratory on to real life situations. This is the practical application of the findings generated by research.

Research is also a way of preparing the mind to look at things in a fresh or different way. Out of such on orientation would come new and innovative observations about everyday events and happenings. This is how originality comes about in research. Some of the most outstanding discoveries have been made in the most serendipitous manner. Some outstanding results have been obtained by researchers who had kept their minds open and free of clutter. This enabled them to see startlingly new connections.

1.1 UNIT OBJECTIVES

After going through this unit, you will be able to:

- Explain the meaning of research
- Learn the objectives and types of research orientations that are possible
- Examine the significance of research
- Know the distinction between research methods and methodology
- Gain insights into the research processes

1.2 MEANING OF RESEARCH

Research in common parlance refers to search for knowledge. One can also define research as a scientific and systematic search for pertinent information on a specific topic. In fact, research is an art of scientific investigation. According to the Advanced Learner's Dictionary of Current English, 'research is a careful investigation or enquiry, especially a thorough search for new facts in any branch of knowledge.' Redman and Mory (1923) defined research as a 'systematized effort to gain new knowledge.' Some people consider research as a voyage of discovery that involves movement from the known to the unknown.

NOTES

Research in a technical sense is an academic activity. Clifford Woody defined research as an activity that comprises defining and redefining problems, formulating a hypothesis; collecting, organizing and evaluating data; making deductions and reaching conclusions; and carefully testing the conclusions to determine if they support the formulated hypothesis. D. Slesinger and M. Stephenson, in the *Encyclopaedia of Social Sciences*, defined research as 'the manipulation of things, concepts or symbols for the purpose of generalizing, extending, correcting or verifying the knowledge, whether that knowledge aids in the construction of theory or in the practice of an art.' Research is thus an original contribution to the existing stock of knowledge making for its advancement.

Principles of a typical research:

- It is based an empirical data
- It involves precise observations and measurements
- It is aimed at developing theories, principles and generalizations
- There are systematic, logical procedures involved
- It is replicable
- The findings of the research need to be reported

A flow chart in terms of how research proceeds is presented in Section 1.9 for better understanding where the criteria of good research have been identified in easy steps.

Finally, one is exposed to some of the specific problems that researchers in India have to face. Measurement scales are introduced for the benefit of the researcher in terms of their mathematical properties. The major scales that have been widely used have been presented and discussed in some detail. The scales are:

- Nominal scale
- Ordinal scale
- Interval scale and
- Ratio scale

Their differences and significance in terms of use, sources of error in measurement have been dealt with in detail the following sections so as to facilitate precise measurements in research work.

The qualities of good or sound measurement have been described for the researcher to be aware of. The three major tests of effective measurements are specified and discussed at length. These are tests of reliability, validity and practicality.

1.3 OBJECTIVES OF RESEARCH

NOTES

The objective of any research is to find answers to questions through the application of scientific procedures. The main aim of any research is exploring the hidden or undiscovered truth. Even though each research study has a specific objective, the research objectives in general can be categorized into the following broad categories:

- Exploratory or formulative research studies: These are aimed at gaining familiarity with a particular phenomenon or at gaining new insights into it.
- Descriptive research studies: These are aimed at accurately portraying the characteristics of a particular event, phenomenon, individual or situation.
- Diagnostic research studies: These studies try to determine the frequency with which something occurs.
- Hypothesis testing research studies: These studies test a hypothesis and determine a causal relationship between the variables.

1.4 TYPES OF RESEARCH

The basic types of research are as follows:

- **Descriptive vs. analytical research:** Descriptive research includes different kinds of fact-finding inquiries and surveys. The main objective of this research is describing the state of conditions as it exits at the present moment. For descriptive research studies in the social sciences, we often use the term ex post facto research method whose main characteristic is the researcher has no control over the variables; he can report only what has happened or what is happening. Used often for descriptive studies, most ex-post-facto research projects, seek to measure such phenomena as preferences of people, frequency of shopping, or similar data. The methods used in descriptive research include all kinds of survey, comparative and correlation methods. On the other hand, in analytical research, the researcher uses the information and facts already available, and analyzes them to make a critical evaluation.
- **Applied vs. fundamental research:** Research can be either fundamental (basic or pure) or action-oriented (applied) research. Fundamental research focuses on finding generalizations and formulating theories whereas applied research aims at finding a solution for an immediate problem facing a society or a business/industrial organization. Gathering knowledge for knowledge's sake is termed pure or basic research (Pauline V. Young). Research studies carried out with a view to making generalizations about human behaviour are examples of fundamental research. Research aimed at particular

conclusion (for a particular problem). Doing research on a current social or business problem is an example of applied research.

- **Quantitative vs qualitative**: Quantitative research is based on the measurement of quantity or amount. This research can be applied to the phenomena that can be quantified. On the other hand, qualitative research focuses on qualitative phenomena such as happiness, beauty, etc. that cannot be strictly quantified. 'Motivation Research' related to human behaviour is a type of qualitative research.
- **Conceptual vs empirical research:** Conceptual research is concerned with some abstract theory or idea(s). Philosophers and thinkers generally use the conceptual research for developing new concepts and for reinterpreting the existing. Empirical research, on the other hand, relies only on real experiences and observations. It is data-based research and its conclusions can be verified by observations or experiments. It is also called experimental type of research. In empirical research all facts are obtained at first hand, at their source, and at times by stimulating the production of desired information. To prove a given hypothesis, the evidence gathered through empirical studies and experiments is considered to be the most powerful and accurate.
- **Some other types of research:** All other different types of research are variants or combinations of the approaches discussed above. Research can also be classified based on the conditions in which research is carried out, for example, the purpose of research, the environment in which research is done, the time required to accomplish research, etc. From the time point of view, research can be regarded as either as one-time or longitudinal research. If the research is confined to a single time-period, it is called one-time research whereas the research carried over several time-periods is considered as longitudinal research. Depending upon the environment in which research is to be carried out, it can be termed as laboratory research or field-setting research or simulation research. Clinical or diagnostic research involves using case-study methods or in-depth approaches into the causes or events. Clinical or diagnostic research methods also use very deep probing data gathering devices and very small samples to obtain the necessary data. Historical Research method uses historical sources like documents, archeological remains, archives, old information, etc., to study past events and ideas, including the ideas, thoughts, philosophy of persons and groups at distant time periods. Conclusion-oriented and decision-oriented research are two other classifications, where in conclusion-oriented research, a researcher is free to pick up a problem, redesign the enquiry as he proceeds and conceptualize as he sees fit. In decision-oriented research, the researcher to take the decision-maker into confidence and take his suggestions at every stage of his research. It has to be much focused.

NOTES

1.5 APPROACHES TO RESEARCH

Quantitative approach and qualitative approach are the two basic approaches to research. These two paradigms are based on two different and competing ways of understanding the world. These competing ways of comprehending the world are reflected in the way the research data is collected (e.g., words vs numbers), and the perspective of the researcher (perspectival vs objective). The perspectives of the participants are very critical.

(i) **Quantitative approach:** If there has been one overwhelming consensus among academic psychologists on a single point over the past few decades, it is that the best empirical research in the field is firmly grounded in quantitative methods. In this approach, data is generated in quantitative form and then that data is subjected to rigorous quantitative analysis in a rigid and formal fashion. Inferential, experimental and simulation approaches are the sub-classifications of quantitative approach. Inferential approach to research focuses on survey research where databases are built studying samples of population and then these databases are used to infer characteristics or relationships in populations. In experimental approach, greater control is exercised over the research environment and often, some variables (independent) variables are controlled or manipulated to record their effects on other variables (dependent variables). In simulation approach, artificial environment is constructed within which relevant data and information is generated. This way, the dynamic behaviours of a system are observed under controlled conditions.

(ii) **Qualitative approach:** This approach to research is concerned with subjective assessment of attitudes, opinions and behaviour. Research in such a situation is a function of researcher's insight and impressions. Such an approach to research generates results either in non-quantitative form or in the forms which are not subjected to rigorous quantitative analysis.

Below table gives types of research, methods employed and techniques used by those types of research.

CHECK YOUR PROGRESS

1. What is scientific research?
2. Differentiate between fundamental and applied research.
3. Differentiate between conceptual and empirical research.
4. Differentiate between quantitative and qualitative approach.

NOTES

Table 1.1 Types of Research

	Type	Methods	Techniques
1.	*Library*	(i) Analysis of historical records (ii) Analysis of documents	Recording of notes, content analysis, tape and film listening and manipulations, reference and abstract guides, content analysis.
2.	*Field Research*	(i) Non-participant direct observation	Observational behavioural scales, use of score cards, etc.
		(ii) Participant observation	Interactional recording, possible use of tape recorders, photographic techniques.
		(iii) Mass observation	Recording mass behaviour, interview using independent observers in public places.
		(iv) Mail questionnaire	Identification of social and economic background of respondents.
		(v) Opinionnaire	Use of attitude scales, projective techniques, use of goniometric scales.
		(vi) Personal interview	Interviewer uses a detailed schedule with open and closed questions.
		(vii) Focused interview	Interviewer focuses attention upon a given experience and its effects.
		(viii) Group interview	Small groups of respondents are interviewed simultaneously.
		(ix) Telephone survey	Used as a survey technique for information and for discerning opinion; may also be used as followup questionnaire.
		(x) Case study and life history	Cross-sectional collection of data for intensive analysis, longitudinal collection of data of intensive character.
3.	*Laboratory Research*	Small group study of random behaviour, play and role analysis	Use of audio-visual recording devices, use of observers, etc.

1.6 SIGNIFICANCE OF RESEARCH

Research involves developing a scientific temper and logical thinking. The significance of research-based answers can never be underestimated. The role of research is specially important in the fields of Economics, Business, Governance, etc. Here research helps in finding solutions to problems encountered in real life. Decision-making is facilitated by applied research. Research is also of special significance in the operational and planning processes of business and industry. Here logical and analytical techniques are applied to business problems to maximize profits and minimize costs. Motivational research is another key tool in understanding consumer behaviour and health related issues. Responsible citizenship concerns can all be addressed through good research findings. Social relationships involving issues like attitudes, interpersonal helping behaviour, environmental concerns like crowding, crime, fatigue, productivity and other practical issues are all capable of being addressed well by scientific research.

NOTES

Social science research is extremely significant in terms of providing practical guidance in solving human problems of immediate nature.

Research is also important as a career for those in the field of academics. It could be a career option for professionals who wish to undertake research to gain new insights and idea generation. Research also fosters creative thinking, and new theorizations.

Research for its own sake and for the sake of knowledge and for solving different problem is all require formal training in scientific methodology.

1.7 METHOD VS. METHODOLOGY

Research methods: They refer to all the methods the researchers use while studying the research problems and while conducting research operations. In general, the research methods can be categorized into the following three groups:

(i) The first group includes the methods that are concerned with the data collection.

(ii) The second includes the statistical techniques needed for mapping relationships between the unknowns and the data;

(iii) The third group contains the methods necessary to evaluate the accuracy of the results obtained.

Research methodology: It is the procedure that helps to systematically proceed in steps to solve a research problem. Research methodology is a broader concept that includes not the research methods; but also the logic behind the research methods in the context of a particular research study; and it explains the reasons for using particular research methods and statistical techniques. Research mythology also defines how the data should be evaluated to get the appropriate results.

CHECK YOUR PROGRESS

5. What is the significance of research?
6. Differentiate between methods and methodology.
7. What are the criteria for a good researchs?

1.8 RESEARCH PROCESS

Research process includes steps or a series of actions and logical sequence of those steps to carry out research effectively. The various steps in a research process are not mutually separate, exclusive or discrete, but they at the same time need not

NOTES

always follow each other. The researcher, at each step, anticipates subsequent steps and the requirements. The tentative order of the steps and the procedural guidelines of the research process are as given below:

(i) **Formulating the research problem:** At the very beginning of research, the researcher must clearly define the research problem, i.e., the area of interest, the matter to be to be inquired into, etc. The problem, before being solved, is initially stated in a broader perspective and then the researcher arrives at the specific question by gradually reducing the ambiguities, if any. Then, immediately after formulating the problem, the feasibility of different solutions is studied before choosing the right solution.

(ii) **Extensive literature survey:** After formulating the research problem, a brief summary of it should be prepared—this is an essential step. While writing a Ph.D. thesis the researcher has to prepare a synopsis of the topic and submit it to the appropriate committee or research board for approval. Synopsis preparation needs extensive survey of the literature connected with the problem.

(iii) **Development of a working hypothesis:** After surveying the literature, the researcher should clearly state the working hypothesis, which is a tentative assumption made before testing it in logical or empirical sequences. Hypothesis must be as specific as possible and should be limited to the intended research. This helps to choose the right process.

(iv) **Preparing the research design:** The next step, after clearly defining the research problem, is preparing the suitable research design. The research design includes the conceptual framework within which research would be carried out. A good and planned research design helps to carry out the study in an efficient manner saving time and resources. It helps to gather the most useful information and assists in arriving at the accurate results. Simply put, a good research design facilitates the collection of relevant evidence with minimal expenditure of money, effort, time and other resources.

(v) **Determining sample design:** A universe or population includes all the items under inquiry. If all the items in the population are inquired then such an inquiry is called census inquiry. In a census survey, all the items are covered and so the highest accuracy is obtained. But this may not be practicable in surveys involving a big population. Census surveys need huge amounts of time, money and energy. Hence, quite often it is wise to select only a few items from the universe for study purposes. Technically, such a small and convenient number of items selected, is called a sample. Specified plan of the size and method of collecting the sample is technically known as sample design.

(vi) **Collecting the data:** In most cases, the data at hand is insufficient and there is always a need of fresh data. There are different ways of collecting the appropriate data which differ considerably in terms of relevance,

NOTES

expenditure, time and other resources. Therefore, the researcher must select the most appropriate method of collecting the data after considering objective of the research, the nature of investigation, time and financial resources available, scope of the inquiry, and the desired degree of accuracy.

(vii) **Execution of the project:** This is an important step in the research process because if the execution proceeds on the correct lines, the collected data would be dependable, adequate and accurate. Therefore, systematic and timely execution of a project plays a crucial role in ensuring the right results at the end.

(viii) **Analysis of data:** After collecting the data, the next step is analyzing the data. The data analysis includes a number of closely related operations like specifying different categories of data, differentiating and tabulating the data into different categories, applying the statistical techniques and formulae to the data, doing the right calculations and then drawing statistical inferences. Various tests, such as chi-square test, *t*-test, *F*-test, etc. help in data analysis.

(ix) **Hypothesis-testing:** After analyzing the data, the researcher should test the working hypothesis against the statistical inferences obtained after analyzing the data. The question now should be answered: Do the findings support the working hypothesis or they contradict?

(x) **Generalizations and Interpretation:** If a hypothesis is tested and upheld sufficient number of times, the researcher can arrive at a generalization. The degree of success of a research is calculated on the basis of how much the arrived generalizations are close to the acceptability. If the researcher starts with no hypothesis, the researcher will interpret his findings on the basis of some existing theory and this is known as interpretation. The process of interpretation often triggers new questions which lead to further researches.

(xi) **Preparation of the report or the thesis:** Finally, the researcher has to prepare the report of what has been studied. Report must be written with great care keeping the following layout in mind:

a. The preliminary pages: The preliminary pages of the report should contain the title, the date, acknowledgments, foreword, table of contents, list of tables, list of graphs and charts (if any).

b. The main text: The main text of the report should have introduction, summary of findings, main report, conclusion and suggestions for future research.

c. The closure: At the end of the report, appendices should be listed in respect of all technical data, followed by bibliography. Index terms should also be given specially in a published research report. All references should be cited as per the research writing formats.

1.9 FLOW CHART

NOTES

RESEARCH PROCESS IN FLOW CHART

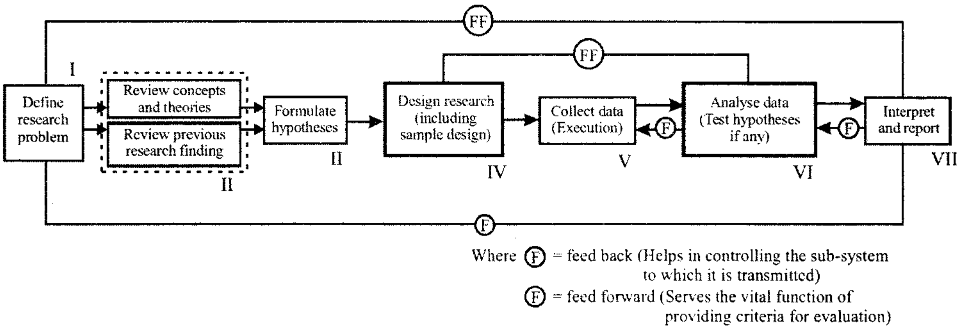


***Fig. 1.1** Research Process*

In figure 1.1 the flow chart indicates the sequential steps to be followed in the research process. One must start with defining the research problem along with reviewing the relevant literature in the field to become familiar with the concepts and theories relevant to the issue to be investigated. The next step is the formulation of the hypothesis, which is followed by the research design and sample selection. Then the collection of data and its analysis is to be attempted. After that the interpretation and the report writing stages complete the research report. These have to be written step by step and then edited and refined several times before preparing the final report.

1.10 CRITERIA OF GOOD RESEARCH

Whatever be the type of research one undertakes, certain common criteria of good scientific methods have to be followed. A good research follows logical methods, is systematic, and structured in accordance with well defined sets of rules and practices to enable the researcher in arriving at dependable conclusions. Both, deductive reasoning and inductive reasoning should be followed for meaningful research.

Good research also implies obtaining reliable data which provides sound validity to the research findings.

The following principles underlie a good research criteria:

- The aim and objective of the research being conducted should be clearly specified;
- The research procedure should be replicable so that if the research needs to be continued or repeated, it can be done easily;
- The research design should be so chosen that the results are as objective as possible.

NOTES

- Interpretation of any research should be done keeping in mind the flaws in the procedural design and the extent to which it has an effect on the results.
- Research should be carried out systematically. It should progress in pre-defined stages, and researchers should avoid using their intuition or guesswork to arrive at conclusions.
- Research should be logical so that it is meaningful, and help in decision making.
- Research should be empirical as far as possible.
- The results of the research should only be used and generalized for the population for which the data provides an adequate basis.
- The validity and reliability of the data used in research should be double checked.
- Further, good research produces results that are examinable by peers, methodologies that can be replicated, and knowledge that can be applied to real-world situations.

1.11 PROBLEMS ENCOUNTERED BY INDIAN RESEARCHERS

There are some common problems faced by researchers in developing countries and India is no exception. Basically, there is a dearth of the tools required for good research. Many of the universities and research institutions are now providing computers with Internet connection to researchers but the facilities provided are not adequate. Luckily, the costs of both hardware and Internet bandwidth are coming down. While Indian researchers have now easy access to these tools, there is still the problem of low visibility of papers published by them. Indian researchers are often de-motivated to do further research. Other factors like lack of scientific training in the methodology of research and a non-existent code of conduct also serve as challenges for the Indian researcher. There is also insufficient interaction between researchers and the end-users. End-users of research are the ones who stand to benefit from research and if they are not made aware of the benefit they can gain, getting sponsors to provide funds for research would be difficult.

There is also a lack of safeguards against any violation of confidentiality in data collection. Research studies that overlap one another are undertaken often and this leads to unnecessary repetition. There is an absence of research culture in our country.

Other problems that Indian researchers face that are common to developing countries are:

- Limited or no access to international research journals
- Lack of infrastructure except in a few metropolitan cities
- Low investment in research due to financial constraints

- Inadequate library facilities and where such facilities exist, they are not easily accessible
- Poor encouragement to do research

These problems need to be surmounted effectively in order to promote research as a professional activity.

NOTES

1.12 MEASUREMENT SCALES

There are three important properties that make one scale of measurement different from another: (i) Magnitude, (ii) Equal intervals, and (iii) Absolute zero.

(i) Magnitude refers to the property of 'moreness'. Any attribute that is being measured can be more, less or equal at one instance as compared to another instance. For example, Vijay is taller than Karan.

(ii) Equal intervals: This means the difference between any two points at any place on the scale has the same meaning as the difference between two other points that differ by the same number of scale units. For example, the difference between 20 kg and 30 kg on a weight measurement is the same as between 80 kg and 90 kg, that is, 10 kg.

(iii) Absolute zero: It is a condition when nothing of the property being measured, exists. For example, if there is no pulse measured the situation is alarming, for the body.

For many psychological tests such a condition does not exist. For example, in case of anxiety, there cannot be a 'zero' for there can be no situation where there is absolutely no anxiety.

Any measurement is a yardstick for evaluation. It involves some form of judgment. Measurement can be made of physical objects, people, situations, etc. It can be quantitative like height, area, etc. or qualitative like personality traits or abstract (like patriotism). Measurement is a process of assigning some value to the observed phenomena. There are clearly accepted rules under which these evaluations are to be made. Some tasks are easy while others are complex and difficult. Some are direct, like weight, while others like motivation, leadership etc., have to be inferred. Some measurements are very precise whereas others are abstract and tentative. The attempt of all measurement is to achieve confidence in the evaluations that are made.

Measurements are generally presented in the form of a scale in a range. For example, if the movie is watched by men, women and children, we assign a 0 to men 1 to women and 2 to children in order to tabulate our observations of the audience. Similarly the people who attend alone, or as a couple or in a group can be allotted alphabets like A, C or G for classification purposes. This is the process of artificially determining categories. These can further be divided qualitatively or descriptively. There are four levels or types of scales for measurements—nominal scale, ordinal scale, interval scale and ratio scale.

NOTES

Nominal scales: This is a system of measurement where a number is given to label an event. These are merely convenient labels and do not have any special meaning or significance. The numbers assigned do not have any quantitative value. They do not indicate any order or distance. For example, one player of a baseball team is given jersey no. 3 and another 5. This is neither a rank, nor a distance between them. It is only a way of categorizing them as a team member. Nominal scale Measures offer a Count of Data. This is a useful way of classification.

Ordinal scales: The ordinal scale places events in an order. It is a ranking method. For example, Ram stood first in the class and Sita came second. Bharat was third. However this scale does not make the intervals equal. The real differences between Ram, Sita and Bharat may not be equal despite their being adjacent in rankings. This is a 'greater than, lesser than' measurement.

Interval scale: Here, the intervals are adjusted to be equal in terms of some established rule. For example, the Fahrenheit scale. If the zero '0' is accepted as the starting point of the scale (water freezes at this point) then measurement can be made from this absolute zero point. So when temperature rises from 30 to 40 it has the same increment as from 60 to 70. But this does not mean that 60 temperature is twice as warm as 30. This is because both the numbers are based on the value of 0. The interval of 30–40 and 60–70 is the same but the ratio of 30 and 60 means nothing because '0' is an arbitrary point. Also in psychological tests the difference in IQ between 70 and 80 is not the same between100 and 110. Here the 10 points do not have the same meaning.

Ratio scale: Here there is an 'absolute zero' for measurement. For example, the '0' on a centimeter scale, '0' time, '0' errors in a typed page, etc. The most precise of the scales is the ratio scale because it is easy to make comparisons. This makes the scale fairly powerful. IQs are also presented in a ratio form, to facilitate comparisons. The nominal scale is the least precise while the ratio scale is the most precise, with the other two scales falling in between. Behavioral scientists have to use the interval scale because of the nature of their discipline.

However, most behavioural disciplines widely use the interval scale.

1.13 SOURCES OF ERROR IN MEASUREMENT

All measurement should be error free in any research, ideally. Since this is not always possible, it is important to be aware of some of the sources of errors:

- **Subject or respondent factors:** Many respondents may be unwilling to participate in a study. So, if they are persuaded to give their responses, they could be negatively disposed or be indifferent in their replies. This could have a significant bearing on the results. Also, a fatigued, anxious or bored subject may not provide the true-to-his-heart responses. Thus the subject could contribute to the source of error in the data collected.

- **Situation:** This involves the context in which any measurement is made. If the context places any type of stress on the individual, then the response would reflect it. Such a response, therefore, would not be a true index of the topic of measurement. Even the mere presence of another person can affect the responses obtained. It is important to keep the situation balanced and free of distorting influences.
- **The interviewer / measurer effect:** The way questions are put forth can reflect the bias of the experimenter / researcher. The style, mannerisms, appearance, language, expression, etc., all could distort the presentation. Any casualness in carrying out the measurement could have large implications in terms of the data obtained.
- **Instrument/ tools:** Malfunctioning of the equipment could instantly lead to errors in the measures obtained. If language is an important component of the measurement, then ambiguous words, inadequate expressions, incomplete questions, etc., could all lead to measurement errors. All these factors must be borne in mind, and carefully neutralized or completely eliminated as for as possible to get accurate information.

1.14 TEST OF SOUND MEASUREMENT

Every good measuring tool must be subjected to the tests of validity, reliability and practicality.

Validity is the most important of these three criteria. It means that the test must measure what it purports to measure. For example, if a test is a measure of speed, then speed is what that test should measure.

Reliability implies the consistency with which a test measures what it seeks to measure.

Practicality refers to the costs involved in administering a test, the time needed, the convenience and the ease of carrying out the measurements as well as the usefulness of the obtained data, besides the interpretation.

1.14.1 Test of Validity

A simple way to determine validity is to ask the question: Is one measuring what is being thought that the test is measuring? Another way to determine the validity of a test is the accuracy with which specific predictions can be made from the test scores. This is determined by comparing the scores obtained with some external test scores as standard. This means one can assess the individual's present status or predict his / her future status with respect to some type of functioning. This comparison with other relevant evidence is one of the better ways of establishing the validity of an instrument.

NOTES

Types of validity:

(i) **Content validity:** This implies the extent to which the measuring instrument has covered adequately all the aspects of the topic that is to be measured. If the measure includes a representative sample of the population, then the content validity would be assumed to be good. This can be determined by a panel of judges who evaluate the contents of an instrument for yielding the measures the tool is intended to measure. Another way of establishing the content validity is intuitive judgments involving the theme of the measuring tool.

(ii) **Criterion-based Validity:** Here the success of the measuring instrument is determined by the ability of the scores obtained to predict some outcomes of a current condition. For example, those who train hard can be predicted to be winners. The extent of training can thus be reflected in the number of winners.

A good research criterion must posses the following characteristics:

a. **Relevance:** If the criterion is judged to be the proper measure

b. **Freedom from bias:** This is reached if every test taker has an opportunity to score well. (It is not biased in favour of any group)

c. **Reliability:** The measure is stable across several administrations.

d. **Availability:** The information specified by the criterion must be accessible and available.

i. Predictive validity: The test must be useful in predicting some future performance from present measures. This is known as predictive validity.

ii. Construct validity: This involves finding out if the individual possesses a trait or a characteristic. The degree of the existence of this trait is also of interest to the assessor. This is the construct on which a test rests; Example: Anxiety, Intelligence etc. So, the instrument is validated against the construct. The greater the ability of the instrument measuring the construct, the higher will be the validity of the tool.

iii. Pragmatic validity: This refers to the usefulness of an instrument as an indicator or as predictor of some behaviour in the future. So, the validity of a measure rises from this ability of the tool. Under this approach to validity, two sub-types are included:

- Concurrent validity: This seeks to see how closely the scores relate to other known measures of validity.
- Predictive validity: Here the object of the scores is to predict the future situation/ behaviours on the basis of the scores obtained from the tool, at the present.

If these criteria are met by a tool used for measurement, then the instrument is thought to be valid. The measures obtained from such a tool can be viewed as a correct estimate of the feature under assessment.

1.14.2 Test of Reliability

The test of reliability of a measuring device is its ability to yield consistent results from one set of measures to another.

NOTES

A valid instrument is always reliable

However, a reliable instrument is not always a valid one. For example, a machine may consistently under-weigh by 2 kg. It is a reliable machine, but it is not valid as a good measure of weight.

If an instrument is reliable, then temporary and extraneous factors would not affect the measures obtained.

There are two aspects to reliability.

(i) **Stability:** The extent to which consistent results are obtained with repeated measurements with the same instrument on the same individual. A measure of stability is obtained by comparing the results of repeated measurements.

(ii) **Equivalence:** This is estimated by comparing the measures obtained by two assessors on the same aspect/situation or individual.

Reliability can be enhanced by three procedures

(i) By standardizing the conditions under which measurements are made. Here, all extraneous factors can be kept under control.

(ii) By systematizing the directions for measurement.

(iii) By training personnel suitably, e.g., technicians who are trained for measuring blood pressure. Also, having larger samples from the person on whom the measurement is done.

Types of reliability

There are three common methods of estimating reliability:

(i) Retest reliability

(ii) Internal consistency reliability

(iii) Parallel forms / or alternate forms / or equivalent forms. Here, a single form of the test is administered twice on the same sample with a reasonable time gap.

(i) **Retest reliability:** Two measures yield independent sets of scores. The two scores when correlated would give the value of the reliability coefficient. Such a coefficient is also known as the temporal stability coefficient. This means how far the examinees retain their relative position as measured in terms of test scores over a given period of time. The ideal time gap between the two administrations is about 15 days.

(ii) **Internal consistency reliability:** This indicates the homogeneity of the test. If all the items of a test measure the same function or trait, then the test

NOTES

is seen to be homogeneous and the internal consistency would be high. The most common way of determining internal consistency reliability is by the split-half method. Here the items to be tested are divided into two equal or nearly equal halves. Another way to split a test is by using the odd and even numbered items. This method is preferred to the regular split-half method because in a power test, the first half would normally be made up of the easier items while the second half would have the tougher items. The odd numbers 1, 3, 5, 7, 9 and so on and even numbers of, 4, 6, 8, and so on would balance the items. Each examiner would receive two scores, i.e., the scores of the odd and those of the even numbers of a given test. Thus, from a single administration of the same test, two sets of scores are generated. A 'product moment correlation' is computed to obtain the reliability of the half test. On the basis of the reliability of this half test, the reliability of the whole test is estimated. The Spearman/Brown formula is used for estimating the reliability of the whole test.

This is a useful method as it eliminates two administrations of the same test. A quick estimate of reliability is possible. This is a kind of on-the-spot reliability measure. The demerit of this method is that it cannot be used for power tests.

The Kuder–Richardson formula is used for determining the reliability coefficient in a test where the terms are scored as 0 or 1 or (right or wrong). This estimates the coefficient alpha which yields a measure of internal consistency.

Alternate forms reliability: These are also known as parallel forms or equivalent forms or comparable forms reliability. This requires that the test be developed in two forms which are comparable or equivalent. Two forms of the test are administered to the same sample either immediately the same day or in a time interval of a fortnight. When the reliability is calculated on the basis of data obtained from the two administrations of the test, it is the alternate form reliability. Pearson coefficient (r) between two sets of scores obtained from two equivalent forms becomes the measure of reliability. Such a coefficient is known as the coefficient of equivalence. Alternate forms reliability measures the consistency of the examinee's scores between two administrations of parallel forms of a single test.

The biggest problem with this procedure is making the two forms of a test, truly equivalent.

Criteria for judging whether the forms are parallel:

- The number of Items in both the forms, should be same
- The item-difficulty levels in both the forms should be similar
- Mode of administration of both forms should be the some.

Scorer reliability: This kind of reliability is important in tests of creativity, projective tests, etc. This is the reliability which can be estimated by having a sample of test independently scored by two or more examiners. The two sets of scores obtained by each examiner are completed and the resulting correlation coefficient is known as scorer reliability.

Test–Retest reliability, internal consistency reliability, and parallel forms reliability all express reliability in terms of the correlation coefficient.

NOTES

1.14.3 Test of Practicality

The practicality characteristic of a measuring instrument can be estimated in terms of convenience, cost effectiveness, ease of administration, scoring, interpretation, etc. These factors have an important bearing on the development of the most suitable measuring instruments needed for research.

1.15 SUMMARY

- Research is done to find the solution to a problem, or to know more about something, or to know new things. Scientific research involves systematic, controlled, empirical and critical examination of a hypothesis or proposition about the relations in a phenomenon.
- The important steps needed for conducting scientific research are as follows:
  - (i) Identifying the problem;
  - (ii) Formulating a hypothesis: once the problem has been identified and the literature review has been completed; the hypothesis has to be formulated;
  - (iii) Identifying the variables to be manipulated:
    (a) The independent variable, (b) The dependent variable, and (c) The extraneous variables;
  - (iv) Formulating a research design;
  - (v) Carrying out the observations and measurements – This involves the utilizing of the tools of the study, for obtaining data; and
  - (vi) Summarizing the results;
  - (vii) Statistical Treatment;
  - (viii) Drawing conclusions on the basis of the study and deciding upon applications, further research, etc.
- There are four different ways of assigning numerals to the attributes of any event:
  - (i) Nominal scale,
  - (ii) Ordinal scale, and
  - (iii) Interval or equal interval scale, which includes the characteristics of the nominal and ordinal scales of measurement, and
  - (iv) Ratio scale.

NOTES

- Some of the common sources of errors in measurement arise from:
  (i) Respondent,
  (ii) Measurer,
  (iii) Situation, and
  (iv) Test instrument.
- Sound measurements require that the criteria of validity, reliability and practicality be met. Types of validity are content validity, criterion related validity and construct validity.
- Reliability refers to the consistency of a score. There are two types of reliability: (i) Test–Re-test method and (ii) Internal consistency reliability
- Test of practicality refers to the aspects of administering a test in terms of costs involved, ease of administration and scoring, convenience of interpretation, etc. Economic and convenience factors are of significance in testing.

1.16 KEY TERMS

- **Fundamental research:** A research which focuses on finding generalizations and formulating theories.
- **Applied research:** A research which aims at finding a solution for an immediate problem facing a society or a business/industrial organization.
- **Conceptual research:** A research that is concerned with some abstract theory or idea(s).
- **Empirical research:** A research that it relies only on real experiences and observations.
- **Quantitative approach:** In this approach the data is in the form of quantities which is then subjected to mathematical and statistical approaches.
- **Qualitative approach:** Approach that deals with data that cannot be strictly quantified, for example, opinions, tastes, attitudes, etc.

1.17 ANSWERS TO 'CHECK YOUR PROGRESS'

1. A scientific research seeks answers in a systematic and objective way to certain questions. It requires inquiry and insight.
2. Fundamental research focuses on finding generalizations and formulating theories. It gathers knowledge for use in distant future and is termed pure or basic or fundamental research. On other hand, applied research aims at finding a solution for an immediate problem facing a society or a business/ industrial organization.

NOTES

3. Conceptual research is concerned with some abstract theory or idea(s). Empirical research, on the other hand, relies only on real experiences and observations. It is data-based research and its conclusions can be verified by observations or experiments.
4. In quantitative approach, the data is in the form of quantities which is then subjected to mathematical and statistical approaches. Qualitative approach deals with data that cannot be strictly quantified, for example, opinions, tastes, attitudes, etc.
5. Research is born out of the spirit of inquiry. Research inculcates a scientific temper and logical thinking. The outcome of research is useful in policy-making, for solving economic and business problems and for promoting industrial growth. Social scientists look for research-based answers for societal problems. Research is the process for stimulating intellectual growth and building a knowledge base.
6. Research methods are the techniques used for conducting a research. Methodology includes not only the methods for doing research but the logic behind the need to use a particular method. The application of certain statistical procedures also forms a part of the methodology. Decisions about size and, type of data required are methodological concerns.
7. For a research to be good, it should be systematic, logical, empirical and replicable.

1.18 QUESTIONS AND EXERCISES

Short-Answer Questions

1. List the different approaches to scientific research.
2. State the objectives of research.
3. What are the criteria of good research?
4. State the problems encountered by researchers in India.
5. State the major measurement scales.

Long-Answer Questions

1. Describe different types of research.
2. Explain the research process in detail.
3. Discuss the objectives types and approaches to scientific research.
4. Examine the concepts of validity and reliability along with the various tests that are available for ensuring them.
5. List and explain the problems faced by researchers in India.

1.19 FURTHER READING

NOTES

Kerlinger, F.N. 1973. *Foundations of Behavioral Research.* New York: Holt Reinhart and Winston, Inc.

Kothari, C.R. 1995. *Research Methodology: Methods and Techniques*, New Delhi: Wishwa Prakashan.

Singh, A.K. 2008. *Tests, Measurements and Research Methods in Behavioural Sciences.* New Delhi: Bharati Bhawan.

UNIT 2 MEANING AND DEFINITION OF A RESEARCH PROBLEM

NOTES

Structure

2.0 INTRODUCTION

In this unit, you will learn how to define a research problem. You will also learn how to select the problem and define it clearly and specifically. The next step is formulating a working hypothesis. A working hypothesis is a tentative proposition that provides the solution, in the researchers view, to the defined problem. The characteristics of a good hypothesis are presented in the unit.

The next step is testing the hypothesis. For testing the hypothesis there are two kinds of tests, namely, parametric and non-parametric tests. This unit will teach you how to choose the right type of test to test a hypothesis, depending upon the data. This unit presents a detailed examination of parametric and non-parametric statistical tests. The different types of tests under each category have been presented and discussed.

2.1 UNIT OBJECTIVES

After going through this unit, you will be able to:

- Explain how to define a problem and formulate a hypothesis
- Examine the different parametric and non-parametric tests in detail
- Know how to choose from different types of tests to test a hypothesis

2.2 RESEARCH PROBLEM AND WORKING HYPOTHESIS

NOTES

Clearly identifying and defining a research problem as the first step before beginning any investigation.

The concept of a research problem: There are a few specifications that have to be met for a research problem to exist:

(i) There must be an entity occupying a given environment and there should exist some uncontrolled values therein.

(ii) There must exist at least two courses of action namely A1 and A2 to be pursued. A course of action is defined by one or more values of the controlled variables, for example, during festival time, the purchase of certain types of items is one course of action.

(iii) There must be at least two possible outcomes, O1 and O2 due to the course of action. At least one of the outcomes should be of interest to the researcher for his study. Then that becomes the objective of research.

(iv) The courses of action must have some bearing on the objective. Further; the person concerned does not know which course of action is the best. Then, a problem can be seen to exist. So, the elements that go into making a research problem are:

  a. There should exist some difficulty or problem.
  b. There must be some objective.
  c. There must exist at least two means to reach the objective.
  d. The relative efficiency of the two alternatives must be considered for proceeding with the inquiry.

(v) There must be a given environment in which the difficulty exists and persists.

A research problem is an existing problem to which the researcher wants to find out the best solution, and the course of action to be taken to attain the objective. The action has to be executed in a given context or environment. Several factors may influence the outcomes and these have to be kept in mind, while carrying out the research.

Selecting the problem: The research problem must be carefully selected to get dependable results. Some tips for choosing a problem.

- Avoid choosing a topic that has already been researched, extensively.
- Avoid controversial topics, while starting a research career.
- Costs, time, competency, etc., need to be kept in mind.
- Carry out a small pilot study, before launching into the final investigation, if possible.

- The boundaries within which a problem to be studied should be clearly defined. This is to be followed by first posing a problem for the study then determining the processes and techniques for the study.

Steps involved in research

(i) Prepare a general statement of the problem to be studied

(ii) Surveying the available literature

(iii) Developing ideas through intuition and discussions

(iv) Formulating a working definition

Statement of the problem: Begin with a broad field and then narrow it down to specifics in order to search a limited area in terms of a specific, single problem.

Understand the problem: This is done in terms of its origin background and other factors.

Survey the available literature: This is to become familiar with the relevant theories in the field of study and all other records of research work that have been carried out in the field so far. Knowing about the field helps one to focus on the research topic, better.

Developing ideas: Ideas can be developed through one's own intuition, reading the background literature and through discussions with others in the related field.

Phrasing the research problem: This means stating the research problem in operational terms. A statement of all technical terms should be made clearly. The basic postulates of research should be identified. A reasonable time period is to be specified, for the study. The scope of the investigation and the limitations within which the study is to be carried out is to be specified.

Example 2.1: Why are goods manufactured in India so much cheaper than those produced in the West?

This is a general problem. Then it could be examined, discussed and formulated as a working hypothesis. A working hypothesis is a supposition based on previous observations that is offered as an explanation for the observed phenomenon and is used as the basis for further observations, or experiments. The task of research is to formulate a research problem, study it systematically and then test the hypothesis for acceptance or rejection.

Example 2.2: What are the factors responsible for the Chinese goods being cheaper than similar Indian products? Here the focus must be on cost factors and pricing methods adopted by the manufacturers in the two countries.

To conclude, research has to be carried out by being well informed, developing a hunch about the topic to be studied by extensive reading about related work, discussions and if possible carrying out a pilot project. Then, the working

NOTES

definition needs to be fully operationalized, the techniques, procedures identified, sample selected and the timeframe, costs, etc., have to be estimated. Then the research has to begin.

A few conditions about a research question:

- It is clear and unambiguous
- It is fairly specific
- If identifies the type of data to be collected
- The questions asked are related to each other
- It is a worthwhile attempt

The problem statement of any research is written in the form of a question.

Examples 2.3:

(i) What is the relationship between stress and performance?

(ii) Does learning proceed better in democratic environment or in an authoritarian structure?

A problem also expresses the relationship between two or more variables. So, one of the variables could be manipulated and its effects on the other, studied.

Example 2.4: Is higher anxiety related to lowered performance?

A problem should be testable by empirical methods. Data obtained should be testable.

It is wise to avoid ethical and moral statements like for example should gay people be allowed to join the defiance services. The problem chosen should not be very trivial, extremely broad or general that cannot be studied effectively. Similarly excessive specificity can also be a limiting factor.

Considerations for selecting a problem: Before a problem for investigation is selected, several questions have to be asked.

(i) Is the problem significant in terms of the variables to be studied and its contribution to research, social theories, practices, etc.

(ii) Can useful data be gathered for the study?

(iii) Is there any duplication that could happen?

(iv) Is the problem operationally workable?

(v) Are there good instruments to measure?

A problem is worth studying, if the answers to most of these questions are positive.

How does a problem show up for study?

Does a gap in the information exist?

Contradictory results occur when certain facts are unwarranted from the existing knowledge.

Research problems are of two types

(i) **Solvable:** For solvable problem an answer is possible. A hypothesis can be formulated.

(ii) **Unsolvable:** For unsolvable problem, no answers are possible.

Example 2.5: Is there life after death?

Meaning and characteristics of a good hypothesis

While conducting research, the first step is to identify the problem. The next step is to researcher should formulate a hypothesis which is a tentative solution in the form of a testable proposition.

This testable proposition is called a hypothesis. Hypothesis is, therefore, a suggestion about a possible solution. Hypothesis seeks to study the relationship between two variables. It is a potential association between variables. It is a conjecture of the relationship between two variables.

Differences between a problem and a hypothesis

Problem:

1. Is not testable
2. Is stated in the form of interrogative sentences.

Example 2.6: What is the relationship between reward and learning?

Hypothesis: It is testable and is stated in the form of a declarative sentence.

Example 2.7: Reward facilitates learning

Criteria of a good research hypothesis

(i) A good hypothesis should be clear and based on operationally defined concepts.

(ii) It must be testable.

(iii) It should be short, clear, specific and simple (in the use of words).

(iv) It should relate to the existing body of theory and fact.

(v) It should display logic and comprehensiveness.

(vi) It should be general in scope.

(vii) It should be in accord with other existing hypotheses in the field of study

Types of hypothesis

- Causal hypothesis
- Descriptive hypothesis

Causal hypothesis explains a tentative influence of one factor on another.

Example 2.8: Frustration is the cause of aggression (causal).

Descriptive hypothesis: Merely describes the behaviours.

NOTES

Example 2.9: Frustration frequently could result in aggression.

Other types of hypothesis are simple hypothesis, complex hypothesis, research hypothesis, null hypothesis, and statistical hypothesis.

Simple hypothesis: It contains only one or two variables

Complex hypothesis: It has more than two variables.

Research hypothesis: It is derived from researcher's theory. It is called a working hypothesis

Null hypothesis: The null hypothesis is in a sense the opposite of the research hypothesis. It is a hypothesis that negates differences. In fact, researchers are interested in rejecting this hypothesis; so that the reverse, i.e., the research hypothesis is accepted.

Statistical hypothesis: This is also known as the alterative hypothesis (H_1). Here the null hypothesis (H_0) and the alternative are expressed in numerical terms.

Example 2.10: On the basis of the existing psycho-social theory, one would predict that two groups would differ on the measure of anxiety. So, the research hypothesis would predict that the two groups would differ.

To test this hypothesis, it is stated in an operational way; as the alternative hypothesis, that is H_1. The operational way to state the alternative hypothesis would be that the mean anxiety scores of these two groups would differ or are unequal. The null hypothesis (H_0) would imply that the mean of the anxiety scores of the two groups is the same. If the data suggests that H_0 is to be rejected then H_1 is true and supports the research hypothesis.

CHECK YOUR PROGRESS

1. What is the meaning of research and how does one define a research problem?
2. How does a researcher select a research problem?
3. How is a research problem stated?
4. What are the different types of research problems?
5. What is a hypothesis?
6. What is a null hypothesis?
7. What is statistical hypothesis.

2.3 SAMPLE SELECTION, TYPES AND SIZE

NOTES

Sample selection

All research in the field of behavioural sciences involves drawing inferences from a specified, identifiable group on the basis of a selected sample. The clearly identifiable and specified group is known as the population or universe. The selected group of persons or objects is called the sample. The conclusions are drawn from the sample, which are deemed to be valid to the entire population. Such conclusions are known as the statistical inferences.

A population can be finite or infinite. A finite population is one where all the members can be counted. An infinite population is one where all the members cannot be counted (e.g., stars in the sky). A population can be imaginary or real.

A measure based upon the entire population is called a parameter. A measure based upon the sample is called a statistic.

A sample is limited number or set of persons or elements that are chosen from a population, according to some plan. It is thought to represent the population. Samples are based on probabilities. Probability is a form of relative frequency. For example, the probability of a seeing a head when coin is tossed once is 1/2 or 0.5. Probability is expressed as a fraction or in decimal numbers.

Need for sampling

Sampling is needed for several reasons:

Sampling is economical in terms of time and money. It saves time as well as cost because only a limited sample is involved, not the entire population;

Sampling is carried out by trained personnel, so it has considerable accuracy in measurement and testing;

Sampling errors can be easily determined from the sampling process. This yields valid information about the population characteristics; and

Sampling is the only available procedure when the population is infinite;

Sampling enables fairly accurate generalizations about the population from the study of the sample.

Steps in sampling

Sampling is the plan for obtaining a sample from a given population. There are several steps that a researcher must keep in mind while selecting a sample:

The type of universe or population to be studied can be finite or infinite.

Sampling unit: This stands for a geographical area like a district or province, taluk, etc., or it could be a social unit like a family, school, club, etc. The researcher has to decide the factors that would be studied in advance. A list of all the sampling units is called a sampling frame. This is the source list and contains all the items of a universe or population.

NOTES

Types of sample

Most samples can be categorized into two types:

(i) Probability sampling

(ii) Non-probability sampling.

(i) Probability sampling: It sample is based on the concept of random selection or chance sampling. Here, every item of the universe has an equal chance of inclusion in the sample. It is a form of lottery method where the units are chosen from the whole group by a mechanical method. This is almost a blind selection. Random sampling ensures that the law of Statistical Regularity is followed. This implies that if the sample chosen is a random one, the chances are that the sample will have the same composition and characteristics as the universe. This is why random selection is considered as the most useful method for obtaining a representative sample. Probability sampling must follow the conditions given below:

a. The size of the parent population must be known to the investigator;
b. Each element of the parent population must have an equal chance of being included in the sample;
c. The sample size needed must be clearly specified.

The major probability sampling methods are:

a. Simple random sampling;
b. Stratified random sampling – (a) Proportionate stratified random sampling, and (b) Disproportionate stratified random sampling;
c. Area or cluster sampling.

(ii) Non-probability sampling: This is a method in which there is no way of determining the probability that each item in the population gets included in the sample. This means that there is no basis for estimating how closely the characteristics of a sample approximate the parameters of the population from which the sample has been drawn. This is due to the absence of random selection procedures. The major techniques used in non-probability sampling are:

a. Quota sampling
b. Accidental sampling
c. Judgmental or purposive sampling
d. Systematic sampling
e. Snowball sampling
f. Saturation sampling
g. Dense sampling

NOTES

Size of the sample

Size refers to the number of items selected from the universe to constitute a sample.

The size of the sample should be neither too large, nor too small. An optimum sample size should be:

a. Efficient
b. Representative
c. Reliable
d. Flexible

The sample size should be decided by the level of precision needed and the estimate of the confidence level, desired. The size of the population variance is an important determiner. If the population variance is large, then a larger sample size is indicated. The size of the population is another factor to be kept in mind. This limits the size of the sample

Other aspects to be considered for determining sample size:

(i) *The parameters of interest:* 1. Estimation of some characteristics or some proportion of persons in a population; 2. Knowing some average measure of a population; and 3. Knowing about some sub-groups of a population. All these types of estimates have a bearing on the sample size.

(ii) *Budgetary constraint:* – From a purely practical point of view, the size of the budget is bound to influence the size of the sample.

(iii) *Sampling procedure:* the choice of the method in selecting the sample would have to provide a balance between the cost involved and the least possible sampling error that would result from the sample size.

2.4 PARAMETRIC AND NON-PARAMETRIC TESTS

These tests are utilized extensively in behavioural research.

A parametric statistical test is one in which certain conditions, about the population parameter from which the sample is drawn, are specified.

The conditions for parametric testing are:

- The observations must be independent, i.e., the selection of one case must be dependent upon the selection of another case.
- The observations must be drawn from a normally distributed population
- The samples drawn from a population must have equal variances. This is more so, when the sample size is small. This is the condition of homogeneity of variance
- The variables must be expressed in interval or ratio scales. Nominal or ordinal scales do not qualify for a parametric test.

NOTES

- The variable studied should be continuous. Examples of paramedic tests are z test, t test and F test.

A non-parametric test is one which does not specify any conditions about the parameters of the population from which the sample is drawn. These are also called distribution free statistics because no assumptions are made about the distribution of the population. For these tests also, the variables must be continuous and the observations should be independent. But these conditions do not apply rigidly.

Examples of non-parametric tests are the chi-square test, the Mann-Whitney U test, Kendall's T, Kendall's coefficient of concordance W, etc.

Conditions that suggest the use of the non-parametric test:

- When the shape of the distribution of the population is not a normal one.
- When the variables have been expressed as nominal measures or frequency counts
- When the variable have been expressed as ordinal measure (ranking)

These tests are less precise and therefore run the risk of not rejecting the null hypothesis, when it is false.

Advantages of non-parametric tests:

- These tests allow the statistics to be computed from simple formulae.
- Greater applicability
- Calculations faster, specifically when the size of sample is small
- Assumptions being easier, less violations occur.
- Nominal and ordinal scales measures are acceptable
- These tests are useful for mall samples. But as the size increases, the statistics become cumbersome.
- Non-parametric tests are more efficient and convenient, specifically when the assumptions of paramedic tests are not met.

Disadvantages: When the sample size is large (above 30) they give lower statistical efficiency. If the assumptions of parametric statistics are met, the use of non-parametric tests means a waste of useful data. It is not easy to find tables for testing significance of data.

CHECK YOUR PROGRESS

8. What is a parametric statistical test?
9. What is a non-parametric statistical test.

2.5 TESTING HYPOTHESIS I: PARAMETRIC OR STANDARD TESTING

The most important parametric statistics are:

(i) Student's t test and z test

(ii) F Ratio

(iii) Pearson r

(i) Student's t test and z test

A researcher can use either the t test (or t ratio) or z test (or z ratio) to test the significance of difference between two means. The computation of t or z involves computing the ratio between the experimental variances (that is, the difference between two means) and the error variance (that is, the standard error of the mean difference). However, there is one basic difference between t ratio and z ratio. If the sample size is less than 30, one uses the t test or student's t for testing the significance of the difference between the two means. This small sample size test was developed in 1915 by William Seely Gosset, a statistician. This statistic is known as student's t. When the sample size is more than 30, the ratio of the difference between the two sample means to the standard error of this difference is calculated by the z ratio which is interpreted through the use of normal probability tables. However, one should note that as the sample size increases, the critical values of t necessary for rejecting the null hypothesis gradually reduce and approach the z values of the normal probability table.

Before discussing the small sample t test in detail, five important concepts, namely, degree of freedom, null hypothesis, level of significance and one-tailed test should be understood.

Degree of freedom

The degree of freedom means freedom to vary and its abbreviated expression is d.f.

In statistical language, the degree of freedom means the number of observations that are independent of each other and that cannot be deduced from each other. Suppose we have five scores and the mean of five scores is 10. The fifth score immediately makes adjustment with the remaining four scores in a way which assures that the mean of all five scores must be 10. For example, suppose we have four scores 12, 18, 5, 12, and for the mean to become 10, the fifth score must be 3. In another distribution, if the four scores are 2, 10, 8, 5, the fifth score must be 25 in order to derive a mean of 10.

The meaning of this is that four scores in the distribution are independent, they may have any values and they cannot be deduced from each other. The size of the fifth score, however, is fixed because the mean in each case is 10. Hence d.f. $= N - 1 = 5 - 1 = 4$.

NOTES

In larger cases, in computing the mean it goes something like this: Suppose we have a set of 101 scores. We compute the mean and in computing the mean, we lose 1 d.f.. We had initially 101 d.f. (because there were 101 scores) but now after computing mean, we have $N - 1 = 101 - 1 = 100$ degrees of freedom. Sometimes, we have paired data and in such cases, the number of degrees of freedom is equal to one less than number of pairs.

Null hypothesis

The starting point in all statistical tests is the statement of null hypothesis (H_0), which is a 'no difference hypothesis.' A null hypothesis states that there is no significant difference between the samples under study. It makes a judgment about whether the obtained difference between the samples is due to some true differences or due to some chance error. The null hypothesis is formulated for the express purpose of being rejected because if it is rejected the alternative hypothesis (H_1) which is an operational statement of the investigators' research hypothesis, is accepted. A research hypothesis is nothing but predictions or deductions drawn from a theory. The tests of the null hypothesis are generally called tests of significance, the outcome of which is stated in terms of probability figures or levels of significance.

If the difference between the experimental group and the control group is very small, the experimenter is likely to accept the null hypothesis, indicating the fact that the small difference between these two groups is due to sampling errors or some other chance fluctuations. On the other hand, if the difference between the experimental group and the control group is too large, the experimenter is likely to refuse or reject the null hypothesis, including the fact that the obtained differences are real differences between or among the samples under study.

Level of significance

The null hypothesis has been developed for the express purpose of rejection. The rejection or acceptance of the null hypothesis is based upon the level of significance, which is used as a criterion. The levels of significance are also known as alpha levels.

In psychological, sociological and educational researches, there are two levels of significance which are commonly used for testing the null hypothesis.

One is the 0.05 or 5 per cent level and another is the 0.01 or 1 per cent level of significance. If the null hypothesis is rejected at the 0.05 level, it means that 5 times in 100 replications of the experiment the null hypothesis is true and was false in 95 other replications, which means that there is 95 per cent probability of the obtained results are due to experimental treatment rather than due to some chance factors. Rejection of the null hypothesis when that hypothesis is true constitutes a Type I error or alpha error. Thus it can be said that at the 0.05 level of significance the experimenter commits a 5 per cent type I errors when he rejects a null hypothesis. Some investigators may want a more stringent test and the 0.01 level is one such level where the investigator commits a Type I error 1 per cent only. The 0.01 level suggest that a 99 per cent probability exists that the obtained results are due to the experimental treatment, and hence, once in 100 replications of the experiment, the null hypothesis would be true. An even more a stringent test of significance is the 0.001 level, which is not commonly used in behavioural researches. This level suggests that only once in 1000 replications of the experiment, the null hypothesis would be a true one and in 999 replications (out of 1000) the obtained results can be attributed to the experimental treatment. In testing the significance of the statistics, sometimes the investigator accepts the null hypothesis when, in fact, it is false. This error is technically known as the Type II error or beta error.

From the above discussion, it can be said that an error of Type I can be reduced by putting the alpha level at the 0.01 or 0.001 level. But as one reduces the chance for making a Type I error, one increases the chance for making a Type II error where one does not reject the null hypothesis when it should be rejected. Therefore, while decreasing the possibility of making one type of error, the investigator also increases the probability of making another type of error. The research workers must be cautious with this situation and should try, as far as possible, to limit the probability for making a Type I error.

One-tailed test and two-tailed test

One-tailed test is a directional test, which indicates the direction of the difference between the samples under study. Suppose the experimenter conducts an experiment in which one takes two groups—one is the control group and another is the experimental group. Only the experimental group is given training for 5 days on various kinds of arithmetic operations. Subsequently, an arithmetical ability test is administered on the two groups and the scores are obtained. In the above situation, there is reason to say that the mean arithmetical score of the experimental group will be higher than the mean arithmetical score of the control group. This is the alternative hypothesis, which indicates the direction of difference. When the alternative hypothesis states the direction of difference, it constitutes a one-tailed test. The null hypothesis would be that the mean of the experimental group is equal (no difference) to the mean of the control group. If this is rejected we accept the above alternative hypothesis.

The above facts can be schematically presented as follows:

a. $H_0 = M_1 = M_2$ (no difference between M_1 and M_2)

Alternative hypothesis:

b. $H_1 = M_1 \neq M_2$

c. $H_2 = M_1 < M_2$

d. $H_1 = M_1 > M_2$

NOTES

where M_1 = mean of the experimental group and

M_2 = mean of control group.

When it is said that the mean of the experimental group will be higher than the mean of the control group, one is concerned with only one end of the distribution. Putting it in terms of a normal curve, one is concerned with only one end of the curve (see figure below). When the alpha level is set at the 0.05 level, a 5 per cent of area of the normal curve is obtained, all in one-tail rather than having distributed it equally into two tails of the curve. Therefore, the directional null hypothesis is called a one-tailed test. A simple inspection of the table of areas of the normal curve given at the end reveals that a z score of 1.64 cuts off 5 per cent of the area under the normal curve in the smaller part, and similarly a z score of 2.33 cuts off 1 per cent of the area in the smaller part. If the null hypothesis is rejected, that is hypothesis 1 is not tenable, we automatically accept the alternative hypothesis. If the experimenter has some reason to believe that the experimental group would have a lower mean score than the control group (alternative hypothesis), he can set up a directional hypothesis that the mean of the experimental group is lower than the mean of the control group (one-tailed test). Rejection of this hypothesis would automatically lead to the acceptance of the null hypothesis. This time the normal curve in which the experimenter is interested is 5 per cent or 1 per cent of the area in the left-hand tail of the normal curve. When the null hypothesis is rejected by using a one-tailed test, one must say that one is rejecting the hypothesis at 1 per cent or 5 per cent points, not levels (see Fig. 2.1).

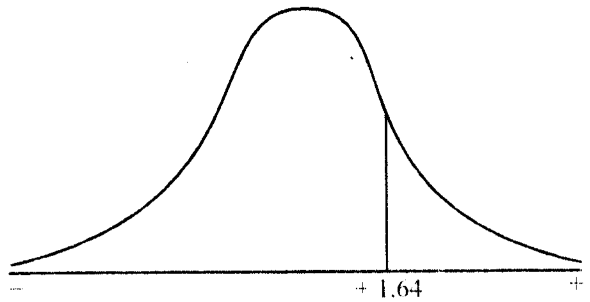


Fig. 2.1 One-Tailed Test at 0.05 or 5% Points

Source: Singh, A.K., Tests, Measurements and Research Methods in behavioural sciences, 2008.

A two-tailed test is one in which the investigator is interested in evaluating the difference between the groups. The direction of difference is of no importance here. The null hypothesis will be that where the mean of the experimental group is equal to the mean of the control group, that is, there is no difference between the means of the experimental group and the control group. The alternative hypothesis would be that the mean of the experimental group is not equal to the mean of the control group. Thus, the investigator's concern is shown with both tails of the distribution. At the 5 per cent level, there is 5 per cent of the area under the normal curve equally divided at both the tails (see figure below). From the table of areas of the normal curve, it can easily be read that a z score of +1.96 cuts off 2.5 per

NOTES

cent area at both extremes. Since a normal curve is bilaterally symmetrical, a z score of -1.96 would also cut off 2.5 per cent area in the left-hand tail. In a two-tailed test a negative z score is interpreted in the same way as a positive z score. When the null hypothesis is rejected by using a two-tailed test, it is said that it has been rejected at 5 per cent or 1 per cent levels and not at 5 per cent or 1 per cent points (see Fig. 2.2).

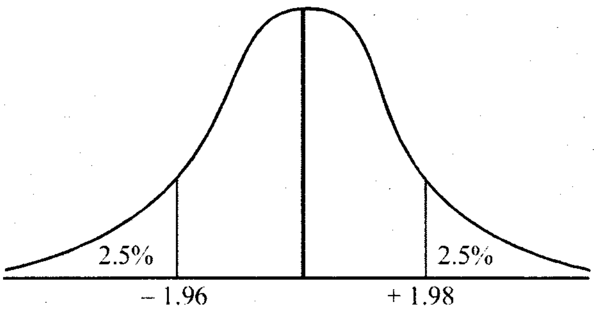


Fig. 2.2 *Two-Tailed Test at 0.05 or 5% Level*

Source: Singh, A.K., Tests, Measurements and Research Methods in behavioural sciences, 2008.

The *t* ratio calculation from different groups

Ordinarily, three types of situations exist while one is calculating the *t* ratio:

a. *t* ratio from independent groups
b. *t* ratio from correlated groups
c. *t* ratio from matched groups

Regardless of the nature of the group, *t* ratio is calculated by the following equation:

$$t = \frac{(M_1 - M_2) - 0}{SE_D} \qquad ...(2.1)$$

where M_1 = mean of the first group
M_2 = mean of the second group
SD_D = standard error of the difference between two sample means.

In each of the above three groups, the formula for SE varies and hence, the following separate calculation of the t ratio from each of the above types:

t ratio from independent groups: Two groups are said to be independent when no correlation exists between them. Suppose one group of boys ($N = 20$) and one group of girls ($N = 22$) were administered a mechanical reasoning test. Their data are summarized as:

$N_1 = 20$ $\qquad N_2 = 22$
$M_1 = 34.56$ $\qquad M_2 = 30.56$
$SD_1 = 5.68$ $\qquad SD_2 = 6.98$

NOTES

Do the two groups differ on the measure of mechanical reasoning test?

The above question can be answered by computing the *t* ratio. For independent groups, *SE* may be calculated with the help of the following question.

$$SE_D = \sqrt{SE_{M_1^2} + SE_{M_2^2}} \qquad ...(2.2)$$

where SE_D = Standard error of the difference between means;

SE_{M_1} = Standard error of the first mean; and

SE_{M_2} = Standard error of the second mean.

Here standard error of mean can be calculated as given below:

$$SE_M = \frac{SD}{\sqrt{N-1}} \qquad ...(2.3)$$

Substituting the value in Equation (2.3), we have

$$SE_{M_1} = \frac{5.68}{\sqrt{19}} = 1.302; \quad SE_{M_2} = \frac{6.98}{\sqrt{21}} = 1.524$$

Now, SE_D becomes equal to:

$$SE_D = \sqrt{(1.302)^2 + (1.524)^2} = 2.004$$

$$\therefore \quad t = \frac{(M_1 - M_2) - 0}{SE_D} = \frac{(34.56 - 30.56) - 0}{2.004} = \frac{4.00}{2.004} = 1.996$$

$$\text{d.f.} = (N_1 - 1) + (N_2 - 1) = (20 - 1) + (22 - 1) = 40$$

Entering the probability table of *t* ratios we find that the obtained *t* at d.f. = 40 is significant at the 0.05 level but not at the 0.01 level. Hence, we conclude that the null hypothesis is rejected and that there is a true difference between the means of the groups of boys and girls.

***t* ratio from correlated groups:** The correlated groups are those which exhibit some correlation with each other on the given measures. One of the most convenient ways of getting the correlated means is to repeat the same test on the same groups twice on two different occasions. In between the initial and the final administration of the test, some experimental treatment is given to the group. Because the group and the test are the same, it is highly probable that there will be a correlation between the initial measures and the final measures. Suppose a group of 20 students of Class IV is administered an English spelling test in January. The obtained mean score and standard deviations are given below. After a year's training in spelling, they are again given the same test. This time their mean is considerably raised and the standard deviation is lowered. The correlation coefficient between the initial and final set of scores was positive and significant.

Does training produce a significant difference between the initial mean and the final mean?

	Initial test	Final test
Ns	20	20
Means	36.28	40.33
SDs	4.28	3.32
Coefficient of correlation	0.80	

NOTES

In case of testing the significance of mean difference between two correlated means, SE_D is calculated with help of the following formula:

$$SE_D = \sqrt{SE_{M_1^2} + SE_{M_2^2} - 2r_{12}SE_{M_1}SE_{M_2}} \quad ...(2.4)$$

where r_{12} = coefficient of correlation between the initial set of scores and the final set of scores. The rest of the subscripts are defined like those in the equation above.

The standard error of mean (SE_M) is calculated by the equation below:

$$SE_{M_1} = \frac{4.28}{\sqrt{19}} = 0.982 \qquad SE_{M_1} = \frac{2.32}{\sqrt{19}} = 0.532$$

$$SE_D = \sqrt{SE_{M_1^2} + SE_{M_2^2} - 0.2r_{12}SE_{M_1}SE_{M_2}}$$

$$= \sqrt{(0.982)^2 + (0.532)^2 - 2(0.80 \times 0.982 \times 0.532)}$$

$$= \sqrt{0.9643 + 0.2830 - 0.8358} = 0.641$$

$$t = \frac{(M_1 - M_2) - 0}{SE_D} = \frac{(36.38 - 40.33) - 0}{0.625} = \frac{4.05}{0.641} = 6.31$$

$$\text{df} = N - 1 = 20 - 1 = 19$$

Entering the probability table of t ratios at d.f. = 19, we find that our obtained value of t exceeds the value of t at even the 0.001 level of significance. Hence, the null hypothesis is rejected and it is concluded that the training has produced significant difference between the mean of the initial set of scores and the mean of the final set of scores.

***t* ratio from matched groups:** Sometimes it becomes necessary for the researchers to match the groups under study. Matching can be done on the basis of numbers or it can be done in terms of mean and standard deviation. When the matching is done on the basis of the number of the subjects, each person has his corresponding match in the other group and therefore, the number of persons in the two matched groups is always equal. When matching is done in terms of mean and standard deviation, the number in the two groups may or may not be equal.

Suppose two groups from two different classes are considered and each group is compared on the basis of numerical reasoning test. Both groups have been matched in terms of mean and standard deviations on the basis of their scores on a General Intelligence Test. Do the groups differ in terms of mean numerical ability? The following is the data obtained:

NOTES

	Class VIII	Class X
N1	100	120
Means of Intelligence tests	70.26	70.25
SDs on Intelligence Tests	9.98	10.02
Means on Numerical Reasoning Tests	55.62	60.34
SDs on Numerical Reasoning Tests	8.67	7.58
Coefficient of Correlation between General Intelligence Test Scores and numerical reasoning test scores		0.45

The equation for calculating SE_D in case of groups matched in terms of means and *SD* is given below:

$$SE_D = \sqrt{\left(SE_{M_1^2} + SE_{M_2^2}\right)(1-r^2)} \quad ...(2.5)$$

where subscripts are defined as usual. Now, we can proceed as under:

$$SE_{M_1} = \frac{8.67}{\sqrt{100}} = 0.867; \ \text{SE}_{\text{M}_2} = \frac{7.58}{\sqrt{120}} = 0.691$$

$$SE_D = \sqrt{(0.867)^2 + (0.691)^2(1-0.45^2)}$$

$$= \sqrt{(0.7516)+(0.4777)(1-0.2025)}$$

$$t = \frac{(M_1 - M_2) - 0}{SE_D} = \frac{(55.62-60.34)-0}{0.991} = 0.991$$

$$= \frac{4.72}{0.991} = 47.62$$

$$\text{df} = (N_1 - 1) + (N_2 - 1) - 1 = (100 - 1) + (120 - 1) - 1$$
$$= (99 + 119) - 1 = 218 - 1 = 217$$

Entering the probability table of *t* ratios at d.f. 217, we find by interpolation that the obtained value of *t* exceeds the *t* value required at even the 0.001 level of significance. Rejecting the null hypothesis, we conclude that the two groups differed significantly in terms of numerical mean ability.

(ii) F ratio

The *t* ratio or *z* ratio is one of the powerful parametric tests through which we can test the significance of the difference between two means. There are two general limitations of the *t* ratio. First, when there are several groups and if we want to test the significance of the mean difference among them, several *t* ratios are required to be computed. For example, suppose there are five groups. Then there is need to compute.

NOTES

$$\frac{N(N-1)}{2} = \frac{5 \times 4}{2} = 10t \text{ ratios} \qquad ...(2.6)$$

Then it becomes a cumbersome task. Secondly, the *t* ratio does not account for interaction effect in its statistical analysis. The variations in the scores may be due to the interactions taking place among groups. Such variations are not accounted for by *t* ratios. In order to remove these two limitations we turn to analysis of variance, originally developed by R.F. Fisher. Analysis of variance is a class of statistical techniques through which we test the overall difference among the two or more than two (normally more than two) sample means.

Analysis of variance is of two common types: Simple analysis of variance or one-way analysis of variance and complex analysis of variance or two-way analysis of variance. Analysis of variance (of whatever types) is often referred to by its acronym, ANOVA.

In simple analysis of variance there is only one independent variable and the samples are classified into several groups on the basis of this variable. Since the basis of classification is only one independent variable, the simple analysis of variance is also known as one-way ANOVA. Such ANOVA is suited to the completely randomized design. In complex ANOVA, there are two or more than two independent variables, which form the basis of classification of groups. Such ANOVA is suited to factorial design.

Statistically, the *F* ratio is calculated as follows.

$$F = \frac{\text{Larger variance}}{\text{Smaller variance}} \text{ or } \frac{\text{Between-groups variance}}{\text{Within-groups variance}} \qquad ...(2.7)$$

Between-groups variance refers to variation in the mean of each group from the total or grand mean of all groups. Within-groups variance refers to the average variability of scores within each group. The theme of the analysis of variance is that if the groups have been randomly selected from the population, these two variances, namely, between-groups variance and within-group variance are the unbiased estimates of same population variance. The significance of difference between these two types of variances is tested through the *F* test.

The following table illustrates the simple analysis of variance, with independent measures. One can see that the scores of the three groups of subjects *A*, *B* and *C* are shown on the educational achievement test, and ANOVA has been calculated from those scores.

Grand sum (ΣX): 561 + 154 + 312 = 1027

Grand sum of squares (ΣX^2): 38611 + 2610 + 16070 = 57291

Step 1: Correction (C): $\frac{(\Sigma X^2)}{N} = \frac{(1027)^2}{30} = 35157.63$

Step 2: Total sum of squares (TSS): $\Sigma X^2 - C$

$= 57291 - 35157.63 = 2213337$

NOTES

Table 2.1 Simple ANOVA based on the Scores of Three Groups

Gr. A X	Gr. B X	Gr. C X	Gr. A X^2	Gr. B X^2	Gr. C X^2
78	10	50	6084	100	2500
58	14	58	3364	196	3364
40	9	65	1600	81	4225
30	12	72	900	144	5184
10	20	10	100	400	100
65	25	9	4225	625	81
88	12	8	7744	144	64
87	18	10	7569	324	100
80	14	16	6400	196	256
25	20	14	625	400	196
Sums 561	154	312	38611	2610	16070

Step 3: Between (or among) sum of squares (*BSS*):

$$\frac{(\Sigma X_1)^2}{n_1}+\frac{(\Sigma X_2^2)}{n_2}\cdots+\frac{(\Sigma X_n)^2}{n_n}n_n-C$$

$$=\frac{(561)^2}{10}+\frac{(154)^2}{10}+\frac{(312)^2}{10}-35157.63 \qquad ...(2.8)$$

$$=\frac{314721+23716+97344}{10}-35157.63$$

$$=43578.10-35157.63=8420.47$$

Step 4: Within sum of squares: *TSS* – *BSS* = 22133.37 – 8420.47 = 13712.90

Table 2.2 Summary: Analysis of Variance

Sources of variation	d.f.	Sum of squares	Mean squares or variance
Between-groups	$K-1=3-1=2$	8420.47	4210.235
Within-groups	$N-K=30-3=27$	13712.90	507.885
Total	$N-1=30-1=29$	22133.37	

$$F=\frac{\text{Between-groups variance}}{\text{Within-groups variance}}=\frac{4210.235}{507.885}=8.289=8.29$$

Since there are 30 cases in the above problem we have $N-1=30-1=29$ d.f. in all. d.f. for groups (between-groups) is equal to the number or groups (K) minus one. Since there are three groups d.f. for between-groups is $K-1=$

3 – 1 = 2. d.f. for within-groups is equal to the total number of cases minus number of groups (K). Hence, it is equal to $N-K=30-3=27$. After putting the number of degrees of freedom, the sum of squares for each of the three sources of variations, we compute mean squares or variances, which are obtained by dividing each of the sum of squares by its respective number of degrees of freedom of freedom. These two types of variances are the estimates of the population variance. We obtain F ratio by dividing the between-groups variance by the within-groups variance.

F ratio is interpreted by the use of the *F* table (Guilford & Fruchter, 1978). In the *F* table the number of degrees of freedom for greater mean square (d.f.$_1$) is written at the top and the number of degrees for freedom for smaller mean square (d.f.$_2$) is written on the left-hand side. For this problem d.f.$_1$ – 2 and d.f.$_2$ = 27. Locating at these d.f.s, one finds that the required *F* ratio at the 0.05 level is 3.35 and at the 0.01 level is 5.49. Since the obtained value of F ratio is 8.29, which exceeds 5.49, we reject the null hypothesis and conclude that there is an overall difference between the three groups of subjects on the educational achievement test.

Tests after the *F* test

A general limitation of the *F* test is that it only tells about the overall difference between the groups under study but tells nothing about the location of the exact difference. For example, in the above problem, the obtained *F* ratio of 8.29 is significant, which definitely indicates that there is significant difference between the groups under study, but whether the significant difference is between *A* and *B* or *A* and *C* or *B* and *C*, cannot be said. Therefore, when F ratio is significant, one needs some additional tests after the *F* test. Winer (1971) has summarized half a dozen different statistical tests, which can be used as additional test. One convenient such statistical test is to apply the t test between the possible pairs of the groups. Scheffe (1957) has introduced one such rigorous test, which reduces the probability of making a Type I error. Following the Scheffe technique, one can locate the difference between three means. Since there are three groups, three comparisons are likely to be made, namely, *A* vs. *B*, *A* vs. *C* and *B* vs. *C*. Following the Scheffe technique, one is required to compute the *F* ratio with the help of equation given for each of three groups.

$$F = \frac{(M_1 - M_2)^2}{SD_W^2 (N_1 + N_2) / N_1 N_2} \qquad \text{...(2.9)}$$

Now, three F_s for three pairs of distribution can be calculated as follows: *F* ratio for distributions *A* and *B*:

$$F = \frac{(56.10 - 15.40)^2}{507.885 \frac{(10+10)}{(10)(10)}} = 16.31$$

NOTES

F ratio for distributions *A* and *C*:

$$\frac{(56.10-31.20)^2}{507.885\frac{(10+10)}{(10)(10)}}=6.10$$

F ratio for distributions *B* and *C*:

$$\frac{(15.40-31.20)^2}{507.885\frac{(10+10)}{(10)(10)}}=2.46$$

As seen out earlier, *F* at the 0.05 level of significance for d.f.1 = 2 and d.f. 2 = 27 is 3.35. This value, if multiplied by $K-1$, yields (3 – 1) (3.35) = 6.70. Only the *F* ratio for distributions *A* and *B* is greater than 6. 70. Hence, it is concluded that there is a significant difference between the means of A and *B* only. The mean difference between *A* and *C* and *B* is not significant.

(iii) Pearson *r*

Of all the measures of correlation the Pearson *r*, developed by Prof. Karl Pearson, is one of the most common methods of assessing the association between two variables under study. It is also known as Pearson product–moment correlation and abbreviated to *r*. The size of Pearson *r* varies from +1 through 0 to –1. In fact, all correlation coefficients have the limit of +1 and –1. A coefficient of +1 indicates a perfect positive correlation, and a coefficient of –1 indicates a perfect negative correlation. The coefficient of correlation indicates two things. First, it indicates the magnitude of a relationship. A correlation coefficient of say, +0.90 or –0.90 gives the same information about the magnitude or size of correlation. The sign makes no variation in the size of the correlation. Second, it gives an indication regarding the direction of the correlation coefficient. A positive correlation indicates a similar trend of relationship between two variables, that is, as one increases the other also increases or as one decreases, the other also decreases. Consider the relationship between intelligence test scores and classroom achievement. Generally, as intelligence test scores become higher, classroom achievement also becomes better. Therefore, the direction of the correlation between these two variables is positive. Similarly, as fatigue increases, output decreases. Here the relationship is negative because as one increases, the other decreases. Sometimes, the relationship is not consistent. And in this situation, coefficient of correlation is likely to be zero.

The Pearson product-moment correlation has two important assumptions:

(i) The relationship between *X* and *Y* variables should be linear. A linear relationship refers to the tendency of the data, when plotted, to follow a straight line as closely as possible. Although there are some statistical treatments which can test whether or not the relationship is linear, generally this is determined by the inspection of the scatter diagram or correlation table.

NOTES

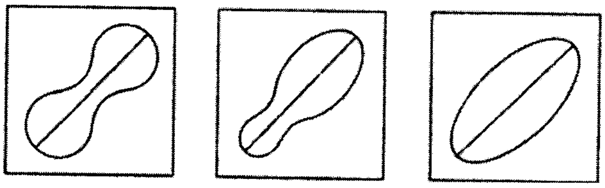


Fig. 2.3 (a) and (b) Non-homoscedastic and (c) Homoscedastic and Linear

(ii) The second assumption is of homoscedasticity (homo means 'like' and scedasticity means scatteredness). Defined statistically, one can say that homoscedasticity refers to the fact that standard deviations (or variances) for columns and rows in the scatter diagram are equal or nearly equal. Figure above illustrates homoscedastic and non-homoscedastic distributions. In this figure, we have three diagrams. In Fig. 2.3 (a), the variance of the distribution near the centre is smaller than the variances near both extremes, and hence, the distribution is non-homoscedastic. In Fig. 2.3 (b), the variances near the bottom extreme are lower than the variances at middle or at the top extreme, and, therefore, the distribution is no-homoscedastic. In Fig. 2.3 (c) the variances are equal throughout as well as linear.

Pearson r can be calculated by several formulas. One can calculate the Pearson r by the raw score formula or machine formula. The equation is:

$$r = \frac{N\Sigma XY - \Sigma X \Sigma Y}{\sqrt{\left[N\Sigma X^2 - (\Sigma X)^2\right]\left[N\Sigma Y^2 - (\Sigma Y)^2\right]}} \qquad ...(2.10)$$

where r = Pearson product–moment correlation coefficient; N = number of scores X = scores in X variance; and Y = scores in Y variable.

The table below presents the scores of 10 students who were administered intelligence test (X) and anxiety test (Y). Pearson r has also been calculated from the ten pairs of scores. The significance of the obtained r is tested with the help of a table (Downie & Health, 1970: 378). The obtained value of r as less than the value required at even 0.05 level of significance. Hence, the null hypothesis is accepted and it is concluded that the scores on intelligence test and the scores on the anxiety test are not correlated. The sign makes no difference in the magnitude of the correlation.

When the data are arranged in bivariate distribution as is the case in the scatter diagram or in the correlation table, Pearson r should be calculated by the following formula:

$$r = \frac{\frac{Ex'y'}{N} = C_x C_y}{(\sigma_{x'})(\sigma_{y'})} \qquad ...(2.11)$$

where r = Pearson r; x' = deviation of scores from mean on x test; y' = deviation of scores from mean on y test; N = sum of frequencies; C_x = correction in x-series scores; and C_y = correction in y-series scores.

NOTES

Pearson r by raw-score method

X	Y	X_2	Y_2	XY
10	8	100	64	80
5	20	25	400	100
6	15	36	225	90
3	13	9	169	39
8	16	64	256	128
12	20	144	400	240
13	13	169	169	169
20	11	400	121	120
15	10	225	100	150
10	12	100	144	120
$\Sigma X = 102$	$\Sigma Y = 138$	$\Sigma X^2 = 1272$	$\Sigma Y^2 = 2048$	$\Sigma XY = 1336$

$$r = \frac{N\Sigma XY - \Sigma X \Sigma Y}{\sqrt{\left[N\Sigma X^2 - (\Sigma X)^2\right]\left[N\Sigma Y^2 - (\Sigma Y)^2\right]}}$$

$$= \frac{(10)(1336) - (102)(138)}{\sqrt{[(10)(1272) - (102)^2][(10)(2048) - (138)^2]}}$$

$$= \frac{-716}{\sqrt{3325776}} = \frac{-716}{1823.671} = -0.392 = -0.39$$

$$df = N - 2 = 10 - 2 = 8$$

CHECK YOUR PROGRESS

10. When is the students 't' test or z test used?
11. What is the meant by degrees of freedom?
12. What is experimental variance?
13. What is error variance?
14. Define the Pearson r.

2.6 TESTING HYPOTHESIS II: NON-PARAMETRIC TESTING

The important non-parametric statistics are as follows:

(i) Chi-square (χ^2) Test

(ii) Mann-Whitney U test

(iii) Rank-difference methods (both ñ and T)

(iv) Coefficient of concordance (W)

(v) Median Test

(vi) Kruskal–Wallis H Test

(vii) Friedman test

NOTES

(i) Chi-square (χ^2) test

The chi-square is one of the most important non-parametric statistics, which is used for several purposes. For this reason, Guilford (1956) has called it the general-purpose statistic. It is a non-parametric statistic because it involves no assumption regarding the normalcy of distribution or homogeneity of the variances. The chi-square test is used when the data are expressed in terms of frequencies of proportions or percentages. The chi-square applies only to discrete data. However, any continuous data can be reduced to the categories in such a way that they can be treated as discrete data and then, the application of chi-square is possible. The formula for calculating X^2 is given below:

$$\chi^2 = \sum \frac{(f_o - f_e)^2}{f_e} \qquad ...(2.12)$$

where χ^2 = chi-square; f_o = obtained or observed frequency; and f_e = expected frequency or theoretical frequency. There are several uses of the chi-square test.

First, chi-square may be used as a test of equal probability hypothesis. By equal probability hypothesis one means the probability of having the frequencies in all the given categories as equal. Suppose, for example, 100 students answer an item on an attitude scale. The item has five categories of response options—strongly agree, agree, neutral, disagree and strongly disagree. According to the equal probability hypothesis, the expected frequency of responses given by 100 students would be 20 in each. The chi-square test would test whether or not the equal probability assumption is tenable. If the value of the chi-square test is significant, the equal probability hypothesis becomes untenable and if the value of the chi-square is not significant, the equal probability hypothesis becomes tenable.

The second use of the chi-square test is in testing the significance of the independence hypothesis. By independence hypothesis is meant that one variable is not affected by, or related to, another variable and hence, these two variables are independent. The chi-square is not a measure of the degree of relationship in these conditions. It merely provides an estimate of some factors other than chance (or sampling error), which account for the possible relationship. Generally, in dealing with data related to independent hypothesis, they are first arranged in a contingency table. When observations on two variables are classified in a two-way table, data are called the contingency data and the table is known as the contingency table. Independence in a contingency table exists only when each tally exhibits a different event or individual.

The third important use of chi-square is in testing a hypothesis regarding the normal shape of a frequency distribution. When chi-square is used in this connection, it is commonly referred to as a test of goodness-of-fit.

The fourth use of chi-square is in testing the significance of several statistics. For example, for testing the significance of the phi-coefficient, coefficient of concordance, and coefficient of contingency, one converts the respective values into chi-square values.

If the chi-square value appears to be a significant one, one should also take their original values as significant. To illustrate this, a 3×3 contingency table, which shows data of 200 students who were classified into three classes on the basis of their educational qualification is used. The students' educational attainments are measured in the course of their study by classifying them as superior, average or inferior.

The use of Chi-square in a 3×3 Contingency Table

	Superior	Average	Inferior	
Master	30 (25)	15 (15)	5 (10)	50
Bachelor	25 (25)	10 (15)	15 (10)	50
Intermediate	45 (50)	35 (30)	20 (20)	100
	100	60	40	200

(The figures in brackets show expected frequency)

The question posed is: Is educational achievement related to educational qualification? The obtained data have been shown above. The first step in calculating X^2 as a test of significance of independence or the relationship between educational achievement and educational qualification is to compute the expected frequency. The null hypothesis is that these two variables are not related or are independent, and if this hypothesis is true, the expected frequencies should be as follows:

Cells of Table	Expected Frequency
Upper left	$(100\times50)/200=25$
Upper middle	$(60\times50)/200=15$
Upper right	$(40\times50)/200=10$
Middle left	$(100\times50)/200=25$
Middle middle	$(60\times50)/200=15$
Middle right	$(40\times50)/200=10$
Lower left	$(100\times100)/200=50$
Lower middle	$(60\times100)/200=30$
Lower right	$(40\times100)/200=20$
	$\Sigma = 200$

After calculating expected frequency for each cell, the chi-square may be calculated as:

f_o	f_e	$f_o - f_e$	$(f_o - f_e)^2$	$\frac{(f_o - f_e)^2}{f_e}$
30	25	+5	25	1
15	15	0	0	0
5	10	–5	25	2.5
25	25	0	0	0
10	15	–5	25	1.67
15	10	+5	25	2.5
45	50	–5	25	0.5
35	30	+5	25	0.83
20	20	0	0	0
$\Sigma f_o = 200$	$\Sigma f_e = 200$	$\Sigma(f_o - f_e) = 0$		$\Sigma = 9.00$
	$df = (r - 1)(K = 1) = (3 - 1)(3 - 1) = 2 \times 2 = 4$			

Entering the probability table of chi-square, one finds that the value of chi-square for d.f. = 4 at the 0.05 level should be 9.488. As the obtained chi-square is below it ($p >> 0.05$), one concludes that the null hypothesis is retained. Hence, the two variables, namely, educational qualification and educational attainment in the present study are found to be independent. For calculating d.f. in a chi-square test, the formula as noted above is $(r - 1)(k - 1)$ where r = the number of rows and k is the number of columns.

Chi-square: When the data has been arranged in a 2×2 contingency table (where d.f. = 1), we need not calculate the expected frequency in the manner described above. In such a situation, the chi-square can be directly calculated with the help of the following equation:

$$\chi^2 = \frac{N[|AD - BC|]^2}{(A+B)(C+D)(A+C)(B+D)} \quad ...(2.13)$$

where A, B, C and D are symbols for frequency of four cells in a 2×2 table; N = total number of frequencies; bars (II) indicate that in subtracting BC from AD the sign is ignored.

Suppose the researcher wants to know whether or not the two given items in the test are independent. Both items have been answered in 'Yes' or 'No' form. The test was administered to a sample of 400 students and the obtained data were as follows:

NOTES

Chi-square in a 2×2 Table

		Item No. 6 Yes	Item No. 6 No	
Item No. 10	No	180 A	120 B	300
	Yes	90 C	10 D	100
		270	130	400

According to the formula:

$$\text{d.f.} = (r-1)(k-1) = (2-1)(2-1) = 1$$

Entering the probability table of chi-square, we find that for d.f., the value of chi-square at the 0.001 level should be 10.827. As the obtained value of the chi-square is much above it, we conclude that item nos. 6 and 10 are not independent, that is, they are related.

Sometimes it happens that with 1 d.f., any one of the expected cell frequencies becomes less than 5. In such a situation a correction called Yates' correction for continuity is applied. Some writers have suggested that Yates' correction for continuity should be applied when any of the expected frequencies goes below 10. Where frequencies are large, this correction makes no difference but where frequencies are small, Yates' correction is significant. Yates' correction consists in reducing the absolute value of difference between f_o and f_e by 0.5, that is, each fo which is larger than fe is decreased by 0.5 and each fo which is smaller than fe is increased by 0.5. The formula for chi-square in such a situation is as given below:

$$\chi^2 = \frac{N\left[|AD-BC|-\frac{N}{2}\right]^2}{(A+B)(C+D)(A+C)(B+D)} \qquad ...(2.14)$$

where subscripts are defined as usual. Suppose, 60 students (50 boys and 10 girls) were administered an attitude scale. The items were to be answered in 'Yes' and 'No' form.

Their frequencies towards 10 items are presented in the table below. The question is: Do the opinions of boys and girls differ significantly?

Chi-square with Yates' Correction in a 2×2 Table

	Yes	No	
Boys	20 A	30 B	50
Girls	3 C	7 D	10
	23	37	60

NOTES

According to Equation (2.14):

$$\chi^2 = 60\frac{\left[|(20)(7)-(30)(3)|-\frac{60}{2}\right]^2}{(50)(10)(23)(37)}$$

$$= \frac{60|140-90|-30]^2}{425500} = \frac{60\times 400}{425500} = \frac{24000}{425500} = 0.056$$

In the above example, the expected frequency (23×10/60) is less than 5. Hence, chi-square has been calculated by Equation (2.14). Entering the table for chi-square, we find that for d.f. = 1, the value of chi-square at the 0.05 level should be 3.841. Since the obtained value is less than it ($P >> 0.05$), one concludes that the opinions of boys and girls do not differ significantly.

(ii) Mann-Whitney *U* test

The Mann-Whitney *U* test is a non-parametric substitute for the parametric *t* test. This test was independently proposed by Mann and Whitney. The Mann-Whitney *U* test is used when the researcher is interested in testing the significance of difference between two independently drawn samples or groups. For applications of the U test it is essential that the data be obtained on ordinal measurement, that is, they must have been obtained in terms of rank. Where the data have been obtained in terms of scores, for application of the Mann-Whitney *U* test, it is essential that those scores be converted into rank without much loss of information. It is not necessary for the application of the Mann-Whitney *U* test that both groups must have unequal size. However, this test can also be applied to groups having equal size.

Here the calculation the Mann-Whitney *U* test, which are concerned with larger sample size of more than 20 cases is given:

$$U = N_1N_2 + \frac{N_1(N_1+1)}{2} - \Sigma R_1 \qquad ...(2.15)$$

$$U = N_1N_2 + \frac{N_2(N_2+1)}{2} - \Sigma R_2 \qquad ...(2.16)$$

Table below presents the scores of two groups on the Lie scale. Group I has 10 subjects and Group II has 21 subjects. The first step is to rank all the scores in one combined distribution in an increasing order of size. The lowest score (taking both sets of scores together) is 7 (second column) and hence, it is given a rank of 1. The nest score is 8, which is again in the second column and it has been given a rank of 2. The third score from below is 10 (in the first column), which has been given a rank of 3. In this way ranking is continued until all scores receive ranks. Subsequently, the two columns of ranks are summed. At this point, a check on arithmetical calculation is imposed. The check is that the sums of these two columns must be equal to $N(N+1)/2$.

NOTES

Check: $\Sigma R_1 + \Sigma R_2 = \frac{(N)(N+1)}{2} = 88.5 + 407.5 = 496; \frac{(31)(32)}{2} = 496$

Hence, we can proceed:

(by Equation 2.15)

$$U = N_1N_2 + \frac{N_1(N_1+1)}{2} - \Sigma R_1 = (10)(21) + \frac{(10)(10+1)}{2} - 88.5 = 176.5$$

(by Equation 2.16)

$$U = N_1N_2 + \frac{N_2(N_2+1)}{2} - \Sigma R_2 = (10)(21) + \frac{(21)(21+1)}{2} - 407.5 = 33.5$$

Table 2.3 *Calculation of the Mann-Whitney U Test from Larger Sample Sizes*

Gr. I (N_1 = 10)	**Gr. II (N_2 = 21)**	R_1	R_2
18	32	7	13
14	40	15	18
30	31	11	12
10	39	3	16.5
39	15	16.5	6
26	8	9	2
27	47	10	19
19	33	8	14
35	52	15	22
11	48	4	20
	7		1
	50		21
	61		27
	58		24
	53		23
	59		25
	60		26
	65		30
	63		28
	67		31
	64		29
		$\Sigma R_1 = 88.5$	$\Sigma R_2 = 407.5$

It is the lower value of *U* test that one wants. For testing the significance of the obtained *U*, its values are converted into *z* score as shown below:

NOTES

$$z = \frac{U - \frac{N_1 N_2}{2}}{\sqrt{\frac{(N_1)(N_2)(N_1 + N_2 + 1)}{12}}} \qquad ...(2.17)$$

$$= \frac{176.5 - \frac{(10)(21)}{2}}{\sqrt{\frac{(10)(21)(10+21+1)}{12}}} = \frac{71.5}{23.664} = 3.02$$

A z score from +1.96 to +2.58 is taken to be significant at the 0.05 level of significance and if the z score is greater than even +2.58, one takes it to be significant at the 0.01 level. Since the obtained z is 3.02, on can take the value of the Mann-Whitney U to be a significant one. Rejecting the null hypothesis, one can conclude that the two groups differ significantly on the measures of the Lie scale.

(iii) Rank-difference methods

The methods of correlation based upon rank differences are very common among behavioural scientists. These are two most common methods which are based upon the differences in ranks assigned on the X and Y variables. One is the Spearman rank-difference method and the other is the Kendall rank-difference method. The Spearman rank-difference method symbolized by ñ (read as rho) is a very popular method of computing the correlation coefficient between two sets of ranks or between two sets of scores converted into ranks. The method has been named after Spearman who discovered it. This method is applicable when the number of pairs of scores or ranks is preferably small, that is, 30 or below.

The equation is:

$$\rho = 1 - \frac{6\sum D^2}{N(N^2 - 1)} \qquad ...(2.18)$$

where ρ = Spearman's rank-difference correlation coefficient; D = difference between rank1 and rank2; and N = Number of pairs of ranks or scores.

To illustrate the calculation of *ñ* consider the data given in table that follows, which show the scores of 12 students on the intelligence test (X) as well as on the educational test (Y). The first step is to rank both sets of scores separately giving the highest score a rank of 1, the next highest score a rank of 2, and so on. Then, keeping the algebraic signs in view, the difference between two sets of ranks is computed. This is noted under column D. Subsequently, each difference is squared and noted under column D_2. Substituting the values in the equation, we get a *ñ* of –0.185. Following the table of Guilford (1956: Table L, p. 549), we can test the significance of the obtained *ñ*. Since the obtained value of *ñ* is less than the value given at the 0.05 level (*ñ* > 0.05) for N = 12, one can accept the null hypothesis and can conclude that X and Y are independent and whatever correlation has been found is due to the chance factor.

NOTES

***Table 2.4** Illustration of the Spearman Rank-Difference Correlation*

X	*Y*	*Rank₁*	*Rank₂*	*D* ($R_1 - R_2$)	D^2
47	68	8.5	1	+7.5	56.25
50	60	5.5	2.5	+3	9.00
70	54	2	7	–5	25.00
72	53	1	8	–7	49.00
46	60	10	2.5	–7.5	56.25
50	55	5.5	6	–0.5	0.25
42	48	11	9	+2	4.00
58	30	3	12	–9	81.00
55	45	4	10	–6	36.00
36	43	12	11	+1	1.00
49	59	7	4	+3	9.00
47	56	8.5	5	+3.5	12.25
				$\Sigma D = 0.0$	$\Sigma D2 = 339.00$

$$\rho = 1 - \frac{6(339)}{12(122-1)} = 1 - \frac{2034}{1716} = 1 - 1.185 = -0.185$$

Another method of computing the rank-difference correlation has been developed by Kendall. The method is known as Kendall's τ for which the formula is as follows.

$$\tau = \frac{S}{(1/2)N(N-1)} \qquad ...(2.19)$$

where τ = Kendall's τ, S = actual total; and N = number of objects or scores which have been ranked.

Suppose 12 students have been administered two tests and their scores are presented in the following table.

Scores of 12 Students on X and Y Test

	A	B	C	D	E	F	G	H	I	J	K	L
X	20	26	17	16	15	23	22	24	19	28	30	10
Y	70	80	40	45	38	49	77	76	72	47	36	35

The first step is to rank both sets of scores giving the highest score a rank of 1, the next higher a rank of 2, and so on. The following table presents the ranks based upon two sets of the scores given in that table above. Subsequently, the ranks of the *X* test are rearranged in a way that they appear in a natural order like 1, 2, 3.

Ranks based upon Two Sets of Scores given in the above Table

	A	B	C	D	E	F	G	H	I	J	K	L
X	7	3	9	10	11	5	6	4	8	2	1	12
Y	5	1	9	8	10	6	2	3	4	7	11	12

NOTES

Accordingly, ranks on the Y test are adjusted. The following table presents the ranks in a rearranged order. Subsequently, the value of S is computed. For this, we start with the rank on the Y test from the left side.

Rearranged Order of Ranks

	A	B	C	D	E	F	G	H	I	J	K	L
X	1	2	3	4	5	6	7	8	9	10	11	12
Y	11	7	1	3	6	2	5	4	9	8	10	12

The first rank on the left side is 11. Count the number of ranks which are above 11 and the number of ranks which are below 11, separately. Only one rank (that is, 12) falling at the right of the first rank on the Y test is above 11 and the remaining 10 ranks fall below it. Hence, its contribution to S would be equal to 1 – 10. Likewise, the second rank on the Y test is 7. The four ranks falling right of 7, are above 7 and 6 ranks are below it. Hence, its contribution to S would be 4 – 6.

Identical procedures are repeated for other ranks on the Y test. Thus:

$S = (1 - 10) + (4 - 6) + (7 - 1) + (4 - 3) + (6 - 0) + (4 - 1) + (4 - 0) + (2 - 1) + (2 - 0) + (1 - 0) = (-9) + (-2) + (9) + (6) + (1) + (6) + (3) + (4) + (1) + (2) + (1) = 33 - 11 = 22.$

Substituting this in the formula given in Equation (2.19):

$$\tau = \frac{S}{\frac{1}{2}N(N-1)} = \frac{22}{\frac{1}{2}12(12-1)} = \frac{22}{66} = 0.333$$

The significance of τ is tested by converting it into a z score, the formula for which is as follows:

$$z = \frac{\tau}{\sqrt{\frac{2(2N+5)}{9N(N-1)}}} \qquad ...(2.20)$$

Hence

$$z = \frac{0.33}{\sqrt{\frac{2[(2)(12)+5)]}{9(12)(12-1)}}} = \frac{0.33}{\sqrt{0.0488}} = \frac{0.3}{0.2209} = 1.4938$$

Since the obtained z score is less than 1.96, one can say that this is not significant even at the 0.05 level. Accepting the null hypothesis, one can say that the given set of scores is not correlated. According to Siegel (1956), τ has one advantage over ρ, and that is that the former can be generalized to partial correlation. If both τ and ρ are computed from the same data, the answer will not be the same and hence, numerically, they are not equal.

NOTES

(iv) Coefficient concordance, *W*

The coefficient of concordance symbolized by the letter W has been developed by Kendall and is a measure of correlation between more than two sets of ranks. Thus, W is a measure of correlation between more than two sets of rankings of events, objects and individuals. When the investigator is interested in knowing the inter-test reliability, W is chosen as the most appropriate statistic. One characteristic of W which distinguishes it from other methods of correlation is that it is either zero or positive. It cannot be negative. W can be computed with the help of the formula given below:

$$W = \frac{S}{\frac{1}{12}K^2(N^3 - N)} \qquad ...(2.21)$$

where W = coefficient of concordance; S = sum of squares of deviations from the mean of R_j; K = number of judges or sets of rankings; and N = number of objects or individuals which have been ranked. Suppose four teachers (A, B, C and D) ranked 8 students on the basis of performance shown in the classroom. The ranks given by the four teachers are presented in the following table. The details of the calculations have also been shown.

Ranks given by Four Teachers to Eight Students on the basis of Classroom Performance

Teachers	Students							
	(i)	(ii)	(iii)	(iv)	(v)	(vi)	(vii)	(viii)
A	3	4	7	5	8	6	2	1
B	2	3	6	4	8	7	1	5
C	1	3	5	6	8	7	2	4
D	3	4	2	5	7	6	1	8
R_j	9	14	20	20	31	26	6	18

$$\text{Mean of } R_j = \frac{9+14+20+20+31+26+6+18}{8} = \frac{144}{8} = 18$$

$$S = (9-18)^2 + (14-18)^2 + (120-18)^2 + (20-18)^2 + (31-18)^2 + (26-18)^2 + (6-18)^2 + (18-18)^2$$

$$= (-9)^2 + (-4)^2 + (2)^2 + (2)^2 + (13)^2 + (8)^2 + (-12)^2 + (0)^2 = 482$$

Now substituting in Equation (2.21):

$$W = \frac{482}{\frac{1}{12}(4)^2(8^3 - 8)} = \frac{482}{672} = 0.717 = 0.72$$

when $N > 7$, the significance of W is tested by converting its value into X^2 with the help of the following equation: $X^2 = K(N-1)W$. Thus, $X^2 = 4(8-1) \times 0.72 = 20.16$ and d.f. in this situation is always equal to $N - 1$. Hence d.f. $= 8 - 1 = 7$.

From the probability table for chi-square we find that the value of chi-square for d.f. = 7 at 0.05 level of significance should be 18.475. Since the obtained value of the chi-square exceeds this required value, one can take this value of W as a significant one. Thus rejecting the null hypothesis, one can say that there is an overall significant relationship in ranking done by the four teachers.

(v) Median test

The median test is used to see if two groups (not necessarily of same size) come from the same population or from populations having the same median. In the median test, the null hypothesis is that there is no difference between the two sets of scores because they have been taken from the same population. If the null hypothesis is true, half of the scores in both the groups should lie above the median and the remaining half of the scores should lie below the median. The following table presents the scores of two groups of students in an arithmetic test. The first step in the computation of a median test is to compute a common median for both distributions taken together.

Scores of 30 Students on an Arithmetic Test

Gr. B (N = 16)	16, 17, 8, 12, 14, 9, 7, 5, 20, 22, 4, 26, 27, 5, 10, 19
Gr. B (N = 14)	28, 30, 33, 40, 45, 47, 40, 38, 42, 50, 20, 18, 18, 19

For computing the common median, both the distributions are pulled together as shown in the following table:

Scores	f
49–53	1
44–48	2
39–43	3
34–38	1
29–23	2
24–28	3
19–23	5
14–18	5
9–13	3
4–8	5
	N = 30

$$\text{Median} = 1 + \frac{(N/2 - F)i}{f_m}$$

$$= 18.5 + \frac{(30/2 - 13)5}{5}$$

$$= 20.5$$

Subsequently, a 2×2 contingency table is built as follows:

Now, the chi-square test can be applied. For computing chi-square from a 2×2 table, we may follow the equation for Chi-square (χ^2). Yates' correction is not needed here because none of the cells contain an expected frequency less than 5.

NOTES

	Above Mdn	Not above Mdn	
Gr. A	A 3	B 13	16
	C 10	D 4	14
	13	17	30

Now substituting the values in Equation (2.13) we get:

$$\chi^2 = \frac{30[|\,3(4)-(13)(10)\,|]^2}{(16)(14)(13)(17)}$$

$$= \frac{30[|\,12-130\,|]^2}{49504} = \frac{417720}{49504} = 8.438 = 8.44$$

$$df = (r-1)(c-1) = (2-1)(2-1) = 1$$

From the probability table for chi-square, we find that for d.f. = 1 the chi-square value at the 0.01 level should be 6.635. Since the obtained value of the chi-square exceeds this value ($p < 0.01$), we can reject the null hypothesis and conclude that the two samples have not been drawn from the same population or from populations having equal medians.

(vi) Kruskal–Wallis *H* test

The primary difference between the *F* test and the Kruskal-Wallis *H* test on the one hand and the Friedman test on the other hand is that the *F* test is a parametric analysis of variance, whereas the *H* test and Friedman test are the non-parametric analysis. The *H* test is a one-way non-parametric analysis of variance and the Friedman test is a two-way non-parametric analysis of variance.

***Table 2.5** H Test from Scores Obtained by Three Groups on Lie Scale*

Gr.A	**($N = 6$)**	**Gr. B**	**($N = 8$)**	**Gr. C**	**($N = 10$)**
15	(14)	17	(15)	20	(17)
10	(9)	9	(8)	25	(19)
8	(3)	8	(6)	13	(12)
5	(4)	14	(13)	11	(10)
6	2)	2	(1)	26	(20)
4		8	(6)	24	(18)
		12	(11)	36	(24)
		18	(16)	30	(23)
				29	(22)
				27	(21)
$R_j = 38$		$R_j = 76$			$R_j = 186$

The H test is used when the investigator is interested in knowing whether or not groups of the independent samples have been drawn from the same population. If the obtained data does not fulfill the two basic parametric assumptions, namely, the assumptions of normality and the assumption of homogeneity of variances, the H test is the most appropriate statistic. The equation for the H test is as given below:

$$H = \frac{12}{N(N+1)}\left[\sum \frac{R_j^2}{N_j}\right] - 3(N+1) \qquad ...(2.22)$$

Where N = number in all samples combined; R_j = sum of ranks in j sample; and N_j = number in j sample.

Data to illustrate the calculation of the H test have been given in above table. Three groups of students were administered a Lie Scale and their scores are presented in table above. The first step is to combine all the scores from all of the groups and rank them with the lowest score receiving a rank of 1 and the largest score by rank N. Ties are treated in usual fashion in ranking subsequently, the sum of ranks in each group or column is found and the H test determines where these sums of ranks are so disparate that the three groups cannot be regarded as being drawn from the same population. The ranks assigned to each score earned by the member of the group are given in brackets. Now, substituting the value in the equation given above, we get

$$H = \frac{12}{(24)(25)}\left[\sum \frac{(38)^2}{6} + \frac{(76)^2}{8} + \frac{(186)^2}{10}\right] - 3(24+1)$$

$$= \frac{53067.192}{600} - 75 = 88.445 - 75 = 13.445 = 13.44$$

When each sample has six or more than six cases, the H test is interpreted as chi-square. In such a situation, d.f. = number of groups or samples minus one. So, here d.f. = 2. Entering the probability table for chi-square, we find hat for d.f. = 2, the value of chi-square at the level of significance should be 9.210. Since the obtained value of H test exceeds this required value, it can be said that the H value is significant. Rejecting the null hypothesis, one concludes that the samples are independent and that they have not been drawn from same population.

(vii) Friedman test

As mentioned earlier, the Friedman test is a two-way non-parametric analysis of variance. When the groups are matched, doubts exist about the two basic parametric assumptions, namely, the assumption of normality and the assumption of homogeneity of variances, the investigator resorts to the Friedman test for testing whether or not the samples have been drawn from the same population. To illustrate this, the calculations of the Friedman test have been presented in table below. The following equation calculates the Friedman test is as follows:

$$X_r^2 = \frac{12}{NK(K+1)}(R_j)^2 - 3N(K+1) \quad ...(2.23)$$

where X_r^2 = Friedman test; N = number of rows; K = number of columns; and R_j = separate sums of ranks of each column.

Table 2.6 *Friedman Test of Two-way ANOVA from Scores of Three Groups on Recall Test*

	I	II	III	IV	V
Gr. A	12(4)	8(2)	4(1)	10(3)	16(5)
Gr. B	10(3)	7(2)	6(1)	11(4)	17(5)
Gr. C	7(2)	83)	4(1)	12(5)	10(4)
R_j	9	7	3	12	14

The first step in calculation of the Friedman test is to rank each score in each row separately giving the lowest score in each row a rank of 1 and the next lowest score in each row a rank of 2, and so on. The ranking can also be done in reversed order, that is, giving the highest score in each row a rank of 1, the next highest score in each row a rank of 2, and so on. Rank assigned to each score in each row is given in parentheses. The Friedman test is applied to determine whether or not the rank totals symbolized by R_j differs significantly. Now, substituting the values in the equation of chi-square, we get

$$X_r^2 = \frac{12}{(3)(5)(5+1)}\left[(9)^2 + (7)^2 + (3)^2 + (14)^2\right] - 3(3)(5+1)$$

$$= 63.866 - 54 = 9.87$$

When the number or rows (N) and the number of columns (K) are too small, the significance of the Friedman test can be ascertained with the help of special tables (Siegel, 1956). For example, when $K = 4$, $N = 2$ to 4 or when $K = 3$, $N = 2$ to 9, the significance of the Friedman test can be done through these special tables. But when the number of rows and the number of columns are greater than those said above, the Friedman test is interpreted as the chi-square test. In the present example, the significance of the Friedman test would be interpreted in terms of chi-square. The d.f. is always equal to $K - 1$ for chi-square applied as a test of significance of the Friedman test. Hence d.f. in the present example would be $K - 1 = 5 - 1 = 4$. Entering the table for d.f. 4, we find that the chi-square should be 9.488 at the 0.05 level of significance. Since the obtained value of the Friedman test exceeds this value ($p < 0.05$), one rejects the null hypothesis and conclude that three matched groups differ significantly.

2.7 CORRELATION AND REGRESSION

Correlational methods are the most commonly used statistical techniques in the testing field. Some important methods of correlation are: Pearson r, Spearman rank difference method and Kendall's T discussed earlier. A correlation coefficient is a mathematical index that describes the direction and magnitude of a relationship.

There is also a related technique called regression which is used to make a prediction about scores on one variable on the basis of known score on another variable. In fact, these predictions are done from the regression line, which is defined as the best-fitting straight line through a set of points in a scatter diagram. It is estimated by using the principle of least squares which, in fact, minimizes the squared deviations around the regression line.

This can be explained through an example. The mean is the point of least squares for any single variable. In other words, the sum of squared deviations around the mean will be less than it is around any value other than mean. For example, the mean of the five scores namely, 2, 3, 4, 5 and 6 is ...$X/N = 20/5 = 4$. The squared deviation of each score around the mean can now be easily determined. For score 6, the squared deviation is $(6-4) \times 2 = 4$; for score 5, the squared deviation is $(5 - 4) \times 2 = 1$. The score 4 is equal to the mean and therefore, the squared deviation around mean will be $(4-4) \times 2 = 0$. Thus, by definition, the mean will always be the point least squares. The regression line is the line of least squares or ruining mean in two dimensions or in the space created by two variables.

As described earlier, a regression line is the best-fitting straight line through a set of several points in a scatter diagram, which is the picture of the relationship between two variables. The regression line is described as a mathematical index called regression equation. The general linear regression equation for the straight line is:

$$Y = bx + a \text{ (figure 2.4)} \qquad ...(2.24)$$

Where a is the intercept, that is, value of Y when X is zero. In other words, the point at which the regression line crosses the Y-axis (a) is found by using the following formula:

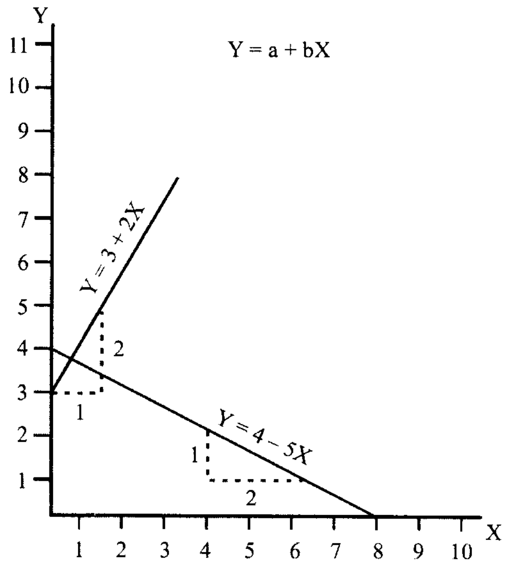


Fig. 2.4: $y = b_2 + a$

$$a = Y - bx \qquad ...(2.25)$$

b is the slope of regression line or it is called the regression coefficient. It is expressed as the ratio of sum of squares for the covariance to the sum of squares for X. In the above figure, two regression lines have been drawn based on equation $Y = a + bx$. The regression equation gives a predicted value for Y, on the basis of X. This predicted value is called Y'. In fact, the actual score and predicted score on Y is rarely exactly the same. Some difference between the observed and the predicted score, that is, $Y - Y'$ occurs and this is technically called the residual. Symbolically, the residual thus, is defined as $Y - Y'$. The best-fitting line actually keeps the residual to minimum. In other words, it minimizes the deviation between the observed and the predicted Y score. Since residuals can be either positive or negative and will cancel to zero if averaged, the best-fitting line is rightly estimated by squaring each residual. In this way, it can be said that the best-fitting line is best obtained by keeping these squared residuals as small as possible.

Difference between correlation and regression

The correlation coefficient has a reciprocal nature whereas regression lacks this property. The correlation between X and Y will always be the same as the correlation between Y and X. For example, if the correlation between test anxiety and test performance is 0.64, the correlation between test performance and test anxiety is also 0.64. One cannot claim similar thing about regression. In fact, regression is used to transform one variable into estimated scores on the variable. In other words, regression is often used to predict raw scores on Y on the basis of raw scores on X. For example, through regression equation, the investigator might be interested in predicting score on the intelligence test on the basis of score on, say, the anxiety test. The coefficient denoting regression of X on Y is usually not the same as the coefficient denoting the regression Y on X. Since regression uses the raw scores of the variables, the trait of reciprocity does not hold true.

Regression analysis tends to show how changes in one variable are related to changes in another variable. In psychological testing, regression is often used to determine whether changes in the test scores are related to the changes in performance. For example, do persons who score higher on mechanical reasoning test perform better as a mechanical engineer? In fact, regression analysis and related correlational methods show the extent to which these variables are linearly related. Besides, they also often create an equation that estimates scores on the criterion (such as performance as mechanical engineer) on the basis of scores on a predictor (such as scores on mechanical reasoning test).

Choosing appropriate statistical tests

It is customary to choose the appropriate statistical tests on the basis of the nature of the obtained data.

If the data fulfills the requirement of parametric assumptions, any of the parametric tests which suit the purpose can be selected. On the other hand, any of

NOTES

the non-parametric statistical tests that suits the purpose can be chosen, if the data does not fulfill the parametric requirements. Other factors to be kept in mind in While selecting the appropriate statistical tests other factors like the number of independent and dependent variables and the nature of the variables—whether they are nominal, interval or ordinal—should be kept in mind.

When both the independent variable and the dependent variable are interval measures and are more than 1, multiple correlation is the most appropriate statistic. On the other hand, when they are interval measures and their number is only one, Person r may be used. As has been noted earlier, with ordinal and nominal measures the non-parametric statistics are the common choice. Sometimes, researchers transform the measures so that the appropriate statistical test may be applied with out loss of much information. For example, if scores of two groups on interval measures are available but the data does not fulfill the requirement of the t test, the researcher can transform the interval measures into ordinal measures and subsequently apply the Mann-Whitney test.

CHECK YOUR PROGRESS

15. What is the chi-square test?
16. Define the Mann-Whitney U test.
17. Define the Spearman rank-difference method.
18. What is coefficient of concordance W?
19. When is the freedman test of two-way analysis of variance used?

2.8 SUMMARY

- While carrying out research, defining the problem is the most important step. Once this is done, the next step is formulating a hypothesis, which is a tentative solution to the problem, in the view of the researcher. Hypotheses are of two types—solvable and unsolvable.
- There are different types of hypotheses which include causal hypothesis, descriptive hypothesis, simple hypothesis, complex hypothesis, research hypothesis, null hypothesis and statistical hypothesis.
- To test the hypotheses, the following kinds of tests are used: parametric tests and non-parametric tests.
- In a parametric statistical test, certain conditions about the population parameter from which the sample is drawn, are specified. Examples of paramedic tests are z test, t test and F test.
- On the other hand, in a non-parametric test, no conditions are specified about the parameters of the population from which the sample is drawn.

NOTES

- The appropriate statistical tests are chosen on the basis of the nature of the obtained data. If the data fulfills the requirement of parametric assumptions, any of the parametric tests which suit the purpose can be selected. On the other hand, any of the non-parametric statistical tests that suit the purpose can be chosen, if the data does not fulfill the parametric requirements.
- Other factors to be kept in mind while selecting the appropriate statistical tests other factors like the number of independent and dependent variables and the nature of the variables—whether they are nominal, interval or ordinal—should be kept in mind.

2.9 KEY TERMS

- **Hypothesis:** A proposition that can be tested and suggests a possible solution.
- **Research hypothesis:** A working hypothesis which is derived from researcher's theory.
- **Null hypothesis:** A hypothesis that negates difference
- **Statistical hypothesis:** A hypothesis where in the null hypothesis (H_0) and the alternative are expressed in numerical terms.
- **Parametric statistical test:** A test in which certain conditions, about the population parameter from which the sample is drawn, are specified.
- **Non-parametric statistical test:** A test which does not specify any conditions about the parameters of the population from which the sample is drawn.

2.10 ANSWERS TO 'CHECK YOUR PROGRESS'

1. Research is an inquiry for gaining knowledge. It is a systematic search for information on a chosen topic. The definition of a research problem has to be carefully scrutinized for the following factors: (i) There should to a person group or institution to wham a give problem can be attributed; (ii) Two sources of action must be possible; (iii) There must be at least two possible outcomes: and (iii) the course of action must have a bearing on the outcome'.
2. While selecting a research problem, a researcher should follow the following principles: (i) Avoid an often researched topic; (iii) Avoid controversial issues; (iii) Keep costs, time and competencies in mind; (iv) Choose a problem of genuine interest and (v) Carry out a pilot study, if possible .
3. A research problem should be stated in operational terms. The time-frame should be set and the scope and limitations of the study should be clearly specified.

4. Research problems are of two types: solvable and unsolvable.
5. A hypothesis is a suggestion about a possible solution. It is a patented association between variables, a conjecture about a relationship, and a testable proposition.
6. A null hypothesis is a hypothesis that negates the research hypothesis. It is a no difference hypothesis.
7. Statistical hypothesis is also known as the alternate hypothesis. This hypothesis states that the two groups that are being studied, would differ.
8. In a parametric statistical test, certain conditions are specified about the population parameter from where the sample is drawn.
9. A non-parametric test does not specify any conditions about the parameters of the population, from which the sample is drawn.
10. When significance of difference between two means is to be tested the students *t* test or t ratio or the *z* test or *z* ratio, is used.
11. Degrees of freedom is the number of observations that are independent of each other and cannot be deduced from each other.
12. Experimental variance is the effect of manipulation of the independent variable on the dependent variable.
13. An error variance is the variability in measurements that occur as a result of factors that are beyond the control of the experimenter. In psychological research individual differences, motivations, attitudes etc are typical factors that contribute to error variance.
14. The Pearson *r* is a common correlation method where the size of the *r* varies from +1 – 0 + – 1 this gives the degree of association.
15. The chi-square is a general purpose statistic. This test is used when data are presented in frequencies, proportions or percentages. It can only be used when the data are discrete.
16. The Mann-Whitney *U* test is a test of difference between two means independently obtained from two groups or samples. The scores have to be in terms of ranks.
17. The Spearman rank-difference method is symbolized as *ñ* (read as rho). It is a common method for calculating coefficient of correlation between sets of ranks
18. Coefficient of concordance *W* was developed by Kendall and is a measure of correlation between more than two sets of ranks, events or individuals—this is used when one wishes to study the inter-test reliability.
19. The Freedman test of two-way analysis of variance is useful when matched groups and their parametric conditions are in question.

NOTES

2.11 QUESTIONS AND EXERCISES

Short-Answer Questions

1. List any two specifications for a research problem.
2. List two considerations while selecting a research problem.
3. Define parametric statistical test.
4. Define non-parametric statistical test

Long-Answer Questions

1. Explain a research problem and hypothesis in detail.
2. Explain in detail the three most important parametric statistics: (i) Student's t test and z test, (ii) F Ratio, and (iii) Pearson r.
3. Explain sample selection, types and size and the steps in sampling in detail.

2.12 FURTHER READING

Guilford, J.P. and B. Fruchter. 1978. *Fundamental Statistics in Psychology and Education*. New York: McGraw-Hill.

Kerlinger, F.N. 1973. *Foundations of Behavioral Research*. New York: Holt Reinhart and Winston, Inc.

Kothari, C.R.1995. *Research Methodology—Methods and Techniques*. New Delhi: Wishwa Prakashan.

Singh, A.K. 2008. *Tests, Measurements and Research Methods in Behavioural Sciences*. New Delhi: Bharati Bhawan.

Winer, B.J. 1971. *Statistical Principles in Experimental Design*. Tokyo: McGraw-Hill Kokakusha.

UNIT 3 SAMPLING DESIGN

NOTES

Structure

3.0 INTRODUCTION

In this unit, you will learn how to design the right sample for a chosen study. All items in a field of inquiry are thought to constitute a universe or a population. A sample is any number of persons selected to represent a population according to some rule or plan. Hence, a sample is a smaller representation of the population. A measure based on a sample is known as a statistic.

To measure the characteristics of population, a small sample that nearly represents the whole population is picked and studied. And then the findings are projected over the entire population to get the approximate overall picture. Sampling methods are basically of two broad types—probability sampling and non-probability sampling. Again each broad type is divided into subtypes.

In this unit, you will also learn different sampling errors and how to minimize those errors.

3.1 UNIT OBJECTIVES

After going through this unit, you will be able to:

- Explain the concept of sampling and sampling designs
- Know the steps in sampling design and know the concepts of universe sampling unit, scores list, size of sample and other matters
- Examine the criteria involved in a sample selection

- Gain familiarity with the probability and non-probability sampling techniques
- Learn how to select right size of the sample and how to minimize sampling errors

NOTES

3.2 SAMPLING DESIGN

A population is the aggregate of all the cases that conform to the researcher's designated set of specifications. Therefore, the term people may mean all the residents of India, or those engaged in factory work, or women, boys under the age of 20 and so on, as defined by the researcher. By specification all the boys under 20 would be included in the population of India, can be referred to as a sub-population or stratum with reference to the main population.

A stratum may be defined by one or more specifications that divide a population into mutually exclusive segments, e.g., a given population may be subdivided into strata of males under the age of 21 and females under age of 21. Similarly, one can have a stratum based on education, income, etc.

A single member of a population is known as an element. Often, one wants to know how certain characteristics of the elements are distributed in a population, e.g., one wants to the age distribution of people who have a particular political preference.

A census is a count of all the elements in a population and a determination of the distribution of their characteristics, based on the information obtained for each of the elements. It is economical in terms of time, effort and money to get the desired information for only some of the elements than for all of them.

When we select some of the elements with the intention of finding out something about the population from which they are taken, we refer to that group of elements as a sample. The expectation here is that what we find out about the sample is true of the population as a whole. This depends on the way the sample is selected.

A measure use based on the entire population is called a parameter. A sample is any number of persons selected to represent the population according to some rule or plan. So, a sample is a smaller representation of the population. A measure based on a sample is known as a statistic.

All items in a field of inquiry are thought to constitute a universe or a population. A complete enumeration of all the items in the population is called a census inquiry. It is presumed that a census enquiry gives the highest possible accuracy. In practice, this may not be so. Even the slightest error may get magnified as the observations increase. However, such a census type of enquiry costs lots of time, money and effort. Therefore, a smaller sample is chosen for study. The methods used to select samples are called sampling techniques. The survey that follows is called a sample survey.

NOTES

Factors influencing decisions while drawing a sample are:

- Size of the population: When the population size is large; the selection of a sample becomes necessary.
- Costs involved in obtaining the elements: If the cost is reasonable; the sampling inquiry is facilitated.
- Convenience of availability of the elements. Each of these factors is important for deciding to select a sample, for study.

Implications of sample design: A sample is obtained according to a 'plan'. A sample design is a technique for selecting the items for a sample. The size of the sample means the number of items to be included in the sample. Sample design should be determined before data collection and the sample should be designed to suit the study.

Sampling is a process of selecting a few from a bigger group for estimating or predicting the prevalence of some outcome / factor regarding the bigger groups. So, a sample is a sub-group of the population, one is interested in. This is the concept of sampling see figure 3.1

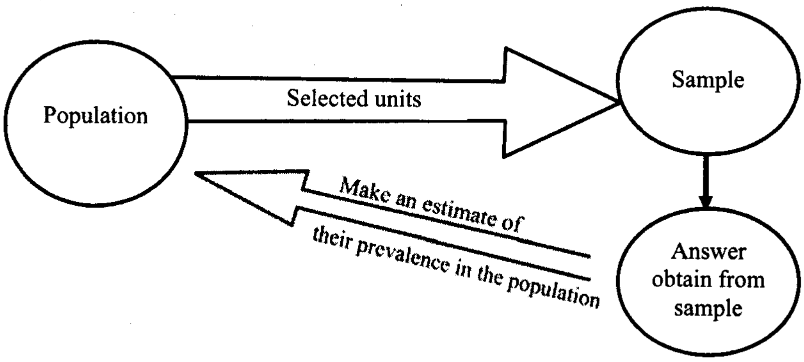


Fig. 3.1 Sampling

3.2.1 Steps in Sampling Designing

The following points should be kept in mind before designing a sample

- **Type of universe:** Define the universe or set of objects to be studied. The universe can be finite or infinite. In the finite universe, the number of items is certain, while in the infinite universe, it is uncertain. An example of the former is the number of industrialists in a country, and the latter, the number of stars in the sky.
- **Sampling unit:** This is the group from which the sample is to be drawn. For example, a population unit can be in terms of people's age, gender, etc and a housing unit like bungalow, flat, etc, or an educational unit like university, college, school, etc.
- **Service list:** The sampling from a list index or other population records from which the sample is to be drawn, e.g., prepare all the items in the universe from which the selection of the sample can be made. It should be

NOTES

comprehensive, correct and reliable, so that the sample becomes representative.

- **Size of the sample:** This refers to the number if items to be selected from the population, to constitute the sample. An optimum sample size should be reliable, flexible and representative. The size could be determined by the precision with which estimations are needed. Cost considerations also come into play, here.
- **Parameters of interest:** This involves the type of measures needed from the sample. What factors you want to study.
- **Budgetary constraints:** This refers to the practical problems about the size of the sample and costs associated with the collection of data from the sample.

3.2.2 Principles for Selecting a Sampling Procedure

There are three principles which guide sampling theory

(i) In majority of cases of sampling there will be a difference between the sample statistics and the true population mean, which is attributable to the selection of the units in the sample.

(ii) The greater the sample size; the mare accurate will be the estimate of the true population mean.

(iii) The greater the difference in the variable of the study in a population for a given sample size, the greater will be the difference between the sample statistics and the true population mean.

3.2.3 Systemic Bias and Sampling Errors

There are two reasons for incorrect inferences arising out of sampling. They are:

(i) Systemic bias

(ii) Errors in sampling procedures

(i) Systemic bias: This bias can arise from one or more of the following reasons:

a. *Inappropriate sampling:* This means there is a bias in the representation of the universe, from where the sample is drawn,

b. *Defective measuring device:* A physical measuring device is faulty or the questionnaire or the interviewer has a bias. This defect would lead to a systematic bias.

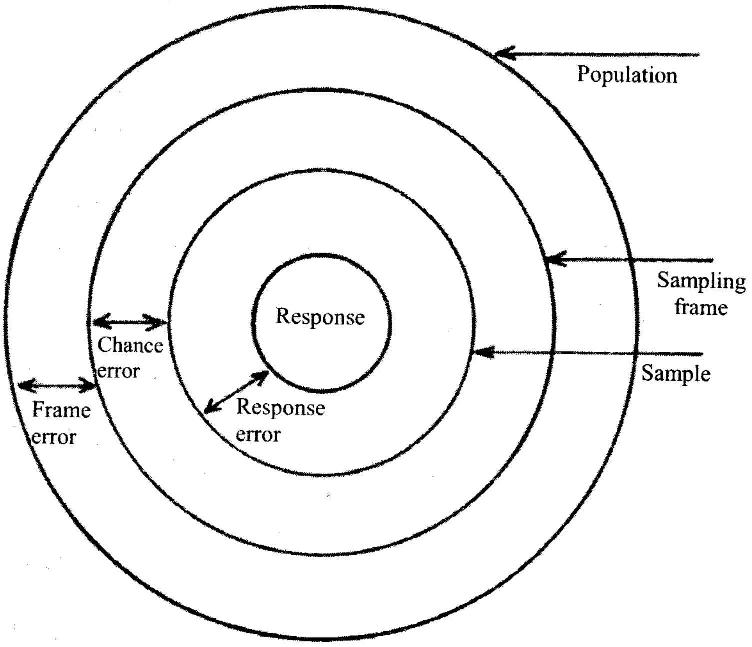
c. *Non-responding:* Inability to sample all the individuals initially included in the sample, could give rise to a bias.

d. *Indeterminacy principle:* Individuals act differently while being observed. This can cause a systematic bias.

e. *Natural bias in the reporting of the data:* For example, people understate their incomes when the government asks for it, but overstate when social status is involved. In psychological surveys there is a tendency to give a 'right' answer, rather than a true one.

(ii) **Sampling errors:** These are random variations in the sample estimate around the true population mean. Sampling errors are errors which arise from inaccurate sampling and they generally happen to be random variations (when sampling is random) in the sample estimates around the true population values. see figure 3.2.

NOTES



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. 3.2 Sampling Error

Source: Kothari, C.R. Research Methodology Methods and Techniques, 1995.

Sampling error = Frame error + Chance error + Response error

Sampling error is compensatory in nature. And the expected value of such error happens to be equal to zero. The magnetite of the sampling error depends on the nature of the universe. A larger sample design is yet another way to reduce sampling error. Select a sample procedure that helps control systematic bias and sampling error.

3.2.4 Criteria of Good Sample Design

The following are the criteria of a good sample design:

(i) Sample design should yield the best possible sampling error.

(ii) Sample design should yield the best possible sampling error.

(iii) Sample design should be chosen judiciously, keeping costs in mind.

(iv) Sample design must attempt to control systematic bias in the best possible way.

(v) Sample findings should be applied with a reasonable level of confidence.

NOTES

3.2.5 Types of Sampling Designs

For selecting elements on the representation basis, the sample may be obtained using either probability sampling or non-probability sampling. Probability sampling is based on random selection whereas non-probability sampling is based on non-random sampling.

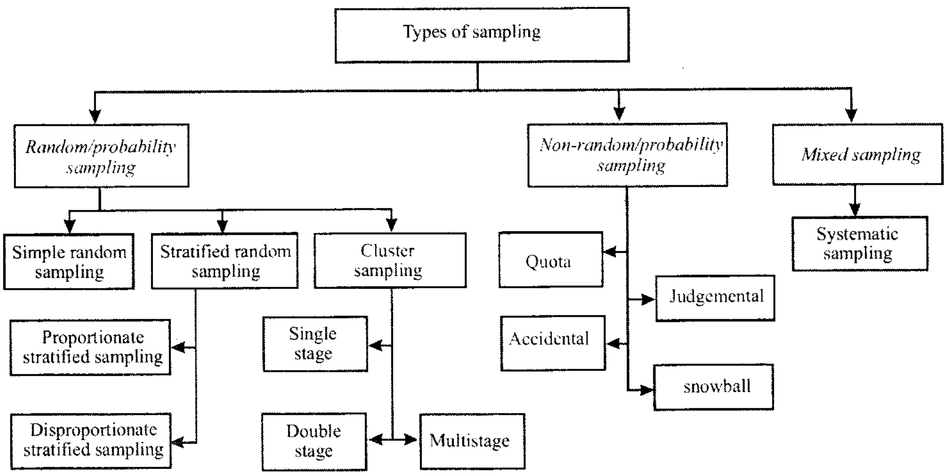
On element selection basis, the sample may be either restricted or unrestricted. When each sample element is drawn individually from the population at large, the sample is known as unrestricted sample. All the other forms of sampling are covered under the term restricted sample.

Basic sampling designs

Element selection technique ↓	Representative basis ←→	
	Probability sampling	Non-probability sampling
Unrestricted sampling	Simple random sampling	Haphazard or convenience sampling
Restricted sampling	Complex random sampling, e.g., cluster sampling systematic sampling stratified sampling etc.	Purposive sampling, e.g., quota sampling judgment sampling

Source: Singh, A.K., Tests, Measurements and Research Methods in behavioural sciences, 2008.

***Fig. 3.3** Sampling Designs*



Source: Kothari, C.R. Research Methodology Methods and Techniques, 1995.

***Fig. 3.4** Types of Sampling*

3.2.6 Types of Sampling

Strategies for sampling can be categorized as:

- Probability sampling designs
- Non-probability sampling designs
- Mixed sampling designs

3.2.6.1 Probability (or random) sampling designs

For a random sampling design it is imperative that each element is included in the sample. Equal element means the probability of each element in the population is the same and is not influenced by other considerations. The notion of independence means that the choice of one element is not dependent upon the choice of another element in the sample. Only if both these conditions are met would the sample be a random / probability sampling. If not, bias would be introduced. As this sample represents the total population; the inferences drawn from such samples can be generalised to the total population.

Methods of drawing a random sample

The three most common methods are

(i) Fish bowl draw method

(ii) Using table of random umbers

(iii) Method of computer determined randomness

The Fishbowl method: This method is useful when the total population size is small. Here, each element is numbered on a small slip of paper and put into a box or bowl. After the slips are mixed thoroughly, the investigator selects one number at a time, until the desired sample size is reached.

Using the table of random numbers: Most research methodology and statistics tools have a table of randomly generated numbers in their appendices. The sample can be selected from these tables by following the procedures described below. The principle of random table of numbers is that the column-row numbers do not appear in any particular sequence, nor does any number appear more frequently than the other one.

The random sample selection process:

a. Specify the number of elements in the study population
b. Number them from 1 to N (N is the total number of elements in the population). Example, if the population size is 500 and one intends to choose 50 cases from the population, now enter the random table at any point.
c. Then move systematically to the right, left, up, down or diagonally skipping the numbers that are too large and also those that have already been drawn.

d. Keep moving through the table and fill until the selected sample has 50 elements.

NOTES

A random selection of 30 cases from a population serially numbered from 1 to 80 is illustrated here. 12 numbers have been omitted. Numbers 85, 84, 97 and 95 have been omitted because they exceed 80 and numbers 03, 74 and 12 have been omitted because they have duplicated the previous selections. The number 00 has been left out because the population number starts from 01.

The following is a sample random number table:

64755	83885	84122	25920	17696
10302	52289	77436	34430	38112
71017	98495	51308	50374	66591
60012	55605	88410	34879	79655

Sample (N = 30 has been taken from a serially numbered population of 80).

64	59	28	12	85	55
75	20	97	71	03	60
58	17	74	01	74	
38	69	36	79	66	
85	61	34	84	59	
84	03	43	95	16	
12	02	03	51	00	
22	52	81	30	12	

The advantage of using the table of random numbers is that it is easily accessible to the researchers and requires no formal training for using it. However, the disadvantage is that it cannot be easily and constructively used when the size of the population exceeds 5 digits.

Method of computer-determined randomness: This method is useful when the population size is large. The data are fed into the computer to obtain a random number of elements corresponding to the elements in the population. This method is easy and fast and therefore is increasingly in use. The only limitation is imposed by the need to have a computer.

Different systems of drawing a random sample

There are two ways of selecting a random sample

(i) Sampling without replacement

(ii) Sampling with replacement

Simple random sample

A simple random sample (also known as an unrestricted random sample) may be defined as one in which each and every individual of the population has an equal chance of being included in the sample and also the selection of one individual is in no way dependent upon the selection of another person. For example, if we are to select a sample of 10 students from the seventh grade consisting of 40 students, we can write the name (or roll number) of each of the 40 students on separate slips of all equal in size and colour and fold them in a similar way. Subsequently, they may be placed in a box and reshuffled thoroughly. A blindfolded person, then, may be asked to pick up one slip. Here, the probability of each slip being selected is 1/40. Suppose that after selecting the slip and noting the name written on the slip, he again returns it to the box. In this case, the probability of the second slip being selected is again 1/40. But if he does not return the first slip to the box, the probability of the second slip becomes 1/39. When an element of the population is returned to the population after being selected, it is called sampling with replacement and when it is not returned, it is called sampling without replacement. Sampling with replacement is wholly feasible except in certain situations where it is seldom used (Cochran, 1963). If sampling with replacement is used, the chance of the same case being selected more than once is increased. In such a situation, the repeated cases may be ignored as is done if a table of random numbers is used in making a selection of cases. Thus random sampling may be defined as one in which all possible combinations of samples of fixed size have an equal probability of being selected.

The major difference between sampling with replacement and sampling without replacement can be illustrated though an example. Suppose the size of the population consists of four persons, who are named as A, B, C and D. Suppose that the investigator wants to select samples of size 2 through the procedure of sampling without replacement. In such a situation, the investigator can maximally draw six samples of size 2 from the population of 4. This could be accomplished with the help of the following equation.

$$\binom{N}{n} = \frac{N!}{(N-n)!n!}$$

where

N = the size of parent population

n = the size of the sample

! = factorial

In the above example where $N = 4$ and $n = 2$, the maximum number of sample size of 2 would be 6 as under.

$${}^4C_2 = \frac{4!}{(4-2)!2!} = \frac{4\times3\times2\times1}{2\times1\times2\times1} = \frac{24}{4} = 6$$

Similarly, where $N = 5$, we can have 10 sampled of size 2 as under.

$$^5C_2 = \frac{5!}{(5-2)!2!} = \frac{5\times4\times3\times2\times1}{3\times2\times1\times2\times1} = \frac{120}{12} = 10$$

NOTES

But from the same population, we can have 5 samples of size 4 as under:

$$^5C_4 = \frac{5!}{(5-4)!4!} = \frac{5\times4\times3\times2\times1}{1\times4\times3\times2\times1} = \frac{120}{24} = 5$$

But if the investigator has decided to proceed with the technique of sampling with replacement, he can derive the likely number of samples from the given population with the help of the following equation.

$$N^n$$

where N and n are again population and the sample size. Suppose the size of the population is 4 and the size of sample is 2. In such a situation the investigator, following the technique of sampling with replacement, can maximally draw 16 samples, that is, $4^2 = 4\times4 = 16$. If the four members of population are named as A, B, C and D, the sixteen samples of size 2 would be

AA	AB	AC	AD
BA	BB	BC	BD
CA	CB	CC	CD
DA	DB	DC	DD

The case of AA, BB, CC and DD combinations reflects the fact that in sampling with replacement, an element or individual once drawn can be drawn again. In actual practice such cases are ignored.

There are some advantages and disadvantages of simple random sampling are as given below:

Advantages of simple random sampling:

(i) It is a representative sample.

(ii) It is assumed that all the characteristics of the population are reflected in the sample.

(iii) This is the easiest and simplest of all probability sampling methods.

(iv) This random sampling can be applied in conjunction with many other probability sampling methods.

(v) The sampling error can be easily estimated.

Disadvantages of simple random sampling:

(i) It is difficult to ensure that the smaller elements that exist in a population are included in the sample. For example, in a population of 500 persons, only

NOTES

12 people are dialectic. The sample size is only 50. The chance that they would be included is very slim.

(ii) The simple random sampling method cannot fully avail of certain information available in the sample. For example, if one knows that there is a population of children, who are bright, dull, those who are artistic, etc., all these elements cannot be brought into the sample.

(iii) Sampling error is greater in the random sample than in a stratified random sample. To minimize the error stratified random sample is used.

3.2.6.2 Mixed (probability and non-probability) sampling methods

Stratification (also called convenience) sampling method

Stratified random sampling is of two types

(i) Proportionate stratified random sampling

(ii) Disproportionate stratified random sampling

(i) Proportionate stratified random sampling

Here the population is first stratified. Then the sample is drawn randomly from the stratum the proportion of persons belonging to each strata of the total population is useful in applied research, e.g., the rank of army officers of one strata in the defence force and so on. Number of cardiologists in the total population of darters, etc., there proportional representation in the sample is a valuable index, for determining further studies.

In stratified random sample, the population is first divided into two or more strata, e.g., on the basis of age-groups of say 30–40, 40–50, 50–60, etc., or as male and female. These divided populations are termed sub-populations. These are non-overlapping parts of the whole population. They are thought to be homogeneous. Then sample elements are selected from each stratum using simple random sample.

Advantages of stratified random sampling:

a. In stratified random sampling even those elements that exist in smaller numbers, get pieced up, and sampling error is minimized because the sample has all the characteristics of the population.

Disadvantages of stratified random sampling:

a. It is a difficult method to start with in terms of assumptions of knowledge of the composition of the population.

b. It is a time consuming method.

c. It is could give rise to classification error.

d. Disproportionate stratified random sampling.

e. It is similar to the earlier me that, but is different in that the substrata are not necessarily distributed according to their proportionate weight in the population from which they were randomly selected. There could be overrepresentation or under representation of a strata in a population.

(ii) Disproportionate stratified sampling means

a. The investigator would give equal weight to each of the substrata

b. Will give greater or representation to some substrata and not enough weight to other substrate in the sample to be brawn.

Advantages of disproportionate stratified random sampling:

a. Comparatively lees time consuming than propitiate stratified random sampling

b. It is possible to give weight to particular elements that are not represented frequently in the population, as compared to other elements.

Disadvantages of disproportionate stratified random sampling:

a. The samples drawn might suffer from cretin stratum being under or overrepresented.

b. Assumption that the knowledge of the composition of the original population limited use in conditions, where this information is not available.

c. Possibility of misclassifying elements into a sub-strata; that they do not belong, legitimately, e.g., element *X* may be put into strata *Y*, but it may belong to strata *Z*.

d. This method is common in the social sciences, despite these limitations.

Area or cluster sampling

This is another important method of probability sampling. This is used often in field research. Here, geographical divisions in terms of territory, zone cities, towns, districts, etc., are made and a number is assigned to each area. Then the random sample is drawn from this area. This method is also called cluster sampling. For example, one wants to study the prevalence of HIV in a state. Then the state is divided into districts, downs and villages, etc. identified. Then finally even households in locality could be identified for studying. This is also called multi-stage sampling.

Advantages of area or cluster sampling:

a. This is useful for large scale studies of area, regions where specific lists are an obtainable useful for pubic opinions polls.

b. It saves time and money

c. The respondents can easily be substituted. It is a flexible method

Disadvantages of area or cluster sampling:

a. Sampling errors rarely come to light.

b. Little control over even the size of the cluster

c. Could result in bias in samples

However this method is used in large behaviour studies.

3.2.6.3 Non-probability sampling methods

Quota sampling

This is an important non-probability sampling method. The population is seen as made up of strata of the population and from each stratum, individuals are chosen randomly, e.g., if the population of students in a school is 5000 made up of high and low socio-economic classes. From this, 500 students can be chosen with 250 from the higher and 250 from the lower class. This is the quota sample.

Advantages of the quota sampling method:

(i) Quota sampling method is quick and easy for gross estimates

(ii) It is a convenient procedure

(iii) Elements from the desired strata get included

Disadvantages of the quota sampling method:

(i) In quota sampling method randomness cannot be established

(ii) In this method, generalizability is poor

(iii) since most credible elements included, it may not be typical of the population

(iv) Possibility of classification error high.

(v) Other variables of significance cannot be controlled.

This method does have some appeal, despite limitations.

Purposive sample

Purposive sample, which is a handpicked sample is typical of the population. It is also called judgmental sample, because the choice is determined by the judgments of the researcher, e.g., attitudes towards corruption would be ascertained by interviewing professionals, academicians, tainted people and politicians. The investigator selects the persons from these select people.

Advantages of purposive sample:

(i) Purposive sample is cost effective and easily accessible

(ii) It is very convenient

(iii) Only relevant individuals get included

Disadvantages of purposive sample:

(i) In purposive sampling randomness representatives not ensured.

(ii) Generalizability is poor.

(iii) Sampling is highly subjective.

(iv) Inability to apply inferential statistics to acceptable levels

Accidental sampling

Accidental sampling is also called incidental sampling. It is based on a non-probability sampling plan. Here, the investigator chooses the sample according to convenience. Convenience and economy guide this method, as a useful option, e.g., students of a particular school chooses because of availability

Advantages of accidental sampling:

(i) Accidental sampling is the most convenient method of sampling

(ii) Economical in terms of time, money, effort and cost.

Disadvantages of accidental sampling:

(i) Accidental sampling has poor generalizability

(ii) Biases and prejudices bound to exist.

(iii) There is a high probability of high sampling error

Accidental sampling is used in psycho-social research due to the convenience factor.

Systematic sampling

This is another method of the non-probability sampling plan. Here, every *n*th person is drawn from a predetermined list for study, e.g., every 5th roll number of class of 50 students or every 10th name from the telephone directory and so on. It is systematic in view of the fact that the selection is made according to a pre-determined plan. The first element selected is random and has non-probability characteristics

Advantages of systematic sampling:

(i) The method is quick and it is easy to obtain the sample.

(ii) It is easy to select, say, every fifth name.

(iii) It is easy to use (simpler than random table chart)

Disadvantages of systematic sampling:

(i) It is not a probability sampling plan

(ii) High sampling error possibility creeps in

(iii) Bias could invade into the sample

This method still finds favour in psycho-social research.

Snow ball sampling

Snow ball sampling is a non-probability sampling method. It is basically a socio-metric method. Here all the persons of a group or sample are identified by friends or other acquaintances. The snowball effect is how one person's contact leads to the others and from there further on. The information gathered is through patterns of friendship, e.g., how rumors spread and how advertisements influence friends, etc. It is useful for small *N* samples, below 100. It is a method to study social change.

Advantages of snow ball sampling

(i) Snow ball sampling is a systematic sampling technique helpful in studying small informal networked social groups.

(ii) Helps in community studies, decision-making

(iii) This technique lends itself to computer-determined use of random numbers

Disadvantages of snow ball sampling

(i) Snow ball sampling method is not feasible with large numbers

(ii) No use of statistical methods, possible

(iii) Bias could enter the sample

(iv) Largely useful for information dissemination, saturation sampling and dense sampling.

Saturation sampling

Saturation sampling involves drawing all elements or individuals having characteristics that are of interest to the researcher, e.g., all psychiatrists below the age of 45 years. Dense sampling is a method that lies between simple random sampling and saturation sampling. Here researcher selects 50 per cent or more from the population and takes a majority of individuals having specific characteristics that are of significance, e.g., 500–600 students from a class of 1000 students. These two methods are convenient. But it is not useful when the N exceeds 1000.

Double sampling

Double sampling means drawing a sample of individuals from a sample that has already been drawn earlier, e.g., from a population of 10,000 people, a sampling of 1000 is drawn. Again from this 1000, a further sample 200 is drawn, for the study, e.g., a questionnaire is sent to 1000 people on the issue of pollution. Say, 50 per cent (or 500) of them respond. From these 500 persons, a sample of 100 is drawn for an in-depth interview. This is double sampling.

If this method is to be meaningful for research, then the representativeness of the sample must be increased. It is a time-consuming and labour intensive method.

Pre-requisites for a sample to be good are as follows:

(i) It must be representative

(ii) It must be of sufficient size

(i) Representativeness: The sample must have approximately all the characteristics of the population that are relevant for the investigation. The relevant characteristics of the population must exist in the same proportion in the sample also, e.g., if the population has an education level ratio of 60:40 for men and women, the sample should also have a similar ratio, to be representative. The population should be clearly defined and the observations to be made should be specified. This can reduce biases.

NOTES

(ii) **The sample size should be adequate:** This implies that the size be sufficient. A larger sample is better for reducing the error. This is the difference between the population value and the sample value. The larger the size of the sample, the lesser is the error. However, too large a sample may not yield better results, as a large sample creates other problems.

Advantages of sampling methods in general

(i) Using sampling methods increases accuracy. Examining a sample becomes both efficient and involves lesser work, so the purpose of a sample is to get maximum accuracy with minimal effort, time, money, etc.

(ii) It reduces the cost as the data is from a smaller number of cases. Statistical calculations for accidental errors are also reduced.

(iii) Since the sample and not the universe is studied, work proceeds fast. This is a great advantage for research.

CHECK YOUR PROGRESS

1. What is a census survey?
2. What is a simple random sample?
3. Why are probability sampling procedures preferred to non-probability sampling procedures?
4. What is random sampling?
5. What is quota sampling?
6. What is systemic bias?
7. Define cluster sampling.

3.3 STATISTICAL DISTRIBUTIONS

The following three types of distributions are of interest to any researcher:

(i) Sample distribution

(ii) Population distribution

(iii) Sampling distribution

Fundamentals of sampling

Sampling is the process of obtaining information about an active population by examining only a part of it. One draws inferences about the parameters of a population from which the samples are obtained. The assumption behind this is that the sample data would enable an estimate of the population. The items selected from the population for observation is called a sample. The method of selecting a sample is called sample design. Any survey that is carried out on the basis of the sample is referred to as sample survey. To draw valid and reliable conclusions, the sample must be truly be representative of the population.

Need for sampling

Sampling reduces time and money. It is less expensive and requires less time. Sampling makes measurements more accurate, due to smaller size.

Sample is the only means to study a large population. When items are to be deconstructed, sample is the only means for study.

Some fundamental definitions: These are to familiarize oneself with the concepts of sampling.

(i) **Universe population:** The total number of items in any field of study is called the universe. The population refers to the total units or items about which information is required. The attributes that are the object of the study are called the characteristics and the units possessing them are known as elementary units. The aggregate of such units is the population.

All units in any field of study constitute the universe. All elementary units are the population often the two terms are used interchangeably, however, research needs this distinction. The population or universe can be of two types: (i) Finite and (ii) Infinite.

A finite population consists of fixed number of elements and the elements can be enumerated totally, e.g., the number of students in a state. The symbol N is used to depict the number of elements or items of a finite population.

An infinite is one where all the elements cannot be observed, at least theoretically, e.g., the number of stars in the sky. In a sense, a very large finite population is an infinite population.

(ii) **Sampling frame:** The elementary units that form the basis of the sampling process are known as sampling units. A list of all such sampling units is the sampling frame. The sampling frame is a count of a list of items from which the sample is drawn. For research, a frame has to be constructed from the population, to draw the sample, e.g., names from the census records or telephone directory, etc., for conducting a study on a sample that is drawn from the frame. Telephone directory is a frame, from which names are drawn to get the sample.

(iii) **Sampling design:** This is the plan for obtaining a sample from the frame. It is the procedure or technique for obtaining those sampling units from which inferences can be made. The sampling design has to be prepared well in advance of undertaking any research.

NOTES

(iv) **Statistic(s) and parameter(s):** A statistic is the characteristic of the sample whereas the parameter is the characteristic of the population. Sampling analysis involves estimating the parameter from the statistic.

(v) **Sampling error:** This is the inaccuracy in the information collected because only a small portion of the population is included in the study. The sampling error is also known as error variance. Sampling errors arise out of sampling and are generally random variations in the sample estimates around the true population values.

Sampling error = Frame error + Chance error + Response error

Sampling errors decrease as the homogeny of the universe increases. Sampling error is usually worked out as the product of the article value at a certain level of significance and the standard error.

Non-sampling errors also occur during the process of the data collection. But these cannot be measured. Since the sample and not the universe is studied, work proceeds fast. This is a large advantage for research.

(vi) **Sample distribution:** For example, say, from a population of 30,000, a random of 300 persons is chosen for a given study. The observed data are arranged in a frequency distribution, e.g., fertility rate. This type of distribution is called a sample distribution. This is the data got from a sample of the population.

(vii) **Population distribution:** If the fertility rates of all the 30,000 people of the population are obtained and arranged in a frequency distribution, it is known as the population distribution. Since the forms and parameters are not ordinarily known, an estimate of these two characteristics of the population is made from the sample distribution. So, if the sample distribution is normal one can assume that the population distribution is also normal.

(viii) **Sampling distributions:** Important sampling distributions that are commonly used are:

a. Sampling distribution of mean
b. Sampling distribution of population
c. Student's *t* distribution
d. F distribution
e. Chi-square distribution

a. *Sampling distribution of mean:* This refers to the probability distribution of all possible means of random samples of a given size taken from the population. If samples are taken from a normal population, then the sampling distribution of mean would also be normal. But when the sample is not from a normal population (a skewed one negatively or positively), even then as per the central limit theorem, the sampling distribution of mean tends to be closer to the normal distribution, provided the sample size is large—more than 30.

b. *Sampling distribution of proportion:* One can work out the proportion of items from a population, e.g., boys under the age of 20 in a town. This would give the sampling distribution of proportion of boys under 20 to the whole population. As the size increases; the distribution tends to become normal.

c. Student's t distribution when the population standard deviation is not known and the sample size is smaller than 30, the t distribution is used.

$$t = \frac{\bar{X} - \mu}{(\sigma_s \sqrt{n})}$$

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

Here, 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 one uses sample standard deviation (σ_s) while calculating t whereas standard deviation of population (σ_p) is used for calculating z. This gives a different t distribution for each possible sample size, that is, for different d.f. The d.f. for a sample of size n is $n - 1$. As the sample size increases, the shape of the t distribution becomes approximately equal to the normal distribution. In fact, if the sizes are more than 30, the t distribution approaches the normal distribution so close that we can use the normal distribution to approximate the t distribution. But when n is small, the t distribution would be far from normal distribution, while when $n = \alpha$, t distribution is identical with normal distribution. The t distribution tables show the critical values of t for different d.f. at various levels of significance. The table value of t for given d.f. at a particular level of significance is compared with the values of t obtained from the sample data, and if the latter is either equal to or exceeds the former, one can infer that the null hypothesis cannot be accepted.

d. *F distribution*: If $(\alpha_{s1})^2$ and $(\alpha_{s2})^2$ are the variances of two independent samples of sizes n_1 and n_2 respectively taken from two independent normal populations, having the same variance, $(\sigma_{p1})^2 = (\sigma_{p2})^2$, the ratio $F = (\sigma_{s1})^2\ (\sigma_{p2})^2$, where $(\sigma_{s1})^2 = \Sigma(\bar{X}_{1i} - \bar{X}_1)^2/n_1 - 1$ and $(\sigma_{s2})^2 = \Sigma(\bar{X}_{2i} - \bar{X}_2)^2/n_2 - 1$ has an F distribution with $n_1 - 1$ and $n_2 - 1$ degree of freedom. F ratio is computed in such a way that the larger variance is always in the numerator. Tables for F distribution that are available give critical values of F for various values of degree of freedom for larger as well as smaller variances. The value of F obtained 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 the inference drawn is that the null hypothesis of the variances being equal cannot be accepted. The F ratio can be used in the context of hypothesis testing and also in the context of ANOVA technique.

NOTES

e. *Chi-square (χ^2) distribution:* If each collection of sample variances is divided by the known population variance and these quotients are multiplied by $(n-1)$ (n stands for the number of elements in the sample), the chi-square distribution is obtained. Thus, $(\sigma_s^2/\sigma_p)^2\,(n-1)$ would have the same distribution as chi-square distribution with d.f. as $(n-1)$. Chi-square distribution is not symmetrical with all the values being positive. For using chi-square distribution, knowledge of the degrees of freedom is necessary.

Distribution is also used for judging the significance of difference between the observed and the expected frequencies and also as a test of the correctness of the fit. The generalized shape of χ^2 distribution depends upon the d.f. and the χ^2 value is worked out in the following manner.

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

There are tables available that give the value of χ^2 for given d.f. and these can be used for calculating values of χ^2 for relevant d.f. at a desired level of significance for testing hypotheses.

Central limit theorem

When a sample drawn from a population which is not normal, the sample size plays a critical role. If the population is normal, the means of samples drawn from that population would themselves be normally distributed, but when sampling is not from a normal population, the sample size plays a significant role. When n is small, the shape of the distribution to a great extent depends on the shape of the parent population, but as n becomes large ($n > 30$), the shape of the sampling distribution will move closer to normal distribution, irrespective of the shape of the parent population. Central Limit Theorem in deed explains this relationship between the shape of the population distribution and the sampling distribution of the mean. Central limit theorem has a great significance in statistical inference, as it expalins how the sampling distribution of the means approaches normal distribution with the increasing sample size. In technical terms, the distribution of means of random samples from a population with mean μ and finite variance σ^2 approaches the normal distribution with mean μ and variance σ^2/n as n tends to infinity.

The central limit theorem's significance is that it permits the use of sample statistics to make inferences about population parameters without knowing anything about the shape of the frequency distribution of that population other than what is available from the sample.

Sampling theory

Sampling theory deals with the relationships between a population and random samples drawn from the population. Population or a universe is an aggregate of items with common traits. A universe constitutes the totalilty of the items about which researcher seeks study. The universe may be finite or infinite. Finite universe contains a definite number or items. In an infinite unverse the number of items is infinite.

NOTES

The universe may be hypothetical or real. In the hypothetical case, the universe does not exist and one only imagines the items constituting it. Tossing a coin and throwing the dice can be cited as examples of a hypothetical universe. The real universe consists of concrete objects. Sample is that part of the universe which is selected at random for the purpose of investigation.

Sampling theory mainly deals with the relationship between a parameter and a statistic. The theory estimates the properties of the population on the basis of the sample and also evaluates the precision of the estimate. This is known as statistical induction or statistical inference as it attempts to draw the inference concerning the universe from the sample. To use this inductive method, first follow a deductive argument—imagine a universe (finite or infinite) and investigate the behaviour of the samples drawn from this universe applying the laws of probability. Such methodology is known as sampling theory. The theory's objectives are listed below:

(i) *Statistical estimation:* This consists of estimating the unknown population parameters from a knowledge of statistical measures based on sample studies. The estimate can either be a point estimate or an interval estimate. Point estimate is a single estimate expressed in the form of a single figure, while interval estimate has two limits, viz., the upper limit and the lower limit within which the parameter value may lie. Interval estimates are the ones which are often used in statistical induction.

(ii) *Statistical inference:* The sampling theory helps in arriving at generalizations about the population / universe from the studies based on samples drawn from it. It also helps in testing the accuracy of such generalizations.

The theory of sampling can also be studied under two heads: (i) The sampling of attributes and (ii) The sampling of variables in the context of large and small samples. A small sample is one that has 30 items or less whereas a large sample has more than 30 items. When studying the qualitative characteristics of all items in a population, one obtains statistics of attributes in the form of two classes—one in which the attribute is present and the second where it is absent. The presence of an attribute may be termed as a 'success' and its absence a 'failure'.

The theory can also be applied in the context of statistics of variables (i.e., data relating to some characteristic concerning population which can be estimated). The objectives are:

NOTES

- To compare the observed and expected values of the sample and to determine if the difference can be ascribed to the fluctuations of sampling;
- To estimate the population parameters from the sample, and
- To find out the degree of reliability of the estimate.

The tests of significance used in dealing with problems arising in studying large samples are different from those used for small samples. This is because the assumptions that one has to make in the case of large samples do not hold good for small samples. It is assumed in case of large samples that the sampling distribution tends to be normal and the sample values are approximately close to the population values. This helps in applying what is known as the z-test. When n is large, the probability of a sample value of the statistic deviating from the parameter by more than three times its standard error is very small (it is 0.0027 as per table giving area under normal curve). The z-test, thus is applied to find out the degree of reliability of a statistic in case of large samples. One, of course, needs to work out appropriate standard errors as they will enable one to give the limits within which the parameter values would lie or would enable one to judge whether the difference happens to be significant or not. For example, $\bar{X} \pm 3\sigma_{\bar{x}}$ would give the range within which the parameter mean value is expected to vary with 99.73 per cent confidence level.

The sampling theory that is applied for large samples is not applicable in the case of small samples because in the case of samples, one cannot assume that the sampling distribution is approximately normal. A different technique is required for handling small samples in particular when the population parameters are unknown. Sir William S. Gosset developed a significance test, known as student's t test, based on t distribution. His was a significant contribution to the theory of sampling, applicable in case of small samples. Student's t test is used when two conditions are fulfilled: the sample size is 30 or less and the population variance is not known. While using t test, one assumes that in the population from which the sample has been drawn:

- The sample is randomly drawn
- Observations are independent
- There is no measurement error
- And that in the case of two samples where equality of the two population means is to be tested, one assumes that the population variances are equal.

To apply the t test, one needs to work out the value of test statistic (i.e., t) and then compare it with the table value of t (based on t distribution) at certain level of significance for given degrees of freedom. If the calculated value of t is either equal to or exceeds the table value, the inference is that the difference is significant. But if the calculated value of t is less than the concerning table value of

t, the difference is not treated as significant. The following formulae are commonly used to calculate the t value:

NOTES

(i) To test the significance of the mean of a random sample:

$$t = \frac{(\bar{X} - \mu)}{\sigma_{\bar{x}}}$$

where $\bar{X}$ = Mean of the sample

μ = Mean of the universe / population

σ = Standard error of mean worked out as under

$$\sigma_{\bar{x}} = \frac{\sigma_s}{\sqrt{n}} = \sqrt{\frac{\Sigma(X_i - \bar{X})^2}{n-1}} \Big/ \sqrt{n}$$

and the d.f. = $(n-1)$.

(ii) To test the difference between the means of two samples

$$t = \frac{\bar{X}_1 - \bar{X}_2}{\sigma_{\bar{x}_1 - \bar{x}_2}}$$

where $\bar{X}_1$ = Mean of sample one

$\bar{X}_2$ = Mean of sample two

$\sigma_{\bar{X}_1 - \bar{X}_2}$ = Standard error of difference between two

Sample means worked out as

$$\sigma_{\bar{x}_1 - \bar{x}_2} = \sqrt{\frac{\Sigma(X_{1i} - \bar{X}_1)^2 + \Sigma(X_{2i} - \bar{X}_2)^2}{n_1 + n_2 - 2}} \times \sqrt{\frac{1}{n_1} + \frac{1}{n_2}}$$

And the d.f. = $(n_1 + n_2 - 2)$.

(iii) To test the significance of the coefficient of simple correlation

$$t = \frac{r}{\sqrt{1-r^2}} \times \sqrt{n-2} \text{ or } t = r\sqrt{\frac{n-2}{1-r^2}}$$

where

r = the coefficient of simple correlation and the d.f. = $(n-2)$.

(iv) To test the significance of the coefficient of partial correlation

$$t = \frac{r_p}{\sqrt{1-r_p^2}} \times \sqrt{n-k} \text{ or } t = r_p\sqrt{\frac{(n-k)}{1-r_p^2}}$$

where r_p is any partial coefficient of correlation and the d.f. = $(n-k)$, n being the number of pairs of observations and k being the number of variables involved.

NOTES

(v) To test the difference in case of paired or correlated samples data (in which case t test is often described as difference test).

$$t = \frac{\bar{D}-\mu_D}{\sigma_D}\sqrt{n} \text{ i.e., } t - \frac{\bar{D}-0}{\sigma_D}\sqrt{n}$$

where hypothesised mean difference (μ_D) is taken as zero (0),

$\bar{D}$ = Mean of the differences of correlated sample items.

σ_D = Standard deviation of differences worked out as under.

$$\sigma_D = \sqrt{\frac{\Sigma D_i^2 - D_i \cdot n}{n-1}}$$

D_i = Differences { i.e., $D_i = (X_i - Y_i)$}

n = number of pairs in two samples and the d.f. = $(n-1)$

Sandler's A test

Joseph Sandler has developed an alternate approach based on a simplified version of t-test. Sandler's A test that serves the same purpose as is accomplished by t-test relating to paired data. Researchers can also use A-test when correlated samples are employed and hypothesized mean difference is taken as zero, i.e., H_0: $\mu_D = 0$. Psychologists generally use this test in case of two groups that are matched with respect to some extraneous variable(s). While using A-test, one has to work out A-statistic that yields exactly the same results as student's t-test. A-statistic is found as follows.

$$A = \frac{\text{the sum of squares of the differences}}{\text{the square of the sum of the differences}} = \frac{\Sigma D_i^2}{(\Sigma D_i)^2}$$

In a A test, the d.f. is the same as in Student's t test, i.e., d.f. = $n-1$, n being equal to the number of pairs. The critical value of A, at a given level of significance for given d.f., is possible to obtain from the table of A statistic.One needs to compare the computed value of A with its corresponding table value for drawing an inference as to the acceptance or rejection of null hypothesis. If the calculated value of A is equal to or less than the table value the A-tatistic is considered significant and thereupon one can reject H_0 and accept H_a. But if the calculated values of A is more than its table value, then A-statistic is taken as insignificant and accordingly H_0 is accepted. This is because the two test statistics, viz., t and A are inversely related. These two statistics can be written in terms of one another in the following manner:

(i) A in term of t can be expressed as

$$A = \frac{n-1}{n \cdot t^2} + \frac{1}{n}$$

(ii) t in terms of A can be expressed as

$$t = \sqrt{\frac{n-1}{A \cdot n-1}}$$

Thus, computing A-statistic is relatively simple. Using of A-statistic helps in saving time and labour considerably, especially when the matched groups are to be compared with respect to a large number of variables. Researchers would be well advised to replace student's t test by Sandler's A test whenever correlated sets of scores are employed. Sandler's A-statistic can also be used 'in the one sample case as a direct substitute for the student t-ratio.' because Sandler's A is an algebraically equivalent to the Student's t. when one uses A-test in one sample case, the following steps are involved:

(i) Subtract the hypothesised mean of the population (m_H) from each individual score (X_i) to obtain D_i and then work out ΣD_i

(ii) Square each D_i and then obtain the sum of such squares i.e., ΣD_i^2

(iii) Find A statistic as under:

$A = \Sigma D_i^2/(\Sigma D_i)^2$

(iv) Look up 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.

(v) Finally, the inference can be drawn as indicated below:
When calculated value of A is equal to or less than the table value, rject H_0 (or accept H_a) but when computed A is greater than its table value, then accept H_0.

Concept of standard error

The standard deviation of the sampling distribution of a statistic is known as its standard error (S.E.) and is considered the key in sampling theory. The utility of the concept of standard error in statistical induction lay in:

(i) The standard error helps in testing whether the difference between the observed and the expected frequencies would arise due to chance. The usual criterion followed is to find out if a difference is less than three times the S.E. The difference is supposed to arise as a matter of chance. If the difference is equal to or more then three times the S.E., the chance fails to account for it, and the conclusion drawn is that it is a significant difference. The criterion is based on the fact that at ± 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. In other

NOTES

words, 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 one can conclude with 95 per cent confidence that the said difference is not due to fluctuations of sampling. In such a case, the hypothesis that there is no difference is rejected at 5 per cent level of significance.

But if the difference is less than 1.98 times the S.E., then it is considered not significant at 5 per cent level. It can then be said with 95 per cent confidence that it is because of the fluctuations of sampling. In such a case, the 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 S.E. is described as 'sampling error' at that particular level of significance. One can test the difference at certain other levels of significance as well depending upon one's requirement.

Criteria for Judging Significance at Various Important Levels

Significance level	**Confidence level**	**Critical value**	**Sampling error**	**Confidence limits**	**Difference Significant if**	**Difference Insignificant if**
5.0%	95.0%	1.96	1.96σ	$\pm 1.96\sigma$	$\bar{>} 1.96\sigma$	$\bar{<} 1.96\sigma$
1.0%	99.0%	2.5758	2.5758σ	$\pm 2.5758\sigma$	$\bar{>} 2.5758\sigma$	$\bar{<} 2.5758\sigma$
2.7%	99.73%	3	3σ	$\pm 3\sigma$	$\bar{>} 3\sigma$	$\bar{<} 3\sigma$
4.55%	95.45%	2	2σ	$\pm 2\sigma$	$\bar{>} 2\sigma$	$\bar{<} 2\sigma$

(ii) The S.E.also provides a good measure of reliability and precision of a sample. The smaller the S.E., the greater the uniformity of sampling distribution and, therefore, greater is the reliability of the sample. Conversely, the greater the S.E., the greater is the difference between observed and expected frequencies. In such a case, the unreliability of the sample is higher. The size of S.E., depends upon the sample size to a large extent and it varies inversely with the size of the sample. If reliability factor is to be doubled, i.e., reducing S.E. to 1/2 of its existing magnitude, the sample size should be increased four-fold.

(iii) The S.E. lso enables one to specify the limits within which the parameters of the population with a specified degree of confidence. Such an interval is known as confidence interval. The following table gives the percentage of samples having their mean values within a range of population mean (μ) $\pm$ S.E.

Range	**Per cent Values**
$\mu \pm 1$ S.E.	68.27%
$\mu \pm 2$ S.E.	95.45%
$\mu \pm 3$ S.E.	99.73%
$\mu \pm 1.96$ S.E.	95.00%
$\mu \pm 2.5758$ S.E.	99.00%

NOTES

Important formulae for computing the standard errors concerning various measures based on samples are as under:

a. In case of sampling of attributes

i. Standard error of number of successes = $\sqrt{n \cdot p \cdot q}$

where

n = number of events inn each sample,
p = probability of success in each event,
q = probability of failure in each event,

ii. Standard error of proportion of successes = $\sqrt{\frac{p \cdot q}{n}}$

iii. Standard error of the difference between proportions of two samples:

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

$$p = \sqrt{p \cdot q\left(\frac{1}{n_1} + \frac{1}{n_2}\right)}$$

$q = 1 - q$
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 and where we cannot have the best estimate of proportion in the universe on the basis of the given sample data. Such a situation often arises in the study of association of attributes.

b. In case of sampling of variables (large samples):

i. 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 even when n is 30 or less.

NOTES

ii. Standard error of mean population standard deviation is unknown:

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

where

σ_2 = standard deviation of the sample and is worked out as under

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

n = number of items in the sample

iii. Standard error of standard deviation when population standard deviation is known:

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

iv. Standard error of standard deviation when population standard deviation is unknown:

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

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

where

n = number of items in the sample.

v Standard error of the coefficient of simple correlation:

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

where

r = coefficient of simple correlation

n = number of items in the sample

vi. Standard error of difference between means of two samples:

c. When two samples are drawn from the same population:

$$\sigma_{\bar{X}_1 - \bar{X}_2} = \sqrt{\sigma_p^2\left(\frac{1}{n_1} + \frac{1}{n_2}\right)}$$

(If σ_p is not known, sample standard deviation for combined samples (σ_{s12}) may be substituted).

NOTES

d. When two samples are drawn from different populations:

$$\sigma_{\bar{X}_1-\bar{X}_2} = \sqrt{\frac{(\sigma_{p_1})^2}{n_1}+\frac{(\sigma_{p_2})^2}{n_2}}$$

(If σ_{p1} and σ_{p2} are not known, then in their places σ_{s1} and σ_{p2} respectively may be substituted.)

e. in case of sampling of variables (small samples):

i. Standard error of mean when σ_p Is unknown:

$$\sigma_{s_{1,2}} = \sqrt{\frac{n_1(\sigma_{s_1})^2+n_2(\sigma_{s_2})^2+n_1(\bar{X}_1-\bar{X}_{1,2})^2+n_2(\bar{X}_2-\bar{X}_{1,2})^2}{n_1+n_2}}$$

where $\bar{X}_{1,2} = \dfrac{n_1(\bar{X}_1)+n_2(\bar{X}_2)}{n_1+n_2}$

Note: (1) All the above formulae apply only for infinite population. In cse the population is finte, sampling is done without replacement and the sample size is more than 5 per cent of the population, the finite population multiplier must be used in the S.E formulae. For example, SE_x for finite population will be as shown below:

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

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. In such a case where the sampling fraction is less than 0.05, the finite population multiplier is not generally used.

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

ii. Standard error of difference between two sample means when σ_p is unknown

$$\sigma_{\bar{X}_1-\bar{X}_2} = \sqrt{\frac{\Sigma(X_{1i}-\bar{X}_1)^2+\Sigma(X_{2i}-\bar{X}_2)^2}{n_1+n_2-2}}\cdot\sqrt{\frac{1}{n_1}+\frac{1}{n_2}}$$

Estimating population parameters

In most statistical research studies, population parameters are usually not known and therefore, have to be estimated from a sample. Here, methods for estimating the population parameters have an important role in statistical analysis.

The random variables (such as and σ_s^2) used to estimate population:

Conventionally, parameters such as μ and σ_F^2 are called as 'estimators', while specific values of these (such as = 105 or $\sigma_s^2 = 21.44$) are referred to as 'estimates' of the population parameters. The estimate of a population parameter may be one single value or a range of values. In case of single value parameter, it

NOTES

is referred to as 'point estimate,' and in case where there is a range of parameters, it is termed as interval estimate. The researcher has to make these two types of estimates through sampling analysis. While making estimates of population parameters, the researcher can give only the best point estimate as otherwise he/she has to speak in terms of intervals and probabilities; For, he/she can never estimate with certainty the exact values of population parameters.

A good estimator possesses the following properties:

(i) An estimator should on the average be equal to the value of the parameter being estimated. This is known as the property of unbiasedness. An estimator is said to be unbiased if the expected value of the estimator is equal to the parameter being estimated. The sample mean ($\bar{X}$) is the most widely used estimator because it provides an unbiased estimate of the population mean (μ).

(ii) An estimator should have a relatively small variance.This property is technically described as the property of efficiency because an estimator is expected to have the smallest variance.

(iii) An estimator is expected to use as much informatrion as possible as is available from the sample. This property is known as the property of sufficiency.

(iv) As the sample size becomes larger and larger, an estimator has to approach the value of population parameter. This property is referred to as the property of consistency.

The researcher must select the appropriate estimator(s) for his/her study by keeping in view the above stated properties.

Estimating the population mean (μ)

In point estimate, the sample mean $\bar{X}$ is the best estimator of the population mean, m and its sampling distribution when the sample is sufficiently large, because this tends to approximate the normal distribution. If one knows the sampling distribution of $\bar{X}$, one can make statements about any estimate that one may make from the sampling information. In a sample of 36 students, if one finds that the sample yields an arithmetic mean of 6.2 i.e., $\bar{X}$ = 6.2, replace these student names on the population list and draw another sample of 36 randomly and assume that one gets a mean of 7.5 this time. Similarly a third sample may yield a mean of 6.9; fourth a mean of 6.7; and so on. One can go on drawing such samples till one accumulates a large number of means of samples of 36. Each such sample mean is a separate point estimate of the population mean. When such means are presented in the form of a distribution, the distribution happens to be quite close to normal. This is a characteristic of a distribution of sample means (and also of other sample statistics). Even if the population is not normal, the sample means drawn from that population are dispersed around the parameter in a distribution that is generally close to normal; the mean of the distribution of sample means is equal to the population mean. This is true in case of large samples as expected from the central limit theorem. The relationship between a population distribution and a distribution of sample mean is critical for drawing inferences about parameters.

The relationship between the dispersion of a population distribution and that of the sample mean can be expressed in the following formula:

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

where σ = standard error of mean of a given sample size

σ_s = standard deviation of the population

n = size of the sample.

How to find σ_p when one has the sample data only for one's analysis? The answer is that one must use the best estimate of σ_p and the best estimate can be the standard deviation of the sample, σ_s. Thus, the standard error of mean can be worked out as under:

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

where $\sigma_s = \sqrt{\dfrac{\Sigma(X_i - \bar{X})^2}{n-1}}$

With the help of the above, one can arrive at interval estimates about the parameter in probabilistic terms (utilising the fundamental characteristics of the normal distribution). Suppose one takes a sample of 36 items and works out its mean () to be equal to 6.20 and its standard deviation (σ_s) to be equal to 3.8: Then the best point estimate of population mean (μ) is 6.20. The S.E of mean (σ) would be $3.8 / \sqrt{36} = 3.8/6 = 0.663$. And if one takes the interval estimate of μ to be $\bar{X} \pm 1.96\,(\sigma_{\bar{X}})$ or 6.20 ± 1.24 or from 4.96 to 7.44, it means that there is a 95 per cent chance that the population mean is within 4.96 to 7.44 interval. This means that if one were to take a complete census of all items in the population, the chances are 95 to 5 that one would find the population mean between 4.96 to 7.44. In case one wants to have an estimate that will hold for a much smaller range, then one must either accept a smaller degree of confidence in the results or take a sample large enough to provide this smaller interval with adequate confidence levels. Usually one thinks of increasing the sample size till one can secure the desired interval estimate and the degree of confidence.

Illustration 3.1: From a random sample of 36 New Delhi civil service personnel, mean age and the sample standard deviation were found to be 40 years and 4.5 years respectively. Construct a 95 per cent confidence interval for the mean age of civil servants in New Delhi.

NOTES

Solution: The given information can be written as under:

$$n = 36$$
$$x = 40 \text{ years}$$
$$\sigma_s = 4.5 \text{ years}$$

And the standard variate, *z*, for 95 per cent confidence is 1.96 (as per the normal curve area table).

Thus, 95 per cent confidence interval for the mean age of population is:

$$\bar{X} \pm z\frac{\sigma_s}{\sqrt{n}}$$

$$\text{or } 40 \pm 1.96\frac{4.5}{\sqrt{36}}$$

$$\text{or } 40 \pm (1.96)(0.75)$$

$$\text{or } 40 \pm 1.47 \text{ years}$$

Illustration 3.2: In a random selection of 64 of the 2400 intersections in a small city, the mean number of scooter accidents per year was 3.2 and the sample standard deviation was 0.8.

(1) Make an estimate of the standard deviation of the population from the sample standard deviation.

(2) Work out the standard error of mean for this finite population.

(3) If the desired confidence level is .9, what will be the upper and lower limits of the confidence interval for the mean number of accidents per intersection per year?

Solution: The given information can be written as under:

$$N = 2400 \text{ (This means that population is finite)}$$
$$n = 64$$
$$= 3.2$$
$$\sigma_s = 0.8$$

And the standard variate (*z*) for 90 per cent confidence is 1.645 (as per the normal curve area table)

Now we can answer the given questions thus:

(1) The best point estimate of the standard deviation of the population is the standard deviation of the sample itself.

Hence

$$\sigma_p = \sigma_s = 0.8$$

(2) Standard error of mean for the given finite population is as follows.

$$\sigma_{\bar{X}} = \frac{\sigma_s}{\sqrt{n}} \times \sqrt{\frac{N-n}{N-1}}$$

NOTES

$$= \frac{0.8}{\sqrt{64}} \times \sqrt{\frac{2400-64}{2400-1}}$$

$$= \frac{0.8}{\sqrt{64}} \times \sqrt{\frac{2336}{2399}} = (0.1)(0.97) = 0.097$$

(3) 90 per cent confidence interval for the mean number of acciedents per intersection per years is as follows:

$$\bar{X} \pm z\left\{\frac{\sigma_s}{\sqrt{n}} \times \sqrt{\frac{N-n}{N-1}}\right\}$$

$$= 3.2 \pm (1.645)(0.97)$$

$$= 3.2 \pm 0.16 \text{ accidents per intersection.}$$

When the sample size happens to be a large one or when the population standard deviation is known, one uses normal distribution fo determining confidence intervals for population mean as stated above. But how to handle estimation problem when population standard deviation is not known and the sample size is small (i.e., when $n \leq 30$)? In such situations, normal distribution is not the appropriate tool and one can use t distribution for the purpose. While using t distribution, one assumes that population is normal or approximately normal. There is a different t distribution for each of the possible degrees of freedom. When using t distribution for estimating a population mean, one works out the degrees of freedom as equal to $n - 1$, where n means the size of the sample and then can look for critical value of 't' in the t distribution table for appropriate degrees of freedom at a given level of significance. This can be illustrated by taking an example:

Illustration 3.3: The foreman of ABC mining company has estimated the average quantity of iron ore extracted to be 36.8 tonnes per shift and the sample standard deviation to be 2.8 tonnes per shift, based upon a random selection of 4 shifts. Construct a 90 per cent confidence interval around this estimate.

Solution: As the standard deviation of population is not known and the size of the sample is small, one can use t distribution for finding the required confidence interval about the population mean. The given information can be written as under:

= 36.8 tonnes per shift

σ_s = 2.8 tonnes per shift

$n = 4$

degrees of freedom = $n - 1 = 4 - 1 = 3$ and the critical value of t for 90 per cent confidence interval or at 10 per cent level of significance is 2.353 for 3 d.f. (as per the table of t distribution).

Thus, 90 per cent confidence interval for population mean is

$$X \pm t\frac{\sigma_s}{\sqrt{n}}$$

$$= 36.8 \pm 2.353 \frac{2.8}{\sqrt{4}}$$

$$= 36.8 \pm (2.535)(1.4)$$

NOTES

$$= 36.8 \pm 3.294 \text{ tonnes per shift.}$$

Estimating population proportion

As for the point estimate, the sample proportion (p) units that have a particular characteristic is the best estimator of the population proportion ($\hat{p}$) and its sampling distribution, so long as the sample is sufficiently large, approximates the normal distribution. Thus, if one takes a random sample of 50 items and finds that 10 per cent of these are defective, i.e., $p = 0.10$, one can use this sample proportion ($p = 0.10$)as best estimator of the population proportion ($\hat{p} = p = 0.10$). In case one wants to construct confidence interval to estimate a population proportion, one should use the binomial distribution with the mean of population (μ) = n.p, where n = number of trials, p = probability of a success in any of the trials and population standard deviation = $\sqrt{npq}$. As the sample size increases, the binomial distribution approaches normal distribution and this can be used for the purpose of estimating a population proportion. The mean of the sampling distribution of the proportion of successes (μ_p) is taken as equal to p and the standard deviation for the proportion of successes, also known as the standard error of proportion, is taken as equal to $\sqrt{pq/n}$. But when population proportion is not known, then one can estimate the population parameters by substituting the corresponding sample statistics p and q in the formula for the standard error of proportion to obtain the estimated standard error of the proportion (shown below):

$$\sigma_p = \sqrt{\frac{pq}{n}}$$

Using the above estimated standard error of proportion, one can work out the confidence interval for population proportion thus:

$$p \pm z \cdot \sqrt{\frac{pq}{n}}$$

where

p = sample proportion of successes;

$q = 1 - p$;

n = number of trials (size of the sample);

z = standard variate for given confidence level (as per normal curve area table).

This formula can be explained as in illustration 3.4.

Illustration 3.4: A market research survey in which 64 consumers were contacted states that 64 per cent of all consumers of a certain product were motivated by the product's advertising. Find the confidence limits for the proportion of consumers motivated by advertising in the population, given a confidence level equal to 0.95.

NOTES

Solution: The given information can be written as under:

$n = 64$

$p = 64$ per cent or 0.64

$q = 1 - p = 1 - 0.64 = 0.36$

and the standard variate (z) for 95 per cent confidences is 1.96 (as per the normal curve area table)

Thus, 95 per cent confidence interval for the proportion of consumers motivated by advertising in the population is:

$$p \pm z \cdot \sqrt{\frac{pq}{n}}$$

$$= 0.64 \pm 1.96\sqrt{\frac{(0.64)(0.36)}{64}}$$

$$= 0.64 \pm (1.96)(0.06)$$

$$= 0.64 \pm 0.1176$$

Thus, lower confidence limit is 52.24 per cent upper confidence limit is 75.76 per cent.

For the sake of convenience, one can summarise the formulae which give confidence intervals while estimating population mean (μ) and the population proportion $(\hat{p})$ s shown in the following table.

The 3.1 table summarizes important formulae concerning estimation:

Table 3.1 *Formulaic for Estimation*

	In case of infinite population	In case of finite population
Estimating population mean (μ) when we know σ_p	$\bar{X} \pm z \cdot \frac{\sigma_p}{\sqrt{n}}$	$\bar{X} \pm z \cdot \frac{\sigma_p}{\sqrt{n}} \times \sqrt{\frac{N-n}{N-1}}$
Estimating population mean (μ) when we do not know σ_p and use σ_s as the best estimate of σ_p and sample is large (i.e., $n > 30$)	$\bar{X} \pm z \cdot \frac{\sigma_s}{\sqrt{n}}$	$\bar{X} \pm z \cdot \frac{\sigma_s}{\sqrt{n}} \times \sqrt{\frac{N-n}{N-1}}$
Estimating population mean (μ) when we do not know σ_p and use σ_s as the best estimate of σ_p and sample is large (i.e., $n \lesssim 30$)	$\bar{X} \pm t \cdot \frac{\sigma_s}{\sqrt{n}}$	$\bar{X} \pm t \cdot \frac{\sigma_s}{\sqrt{n}} \times \sqrt{\frac{N-n}{N-1}}$
Estimating the population proportion $(\hat{p})$ when p is not known but the sample is large.	$p \pm z \cdot \sqrt{\frac{pq}{n}}$	$p \pm z \cdot \sqrt{\frac{pq}{n}} \times \sqrt{\frac{N-n}{N-1}}$

In case of finite population, the standard error has to be multiplied by the finite population multiplier, viz., $\sqrt{(N-n)/(N-1)}$.

NOTES

Sample size and its determination

The most intriguing question in sampling analysis is: What should be the size of the sample or how large or small should be *n*? If the sample size (*n*) is too small, it may not help achieve the objectives and if it is too large, one would have to incur huge cost and resource waste. As a general rule, the sample must be of an optimum size, i.e., neither be excessively large nor too small. Technically, the sample size should be large enough to give a confidence interval of desired width. This involves choosing the size of the sample through a logical process before sample is taken from te universe.

The size of the sample should be determined by a researcher, keeping in view the following points:

- **Nature of universe:** The universe is either homogenous or heterogenous in nature. If the items in the universe are homogenous, a small sample can serve the purpose. But if the items are heterogenous, a larger sample would be required, Technically, this is termed as the dispersion factor.
- **Number of classes proposed:** If many classes (groups and sub-groups) are to be formed, a large sample becomes necessary because a small sample might not be able to accommodate a reasonable number of items in each class-group.
- **Nature of study:** If items are to be studied intensively and continuously, the sample should be small. For a general survey, however, the size of the sample should be large. 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 certain to be much superior to a large but badly selected sample.
- **Standard of accuracy and acceptable confidence level:** If the standard of accuracy or the level of precision is to be of a high order, one would require relatively larger sample. For doubling the accuracy for a fixed significance level, a fouir-fold increase in the sample size is required.
- **Availability of finance:** In practice, size of the sample depends upon the amount of money available for the study purpose. This is an important factor to be kept in mind while determining the size of sample because large samples result in increasing the cost of sampling estimates.
- **Other factors:** Other considerations include the 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..

A researcher has two alternative approaches for determining the size of the sample:

NOTES

The first approach is to specify beforehand the precision of estimation that is desired and then to determine the sample size necessary to ensure it. The second approach uses Bayesian statistics to weigh the cost of additional information against the expected value from that additional information. The first approach is capable of yielding a mathematical solution, and therefore, is a frequently used technique for determining n. The limitation of this technique is that it does not analyse the cost of gathering information vis-à-vis the expected value of information gained.

The second approach is theoretically optimal, but is seldom used because of the difficulty involved in measuring the value of information. Hence, one needs to mainly concentrate on the first approach.

Determining sample size through the approach based on precision rate and confidence level

It must be recognized at the outset that that whenever a sample study is made, there is every possibility of some sampling error occuring. This is usually controlled by selecting a sample of adequate size. A researcher will have to specify the precision that he.she wants in respect of his/her estimates regarding the population parameters. For instance, a researcher may like to estimate the mean of the universe within ±3 of the true mean with 95 per cent confidence. In that case the desired precision will be ±3, i.e., if the sample mean is ₹ 100. the true value of the mean will be no less than ₹ 97 and no more than ₹ 103. In other words, all this means that the acceptable error, e, is equal to 3. Keeping this in view, one can explain the determination of sample size so that specified precision is ensured.

(a) Sample size when estimating a mean: The confidence interval for the universe mean, μ, is given by

$$\mu = \bar{X} \pm z \frac{\sigma_p}{\sqrt{n}}$$

where

$\bar{X}$ = sample mean;

z = the value of the standard variate at a given confidence level (to be read from the table giving the areas under normal curve as shown in appendix) and it is 1.96 for a 95 per cent confidence level.

n = size of the sample

σ_p = Standard deviation of the population (to be estimated from past experience or on the basis of a trial sample). Suppose, one has $\sigma_p = 4.8$ for the purpose.

If the difference between μ and $\bar{X}$ or the acceptable error is to be kept with in ±3 of the sample mean with 95 per cent confidence, then one can express the acceptable error, e as equal to

NOTES

$$e = z \cdot \frac{\sigma_p}{\sqrt{n}}$$

$$3 = 1.96\frac{4.8}{\sqrt{n}}$$

$$n = \frac{(1.96)^2(4.8)^2}{(3)^2} = 9.834 = 10.$$

In general, if one wants to estimate μ in a population with standard deviation σ_p with an error no greater then e by calculating a confidence interval with confidence corresponding to z, the necessary sample size, n, equals as under:

$$n = \frac{z^2\sigma^2}{e^2}$$

This is true if the population is infinite, but if the population is finite, the above stated formula for determining sample size will become.

$$n = \frac{z^2 \cdot N \cdot \sigma_p^2}{(N-1)e^2 + z^2\sigma_p^2}$$

In case of finite population the confidence interval for μ is given by

$$X \pm z\frac{\sigma_p}{\sqrt{n}} \times \sqrt{\frac{(N-n)}{(N-1)}}$$

where $\sqrt{(N-n)/(N-1)}$ is the finite population multiplier and all other terms mean the same thing as stated above. If the precision is taken as equal to 'e', then one has

$$e = z\frac{\sigma_p}{\sqrt{n}} \times \sqrt{\frac{N-n}{N-1}}$$

$$e^2 = z^2\frac{\sigma_p}{n} \times \frac{N-n}{N-1}$$

$$e^2(N-1) = \frac{z^2\sigma_p^2 N}{n} \times \frac{z^2\sigma_p^2 n}{n}$$

$$e^2(N-1) + z^2\sigma_p^2 = \frac{z^2\sigma_p^2 N}{n}$$

$$n = \frac{z^2 \cdot \sigma_p^2 \cdot N}{e^2(N-1) + z^2\sigma_p^2}$$

$$n = \frac{z^2 \cdot N \cdot \sigma_p^2}{(N-1)e^2 + z^2\sigma_p^2}$$

where

N = size of population

n = size of sample

e = acceptable error (the precision)

σ_p = standard deviation of population

z = standard variate at a given confidence level.

NOTES

This is how one obtains the above stated formula for determining n in the case of infinite population given the precision and confidence levels.

Illustration 3.5: Determine the size of the sample for estimate the true weight of the cereal containers for the universe with $N = 5000$ on the basis of the following information:

(1) The variance of weight = 4 ounces on the basis of past records.

(2) Estimate should be within 0.8 ounces of the true average weight with 99 per cent probability.

Will there be a change in the size of the sample if one assumes infinite population in the given case? If so, explain by how much?

Solution: In the given problem we have the following:

N = 5000;

σ_p = 2 ounces (Since the variance of weight = 4 ounces);

e = 0.8 ounces (Since the estimate should be within 0.8 ounces of the true average weight);

z = 2.57 (as per the table of area under normal curve for the given confidence level of 99 per cent)

Hence, the confidence interval for the universe mean, m, is given by

$$\bar{X} \pm z \cdot \frac{\sigma_p}{\sqrt{n}} \cdot \sqrt{\frac{N-n}{N-1}}$$

And accordingly the sample size can be worked out as under:

$$n = \frac{z^2 \cdot N \cdot \sigma_p^2}{(N-1)e^2 + z^2\sigma_p^2}$$

$$= \frac{(2.57)^2 \cdot (5000) \cdot (2)^2}{(5000-1)(0.8)^2 + (2.57)^2(2)^2}$$

$$= \frac{132098}{3199.36 + 26.4196} = \frac{132098}{3225.7796} = 40.95 = 41$$

NOTES

(a) **The sample size (or n) = 41 for the given precision and confidence level for a finite population:** If one takes the population to be infinite, the sample size will be calculated as under:

$$n = \frac{z^2\sigma_p^2}{e^2} = \frac{(2.57)^2(2)^2}{(0.8)^2} = \frac{26.4196}{0.64}$$

$$= 41.28 = 41$$

In the given case the sample size remains the same even if one assumes infinite population.

In the illustration given above, the standard deviation of the population was available, but in many cases the standard deviation of the population may not be available. Since one has not yet taken the sample and is still at the stage of deciding how large it should be, one cannot estimate the population standard deviation. In that case where one has an idea about the range (i.e., the difference between the highest and lowest values) of the population, one can use that to get a crude estimate of the standard deviation of the population for getting a working idea of the required sample size. One way of getting the estimate of standard deviation is to follow these steps:

Since 99.7 per cent of the area under normal curve lies within the range of ±3 standard deviations, it can be considered that these limits include almost all of the distribution. Accordingly, one can assume that the given range equals 6 standard deviations because of ±3. Thus, a rough estimate of the population standard deviation would be:

$$6\hat{\sigma} = \text{the given range}$$

or $$\hat{\sigma} = \frac{\text{the given range}}{6}$$

If the range happens to be, say ₹ 12, then

$$\hat{\sigma} = \frac{12}{6} = ₹\,2.$$

And this estimate of standard deviation, σ, can be used to determine the sample size in the formulae given above.

(b) **Sample size when estimating a percentage or proportion:** One can find the sample size for estimating a proportion too. The reasoning put forth would be similar to the one stated in the context of estimating the mean. First of all, one has to specify the precision and the confidence level and then work out the sample size as under:

Since the confidence interval for universe proportion, is given by

$$p \pm z \cdot \sqrt{\frac{p \cdot q}{n}}$$

NOTES

where

p = sample proportion, $q = 1 - p$;

z = the value of the standard variate at a given confidence level and to be worked out from table showing area under Normal Curve;

n = size of sample;

Since p is what one is trying to estimate, what value should one assign to it? One method may be to take the value of $p = 0.5$ in which case n will be the maximum and the sample will yield at least the desired precision. This will be the most conservative sample size. The other method may be to take an initial estimate of p which may either be based on personal judgment or may be the result of a pilot study. In this case, it is suggested that a pilot study of something like 225 or more items may result in a reasonable approximation of p value. Then with the given precision rate, the acceptable error, e, can be expressed as under:

$$e = z \cdot \sqrt{\frac{pq}{n}}$$

or $$e^2 = z^2 \frac{pq}{n}$$

or $$n = \frac{z^2 \cdot p \cdot q}{e^2}$$

The formula gives the size of the sample in case of infinite population when one has to estimate the proportion in the universe. But in case of finite population the above formula will change to:

$$n = \frac{z^2 \cdot p \cdot q \cdot N}{e^2(N-1) + z^2 \cdot p.q}$$

Illustration 3.6: What should be the size of the sample if a simple random sample from a population of 4000 items is to be drawn to estimate the per cent defective within 2 per cent of the true value with 95.5 per cent probability? And what would be the size of the sample if the population is assumed to be infinite in the given case?

Solution: The answer to the questions will be as follows:

N = 4000;

e = 0.02 (since the estimate should be within 2 per cent of true value);

z = 2.005 (as per table of area under normal curve for the given confidence level of 95.5 per cent).

Since the investigator has not been given the p value being the proportion of defectives in the universe, he/she has to assume it to be $p = 0.02$ (this may be on the basis of one's experience or on the basis of past data or may be the result of a pilot study).

NOTES

The size of the sample can now be determined using all this information for the given question as follows:

$$n = \frac{z^2 \cdot p \cdot q \cdot N}{e^2(N-1) + z^2 \cdot p \cdot q}$$

$$= \frac{(2.005)^2(0.02)(1-0.02)(4000)}{(0.02)^2(4000-1) + (2.005)^2(0.02)(1-0.02)}$$

$$= \frac{315.1699}{1.5996 + 0.0788} = \frac{315.1699}{1.6784} = 187.78 = 188$$

And if the population is infinite, then the sample size will be as under:

$$n = \frac{z^2 \cdot p \cdot q}{e^2}$$

$$= \frac{(2.005)^2 \cdot (0.02)(1-0.02)}{(0.02)^2}$$

$$= \frac{0.0788}{0.0004} = 196.98 = 197$$

Illustration 3.7: Let us assume that a certain hotel management is interested in determining the percentage of the hotel's guests who stay for more than 3 days. The reservation manager wants to be 95 per cent confident that the percentage has been estimated to be within ±3 per cent of the value. What will be the conservative sample size needed to tackle this problem?

Solution: If the population is infinite:

e = 0.03 (since the estimate should be within 3 per cent of the true value);

z = 1.96 (as per table of area under normal curve for the given confidence level of 95 per cent)

As one wants the most conservative sample size one can take the value of $p = 0.5$ and $q = 0.5$. Using all this information, one can determine the sample size for the given problem as under:

$$n = \frac{z^2 pq}{e^2}$$

$$= \frac{(1.96)^2 \cdot (0.5)(1-0.5)}{(0.03)^2} = \frac{0.9604}{0.0009} = 1067.11 = 1067$$

Then, the most conservative sample size needed for the problem is = 1067.

Determination of sample size through the approach based on Bayesian Statistics:

NOTES

Another approach to determine n is to use the Bayesian statistics known as the Bayesian approach. The procedure for finding the optimal value of n or the size of sample under this approach is as under:

(i) Find the expected value of the sample information (EVSI) for every possible n;

(ii) Workout reasonably the approximate cost of taking a sample of every possible n;

(iii) Compare the EVSI and the cost of the sample for every possible n. In other words, work out the expected net gain (ENG) for every possible n as stated below:

For a given sample size (n): (EVSI) – (Cost of sample) = (ENG)

(iv) From (iii) above, the optimal sample size, that is the value of n which maximizes the difference between the EVSI and the cost of the sample, can be determined.

One disadvantage of this approach is that the computation of EVSI for every possible n and then comparing the same with the respective cost is often a very cumbersome task and is generally feasible with the help of computers only. Therefore, although theoretically optimal this approach is rarely used in practice.

CHECK YOUR PROGRESS

8. What is the concept of standard error?
9. What is the central limit theorem?
10. What does the Bayesian statistic suggest?

3.4 SUMMARY

- Universe or population in terms of statistics includes all items in the researcher's field of inquiry. A sample is the small number of elements selected to represent the entire population, according to some rule or plan. A measure based on a sample is known as a statistic.
- To measure the characteristics of the universe, a small sample that is assumed to nearly represent the entire population is studied. And then the results are projected over the entire population to get the approximate overall picture.
- Sampling methods are basically of two broad types—probability sampling and non-probability sampling. Again each broad type is divided into subtypes.
- It is necessary to learn different sampling errors and how to minimize those errors. There are two types of errors that can affect the inferences made in findings—systemic bias and sampling errors.

NOTES

- Systemic bias can occur due to inappropriate sampling, defective measuring device, non-responding, indeterminancy and natural bias in the reporting of the data. Sampling errors arise from inaccurate sampling and they generally happen to be random variations (when sampling is random) in the sample estimates, around the true population values.
- To get the right inferences it is necessary to keep the sampling and systemic errors at the minimum possible limits.
- Probability sampling methods include simple random sampling, stratified random sampling, cluster sampling, quota sampling, etc. Non-probability sampling methods include judgmental sampling, accidental sampling, systematic sampling, snow-ball sampling, etc.
- A researcher is basically interested in three types of distributions—sample distribution, population distribution and sampling distribution.
- To get the sample distribution the observed data are arranged in a frequency distribution, e.g., fertility rate. Then, for example, if the fertility rates of all the 30,000 people of the population are obtained and arranged in a frequency distribution, it is known as the population distribution. Since the forms and parameters are not ordinarily known, an estimate of these two characteristics of the population is made from the sample distribution. So, if the sample distribution is normal, one can assume that the population distribution is also normal.
- In statistics, a sampling distribution or finite-sample distribution (which is different from sample distribution) is the distribution of a given statistic based on a random sample of size n. It may be considered as the distribution of the statistic for all possible samples from the same population of a given size. The sampling distribution depends on the underlying distribution of the population, the statistic being considered and the sample size used. The sampling distribution is frequently opposed to the asymptotic distribution, which corresponds to the limit case $n \rightarrow \infty$.
- Sampling distributions: Important sampling distributions that are commonly used are: (i) sampling distribution of mean, (ii) sampling distribution of population, (iii) Students t distribution, F distribution and chi-square distribution.
- While conducting a research, the researcher should aim at getting the most accurate results possible, subject to the constraints of money, time, facilities and effort.

3.5 KEY TERMS

- **Systemic bias:** A bias of a measurement system or estimate method.
- ***F* distribution**: A probability distribution of the ratio of two variables.
- **Chi-square (χ^2)** A probability distribution used in inferential statistics.

- **Student's *t* test:** This test is used when two conditions are fulfilled: the sample size is 30 or less and the population variance is not known.
- **Sandler's *A* test:** A test that serves the same purpose as is accomplished by *t* test relating to a paired data.

3.6 ANSWERS TO 'CHECK YOUR PROGRESS'

1. When all items of a universe are enumerated, it is called a census inquiry.
2. A simple random sample is also known as an unrestricted random sample. Here every individual of a population has an equal chance of being included in the sample.
3. Probability sampling is preferred because it ensures that every element of the population has an equal chance of being included.
4. Random sampling means all the elements of a population have an equal chance of being included in the sample. The methods of drawing random sample are: (i) Fish bowl draw method; (ii) Using table of random numbers; and (iii) Computer determined randomness.
5. Quota sampling is a non-probability sampling method. Here different strata are identified and from each stratum the number of individuals needed is selected for the study.
6. In behavioural research people tend to give 'right' answers, as opposed to true 'answers.'
7. Cluster sampling is a form of probability sampling. Generally large geographical regions, territories, community, neighborhood, etc. are identified on a map by a number and then from these, a sample is drawn for the study.
8. Standard error helps in testing whether the difference between the observed and the expected frequencies arise due to chance.
9. When sampling is not from a normal population and the size of the *n*, the shape of the distribution will depend on the parent population. But as *n* gets larger, the sampling distribution would resemble a normal distribution, irrespective of the shape of the population distribution.
10. The Bayesian statistic suggests weighing the cost of additional information obtained against the expected value of the additional information. The second approach is difficult to determine, so the first is used often.

3.7 QUESTIONS AND EXERCISES

Short-Answer Questions

1. Write two points to be kept in mind while designing a sample.
2. Define bias and sampling errors.

3. What are probability sampling methods?
4. What is sampling distribution of mean?
5. What is sampling distribution of population?

NOTES

Long-Answer Questions

1. Explain the different types of sampling designs and their advantages and disadvantages.
2. Explain the three types of statistical distributions in detail.
3. Explain student's *t* distribution.
4. Write a note on *F* distribution.
5. Explain Chi-square distribution.

3.8 FURTHER READING

Cochran, W.G. 1963. *Sampling Techniques*. New York: John Wiley and Sons

Kerlinger, F.N. 1973. *Foundations of Behavioral Research*. New York: Holt Reinhart and Winston, Inc.

Kothari, C.R. 1995. *Research Methodology: Methods and Techniques*. New Delhi: Wishwa Prakashan.

Singh, A.K. 2008. *Tests, Measurements and Research Methods in Behavioural Sciences*. New Delhi: Bharati Bhawan.

Siegel, S. 1956. *Nonparametric Statistics for the Behavioral Sciences*. New York: McGraw-Hill Publishing Co., Inc.

UNIT 4 RESEARCH DESIGN

NOTES

Structure

4.0 INTRODUCTION

In this unit, you will learn the meaning and the purpose of research design. A research design is a plan and a systematic procedure for collecting the data and performing analysis on that data for the purpose of research. In other words, a research design is a conceptual framework for conducting research. It is a blueprint for collecting, measuring and analyzing data. Research designs tell us what, where, when and how any inquiry is to be made.

The unit explains the features and important concepts related to research design. It throws light on different types of research designs. It explains the basic principles and different types of experimental designs. It also discusses quasi-experimental design.

The later part of the unit teaches you the principles and different types of qualitative research. It will also make you familiar with different types of interviews, group discussions and observations. At the end of the unit, you will also learn data documentation and analysis.

NOTES

4.1 UNIT OBJECTIVES

After going through this unit, you will be able to:

- Get familiar with the concept and types of research designs
- Understand sampling design and its importance
- Learn the basic principles and different types of experimental design
- Know the concept and types of qualitative research
- Examine content analysis in qualitative research

4.2 RESEARCH DESIGN

A research design is a conceptual framework for conducting research. It is a blueprint for collecting, measuring and analyzing data. Research designs deal with the What, Where, When and How of an inquiry. Several questions in the following list have to be answered before starting a research work.

- What is the study and who is doing the research?
- Why is the study being conduced?
- Where will it be done?
- What type of data is to be collected?
- What is the timeframe for the study?
- What techniques will be used for data collection?
- Hon will it be analyzed?
- What kind of report will be prepared?
- What will be the costs involved?
- Who are the personnel for doing the study?

Keeping these in mind, the design can be sub-divided into specific areas for clarity:

- **The sampling design:** Methods of selecting the items for observation for the study
- **The observational design:** The conditions under which the observations are to be carried out.
- **The statistical design:** The data gathering and analysis methods.
- **The operational design features:** The entire procedure for carrying out the study.

Important features of a design

The important features of a design are that it is:

- A plan that identifies the resources and the type of information needed
- The strategy for gathering data
- An estimate of time, costs, etc.

So, the research design needs to have a clear research problem procedure for data collection, the population to be studied and the type of data analysis that has to be carried out.

Need for a research design

For smooth conduct of the study, for efficient data gathering and analysis and for economy and effectiveness, research designs must be planned well in advance and with great care. Good designs help to obtain reliable results.

Features of good design

- It should seek to minimize bias and maximize reliability of the data obtained
- It should give the least possible experimental error
- It should be as objective as possible

No single design should be used for all types of research problems. Reporting on the purpose of the study, the type of data needed and other considerations determine the design to be chosen.

Important concepts relating to research design

1. **Dependent and independent variables:** A variable is any unit that can have different quantitative values, e.g., height, loudness etc. Qualitative units are attributes, e.g., honest, extroverted, shy, etc. Largely quantitative variables are continuous, e.g., age is a continuous variable while 'students' is a non-continuous variable.
   (a) An antecedent variable is an independent variable.
   (b) A consequent variable is a dependent variable.
   (c) Height is a variable dependent on age where age is an independent variable
   (d) Height is sex related, so height is a dependent
   (e) Variable, age and sex are independent variables.
   (f) Behaviour changes as a function of the manipulation is an independent variable.
   (g) The independent variable (IV) is the one that is manipulated. It is under the control of the experimenter, generally. This variable is also called the experimental variable. The effect of the experimental variable is reflected on the dependent variable, e.g., knowledge of results improves learning. Here providing knowledge of results are the

NOTES

independent variables (DVs) and the improvement in learning is the dependent variable.

2. **Extraneous variables:** Variables unrelated to the study but having an influence on the dependent variable are called extraneous variables. Examination results studied to be a function of the methods of studying. Results are DVs and the methods of study are IVs. However, intelligence also plays a part in the results. This becomes an extraneous variable, and affects the outcome. Such an influence is known as the experimental error.
3. **Control:** This implies any attempt to minimize the influence of the extraneous factor(s) or variable(s). The attempt is to keep the experimental conditions, well controlled.
4. **Confounding or compounding influence**: When the external factors cannot be controlled, they are thought to compound the outcome(s).
5. **Research hypothesis:** When a hypothesised relationship is to be tested scientifically, it is called hypothesis testing. A hypothesis is a hunch. It should have one IV and one DV, as part of the design.
6. **Experimental and non-experimental hypotheses testing research**: When a hypothesis is to be tested, it is known as a hypothesis testing research. Here the experimental variable is manipulated. When the IV is not manipulated, it is called as non-experimental hypothesis testing research. For example, how age affects memory. Here people of different ages are tested on a memory task. The memory is determined by calculating a coefficient of correlation between the sets of obtained scores. This is a non-experimental hypothesis testing research.
7. **Experimental and control groups: The control group is exposed to the regular conditions.** The experimental group is exposed to the experimental variable. The experimental group receives the experimental variable. This enables comparison in terms of the outcome on the dependent variable.
8. **Treatments:** The conditions under which the two groups are studied is the treatment procedure.
9. **Experiment:** The procedure for testing a hypothesis or attempts to establish the veracity of a relationship is known as an experiment.
10. **Experimental unit(s):** The pre-determined block or conditions where different treatments are applied is called an experimental unit.

Different types research designs

Designs can be categorized as:

(a) Exploratory study design

(b) Descriptive studies design

(c) Diagnostic study design

(d) Hypothesis-testing study design

NOTES

Exploratory study design: The purpose here is to do a preliminary study to be able to formulate a research design later. Such a design needs to be flexible to be altered, depending on what the explorations yield. These designs are based on survey of the literature, survey of the experiences encountered, and analysis of insights or intuitions.

Research design for descriptive and diagnostic research: These designs must be tight and well planned.

Steps in this type of design:

- State objective(s) and design methods of data collection
- Select samples, suitably collect data, analyze and report results.
- Questionnaires, interviews, case-studies and observations are used extensively, here.

Table 4.1 *Summary of Research Designs*

Research Design	**Type of Study**	
	Exploratory or Formulative	Descriptive/Diagnostic
Overall Design	Flexible design (design must provide opportunity for considering different aspects of the problem)	Right design (design must make enough provision for protection against bias and must maximize reliability)
(i) Sampling design	Non-probability sampling design (purpose or judgment sampling)	Probability sampling design (random sampling)
(ii) Statistical design	No pre-planned design for analysis	Pre-planned design for analysis
(iii) Observational design	Unstructured instrument for collection of data	Structured or well thought out instruments for collection of data
(iv) Operational Design	No fixed decisions about the operational procedures	Advanced decisions about operational procedures

Source: Kothari, C.R., *Research Methodology Methods and Techniques*, 1995.

Research design in hypothesis testing: Here the researcher tests the causal relationship between variables. Studies have to help in reducing bias and increase reliability.

Basic principles of experimental design:

- The principle of replication
- The principle of randomization
- The principle of local control

The principle of replication: It means that any experiment should be repeatable. No single performance of a study is sufficient for results to be reliable. Also, the precision of a study is enhanced with replication.

The principle of randomization: This method is to protect against extraneous factors creeping into and contaminating the outcome(s). The role of chance factors is sought to be neutralized through randomization.

The principle of local control: The design is so structured that the extraneous factor influences are controlled. Here, the field of study is divided into homogeneous parts equal to the number of treatments. The treatments are then randomly assigned to these blocks. A block is a division of the field into homogeneous parts. The blocks are the levels at which the extraneous factor is fixed. This can lead to the elimination of the variability due to the extraneous factor(s) that cause the experimental error.

There are several research designs possible. The researcher has to choose a design keeping in mind the objective of the study, the type of data needed, the sampling methods, the desired level of accuracy and other constraints under which the study is done.

Developing a research plan

The research plan is the writhen format of the research that is to be undertaken. This is useful:

- For organizing thoughts, ideas and concepts, clearly and well
- For develop a checklist of all materials needed for research.
- As a preparatory document that is available for scrutiny.

The plan must consist of:

(i) The research objective
(ii) The problem to be investigated
(iii) The major variables and issues to be studied/measured
(iv) The methods / techniques to be used
(v) The procedures to be followed
(vi) The sample: from where to be drawn, the size, etc.
(vii) Statistical and other tools to be used
(viii) Budget and time constraints involved.

CHECK YOUR PROGRESS

1. What is a design?
2. What is a variable?
3. Name are the different research designs?
4. What is meant by randomization?

NOTES

4.3 EXPERIMENTAL DESIGN

4.3.1 Important Experimental Designs

Experimental designs are the structures of an experiment, which are of two types—informal and formal. Informal designs have less sophistication and less control while the formal designs offer more control and lend themselves to the use of precise statistical procedures for analysis.

The informal experimental designs:

(i) Before-and-after without control design
(ii) After-only, with control design
(iii) Before-and-after with control design

The formal experimental designs:

(i) Completely randomized design (C-R design)
(ii) Randomized Block design (R-B design)
(iii) Latin- square design (L-S design)
(iv) Factorial designs

4.3.1.1 The informal experimental designs

(i) **Before-and-after without control design:** Here a single test group or area is selected and the dependent variable is measured before the treatment or intervention is introduced. Then the independent variable or the intervention is introduced and again the dependent variable is measured. The effect of the treatment would be equal to the level of a given phenomenon after the treatment minus the level of the phenomenon before the treatment.

The main difficulty with such an experiment is that with the passage of time, some extraneous variations could enter into the treatment effect, e.g., fatigue, boredom, familiarity, etc.

(ii) **After-only with control design:** In this design, two conditions are involved (experimental condition and control condition). The treatment or intervention is introduced into the experimental condition only. The dependent variable is then measured in both the conditions at the same time. The treatment effect is obtained by subtracting the value of the dependent variable in the control condition from the value in the experimental condition.

NOTES

Experimental condition	Treatment introduced	Value of the dependent variable (*x*) after treatment
Control condition	Treatment withheld	Value of the dependent variable without treatment (*y*)
Treatment effect = (*x*) – (*y*)		

Source: Kothari, C.R., *Research Methodology Methods and Techniques*, 1995.

The basic assumption of this design is that the groups are identical with respect to their behaviour towards the phenomenon under consideration. If this is not true, there are chances of extraneous factors entering into the treatment Problems with respect to time lapses do not enter into this design. This design is superior to Before-and-After without control design.

(iii) before-and-after with control design: In this design, two groups are selected and the dependent variable is measured in both, for an identical length of time prior to introducing the independent variable or treatment. Then the independent variable or the treatment is introduced only in the experimental condition (as shown in the table below). After an identical time period lapses, the dependent variable is measured in both the conditions. The treatment effect is determined by subtracting the change in the dependent variable in the control condition from the change in the dependent variable, in the experimental condition.

Table 4.2 *Before-And-After with Control Design*

	Time Period I	**Time Period II**
Test area:	Level of phenomenon before treatment (*X*) introduced	Level of phenomenon after treatment (*Y*)
Control area:	Level of phenomenon Without treatment (*A*)	Level of phenomenon without treatment (*Z*)
	Treatment effect = (*Y* – *X*) – (Z – A)	

This design is superior to the earlier two designs in that it overcomes extraneous influences arising from time lapses and also from the non-comparability of the experimental and control conditions. However, when suitable control groups are not available, one of the other two earlier designs could be considered.

4.3.1.2 The formal experimental designs

(i) Completely randomized design (CR design): This design rests on the principle of randomization. The subjects are assigned randomly to the experimental conditions.

Two-group simple randomized design: In this design, the population is identified and then a sample is chosen randomly. Then the subjects are again randomly assigned to the experimental and control group conditions. This ensures the principle of randomization, so both the groups are representative of the population chosen for the study.

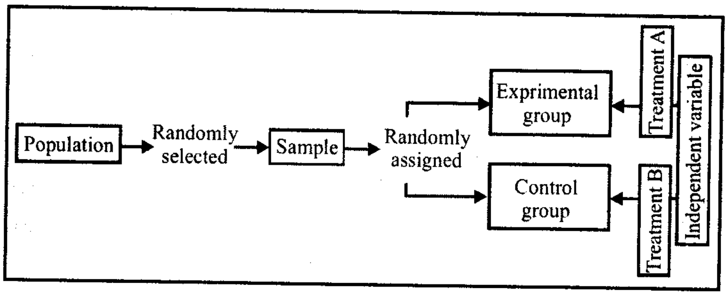


Fig. 4.1 Two-group Simple Randomized Experimental Design (in Diagram form)

Since the sample is drawn randomly from the population and then again randomly assigned to the two conditions (experimental and control) and then each of these conditions receive different treatments (x and y), the independent variables; the conclusions drawn from the samples are applicable to the population. Such a design has the merit of randomizing the sample. One could test each group before and after the treatment after ensuring the equivalence of the two groups.

Individual differences that exist in the two conditions and the experimenter influences (e.g., the teacher or trainer differences in applying the methods) can be further controlled by the random replication design. Such differences get minimized with such a design.

(ii) **Randomized block design (RB design):** This is an improvement over the completely randomized design. Here, the subjects are first divided into two groups known as blocks (married men, single men, high or low income, etc.). So, the blocks are homogeneous with respect to the chosen variable. The number of subjects in a given block would be equal to the number of treatments. One subject in each block would be randomly assigned to each treatment. The blocks are the levels at which the extraneous factor is fixed. This enables one to know its contribution to the total variation. The main feature of this design is that each treatment appears the same number of times in each block. The RB design is analysed by a two-way ANOVA (The ANalysis Of Variance)

(iii) **Latin-Square design (L-S design):** A Latin square is a multi-faceted experimental design. The Latin square can be used in many different situations. It is an experimental design that consists of allocation of treatments to experimental units. The set of experimental units and the set of treatments may be structured in some way.

Below is a Latin square of order 5. In a number of different experiments this could be the basic structure for the design. The following three examples will help the student understand the concept of Latin squares better.

NOTES

First example

Suppose we want to test five drugs A, B, C, D and E for their efficacy in alleviating the symptoms of a chronic disease. Five patients are available for a trial and each will be available for five weeks. Testing a single drug requires a week. Thus an experimental unit is a 'patient week.'

The structure of the experimental units is a rectangular grid (which happens to be square as in Fig. 4.2). There is no structure for the set of treatments. We can use the Latin square to allocate treatments. If the rows of the square represent patients and the columns are weeks, then for example, the second patient in the third week of the trial, will be given drug D. Now each patient receives all five drugs and thus in each week, all the five drugs are tested.

A	*B*	*C*	*D*	*E*
B	*A*	*D*	*E*	*C*
C	*E*	*A*	*B*	*D*
D	*C*	*E*	*A*	*B*
E	*D*	*B*	*C*	*A*

Fig. 4.2 *Experimental units*

Second example

The second example is very similar to the first. Suppose we are testing five varieties of pesticides on a square orchard with 25 apple trees. There may be differences between the rows, and differences between the columns, but we assume that rows and columns have the same status (e.g., the orchard is not on a hillside). Each plot occurs in one row and one column, just as each experimental unit in the preceding example occurs in one patient and one week. In a sense, the rows are 'equivalent' to the columns, but the patients are not equivalent to weeks in the same way as in the earlier example, because of the influence of the drug. The structure on the plots is a Hamming scheme in this case.

Again the Latin square gives an allocation of varieties to plots in the same way as it does in the earlier case.

Third example

Suppose we have to test five different cake recipes and five baking temperatures. We can perform experiments on five days, and five timeslots are available each day. The 25 experimental units are divided into five blocks of five, while the treatments have a 5×5 factorial structure.

Again one can assign treatments to units using the Latin square. Let us number the cells of the square from 1 to 25 as shown in Fig. 4.3.

1	2	3	4	5
6	7	8	9	10
11	12	13	14	15
16	17	18	19	20
21	22	23	24	25

Fig. 4.3 *Latin Square*

NOTES

We assume that the rows of the table represent days, so that units 1 to 5 are the timeslots on the first day. Now we use the five columns of the square to assign the five temperatures, and the letter to assign the recipes A, B, C, D, E.

(iv) Factorial design: This is used when to study the effects by varying more than one factor. This is especially useful in social science research where a number of factors do influence a particular phenomenon. Factorial designs are of two types: 1. Simple factorial designs, and 2. Complex factored designs.

A simple factorial design is one where the effects of varying two factors on the dependent variable are studied. If more than two factors are involved, the complex factorial design is used. A simple factorial design is also known as a two-factor factorial design while a complex factorial design is called a multifactor factorial design.

	Experimental variable	
Control variable	Treatment A	Treatment B
Level I	Cell 1	Cell 3
Level II	Cell 2	Cell 4

Fig 4.4: *2 × 2 Simple Factorial Design*

Here there are two treatments or experimental conditions and two control conditions. This sample is divided into four cells. Each of these has one treatment condition. Subjects are assigned at random. The means for different rows and columns can be obtained. The column means represent the main effects of treatments. The row means are the main effects for level without regard to treatment. So, the main effects of treatment as well as levels can be studied by this design. Further, the interaction between treatments and levels can also be studied. This enables one to see whether the treatments and levels are independent of each other or not.

Complex-factorial design: A complex-factorial design has more than two factors at a time. It is a design with three or more than three independent variables simultaneously. In the following design, there are (3 – 1) factors. Experimental variable has two treatments and two control variables each with two levels. Such a design is a 2×2×2 complex factorial design. There are a total of eight cells. The design is as given below:

Table 4.3 *2 × 2 × 2 Complex Factorial Design*

NOTES

		Experimental Variable			
		Treatment A		Treatment B	
		Control Variable 2 Level I	Control Variable 2 Level II	Control Variable 2 Level I	Control Variable 2 Level II
Control Variable 1	Level I	Cell 1	Cell 3	Cell 5	Cell 7
	Level II	Cell 2	Cell 4	Cell 6	Cell 8

Source: Kothari, C.R., *Research Methodology Methods and Techniques*, 1995.

From this design, it is possible to determine the main effects of these variables: one experimental and two control variables. The interaction between these pairs of variables can also be determined. Such an interaction is called a first order interaction. And when the interaction between variables, taken in triplets, it is called a second order interaction.

Experimental variable with control 1 (or EV×CV1);

Experimental variable with control variable 2 (or EV×CV2);

Control variable 1 with control variable 2 (or CV1×CV2):

There will be one second order interaction in this design. It involves all the three variables, i.e., EV×CV1×CV2).

To determine the main effect for the experimental variable, one needs to compare the combined mean data in cells 1, 2, 3 and 4. For treatment A and the combined mean data in cells 5, 6, 7 and 8 for treatment B. In this way, the main effect of the experimental variable is obtained without the influence of control variables 1 and 2. Similarly, the main effect of control variable 1, independent of the experimental variable, and control variable 2 can be obtained if one compares the combined mean data in cells 1,,3,,5 and 7 with the combined mean data in cells 2,,4,,6 and 8 in the 2×2×2 design. Similarly, one can determine the main effect for control variable 2 also independent of experimental variable and control variable 1, if the combined mean data of cells 1, 2, 5 and 6 are combined with mean data of cells 3, 4, 7 and 8. To obtain the first order interaction for EV×CV1 in the above research design, the researcher can ignore control variable 2 and generate a 2×2 design from the two 2×2×2 design by choosing to combined data from the relevant cells as shown below:

NOTES

Table 4.4 Expemental Variables

		Experimental Variables	
		Treatment A	Treatment B
Control Variable 1	Level I	Cells 1, 3	Cells 5, 7
	Level II	Cells 2, 4	Cells 6, 8

Source: Kothari, C.R., *Research Methodology Methods and Techniques*, 1995.

The same can be done for other first order interactions. The complex factorial design need not be a 2×2×2 design only. Any number of combinations of experimental and control variables, which are independent, can be used for generating data and analysis. The more the variables, the more complex the design

Factorial designs have two crucial advantages: 1. Equivalence accuracy: One can determine the main effects of two or more factors in one single experiment. It is economical in terms of effort, 2. This design also permits various comparisons of interest to the researcher, i.e., certain effects that cannot be as ascertained by treating one factor at a time. The determination of interaction effects is also possible with the use of this design. The researcher must decide in advance what the most appropriate design is for a given study.

4.3.1.3 Other important types of experimental designs

The design of an experiment is a blueprint of the procedures that are to be followed for testing a hypothesis. If the experimenter decides that there would be only one group of subjects who will be tested under the different values or conditions or treatment of the independent variable; the ensuing design would be known as the 'within groups design' or 'repeated treatment design'. If the decision is to use separate groups for each value of the independent variable; they are also know as between subjects or within subjects designs.

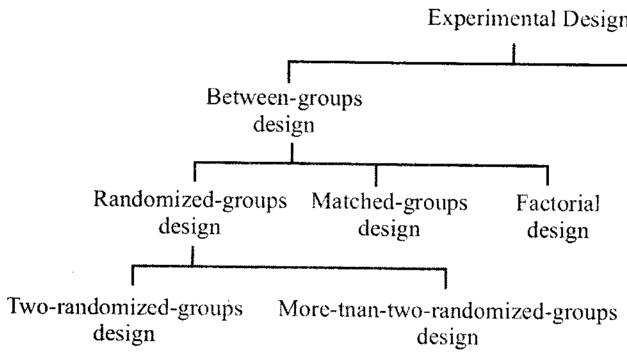


Fig 4.5: *Experimental Design*

Source: Singh, A.K., *Tests, Measurement and Research Methods in Behavioural Sciences*, 2008.

Based on the number of groups, the two designs are: 1. Between groups and 2. Within groups. In psychological and educational studies the 'between groups design'

NOTES

is used often. 'Between group designs' are divided into the following three types:

a. Randomized groups design
b. Matched-group design
c. Factored design

Between groups designs

Randomized groups design: It is one in which subjects are assigned randomly to different groups meant for the different conditions or values of the independent variable. The assumption here is that the random assignment makes these groups statistically equivalent. This means the variations in the dependent variable become easily identifiable. When the subjects are randomly assigned to only two groups; the design is called as 'two-randomized groups design. And when the subjects are assigned to more than two groups, it is called a multi-group design'.

Matched-groups design: It is another between groups design in which the subjects are matched on mean, standard deviation, pairs, etc.

Factorial design: In this design two or more independent variables are studied in various possible combinations. Here their independent and interactive effects on the dependent variable can also be studied.

There are two primary ways through which unbiased groups or random groups subjects can be formed: 1. Captive assignment, and 2. Sequential assignment.

Captive assignment: Here all the subjects are individually known and for the duration of the experiment they are made captive, so that they are randomly assigned to the different conditions or groups and so the term captive assignment. The random procedure is followed. The only pre-requisite is that the '*N*' of the groups, be equal.

Sequential assignment: The technique of sequential assignment is one where the experimenter does not know the subjects in advance. As the experiment progresses; the experimenter may use three objects on day one, six on day two and so on. So the experiment follows a pre-arranged schedule or sequence.

The sequential assignment of pre-arranged schedule or sequence or simply the sequential assignment of subjects can be done by complete randomization or block randomization. Any biases that could arise in the sequence can be overcome by block randomization. The concept of block randomization is that the independent variable occurs in each successive block of trials once, but the order of conditions within a block is random and different from every other condition. All conditions must occur equal number of times for the block randomization to be effective.

Social sciences researches have to establish which of the several independent variables influences the dependent variable most, and also the type of relationship that exists between the most influential of independent variables and the dependent variable. A two-group randomised design satisfies the first need, but not the second one. Here two independent variables are involved so it is difficult to ascertain the

NOTES

most significant independent variable on the dependent variable. It is important to study more than two values of the independent variable for this purpose. This is the 'more than two randomized group deign'.

Two-group randomized design: The subjects are assigned randomly to the two groups. The independent and dependent variables have to be defined clearly. The two values of the independent variable have to be identified. The values are conditions or treatments of the experiment. The objective of the study is to see whether the two conditions affect the dependent variable.

The population has to be specified and the sample drawn. The subjects are then randomly assigned to the two groups. (Method of random selection to be followed)

Here, two equal *N* groups are formed. A coin toss could decide which the experimental group is and which the control group is. The experimental group receives the experimental treatment, while the control group does not receive it. After the experiment, the scores obtained by the subjects in both the groups on the dependent variable are recorded. Usually the *t* test or the non-parametric Mann Whitney *U* test is used in a two-group randomized design. If the two groups differ, then the independent variable is seen to affect the dependent variable. So, the independent variable manipulation is seen to influence the resultant dependent variable.

Central to the two-group randomized design is the random assignment of subjects into the groups, so that no systematic relationship between the characteristics of the subjects and a particular group to which they are randomly assigned emerges. To achieve this, (Underwood, 1966) suggests two ways in which random groups can be formed.

i. Captive assignment
ii. Sequential assignment

More-than-two-randomized-groups design: This is also called a multi-groups design. Here, there are three or more values or conditions of the independent variable. So, three or more groups of subjects participate. All the subjects are randomly assigned to the three or more groups. The process of captive assignment or sequential assignment using the random procedure can be used here. For example, for hunger drive levels on responses, three groups of experimental variables with three conditions of deprivation—mild, moderate and extreme deprivation of food can be chosen. If the two-groups design was chosen and the conditions were mild and moderate deprivation and the number of correct responses (bar press) was used as the dependent variable, perhaps the two conditions mild and moderate would not yield a significant difference. So, the use of the third group (extreme deprivation) would have a clear effect on the dependent variable. This result would have been missed in the two-group experiment.

Multi-groups design helps in determining the relationships between the independent and the dependent variables that cannot be established by the two-group design.

NOTES

Matched groups design: This is also known as the randomized block design. It can have two or more groups. All the subjects are given a common pre-test on a task (matching variable) and then are formed into groups on the basis of their performance. This ensures equivalence of groups. Then the different values of the independent variable are introduced to each group. If these groups have had equivalent means on the dependent variable prior to the experimental treatment is given, then if the significant difference occurs, after the given, and then if significant difference occurs, after the introduction of the treatment, the resulting differences may be confidently be attributed to the experimental treatment.

This design assumes that the subjects in a group are homogeneous. They form a block. So it is also known as a randomized block design.

Two issues connected to the matching task:

(a) Selection of the matching variables

(b) Ways of matching

The subjects are pre-tested on a matching variable in this design. The matching variable is chosen on the basis of its ability to yield a high correlation with the experimental task, or the dependent variable. To obtain this, the dependent variable itself is used as the matching variable, for example, to study knowledge of results on memory, the subjects are all given a memory task and a set of scores are obtained—one with knowledge of results and the other without it. After the pre-test, they are paired off. The method for pairing into the two groups is through randomization. Then the experiment is carried out, as per plan. This study design can establish the relationship, suitably.

Methods of matching: (a) matching by pairs; and (b) Matching by mean and standard deviation.

(a) *Matching by pairs:* On the basis of scores obtained, the subjects are matched and paired. A subject who has a score of 80 is paired with another of a similar score on a test of memory. In this way, pairs are formed. Each pair can be seen to be a block. Another block would have pairs with scores of 90 and so on. Here subjects with deviant scores on the matching variable would get eliminated.

(b) *Matching by mean and standard deviation:* Here the measures of central tendency and variabilities in the distribution of scores on the matching variable are used.

Three matching methods are employed:

(i) Random blocks methods

(ii) Method of counter balancing order and

(iii) Block-repetition method.

(i) *Random blocks method:* In this method, the blocks are first created. Then with the number of subjects in each block being kept the same, as determined by the values of the independent variable, the subjects from each block are randomly assigned to the different groups as per conditions or values of the independent variable.

NOTES

(ii) *Method of counter balancing order:* Counterbalancing method is used to avoid confounding among variables. Consider an experiment where the subjects are tested on both an auditory and visual reaction time task (here the subjects respond to an auditory stimulus) and a visual reaction time task (in which subjects respond to a visual stimulus). If each subject is tested first on the auditory reaction time task and second on the visual reaction time task, then the type of task and the order of presentation would be confounded. If visual reaction time were lower, there would be difficulty to know if the reaction time to a visual stimulus is 'really' faster than to an auditory stimulus or if the subjects learned something while performing the auditory task that improved their performance on the visual task. *This* is the compounding effect on the reaction time of the subject. The experiment could be designed better using the counter-balancing procedure. Here half the subjects would be given the auditory task, first. In this manner, the order effects would be neutralized. So the design counter-balances the sequence of task presentation. This design enables one to conclude that the effects of the reaction time obtained are pure and not influenced by the order of presentation. Half of the subjects should have been given the visual task first and the other half of the subjects should have been given the auditory task first. That way, there would have been no confounding of order of presentation and task, as the order of presentation and task would be 'counterbalanced.'

(iii) *Block-repetition method:* In this method, the block is first created and then successively repeated in which the natural sequence and the same order of the natural sequence of conditions is reported for each block. For example, ABC is repeated for each block. This means the first subjects would be to group A, the second to groups B and the third to group C.

The larger mean differences occur in the random blocks method, while the least occur in the random blocks method.

Factorial design: When the researcher is needed to manipulate two or more independent variables simultaneously; the most suitable design is the factorial design. The benefit from this is that their independent as well as their interactive effects on the dependent variable can be studied. A factorial design has three main features.

1. Two or more independent variables manipulated in all possible culminations.
2. Different sub-groups or subjects can serve every possible combination of the independent variables—equal number of subjects in all groups is preferred, though this is not necessary.
3. Independent as well as interactive effects can be studied, e.g., noise and illumination on the rate of learning. Noise and illumination are two independent variables. Further, noise has two levels—high and low and

NOTES

so also illumination, high and low. The resultant factorial design is a 2×2. There will be four possible combinations of the experimental editions.

A1 B1 for high noise high illumination

A2 B1 for low noise high illumination

A1 B2 for high noise low illumination

A2 B2 for low noise, low illumination

Then the subjects are randomly assigned to the different treatment combinations. In behavioural research ergonomic factors like age, sex, anxiety, etc. are attributes that have to be studied. Here randomization is to be used keeping the situational factors in mind (women, sometimes are ill-advised to be exposed to high noise levels).

Factorial designs with two variables 2×2 design is analysed using the ANOVA developed by R.A. Fisher (as shown in the table below).

Table 4.5 *Fictitious Data Obtained in an Experiment Based upon a 2×2 Factorial Design*

Group I (n1 = 10)	Group II (n2 = 10)	Group III (n3 = 10)	Group IV (n4 = 10)
High noise and high illumination	Low noise and high illumination	High noise and low illumination	Low noise and illumination low
15	10	16	10
14	12	18	9
20	10	22	8
22	9	25	7
16	8	26	6
18	12	20	10
20	11	20	11
21	10	18	12
18	10	17	10
17	10	16	10
.. X^2= 181	102	198	93
.. X^2= 3339	1054	4034	895
Mean = 18.1	10.2	19.8	9.3

Source: Singh, A.K., *Tests, Measurement and Research Methods in Behavioural Sciences,* 2008.

Table 4.6: *Means for the Four Experimental Conditions*

	Noise (A) High (A_1)	Noise (A) Low (A_2)	Means
High (B_1)	(A_1B_1) $\Sigma X = 181$ Mean = 18.1 $n_1 = 10$	(A_2B_1) $\Sigma X = 102$ Mean = 10.2 $n_2 = 10$	14.15
Low (B_2)	(A_1B_2) $\Sigma X = 198$ Mean = 19.8 $n_3 = 10$	(A_2B_2) $\Sigma X = 93$ Mean = 9.3 $n_4 = 10$	14.55
Means	18.95	9.75	

ΣX = Total score on the dependent variable

Illustration of the details of the Calculation of ANOVA in a 2 × 2 Factorial Design (data from Table on the previous page)

NOTES

Step 1: Correction $= \frac{(\Sigma X_1 + \Sigma X_2 + \Sigma X_3 + \Sigma X_4)^2}{N} = \frac{(181 + 102 + 198 + 93)^2}{40}$

$= \frac{(574)^2}{40} = 8236.9$

Step 2: Total SS $= (\Sigma X_1^2 + \Sigma X_2^2 + \Sigma X_3^2 + \Sigma X_4^2) - C = (3339 + 1054 + 4034 + 895) - 8236.9$

$= 9322 - 8236.9 = 1085.1$

Step 3: Among SS $= \frac{(\Sigma X_1)^2}{n_1} + \frac{(\Sigma X_2)^2}{n_2} + \frac{(\Sigma X_3)^2}{n_3} + \frac{(\Sigma X_4)^2}{n_4} - C$

$= \frac{(181)^2}{10} + \frac{(102)^2}{10} + \frac{(198)^2}{10} + \frac{(93)^2}{10} - 8236.9$

$= \frac{32761 + 10404 + 39204 + 8649}{10} - 8236.9 = 9101.8 - 8236.9 = 864.9$

Step 4: Within SS = Total SS − Among SS $= 1085.1 - 864.9 = 220.2$

Step 5: SS between two levels of *A* (or the first independent variable):

$$\frac{(\Sigma X_1 + \Sigma X_3)^2}{n_1 + n_3} + \frac{(\Sigma X_2 + \Sigma X_4)^2}{n_2 + n_4} - C$$

$$= \frac{(181 + 198)^2}{10 + 10} + \frac{(102 + 93)^2}{10 + 10} - 8236.9 = \frac{(379)^2}{20} + \frac{(195)^2}{20} - 8236.9$$

$$= 9083.3 - 8236.9 = 846.4$$

Step 6: SS between the two levels of *B* (or the second independent variable)

$$\frac{(\Sigma X_1 + \Sigma X_2)^2}{n_1 + n_2} + \frac{(\Sigma X_3 + \Sigma X_4)^2}{n_3 + n_4} - C = \frac{(181 + 102)^2}{10 + 10} + \frac{(198 + 93)^2}{10 + 10} - 8236.9$$

$$= \frac{(283)^2}{20} + \frac{(291)^2}{20} - 8236.9 = 8238.5 - 8236.9 = 1.6$$

Step 7: Interaction SS:

Among SS − between SS for *A* − between SS for *B* $= 864.9 - 846.4 - 1.6$

$= 864.9 - 848 = 16.9$

In the above table the number of trials taken by each subject in learning a list of 15 consonant syllables with the criterion of one perfect recitation is shown. This data is rearranged in the next table so that the means of the four groups are placed in their appropriate cell.

Then the first step is to compute the total sum of squares (SS) and then divide it among (SS). (Step 3, next table) and within (SS) (step 4 of the table). Before this, it is important to compute the correction value (step 1 of the table).

The among (SS) indicates whether the groups differ or not. The purpose of this study was to find out whether or not variation in each independent variable affects the dependent variable score and whether there was any significant interaction.

NOTES

So, the among (SS) is divided into three components:

(a) Between SS for A
(b) Between SS for B
(c) Interaction A×B

In order to know whether there is a significant difference or not between high noise condition A, and low noise condition A2, the SS between these two values is computed. Again to know the existence of differences between high level of illumination B1, and low level of illumination B2, the SS between these two values is computed. The A×B interaction is easily calculated by subtracting the values of between SS for A and the between SS for B from among SS (step 7). All the numerical values from step 1 to step 7 must be positive. Any negative value must be corrected before proceeding further.

Within groups design

Within groups designs are of two types 'complete within groups design' and 'incomplete within groups design'. A complete within groups design is one in which practice effects are balanced by exposing the subject to the conditions repeatedly in a different order each time. This enables the researcher to interpret the results for each subject. An incomplete 'within groups design' is one in which each condition is administered to each subject only once while varying the order of administering the conditions across the subjects in such a way that the effects of practice are neutralized when the results of all the subjects are combined. The balancing procedures for 'complete within groups design' are block-randomization and ABBA counter balancing while in the 'incomplete within group designs' one common procedure is to use all possible orders of the treatment and subsequent assignment of each subject to one of the orders.

In within groups design or within subjects design, the same group of subjects is treated differently in different experimental conditions and their influence on the dependent variable is studied. This is also known as repeated-treatments design. For example, a subject tested for non-drug and later for a drug condition. We can then note the change in the dependent variable due to the drug intake. This type of research is common in the fields of learning, memory reaction time and other psychophysical studies.

The within groups design has two types:

(a) Two conditions and many subjects
(b) More than two conditions and many subjects

In the first case, the same group is tested under two conditions only. The mean difference between the two conditions is determined through a matched *t* test or Sandler's *A* test.

Comparison of between groups design with within groups design:

NOTES

S. No.	Between groups	Within groups
1.	Separate groups used in experimental conditions	Same subjects used in experimanental conditions
2.	No economy of subjects	Economy of subjects
3.	Individual differences not controlled	Individual mill controlled
4.	High error variance	Lower error variance
5.	Preparation time and effort of subjects to be tested, high	Preparation time and effort are reduced considerably
6.	Can be used in a wide range of experiments	Cannot be used in all types of experiments (knowledge of results, fatigues, drug effects, etc.)

Generally the between groups design is considered, but more powerful within groups design is preferred:

(a) when only a small number of subjects are available;

(b) when limited matching tasks exist for equating the independent variable;

(c) when convenience is a consideration; and

(d) longitudinal studies and other psychological phenomena can be studied only by this method of design.

Mixed design model

Here, one variable is fixed; the other is random

Example 4.1:

Mixed model (*A* fixed, *B* random)

$$F \text{ for } A = \frac{MS \text{ for } A}{MS \text{ for interaction}}$$

$$F \text{ for } B = \frac{MS \text{ for } B}{MS \text{ for within-groups}}$$

$$F \text{ for } A \times B = \frac{MS \text{ for interaction}}{MS \text{ for within-groups}}$$

Mixed model (*B* fixed, *A* random)

$$F \text{ for } A = \frac{MS \text{ for interaction}}{MS \text{ for within-groups}}$$

$$F \text{ for } B = \frac{MS \text{ for } B}{MS \text{ for interaction}}$$

$$F \text{ for } A \times B = \frac{MS \text{ for interaction}}{MS \text{ for within-groups}}$$

Source: Singh, A.K., *Tests, Measurement and Research Methods in Behavioural Sciences*, 2008.

NOTES

Important symbols: Campbell and Stanley (1963) have described and discussed 16 designs ranging from the poorest to the most powerful that are useful for psychological and educational researcher. They have also identified certain symbols that one must be familiar with. This is a useful key for researchers.

R: Random selection of subjects or random assignment of treatment to experimental groups.

X: Treatment or experimental variably, that is manipulated. When treatments are compared, they are labeled as X1, X2 and X3 and so on.

O: Observation or measurement or test. Where there is more than one measurement, the subscript is used O1, O2 and O3 and on.

When X and O occur in the same row, it indicates that they are being applied to the same persons.

When X and O occur from left to right, it indicates the temporal order. When X and O occur in vertical order X, it indicates that the two are simultaneous.

When parallel rows of symbols are separated by dashed lines (-----) it shows that the groups have been unequated by randomization.

When parallel rows of symbols are separated by dashed lines, it indicates that groups have not been equated by randomization.

Three samples of non-designs: Here, there is no control group or on equivalent group.

1. One shot case study: XO is the symbolic presentation. A treatment is given to one group and then the observation O is made to assess the treatment. This is a way of determining the results in relation to earlier behaviours / activities.
2. One group Pre-test–Post-test design: Here X are judged by comparing the pre- and post-test scores. There is no control group present.

   O1 X O2 is the representation of this design. Attitudinal changes can be studies by this method.
3. Static-group or intact-group comparisons. Here, two groups are used. One group (O1) receives the experimental treatment (X) and the other group does not (O2). Subsequently the two groups are compared.

   Design is $\frac{XO1}{XO2}$

The control group is used for comparing the experimental group. For groups that not equivalent, no random assignment of subjects to the groups is made.

True experimental designs: The three designs mentioned below qualify as true experimental designs, because the control and experimental groups are formed, their equivalence is established by randomization. Relevant factors are controlled.

NOTES

These are the strongest types of design. But under certain situations, they cannot be conducted.

1. Post-test only equivalent group design: This design has all the conditions of a good experimental design, yet it reduces the threat to the experimental validity.

Design	R	X	O1
	R	O2	

   All the other conditions of an experiment are followed. The most suitable statistical treatment methods are a *t* test or ANOVA.

2. Pre-test Post-test control group design: This design is similar to the earlier one except that a pre-test for both the groups before experimental and control treatments are administered.

Design	RO1	X	O2
	RO3		O4

   One group receives the treatment X, and the other does not. The mean gain score of O2 – O1, is compared with O4 –O3. The post-test means of O2 and O4 can also be compared for ascertaining the impact of the treatment on the groups.

3. Solomon four-group design: This is a combination of the above two designs, i.e., the post-test only and the pre-test post-test equivalent groups design. This controls the threats of external validity.

R O1	X	O2
RO3		O4
R	X	O5
R		O6

   Here the four groups are randomly set by the experimenter. There is also the advantage of replication here, because two simultaneous experiments are conducted. The effect of treatment X is replicated in four ways.

   O2 > O1

   O2 > O4

   O5 > O6

   O4 > O1

This design makes it possible to ascertain the testing effects as well as the interaction of testing, with other factors. The factorial analysis of variance is a suitable statistical method.

4.3.2 Quasi-Experimental Designs

This is a situation where an experimental interpretation is applied to results that do not meet the criteria of a true experiment. There is some control over the manipulation of the independent variable but not in creating the basic equivalent

NOTES

groups. This is a means to simulate a true experiment. So it is called a compromise-design. These are like experiments because the control group is used. But the subjects are not assigned randomly, so the validity of the experiment is questionable. The independent variable is not truly manipulated. So, the study is called a quasi-experiment. Subjects are assigned to a particular condition that already exists like age, sex, personality traits IQ, etc.

These designs lie between pre-experimental and true experimental design.

Time-series design: The design is

O1 O2 O3 O4 X O5 O6 O7 O8

Here a series of pre-test are administered to the group. Later the treatment X is given followed by a series of post-tests all for the same group. Extraneous variables are controlled here. The background variable is not controlled as the subjects are exposed to different stimulations daily. The results are analyzed by simple statistical comparisons. O4 and O5 comparisons are avoided. All other pairs are compared.

Equivalent time-samples design: It is like the time-series design. A single group is used and the group is administered repeated treatments in some systematic way.

Designs is

X, O1, XO O2 X1O3 XO O4

The treatment X1 is introduced and re-introduced with some other experience XO which is presented in the absence of the treatment. The background history is well controlled in this experiment by presenting X several times, so the external variables get well-controlled. This design is an improvement over the time series design.

Non-equivalent control-groups design: This is useful when the researcher has to work with intact groups whose membership is already existing (e.g., students of first year of college) the subjects cannot be reconstituted for some administrative reasons. In dealing with intact groups the design is

$$\frac{O1\ X\ O2}{O3\quad O4}$$

This shows that the random assignments of subjects has not been possible to the experimental and control conditions. So the equivalency is absent. The intact groups are compared on the pre-tests O1 and O3. The statistical analysis involves comparing the mean gain score of the treatment group (O2–O1) to that of the non-treatment group (O4–O3). This non-equivalent group design cannot be used when the intact groups are dissimilar.

Counterbalanced design: Here the experimental control is achieved by randomly applying the experimental treatments. This design is also known as cross-over design or switch-over design and rotation experiments. For counter balancing each treatment appears only once in each column and in each row. A counter balancing design involving four treatments and four groups on four different occasions is given below:

Groups A	X1O	X2O	X3O	X4O
Groups B	X2O	X4O	X1O	X3O
Groups C	X3O	X1O	X4O	X2O
Groups D	X4O	X3O	X3O	X1O

NOTES

Almost all the variables threatening the internal validity are well controlled. The only limitation of this design is that of multiple X interference. This reduces generalizability.

Separate sample pre-test-post-test design: This design is used when the experimenter cannot assign treatments to all subjects at the same time. So, a sample is chooser to receive the treatment. Then another sample is chosen and given the same treatment. For example, 5000 newly recruited persons are to be trained by the police. Only 200 persons can be tarried at a time and all people have to be trained. S0 the division of people into training conditions and non-training conditions does not occur. So, one group is chosen as the pre-test post-test one, and another group for repetition to enable comparison.

Design separate sample Pre-test – Post-test

O1 X O2

O3 X O4

If O2 exceeds O1 and O4 exceeds O3, we can conclude the effect of the treatment X. Since the two groups are separate and the independent variable is administered, some amount of power to the conclusion is enabled.

Patched-up design: Here the experimenter starts with an inadequate design and adds more features, as one proceeds.

Groups A	X	O1		
Groups B		O2	X	O3

This is used to overcome the weakness of a given design. The comparisons made are between O3 and O2 and O2 and O1. The groups are tested in sequence and then compared.

Quasi-independent variable of time is a well studied variable. Developmental psychology uses this extensively.

For the study of time, three designs are used:

1. Longitudinal design
2. Cross-sectional design
3. Cohort design

NOTES

Longitudinal design: Here a group of subjects are studied to see the effect of time. There are a lot of confounding factors present here. Motor development, verbal developments etc., are all examples of this. Repeated observations of each subject at different ages are carried out.

Subjects 1	5	6	7	8	9	10
Subjects 2	X	X	X	X	X	X
Subjects 3	X	X	X	X	X	X
Subjects 4	X	X	X	X	X	X
	X	X	X	X	X	X

X and X show the scores and means. Adding vertically the mean scores for the conditions reflects the changes occurring as a function of age.

Advantages of longitudinal study:

1. Subject variables are held constant between the conditions, e.g., genetic factors
2. Sample equivalence problems do not arise
3. Cause–effect connection established, well
4. It can trace growth increments and patterns

Disadvantages:

1. Extraneous variables come into play
2. Societal and cultural factors do influence
3. Due to repeated measures, carry over effects may occur.
4. Environment of the subjects cannot be controlled between testing periods
5. The researchers get stuck into an early determined design.
6. Expensive in terms of time and cost

Cross-sectional study: This is a form of between subjects quasi-experiment. The researcher observes the subjects at different ages or at different points in temporal sequence. A cross-section of ages is selected. For example, studying vocabulary, a group of subjects of 5 year olds, another of 6 year olds and so on are observed at the same time. Then comparisons are made:

5	6	7	8	9	10	years
X	X	X	X	X	X	
X	X	X	X	X	X	
X	X	X	X	X	X	
X	X	X	X	X	X	
X	X	X	X	X	X	

X represents the scores and X represents the mean of different age groups.

Advantages:

- Cross-sectional study can be conducted quickly and easily
- It is cost effective
- Data are collected at one go.

Disadvantages:

- Directional change in not indicated by the study
- Chronological age is the same, but maturity level, not controlled
- Developmental milestones ignored
- Comparability of groups, unsure

It is a form of snap-shot study

Cohort design: Here, longitudinal studies of several groups are done and this is usually a form of gerentological study or developmental study. See figure 4.6

Diagram of a cohort study

Repeated measures over ages

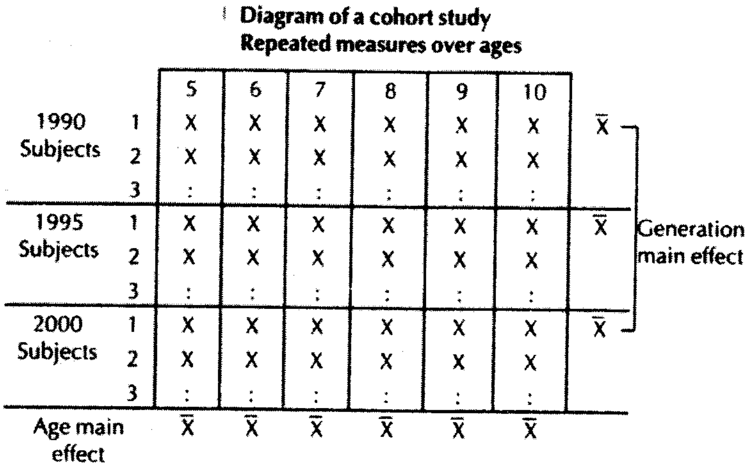


Fig. 4.6 Cohort Design

Source: Singh, A.K., *Tests, Measurement and Research Methods in Behavioural Sciences*, 2008.

The columns represent the repeated measures over the age. The rows denote three respective generations. Age effects and generation effects, can both be ascertained

Ex post facto design: Here the experimenter does not introduce a treatment, but evaluates a naturally present treatment, or one that has already occurred. The dependent variable is related to the conditions that already exist.

Two types of common ex post facto designs

1. Correlational design
2. Criterion-group design

NOTES

Correlational Design: This is also called a psychometric approach. Here two sets or more of data are collected from the same group of subjects so that the relationship between the two subsequent sets of data can be determined. The correlations form the main statistics for analyzing the data. Other variants of correlational techniques can also be utilized.

Diagram of design

O1 O2

Study: To examine the relationship between anxiety and performance of a group of children of Class XII. First a measure of anxiety is obtained and later a test of performance. So, two sets of data are available—O1 and O2. The correlation coefficient can be calculated. It can be positive, negative or absent. Correlations only signify the magnitude of the relationship between the two variables. It is not a cause–effect relationship between the variables.

Criterion group design: Here groups with contrasting characteristics are selected and the experimenter tries to ascertain what has caused the particular state by examining groups which have certain criterion with those which do not posses it.

Diagram of the design

C O1 O1 C O2 O1

Or ———— or ————

O2 O3 O4 O2

C-selected criterion: It is a kind of static group comparison. For example, in a study of creativity in children, 50 students are selected randomly. Then, a test of creativity is administered. Then the students are divided into two groups—high and low on creativity. Later the experimenter tries to relate this to the child rearing practices in the family. This can be done by the structured interview methods. The results could point to a set of potential factors leading to creativity, but not a definitive conclusion.

Multivariate analysis: In modern times, with the availability of the computer and various types of software packages, several multivariate analyses have become possible to treat data. Multivariate analysis includes all statistical methods which can simultaneously analyze more than two variables on a sample of observations.

Multiple regression analysis: This analysis is suitable when there is a dependent variable that is influenced by two or more independent variables. In a simple regression model, the dependent variable is a function of one independent variable, e.g., sales as a function of advertisement expenditure. Suppose the regression yields R^2 =.60 and the regression coefficient is significant. This means only 60 per cent of the variation has been explained. The other 40 pr cent is due to some other influence besides expenditure. So, more explanatory variables need to be added. One such variable may be the per capita income of the trading area. The regression analysis that uses two or more independent variables is termed multiple regression analysis.

NOTES

Multiple discriminant analysis: This method is useful when there is a single dependent variable that cannot be measured, but can be classified into two or more groups on the basis of attribute. The object of this analysis is to be able to predict that a particular entity belongs to a particular group, based on several predictor variables.

Multiple ANOVA: This is an extension of the two-way ANOVA.

Canonical analysis: This is a useful method when both the measurable and non-measurable variables are present. This method can help in simultaneously predicting a set of dependent variables from their joint covariance with a set of independent variables.

Inferential analysis: This is concerned with various tests of significance for hypothesis testing. It can also help in estimation of population values. Conclusions are based on this type of analysis.

CHECK YOUR PROGRESS

5. What is matched groups design also known as?
6. Name the different types of matching.
7. When is a factorial design used?
8. Define true experimental designs.
9. What is a correlational design?
10. What is the cohort-design?

4.4 QUALITATIVE METHODS

Qualitative research is gaining a lot of importance in the social sciences. There are specific methods with different assumptions and objectives. These are intrinsic to all types of research. The steps of the research process and the procedures for collecting data and analyzing and interpreting it are different.

Today, inductive strategies are increasingly used in research instead of starting from the basics and testing them empirically. Traditionally psychology and other social sciences have used the model presented by the natural sciences for developing quantitative and standardized methods. The guiding principles have been identifying the causes and effects and the measures have been quantity phenomena. These have permitted generalizations and the formulation of laws. Causal relations are studied under controlled conditions and observations have been classified in terms of their frequency and distribution. Studies have been made as objective as possible. For a long time, psychological research has depended on experimented designs for study as the mainstay.

NOTES

The results of social sciences are not applied in daily life as the methodological precisions are far too removed from every day problems and situations. The goal of subject and situation-related research can be achieved with qualitative research in a meaningful way.

It is important to recognize that not all questions of research can be studied empirically. Sometimes methods are not available for study purposes. Complex social science problems are not easy to be put into the mould of empirical investigations. It is not easy to reduce complex situations to simple and single variables for the sake of studying. The entire complex situations are rendered rather simplistic and therefore, often inappropriate. The subjectivities of the researcher and those being studied are part of the qualitative research process. The qualitative researchers are soft, understanding, open, and descriptive strategies. This is in opposition to the hard experimental, standardizing and quantifying approaches.

4.4.1 Definitions of Qualitative Research

It is a method of inquiry used in many different academic disciplines, traditionally in the social sciences. It is a set of research techniques in which data is obtained from a relatively small group of respondents and not analyzed with statistical techniques.

Qualitative research follows an inductive research process and involves the collection and analysis of qualitative (i.e, non-numerical) data to search for patterns, themes, and holistic features.

There is no formal quantitative framework used to generate projections in qualitative research. It is concerned with understanding the processes, which underlie various behavioural patterns. 'Qualitative' is primarily concerned with the 'why' of investigations. Case study, observation and ethnography are considered forms of qualitative research. Results are not usually considered generalizable, but are often transferable. A subjective form of research that relies on analysis of controlled observations of the researcher. This type of research yields extensive narrative data, which include detailed descriptions of what has been observed.

Here, collection of non-numerical data using interviews, observations and open-ended questions, to gather meaning from non-quantified narrative information, are involved. It is a form of research that derives data from observations, interviews or verbal interactions and focuses on the meanings and interpretations of the participants.

This is a way to study people or systems by interacting with and observing the subjects regularly. It is a free-form research technique that is used to gain insight into the underlying issues surrounding a research problem by gathering non-statistical feedback and opinions rooted in people's feelings, attitudes, motivations, values and perceptions, often from small samples, also called soft data.

NOTES

The gathering and analysing of data based on interviewees' own perceptions or experiences in order to provide insight into their beliefs about their circumstances rather than measurable data. Subjective information is obtained from groups and in-depth interviews and participant observation. Qualitative research is generally conducted on a small group. It looks at phenomena like attitudes, perceptions, etc.

4.4.2 Distinctions between Qualitative and Quantitative Research

Qualitative methods are used for exploration, especially for generating hypothesis. Quantitative methods are used for hypothesis testing. Establishing the content validity of any measure is the strength of the qualitative research. Qualitative research is used for evaluating any policy, program, etc., more effectively than quantitative research. The how and why of phenomena are better understood by qualitative research. Certain questions like relevance of a particular topic is better studied qualitatively. This type of research allows for more diversity in responses and it can adapt to new developments, as they happen. The count is not important as in quantitative research, but the material got through study is seen as valuable.

The aim of qualitative research: The primary aim of qualitative research is to gain insights into any phenomena of interest to the researcher. For example, one is interested in studying mental disorders. Epidemiological studies show the frequency of schizophrenia and the distribution of this disorder in a population. The finding is it is higher in the lower socio-economic classes. These correlations have been well-established by empirical quantitative data. However, it is not clear whether the conditions of living in the lower economic conditions foster the development of schizophrenia or people with schizophrenia get marginalized and drop into the lower economic strata of society. Often studies fail to capture the contextual perspective of the illness. What is the subjective experience of being schizophrenic? How do the family members deal with this illness? What are the benefits of institutionalizing such a patient? How do professionals view women with schizophrenia? These are some of the relevant questions that qualitative research seeks to find answers to.

For qualitative research, the mental illness starts with the patient, his / her relatives and the professionals involved and their perspectives. It also analyses the interactions in dealing with the illness for all concerned. The different viewpoints are significant in this type of research. Communications between, the investigator, patients, psychiatrists and others in the field, are part of the knowledge about the disease. The subjectivities of the investigator a well as that of the patient, constitute part of the research process. The researcher's reflections frustrations, feelings, impressions and so on all become part of legitimate data in their own right in the context of dealing with patients involving the illness.

Qualitative research is:

- A return to narratives, language and communication
- Questions about the specific, particular and concrete problems

NOTES

- A study of the local contexts, instead of universals
- Putting problems to be studied in their temporal, historical context and describe and explain them are different approaches to qualitative research.

The approaches have different theoretical assumptions:

1. Subjective meaning: This is the meaning individuals attribute to their activities and environments.
2. This meaning is derived from and arises out of social interaction that one has with ones fellows.
3. These meanings are handled and modified through interpretations by people while dealing with various interactions.

'The researcher has to see the world from the angle of the subject, he/she studies' (Toulmin, 1990). This means that methodologically the researcher has to reconstruct the subject's view point in various ways. This is in the form of subjective theories people have in explaining different aspects of the world. For example, The role of punishment in disciplining children; How the universe was for med?, etc. Here autobiographical narratives, biographical reports, etc., are all significant, as methods.

Making of social reality: Ethno methodology is one of the techniques used for studying how people construct their own social reality. Garfunkel (1987) proposed this method. Here, every day activities are analyzed in terms of their social context as to why some activities occur everyday. All these are viewed as visibly rational and reportable for all practical purposes. This research is based on conversation analysis.

Conversation analysis is based on three basic assumptions:

1. Interaction is structurally organized
2. They are context shaped and context renewed
3. Conversational details cannot be dismissed as disorderly or irrelevant.

Cultural framing of social and subjective reality and cultural systems determine the perception and making of social reality.

4.4.3 Research Perspective

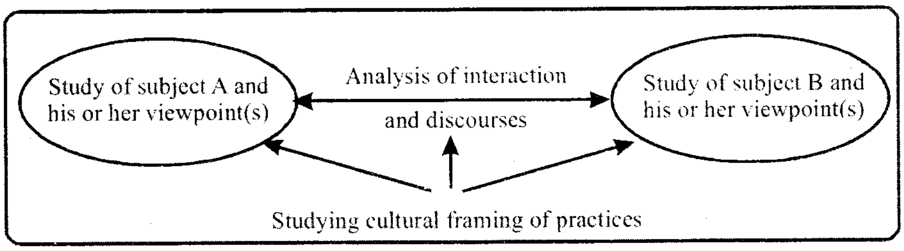


***Fig 4.7** Research Perspectives in Qualitative Research*

Source: Flick, *An Introduction to Quantitative Research*, 1998.

NOTES

A social representation is understood as a system of values, ideas and practices with a twofold function: First to establish an order which will enable individuals to orient themselves properly in their material and social world and to master it; and secondly to enable communication to take place among the members of a community by providing them with a code for social exchange and a code for naming and classifying unambiguously the various aspects of their world and their individual and group history.

Qualitative research aims at understanding an event from the inside. It is the view of one subject or subject(s), the social situation or the cultural or social rules relevant for a situation that have to be understood. A single case is analyzed before a general statement is made.

1. First a single subject theory
2. Then a single conversation
3. And then a single cases is reconstructed
4. Later other case studies are used as a comparison.
5. Then a typology or different subjective theories are developed

The reality studied by qualitative research is not a given reality, but as constructed by different 'actors'. How crucial an actor is, is determined by the theoretical position adopted to study the process of construction.

In the process of reconstruction, texts are used for empirical analysis. The view of the subject is presented as his/her subjective theory. For example, belief in astrology where the course of interaction is recorded and transcribed and then the reconstructions of latent structures of manning are formulated from the texts (i.e., what is told and the actions that follow, e.g., matching of horoscopes). The status of the texts will be determined by the theoretical position held.

Table 4.7 *Theoretical Position in Qualitative Research*

	Subjects' points of view	**Making of social realities**	**Cultural framing of social realities**
Traditional theoretical background	Symbolic interactionism	Ethnomethodology	Structuralism, psycholanalysis
Recent development in social sciences	Interpretive interactionism	Studies of work	Poststructuralism
Recent developments in psychology	Research programme 'subjective theories'	Discursive psychology	Social representations
Common features	• Verstehen as epistemological principle • Reconstructing cases as starting point • Construction of reality as basis • Text as empirical material		

Source: Flick, *An Introduction to Quantitative Research*, 1998.

4.4.4 Texts and Realities in Qualitative Research

NOTES

Texts serve three purposes in qualitative research.

1. They are not only the essential data, but the basis for interpreting and communicating the findings
2. Texts are seen as an instrument for interpretation of social reality
3. Interviews comprise the data that are transformed into texts.

Hereafter the text is used as a substitute for the reality

Text making of the world involves first degree and second degree constructions. Realities are actively produced by the participants through the meanings ascribed to certain events and objects. This is what qualitative researchers seek to study. These ideas of social events in a social field may compete, conflict or succeed and are shared and taken for real. This is how reality is constructed.

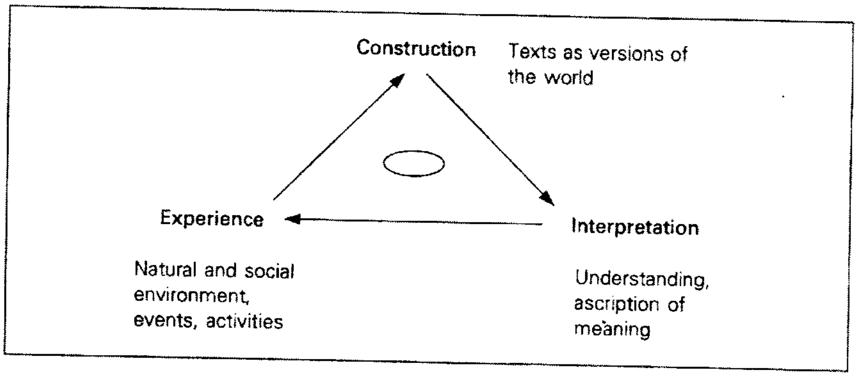


Fig. 4.8 *Realities*

Source: Flick, *An Introduction to Quantitative Research*, 1998.

Strictly speaking there are no such things as facts, pure and simple. All facts are from the outset selected from a universal context by the activities of our mind. They are, therefore, always interpreted facts, either facts looked at as detached from their context by an artificial abstraction or facts considered in their particular setting. In either case, they carry their interpretational inner and outer horizons.

For Goodman (1978), the world is socially constructed through different forms of knowledge—from everyday knowledge to science and art as ways of making the world. Social research is an analysis of such ways of world making and the constructive efforts of the participants in everyday life. A central idea in this context is the distinction between first degree and second degree constructions—first degree is the construct made by an actor. The constructs of the constructs made by the actor in the social scene is the second degree. The exploration of the general principles, according to which man organizes his experiences in daily life and those of the social world, is the first task of the methodology of social sciences.

So, there could be multiple social realities. Social science research encounters the world it wishes to study only in those versions as constructions by the subjects.

So, there are subjective constructions by the participants and subjective constructions by the researchers. Knowledge of the world is not just found in the world, rather it is built into it. Worlds are made from other worlds. A big part of research involves reconstructing life stories or biographies in interviews

4.4.5 Theories in Qualitative Research

To begin with, qualitative research is circular, not linear like in quantitative research. In this type of research, theories undergo revisions, evaluation, construction and reconstruction. They are versions of perspectives through which the world is seen.

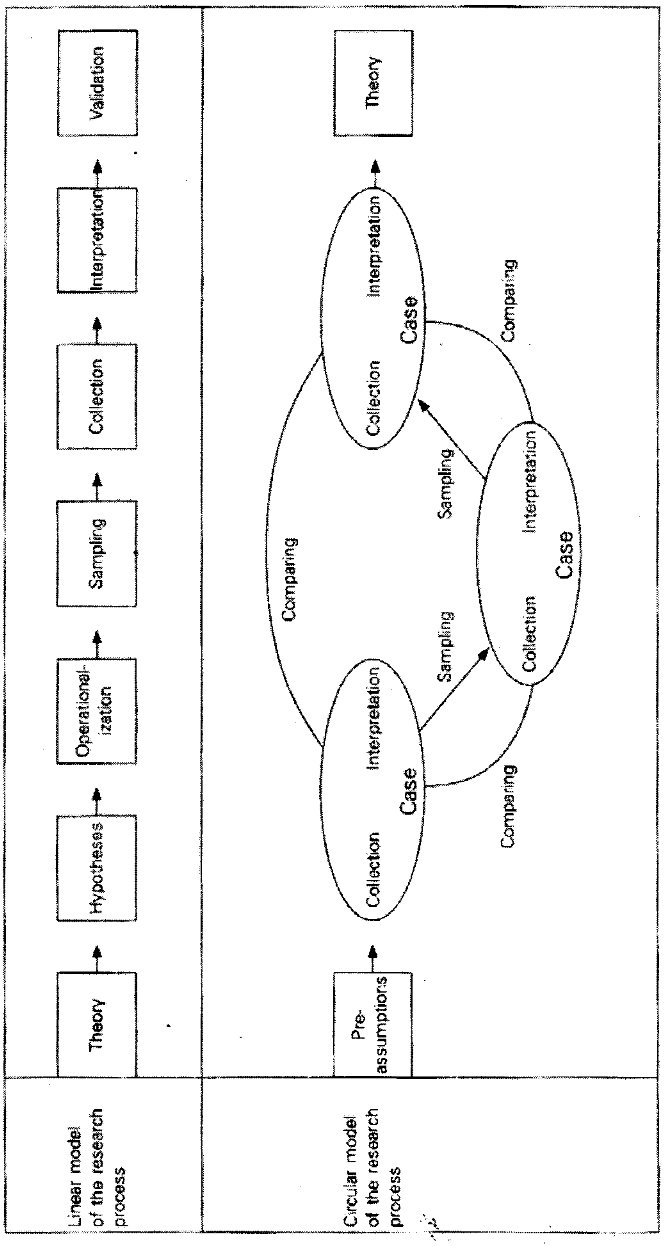


***Fig. 4.9** Models of Research*

Source: Flick, *An Introduction to Quantitative Research*, 1998.

NOTES

Each formulation has its own perspective hidden within it. The perception in turn feeds back into the social construction of this perspective and thereafter the world around us. So, theories are mere preliminaries and relative. When the version is developed further additional interpretations become possible of the new material and so on. So, there is a pre-understanding from which, all study starts. Glasser and Strauss suggest a circular model for qualitative research.

Formulating research question in qualitative research

Research questions exist in the investigators personal biography and social context. Certain issues are brought to the fore and others ignored.

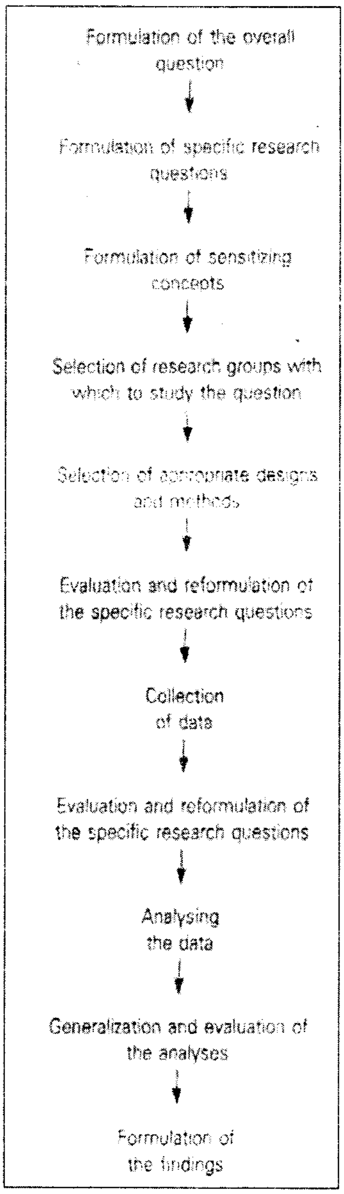


Fig. 4.10 Research Questions in the Research Process

The research questions can be of two types: Those oriented towards describing states (which kind, how often, etc) and those oriented towards describing process (causes, strategies, etc)

TABLE 4.8 Types of research questions

	Questions						
Units	What type is it?	What is its structure?	How frequent is it?	What are the causes?	What are its process?	What are its consequences?	What are people's strategies?
Meanings							
Practices							
Episodes							
Encounters							
Roles							
Relationships							
Groups							
Organizations							
Settlements							
Worlds							
Lifestyles							

Source: Lofland and Lofland © 1984, p. 94. (Reprinted with kind permission of Wadsworth, Inc. Belmont, CA)

Source: Flick, *An Introduction to Quantitative Research*, 1998.

NOTES

A researcher can enter the field of study as: a stranger, visitor or as an insider.

The best role to adopt is that of an insider. The set of realities presented would be most similar in the role of an insider and ideal for qualitative research.

NOTES

4.4.6 Sampling Strategies in Qualitative Research

Instead of selecting a sample, the complete collection method is used in qualities research. The sample is pre-determined by certain criteria, e.g., a certain disease, age, gender, region, marital status, etc.

Sample is also defined gradually. They are made on the basis of the groups to be compared or may focus on specific persons. The sample is chosen on the basis of new insights for developing a theory. Groups or persons chosen for the sample are stopped when saturation is reached, i.e., nothing new could emerge hereafter.

Samples of gradual selection

Table 4.9 *Theoretical Versus Statstical Sampling*

Theoretical sampling	Statistical sampling
Extension of the basic population is not known in advance	Extension of the basic population is known in advance
Features of the basic population are not known in advance	Distribution of features in the basic population can be estimated
Repeated drawing of sampling elements with criteria to be defined again in each step	One-shot drawing of a sample following a plan defined in advance
Sample size is not defined in advance	Sample size is defined in advance
Sampling is finished when theoretical saturation has been reached	Sampling is finished when the whole sample has been studied

Source: Wiedemann 1991, p. 441

Source: Flick, *An Introduction to Quantitative Research*, 1998.

Sampling proceeds according to relevance, not representations; and width and depth as the aim of a sample, not representativeness.

Table 4.10 *Sampling Strategies*

Sampling strategies in qualitative research

- A *priori* determination
- Complete collection
- Theoretical sampling
- Extreme case sampling
- Typical case sampling
- Maximal variation sampling
- Intensity sampling
- Critical case sampling
- Sensitive case sampling
- Convenience sampling
- Primary selection
- Secondary selection

Source: Flick, *An Introduction to Quantitative Research*, 1998.

The object of these sampling strategies is to provide richness of information. It also helps in obtaining relevant information. It is an intense attempt at data sampling. Interviews should be non-directive, specific, average, in-depth and should have personal context and content.

NOTES

Sample selection is based on the principle of gradual selection. It is always purposive sampling, not random sampling. Palton gives the following concrete suggestions:

- Integrate purposively extreme or deviant cases. Cases of successes or failures are chosen and analyzed. This is thought to help in understanding the whole field.
- Another feature is to select typical cases. These are the average for the group. This means looking at the field from the centre and from the inside.
- To choose a maximal variation sample—here, the idea is to integrate only a few cases, but which are as different as possible. This is aimed at obtaining the range of variation and differentiation in a field.
- Choosing cases on the basis of intensity—here, the expectation is that the interesting features, processes and experiences that go with this would be integrated in the study and compared.
- Selection of critical cases is done to enhance the functioning of the study, by virtue of its richness.
- Choice of sensitive cases—these are to be included due to their explosive force and, therefore, the unique prospect of studying them making it special.
- Finally, the concept of convenience in selection—these are included because of easy access, reduced effort and decreased time when people, money, etc., are limited.

Cases are considered as samples

Each case is thought to have five representative aspects:

(i) A single case is understood as an 'individualized universal.' It is viewed as a result of specific individual socialization against the general background, e.g., a physician or psychologist trying to understand a disorder. Such a socialization can be understood in the social context, leading to different subjective opinions, attitudes, etc. This gets to be displayed in the actual interview situation.

(ii) To understand the 'individualized universal', the case is seen in a specific context in which the individual acts and which he/she represents to others. So, the doctor or social worker orients his or her practices to the institutions under which these difficulties arise, e.g., a dysfunctional family, night shift work, etc.

NOTES

(iii) An individual case is also seen as a specific form of professionalisation, e.g., information engineer, social worker, etc. Therefore, the training of the professional individual and the actions that arise from that context got to be studied and analyzed. Therefore, different professionals approach the same problem differently. This makes a given case a unique sample.

(iv) The case also develops its subjectivity because of the specific stock of knowledge and the corresponding ways of acting and perceiving come into play.

(v) The case also represents an interactive mode of context and activity.

Sampling decisions cannot be made in isolation. The appropriateness of the content and the strategy together determine the selection. It may be difficult to make generally valid statements on the basis of a single study, but it is also equally difficult to give deep descriptions and explanations of a case based on the principles of random sampling. This is the strongest argument of the qualitative researcher. Since sampling strategies are meant to disclose a given field, it can start from the extreme, the negative, the critical and the deviant cases and thereafter form the extremities of the field. It can also start from the inside, which is the case of the typical or the developed cases. There, sampling can start from those cases which are as far different as possible in their variation. This precludes the need for homogeneity in the sampling. Sampling decisions determine what becomes the empirical material in the form of texts and what is taken from the texts and how it is used.

4.4.7 Different Types of Interviews for Qualitative Research

4.4.7.1 The focused interview

First part: It has unstructured questions, like what did you like most about?

Second part: Semi-structured questions—they are: left open-ended for being able to get a feel about a given rule.

Third part: Structured questions like was the lecture informative?

4.4.7.2 Problem-centred interview

It involves asking questions like 'How open do you show the allergy?'

Object-orientation to determine a problem: Do you see yourself as healthy or unhealthy?

Process orientation: Do you search the Net to know more about your allergy?

4.4.7.3 Expert interview

This type of interview is directive, stronger and guided.

4.4.7.4 Ethnographic interview

Here the project is explained in descriptive and structural questions.

4.4.7.5 Narrative interview

Here knowledge and experiences are presented. This is called a generative narrative. This is to obtain or elicit answers on a theme of study.

NOTES

4.4.7.6 Episodic interview

Here episodes are presented as a source of knowledge in terms of circumstances or events. Then the semantic knowledge (the meaning) attached to the events is studied through the narration.

4.4.7.7 Group discussions

Observing the responses of a small group at one time is called a group discussion. An explanation is given about a topic, procedure, etc. Members are introduced as a warm up act and then asked to discussion the topic. The opinions given by the members are got from verbal data. Statements and thoughts are gathered. This is a dynamic process in operation.

4.4.8 Focus Groups

Here the focus is on the interaction among members. A focus group is a form of qualitative research in which a group of people are asked about their perceptions, opinions, beliefs and attitudes towards a product, service, concept, advertisement, idea, or packaging.

All the above methods are used for collecting verbal data. The method to be chosen is on the basis of its appropriateness.

4.4.9 Visual Data

4.4.9.1 Observation

Observations can be of the following types.

- Covert vs. overt observation
- Non-participant vs. Participant observation
- Systematic vs. unsystematic observation
- Natural vs. artificial situation observation
- Self-observation vs. observing others

4.4.9.2 Phases of Observation

Authors like Adler and Adler (1994), Denzin (1989b) and Spradley (1980) name the phases of an observation as:

The selection of a setting, i.e. where and when the interesting processes and persons can be observed;

The definition of what is to be documented in the observation and in every case;

NOTES

The training of the observers in order to standardize such focuses;

Descriptive observations which provide an initial general presentation of the field;

Focused observations which concentrate more and more on aspects that are relevant to the research question;

Selective observations which are intended to purposively grasp only central aspects;

The end of the observation, when theoretical saturation has been reached (Glaser and Strauss, 1967), i.e., further observations do not provide any fresh knowledge.

4.4.10 Ethnography

The ethnographer participates, overtly or covertly, in people's daily lives for an extended period of time, watching what happens, listening to what is said, asking questions. He collects whatever data are available to throw light on the issues with which he or she is concerned. He may show photos to the interviewee while interviewing and sometimes he may use film analysis questions or micro analysis of claps done.

4.4.11 Coding and Categorizing

The interpretation of data lies at the core of qualitative research. For coding and categorizing are crucial. The texts obtained can either be augmented by alteration or paraphrased, summarised or categorized.

The process of coding

There are three types of coding:

(i) Open coding

(ii) Axial coding and

(iii) Selective coding

These are ways of handling the textual material. Coding is the process by which data are broken down, conceptualized and put back together. The coding is initially is done closely to the text as possible and then made more abstract. Categorizing refers to the summarizing of such concepts. Another way of coding is by themes. These are useful for comparisons.

4.4.12 Content Analysis

This is a classical procedure for analyzing textual material. This includes both visual and interview data. The categorizing done earlier are used and modified, as required.

First step is to select the relevant parts of a text;

Second is to analyze the data with who, what, when, whom, how etc., questions.

The coding unit – presents the smallest element. The categories – provide the larger unit

The contextual unit determines the largest elements for analysis.

Techniques of content analysis

Three basic techniques are:

1. Paraphrasing the material—here the less relevant details are skipped (first reduction)
2. Bundling and summarizing—similar paraphrases are grouped together (second reduction)
3. Generalization—summarizing at higher level of abstraction.

4.4.12.1 Sequential analyses

In order to analyze the elements or statements, it is necessary to put things contextually. Here sequential analysis is useful. Any idea is put in social order, which also gives the understanding of the interactions that occur. So, the contents are presented in a reliable way in the best possible context.

Conversations, discourses, narratives are all interpreted using sequential analyses.

Making qualitative research reliable and valid: The legitimacy of this type of research has always been criticized. One of the ways to achieve reliability is to train the observers before they enter the field and evaluate the observations. This is to improve comparability. Certain conventionalized procedures for field notes are an example of increasing reliability.

Table 4.11 *Symbols for Reliability*

Sign	Convention	Use
" "	Double quotation marks	Verbatim quotes
' '	Single quotation marks	Paraphrases
( )	Parentheses	Contextual data or fieldworker's interpretations
< >	Angled brackets	Emic concepts (of the member)
/ /	Slash	Etic concepts (of the researcher)
——	Solid line	Beginning or end of a segment

Source: Adapted from Kirk and Miller 1986; Silverman, 1993.

Validity: This has received more attention than reliability. Certain guidelines have been proposed with this in mind.

4.4.13 Documentation of Data in Qualitative Research

For interview data—the spoken words are edited and transcribed. For observational data—the documentation involves recording the actions and interactions. In both these the contextual enrichment of the data is an important aspect of the documentation process. The procedures involve the texts, which becomes the basis for analyses. There are four steps in documenting data.

NOTES

(i) Recording the data

(ii) Editing the data

(iii) Constructing a new reality in and from the produced text.

(iv) This is the way in which the construction of reality in qualitative research takes place.

Audio-visual acoustic and visual recordings are all part of the new recording possibilities of data collection. However, what is worrisome is that the presence of the recorder and the recording instruments could affect the recording. Here the least intrusive of equipment are to be used and the minimal presence of people is suggested in the observation areas.

For field notes, the qualitative researcher is told to take notes and record the observations immediately. Since the researcher is often one who partakes in action research to a large extent, it is also suggested that notes be completed after ending the study. The researcher should spend as much time on noting the observations as was spent for carrying out the observation—the general rule of thumb for effective documentation. This is how reality is to be produced from the field notes. Along with this the protocols of diaries should also be maintained, for corroboration.

Field notes documentation

- Condense conversations into single words, sentences etc.
- Expand the anecdotes from interviews and observations
- The field -- work journal must have an account of experiences, mistakes, problem, insights etc.
- Notes and interpretations can start as soon as the field study commences and till the completion of the study.

Research diary

- Keep diaries updated, as a documentation procedure
- Compare, if more than one researcher is involved
- Catch interesting events in the diary
- Only document that which is essential
- Make recordings immediately after an event.

Transcription

When language analysis is involved, the focus in transcription should be to obtain the maximum exactness in classifying the statements, pauses, hesitations, etc.

NOTES

Table 4.12 Transcription

[	Overlapping speech: the precise point at which one person begins speaking whilst the other is still talking, or at which both begin speaking simultaneously, resulting in overlapping speech.
(0/2) 'Aw:::':	Pauses: within and between speaker turns, in seconds. Extend sounds: sound stretches shown by colons, in proportion to the length of the stretch.
Word: 'fishi-': '.hhhh':	Underlining shows stress or emphasis. A hyphen indicates that a word/sound is broken off. Audible intakes of breath are transcribed as '.hhhh' (the number of h's is proportional to the length of the breath).
WORD: (words...):	Increase in amplitude is shown by capital letters. Parentheses bound uncertain transcription, including the transcriber's 'best guess'.

Source: Flick, *An Introduction to Quantitative Research*, 1998.

Texts produced in this way are the nearly constructed realities

- The research should talk less and listen as much as possible
- Produce notes as exactly as possible
- Bring data to unify as early as possible
- The readers of the notes must be able to see for themselves clearly which persons are providing enough data
- The reports should be open and clear
- The researcher should seek feedback on findings/representations
- Presentations should be balanced between the various aspects observed/ recorded
- Writing should be as accurate, as possible

These steps seek to improve validity by sensitizing the researcher to qualitative research.

Procedural validity is sought to be achieved by the different relationships at work in research

- The relationship between what is observed (behaviours, rituals, meanings, etc.) and the lager cultural, historical and organizational contexts within which the observations are made (the substance)
- The relationship among the observed, the observer and the setting (the context)
- The issue of perspective (or point of view), of the observers' or the members', used to render an interpretation of the ethnographic data (the interpretation);
- The role of the reader in the final product (the audience);

- The issue of relationship, rhetorical, or authorial style used by the author(s) to render the description and / or interpretation (the style).

Here validation involves the entire research process

NOTES

Triangulation

This is a term used for combining the methods in qualitative research. Four types of triangulations are suggested:

- Data triangulation: It involves using different data sources in rating persons, places and situations
- Investigator triangulation: Here different interviewers or observers are used with a view to minimizing errors/biases.
- Theory triangulation: It involves approaching data with multiple perspectives and hypotheses in mind. Various types of orientations are placed side by side to see their usefulness for producing knowledge
- Methodological triangulation: It involves combining different methods such as combing questionnaire with an interview or using different sub-scales for measuring a phenomenon.

Triangulation is seen as a concept for validating results obtained with individual methods. These are thought to enrich and complete knowledge and lessen the limitations of individual methods used singly. These are the ways social realities are sought to be studied systematically. Triangulation is seen as a means to increase the scope, depth and consistency of knowledge through methodological means.

Analytic induction: Here the attempt is to understand and explain the exception that is deviant to a hypothesis in a systematic way to interpret results. It is a case of looking at negative data to be able to substantiate the general.

New criteria to evaluate qualitative research:

1. Trustworthiness
2. Credibility
3. Dependability
4. Transferability
5. Conformability

Credibility can be increased by persistent observation, and triangulation of methods, researchers and data.

Peer meetings is another process

Communication validation

The raw data are subjected to audits in order to increase dependability

Democratization of qualitative research is achieved by using constant comparative methods for interpreting texts. This is carried out by

1. Comparing incidents of each category
2. Integrating categories, by properties/time, etc.

3. Delimiting the theory
4. Writing the theory

This is a continuous growth process. Contrasting cases and ideal type analysis are carried out so that pure cases can be tracked and the understanding of the individual case be made more systematic.

NOTES

Comparisons of old and new criteria for the qualitative field:

Old	New
Objective	Conformability
Reliability	Dependability / Auditability
Internal validity	Credibility / authenticity
External validity	Transferability / Fittingness
Utilization / application	Action / Orientation

4.4.14 Presenting the Findings/ Results Following the Qualitative Research

These are to be seen as 'tales from the field' (Flick, 1998)

Realist tales, confessional tales and impressionist tales (dramatic recall)

All these are aimed at providing knowledge

Other forms of writing are critical stories and formal stories.

Writings never exist in pure form and are never communicable in this form. The readers have to be kept in mind—who need to have a clear understanding of what is presented. What the readers need to know in order not to misunderstand the writing. Therefore, not only how something is presented, but how that is presented is important. The version, interpretation, and content of the text are important factors in terms of writing the findings / results. Here is a diagram of how a text is to be presented:

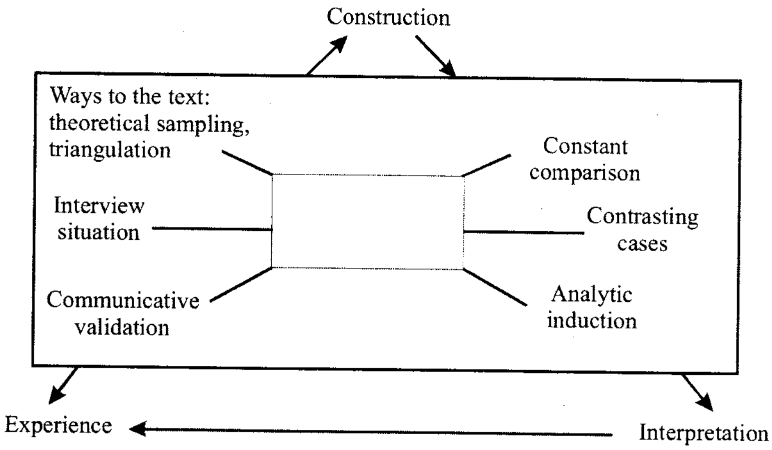


Fig 4.11 *Grounding the Text*

Source: Flick, *An Introduction to Quantitative Research*, 1998.

4.4.15 Computers in Qualitative Research

NOTES

Computers are widely used in the qualitative research for data analysis. Special programs are available for analyzing data, combing qualitative and quantitative research possibilities, and transforming one type of data into another—qualitative to quantitative and vice versa. Triangulation of research can also be done with the help of computers.

CHECK YOUR PROGRESS

11. When is multiple-regression analysis useful?
12. Give one purpose of texts in a qualitative research.
13. Define the theory of qualitative research.
14. What does visual data constitute?
15. Define content analysis.
16. What are the steps in documenting data in a qualitative research?

4.5 SUMMARY

- A research design is a conceptual framework for conducting research. It is a blueprint for collecting, measuring and analyzing data. Depending on the topic, time and cost, different types of research designs are used.
- A good research design should seek to minimize bias and maximize reliability of the data obtained. Given the least possible experimental error, it should be as objective as possible. No single design should be used for all types of research problems.
- Quantitative units or values are quantifiable factors like height, distance, loudness, etc. Qualitative units are attributes, e.g., honest, extroverted, shy, etc. Experimental research usually involves quantitative factors, whereas qualitative research deals with qualitative factors.
- Qualitative research is highly helpful in gaining insights into mental disorders and their patterns. Epidemiological studies show the relationship between the frequency of schizophrenia and the distribution of this disorder in a population. Schizophrenia is higher in the lower socio-economic classes. These correlations have been well-established by empirical quantitative data.
- Different types of interviews are used for different kinds of qualitative researches. Some well-known types of interviews include focused interview, problem-centered interview, expert interview, ethnographic interview, narrative interview and episodic interview.

NOTES

- Observing the responses of a small group at one time is called a group discussion. In a focus group, people are asked about their perceptions, opinions, beliefs and attitudes towards a product, service, concept, advertisement, idea, or packaging.
- Once data is collected it has to be categorized, coded and tabulated for analysis.

4.6 KEY TERMS

- **ANOVA:** In statistics, A collection of statistical models and their associated procedures.
- **Multiple ANOVA:** This is an extension of a two-way ANOVA.
- **Canonical analysis:** A useful method when both the measurable and non-measurable variables are present.
- **Inferential analysis:** Analysis concerned with various tests of significance for hypothesis testing.
- **Analytic induction:** An attempt to understand and explain the exception that is deviant to a hypothesis

4.7 ANSWERS TO 'CHECK YOUR PROGRESS'

1. A design is a plan for identifying the sources for obtaining information.
2. A variable is a unit with different quantitative values like height, loudness, etc.and qualitative traits like dependable, shy etc.
3. Different research designs include exploratory, descriptive, diagnostic and hypothesis testing designs.
4. Randomization protects against extraneous factors creeping into a study. The role of chance factors is also neutralized by randomization.
5. Matched group design is also known as a randomized-block design.
6. There are two types of matching: (i) Matching by pairs—on the basis of scores obtained, subjects are matched and paired: and (ii) Matching by mean and standard deviation.
7. When more than two independent variables are manipulated simultaneously, the factorial design is used.
8. A true experimental design means: (i) Equivalent groups are formal, (ii) Randomization occurs, and (iii) Relevant factors are controlled.
9. In a correlational design, two sets of data are collected from the same group of subjects so that relationship between two subsequent sets of data can be determined.

NOTES

10. In a cohort-design, longitudinal studies of several groups are done. It is useful for gerontological and development studies.
11. Multiple-regression analysis is a useful method when the dependent variable is influenced by two or more independent variables.
12. One purpose of texts in qualitative research is: tests constitute essential data and are also the instruments for interpreting data.
13. The theories of qualitative research are circular, not linear. Theories undergo revisions with new versions of constructions and reconstructions.
14. Visual data constitutes observations, photos and films.
15. Content analysis is a classical method for studying any textual material. It includes verbal and visual data.
16. The steps in documenting observational data in qualitative research are: (i) recording the data, (ii) editing the data, and (iii) constructing a new reality from the data.

4.8 QUESTIONS AND EXERCISES

Short-Answer Questions

1. Give any two basic principles of experimental designs.
2. Give one difference between group and within-group designs.
3. State one concept related to research design.
4. Why are research designs needed?
5. Why is a research plan developed?

Long-Answer Questions

1. Mention and explain different research designs.
2. Describe an experimental design.
3. Describe non-experimental designs and their usefulness in research.
4. Describe correlational designs. What is discriminant function analysis?
5. Define qualitative research. What are its aims?

4.9 FURTHER READING

Flick, U. 1998. *An Introduction to Qualitative Research.* New Delhi: Sage Publications.

Garfunkel, H. 1967. Studies in Ethnomethodology. Englewood Cliffs, New Jersey: Prentice-Hall.

Glaser, B. and Strauss, A.L. 1967. *The Discovery of Grounded Theory: Strategies for Qualitative Research.* New York: Aldine.

Goodman, N. 1978. *Ways of World Making.* Indianapolis: Hackett.

Kerlinger, F.N. 1973. *Foundations of Behavioral Research.* New York: Holt Reinhart and Winston, Inc.

Kirk, J.L. and Miller, M. 1986. *Reality and Validity in Qualitative Research.* Beverley Hills, CA: Sage.

Kothari, C.R. 1995. *Research Methodology: Methods and Techniques,* New Delhi: Wishwa Prakashan.

Silverman, D. 1993. *Interpreting Qualitative Data: Methods for Analyzing Talk, Text and Interaction.* London: Sage.

Singh, A.K. 2008. *Tests, Measurements and Research Methods in Behavioural Sciences.* New Delhi: Bharati Bhawan.

Toulmin, S. 1990. *Cosmopolis: The Hidden Agenda of Modernity.* New York: Free Press.

NOTES

UNIT 5 ROLE OF COMPUTERS IN RESEARCH, INTERPRETATION AND REPORT WRITING

Structure

5.0 INTRODUCTION

In this unit you will learn how computers have revolutionized research work. The high speed electronic digital computer has a major impact on every phase of behavioural research. Problem-solving and lengthy statistical and mathematical calculations, done manually, are things of the past. What took days, weeks and even months earlier, is now done in a matter of minutes. Research studies and calculations, which looked impossible earlier can now be tackled with the aid of computers, in minutes or hours. Computers are now dominating almost every walk of life, which not only makes them important but indispensable. Here, we will look at the applications of the computer, some important characteristics, the binary number system and the role of computers in research.

NOTES

5.1 UNIT OBJECTIVES

After going through this unit, you will be able to:

- Understand the essential characteristics of a computer.
- Understand the binary number system and learn important conversions
- Realize the importance of computers in research and learn how to use them
- Understand and interpret report writing and the relevant techniques
- Know about the different types of reports and their features

5.2 COMPUTER APPLICATIONS

Today, computers are used in all possible fields and for various purposes. Every sector, be it education, commerce, management, industry or communications, relies on computers for its smooth functioning. Even if an individual is not directly involved with the functioning of computers, his / her life is affected by them in everyday life and work.

Computers are not only used in numeric applications like carrying out complex research and data analysis, but also for non-numeric uses like assisting in teaching and learning processes, providing a large databank of information, handling payrolls, record keeping, financial forecasting, making clinical diagnosis, providing entertainment like playing games, watching movies and listening to music, besides sports viewing.

Computers are used in applications ranging from running a farm to monitoring all environmental effects. Computers have made the research and the development of various diagnosis and prevention methods, easy and cost-effective. For example, a computer can accurately provide better forecasts when and where an earthquake or tsunami is likely to occur, the effects of drugs on the human system, dissections for study purposes, creating three-dimensional models for buildings or airplanes, and a host of other applications in media that have brought about a revolution in communications.

5.3 THE COMPUTER SYSTEM

The computer is a programmable machine that receives data, analyses it and gives a useful output. A program is a set of instructions given to the computer for a particular kind of functioning. New and better programs are being written everyday for easier and greater uses. The computer system includes not only the hardware with an operating system, but also the software that is necessary to make the computer function. Complex programs are written for handling a variety of tasks.

Every computer system requires an operating system. An operating system is a software program that enables the computer hardware to communicate and

NOTES

operate with the computer software. The following diagram illustrates how the computer system works:

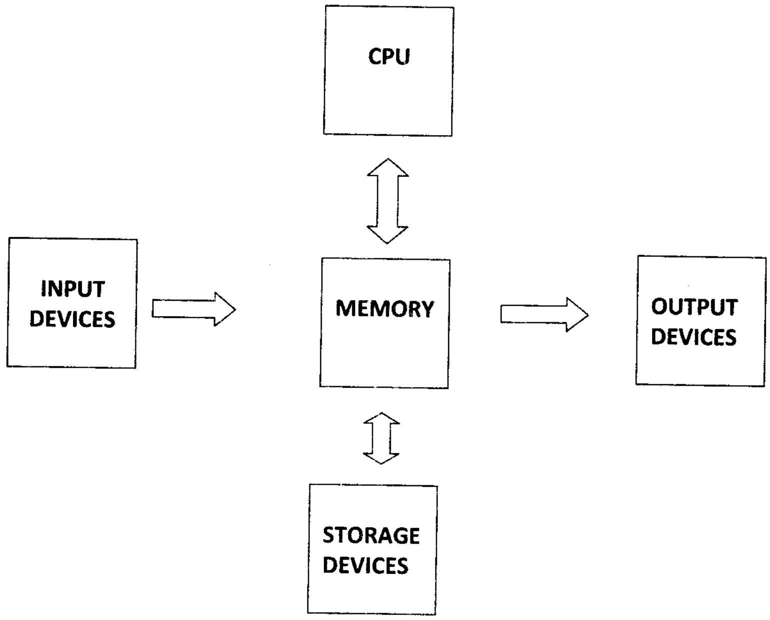


Fig. 5.1 Computer Architecture

CPU: The central processing unit or CPU is responsible for running programs and applications on the computer. The CPU has three segments: 1. Internal storage, 2. Control unit, and 3. Arithmetic logic unit

A computer program is written into the internal storage and then transmitted to the control unit. This data is available for processing by the arithmetic logical unit, which conveys the results back to the internal storage, and thus the output is obtained from the internal stage of the CPU.

The four primary components of a computer system are:

Input: Input devices send data and instructions to the CPU. They translate the data into binary language which the CPU understands.

Output: Output devices make processed data available to the user. The output devices reiterate the binary characters into familiar language characters.

Processing: CPU executes instructions supplied by the program.

Secondary storage: Primary memory temporarily holds limited amounts of data and programs whereas secondary storage devices permanently store data and programs of great amounts.

The four essential hardware components of a computer system are: input, processing, output and storage systems. Physical components of computer system, including the electronic and mechanical parts, are called computer hardware. The intangible components like data and programs are called computer software.

Software components are stored on hardware components like cache in CPU, hard disks, CDs, DVDs, tape, pen drives.

5.3.1 Important Characteristics

NOTES

- **Speed:** The computer is a very high speed electronic device. Computer can perform multiple operations on data in seconds, depending on the speed with which it operates. Their speed is usually measured in megahertz (MHz). Different computers have different speeds.
- **Diligence:** Being a machine, the computer does not suffer from fatigue, boredom or lack of concentration. It can work continuously for hours without creating any error. It performs its operations with the same accuracy and speed throughout.
- **Storage**: A computer's secondary (external) storage can hold large amounts of data and programs which can be retrieved as and when required can be stored for further use. The stored data is available any time for processing. A cache is a storage device inside the CPU (internal) which can store small amounts of essential data for temporary use.
- **Accuracy:** The computer's accuracy is consistently high. Output is totally dependent on the given instructions and input data. If the data provided is right, the output will be accurate. In the event an error occurs, the error detecting techniques are so efficient that it seldom leads to false results.
- **Automation:** A computer can perform automatically once the program is in the computer's memory. It controls different devices attached to it and executes the program instructions one by one.
- **Binary digits:** All the instructions provided and data stored in computers are in the binary number format which consists only two digits 1 and 0. This makes everything simple, accurate and fast.
- **Arithmetical and logical operations:** Arithmetic operations include performing addition, subtraction, multiplication and division on the numerical data. Logical operations compare the numerical and also alphabetical data.
- **Reliability:** The extremely low failure rate of modern day computers has made their results consistent. In general they are reliable.
- **Retrieving data and programs:** The stored data and programs can be retrieved very quickly for further processing.
- **Versatility:** Computers can perform different kinds of tasks. One side you can play games, on other side you can send emails, etc.
- **Communications:** Today computers are emerging as most powerful mean of communications. You can communicate via computers using emails via the LAN, WAN and Internet. You can also talk through Internet phone like Google Talk.

CHECK YOUR PROGRESS

1. List the various uses of a computer.
2. Give an example of binary to decimal conversion.
3. Give an example of decimal to binary conversion.
4. Give one important characteristic of a computer?
5. What is the function of CPU?

5.4 THE BINARY NUMBER SYSTEM

Binary means having two enitities. In binary mathematics system there are only two digits 0 and 1. In this sytem, all the numbers, alphabets, instructions, etc are expressed in terms of 0s and 1s. On the other hand, in decimal system there are ten possible digits, 0 to 9. Since computers use only binary system, the operations are simple, quicker and accurate.

We all know that computers can calculate complex equations and perform complex mathematics at lightning speed. Although a computer will only process 1s and 0s, there comes a point when the **1s** and **0s** have to be converted into the usual decimal numbers that we all are familiar with.

Let us consider a number 1234:

Thousands	Hundreds	Tens	Ones
1	2	3	4

Which means,

$$1234 = 1 \times 1000 + 2 \times 100 + 3 \times 10 + 4 \times 1$$

The binary system operates with base 2 or radix 2 (*bi* is the Latin prefix for two) or it uses 0s and 1s to represent numbers. The simple comparison between decimal and binary are given in the following table:

	Thousands	Hundreds	Tens	Ones
Decimal	10^3	10^2	10^1	10^0
Binary	2^3	2^2	2^1	2^0

In base 10, we put the digits 0-9 in columns 10^0, 10^1, 10^2, and so on. To put a number that is greater than 9 into 10^n, we must add to 10^(n+1). For example, adding 10 to column 10^0 requires us to add 1 to the column 10^1.

In base 2, we put the digits 0-1 in columns 2^0, 2^1, 2^3, and so on. To put a number that is greater than 1 into 2^n, we must add to 2^(n+1). For example, adding 3 to column 2^0 requires us to add 1 to the column 2^1.

5.4.1 Decimal to Binary Conversion

A positive decimal integer can be converted into an equivalent binary form by dividing the decimal number repeatedly by 2.

Start by dividing the given decimal integer by 2.

Let r1 be the remainder and q1 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}....r_1

(rk and r1 are the last and the first remainders respectively, obtained by the division process.)

Illustration 5.1: Finding the binary equivalents of 26 and 45.

Solution:

Number to be divided by 2	Quotient	Remainder
26	13	0
13	6	1
6	3	0
3	1	1
1	0	1

26 (decimal) = 11010 (binary)

Similarly, we can find the binary equivalent of 45 as illustrated below:

Number to be divided by 2	Quotient	Remainder
45	22	1
22	11	0
11	5	1
5	2	1
2	1	0
1	0	1

Thus the binary equivalent of 45 is 101101

NOTES

Here are some decimal numbers represented in binary.

Decimal	Binary
1	1
2	10
3	11
4	100
5	101
6	110
7	111
8	1000
9	1001
10	1010

5.4.2 Binary to Decimal Conversion

This particular method is known as the double-babble method.

To convert a binary number into decimal number, double the leftmost bit of the given binary number and add to it the bit to its right.

Then, double the sum and add to it the third bit from the left.

Proceed in this manner till all the bits have been considered.

The final sum obtained by repeated doubling and adding is the desired decimal equivalent.

Illustration 5.2: Converting the binary number 1101 into its decimal equivalent using the method given above.

Solution:

1. Doubling the leftmost bit we get 2
2. Adding the bit on its right 1, we get $2+1 = 3$
3. Doubling again this number obtained (3), we get 6
4. Adding to it the next bit 0, we get $6 + 0 = 6$
5. Again doubling we get 12
6. Finally adding the last bit (1), we get $12 + 1 = 13$

The decimal equivalent of binary 1101 is 13.

5.4.3 Computations in Binary System

Binary addition: binary addition is just like decimal addition. It follows as:

$0 + 0 = 0$

$0 + 1 = 1$

$1 + 0 = 1$

$1 + 1 = 10$

Sum of 1 and 1 is written as '10' (0 as sum and 1 as carry), which is the equivalent of decimal digit '2'.

Illustration 5.3: Adding the binary numbers 1010 and 101.

Solution: Add 1010 and 101

Binary	Decimal equivalent
1010	(10)
+ 101	+ (5)
= 1111	= 15

The computer performs all the other arithmetic operations, namely multiplication, subtraction, etc. by treating it as repeated addition a form of addition. For example, the multiplication 6*8 is evaluated by repeated addition with the necessary carryovers, 8+8+8+8+8+8. Though this process seems to be a longer method, the computer is well-suited to carry out these operations at great speed.

Converting Binary Fractions into Decimal Fractions

In decimal system, we use a decimal point to separate the whole number and the decimal fraction parts of a given decimal number. In the same way, we can use a binary point to separate the whole and fractional parts in binary numbers. The binary fraction can be converted into decimal fractions as shown in Illustration 4:

Illustration 5.4: Converting the binary fraction 0.101 into its equivalent decimal form.

Solution:

$$\begin{aligned} 0.101 \text{ (binary)} &= (1)\,1/2 + (0)1/4 + (1)1/8 \\ &= 0.5 + 0.0 + 0.125 \\ &= 0.625 \text{ (decimal)} \end{aligned}$$

To convert the decimal fraction to binary fraction, the following rules are applied:

1. Multiply the decimal fraction repeatedly by 2. The whole number part of the first 1 or 0 of the binary fraction;
2. The fractional part of the result is carried over and multiplied by 2;
3. The whole number part of the result gives the second 1 or 0 and so on.

Illustration 5.5: Convert 0.625 into its equivalent binary fraction.

Solution: Let us convert 0.625 into its equivalent binary fraction.

$$\begin{aligned} 0.625*2 &= 1.250 - 1 = 0.250 \\ 0.250*2 &= 0.500 - 0 = 0.500 \\ 0.500*2 &= 1.000 - 1 = 000 \end{aligned}$$

Hence, 0.101 is the required binary equivalent.

We can now see how computer arithmetic is based on addition. While performing such calculations, the number of individual steps increases enormously because all computer arithmetic is reduced to addition. However, since the computer is capable of carrying out millions of binary additions per second, it will not be a disadvantage.

5.5 COMPUTERS AND RESEARCHERS

The computer has emerged as one of the most useful research tools in modern times. It does a great variety of jobs with tremendous speed and efficiency. Computer usage has become a subject of study at all the schools and today computer is an indispensable part of any profession.

Computers have become highly useful tools in the research process, particular when the data is large and complex involving complicated mathematics and statistics. Researchers often deal with huge amounts of data, needing timely storage and quick retrieval when required and data processing using various techniques. Apart from speeding the research process, computers have also added quality to the research process. Computers assist the researchers throughout the different phases of the research process.

Phases of research process:

1. Conceptual phase
2. Design and planning phase
3. Empirical phase
4. Analytic phase
5. Dissemination phase
6. Concluding phase

1. **Role of computers in conceptual phase:** Conceptual phase is when the researcher formulates the research problem, reviews literature and formulates the hypothesis.

   Computers help in searching for the literature through the Internet from databases stored in servers all over the world. The desired information can then be downloaded and stored in the computer for future use. It thus saves the time spent on visiting libraries and collecting the data by writing down the relevant material.

2. **Role of computers in design and planning phase:** Design and planning phase consist of research design, defining the population, identifying research variables, determining sampling plan, reviewing research plan and performing pilot study.

   Several software programs like NCSS-PASS-GESS are available to calculate the sample size required for a proposed study. The standard

NOTES

deviation of the data from the pilot study is required for the sample size calculation.

3. **Role of computers in empirical phase:** This phase consists of collecting and preparing the data for analysis. The data obtained is stored in computers. This can be edited, represented graphically thereby helping the researcher with data entry, data editing and data management.
4. **Role of computers in data analysis:** This phase consists of statistical analysis of the data and interpretation of the results. Many programs are available to perform advanced calculations using statistical methods. They can be used to calculate the sample size, test the hypothesis and calculate. Computers not only analyse but also monitor the accuracy and completeness of the data. Computation of the 'mean,' 'standard deviation,' 'correlation coefficients,' factor analysis and multiple non-parametric analysis are just a few of the statistical calculations that the computers can perform with ease and accuracy, besides speed.

A note of caution

It is wise to remember that a computer is just a tool and a resource. Computers can only compute but cannot think. Computer-based analysis is not usually economical in case of small research projects. Certain important details which are saved in the computer may get lost.

Computer functions on the basis of the program written. Adequate planning and suitable organization is necessary for the efficient handling of data.

To sum up, computers make the research process easier, faster, accurate, reliable and less tedious. However, the experimenter should have thorough knowledge about the abilities and the limitations of the software used for the better use of computers.

The use of computers enables the researcher to use trial and error processes which involve a lot of calculations and repetitive work. Not only does it produce the results rapidly but the different options are also made available to researchers. These would otherwise take days or months.

Researchers must be familiar with data organization and coding, storing the data in the computer, selection of appropriate statistical techniques, selection of appropriate software packages and execution of the computer program.

5.6 INTERPRETATION AND REPORT WRITING

Once the study has been completed, the analysis is carried out and the conclusions drawn, the report of the entire study and the findings have to be written. It is only through the process of interpretation that the researcher can tell the scientific community about the relationships that were studied and the theoretical concepts underlying the findings.

The types of analysis that the data was subjected to and the inferences that were drawn also need to be readily communicated in the report writing. Other researchers look forward to such an endeavour from the investigator.

The meaning of writing a report, the importance of interpretation, the techniques of interpretation and the cautions that need to be exercised, are presented in this section. The significance of scientific writing as well as the steps involved in report writing is suggested. This is followed by the layout plan of a research report.

5.6.1 Meaning of Interpretation

The purpose of writing a research report is to tell the scientific community what the study was, how it was done, the results obtained and the conclusions drawn. All interpretations have two major functions: 1. To establish the findings of the current study and to show its links with other prior studies in the related field. This is to help in the continuity of research; and 2. To provide explanations in terms of relationships studied as suggested by the data obtained from the study. These twin attempts have the goal of advancing knowledge and stimulating further research.

5.6.2 The Need And Importance Of Interpreting The Findings

Interpretation provides a satisfactory explanation to the findings and the usefulness of these for any research. It is through this process that the abstract principles underlying any findings can be clearly understood. Interpretation serves as a framework for linking the current study with other studies in the same field.

Interpretations provide the impetus for further future research. It is also a means for stimulating new avenues in the quest for knowledge. Interpretations also throw out the real significance of the findings of the study. In exploratory studies interpretations become the basis for hypothesis formulation for experimental research. These are termed post-factum interpretations.

All these indicate the need for interpretation of findings following research studies.

Interpretation means drawing inferences from the research data. It establishes continuity in research studies. It acts as a guide for further research studies and explains the concepts that have been observed. It is a skill that can be best learned through practice and experience. Researcher must explain his concepts with the help of theory and keep extraneous information and all relevant factors in consideration while interpreting research studies. The researcher must try and make sure that the source of data is trustworthy and accurate, and that the data reflects good homogeneity. It is important to use proper statistical methods for analysis. It is important to be aware and take precautions against all possible sources of error. The researcher also has to keep a check on his tendency to over-generalize or be biased about the results. Also there should be constant interaction between initial hypothesis, empirical observation and theoretical conceptions.

NOTES

There are certain steps that a researcher should follow while writing a report. Firstly, the material should be analysed either logically or chronologically. After this, an outline should be prepared of the contents that are to be included in the report. After this, a rough draft containing the procedure adopted to collect data, technique of analysis, findings and generalizations is put down. This is then revised, polished and rewritten to make it more clear and comprehensive for the reader. After this step, the bibliography is prepared in an alphabetical manner.

Basically the report can be divided into three broad categories.

(i) **Preliminary pages:** This section carries the title, date, acknowledgements, preface, contents, and a list of tables and illustrations used throughout the report.

(ii) **Main text:** The main text includes the introduction, statement of findings and recommendations, results, implications of the result and summary.

(iii) **End matter:** Here the appendices are listed along with the bibliography of all sources consulted or relevant to the research study.

Research reports vary in length and format. Research reports can be in the form of a technical report, popular report, an article, a manuscript or even an oral presentation.

A technical report is a detailed report required for record keeping or public dissemination. It details the methods used, assumptions made during the course of the study and focuses on detailed presentations of the findings including their limitations and supporting data. It should be substantiated with charts and graphs so that it is clear and easily understood.

A popular report is often simple and attractive. It does not use too many mathematical or technical terms; instead it uses charts and diagrams, large fonts, many subheadings and visuals which make it easy to read. It emphasizes on practical findings and their implications.

Along with the oral presentation, there should also be a written document for record keeping and can also be circulated before the presentation to acquaint the audience. It should be supplemented with visuals, slides, wall charts and blackboards. A brief outline should also be given.

All these factors highlight the need for interpretation of findings of the research studies.

5.6.3 Techniques of Interpretation

The task of interpreting the findings of a study are demanding and involve considerable skills acquired through exposure and practice. It is appropriate to seek expert advice as and when needed to interpret one's findings. The techniques involved in the process of interpretation have several steps as list below:

- All relationships that have been found in the course of the study must be explained adequately. Any common link that ties the findings together must be identified, clearly. This is the beginning of the process of generalization, following a study;

- Data that have been thrown up by the study incidentally or accidentally also needs to be given sufficient importance while interpreting the findings that one was studying. These bits of information may have far reaching consequences for the current study, as well as those in the future.
- One can always open the study to an expert or a practitioner in the given field to act as the devil's advocate, so that any omissions or fallacies in the logic of interpretation can be detected. This would enable correction opportunities.
- All facts must be obtained and collated in the most suitable manner before interpretations are carried out. This would rule out the chances of hasty or premature conclusions.

These techniques are suggested to ensure truly valid and generalizable conclusions from the study.

5.6.4 Precautions during Interpretation

Merely carrying out a well-designed study and obtaining properly collected data, does not preclude the possibility of erroneous interpretations. To prevent drawing of wrong conclusions; a few precautionary steps are indicated and suggestions for correct interpretations are given below:

Some of the precautions to be exercised are:

- Ensure that the data are reliable and valid;
- Ascertain the homogeneity of the data, collected.
- Make sure that the data have been subjected to proper statistical analysis.
- Double check to see if the interpretation of the statistical significance of data is well within the prescribed range.
- It is important to note that correlational data are not only statistically significant, but also in terms of the meaningfulness of the study results.
- Where causal effects are determined, it is absolutely necessary to ensure that causation has indeed occurred. These are critical for avoiding false conclusions and generalizations.
- A second round of computational checks, verification of significance need to be carried out to prevent errors from creeping in.
- Restrict the generalizations and interpretations to the study undertaken and never over-extend. Limiting the scope of the interpretations of one's findings is a safe option to prevent inaccurate conclusions.
- All interpretations must ultimately go back to the starting point of the study in terms of the research problem that is being investigated. Findings must always be made keeping with this connection in mind. The theoretical concepts, the empirical observations and the analysis carried out on the

data obtained should ideally be in synchrony with the findings and interpretations offered.

These alone would ensure error free interpretations

NOTES

5.6.5 Steps In Report Writing

Since research is a demanding and painstaking effort, sufficient time and attention should be devoted to writing the report of the work carried out. Some of the steps involved in the planning of writing a good report are indicated below for the benefit of researchers:

- The entire study should be analyzed very logically in a thorough manner.
- A rough-or preliminary draft of the outline of the proposed research, should be prepared with thought and care;
- Then the final outline should be developed.
- This can then be polished and refined, before being re-written.
- This can be followed by the compiling of the bibliography.
- Then the final draft can be written.

The logical analysis process

This can be carried out in one of two ways:

(i) Advance the subject of the study logically. This means one topic follows the other, rationally.

(ii) Chronological: This involves writing a report on the basis of the sequence of happenings.

Sometimes, a combination of these two strategies is useful.

Preparation of the rough-draft: This is the step of putting down what one did in the study, in the order in which it was carried out. This could include the motivations, the insights, the struggles and the limitations. Also, the process of designing the study, the method of data collection, the controls and variables should be utilized and the broad findings and directions of generalizations possible should be identified.

Preparation of the final outline: This step involves careful reporting of the study along with the theoretical framework for the study. List all the procedures followed

The preparation of the bibliography: This refers to all the books, journals, magazines, reports and other sources that have been consulted for undertaking the study. The format for listing these references is to be written in keeping with the style of 'Manual of the American Psychological Association'.

5.6.6 The Layout of the Report

This is the structure or format of a research report. There is a standardized convention that has been universally accepted for the purpose of writing a research report. It is clearly laid down in the publication, 'Manual of the American Psychological Association.' This manual also gives suggestions on how to present concepts with economy and clarity. The APA manual has divided research report writing into eight parts.

APA format for writing reports

(i) **Title page:** This page has three main sections:

a. **Title and running head.** The title is the problem investigated by the study while the running head is a short description of the Title. It can have a maximum of 50 characters;

b. **Author's name and affiliation:** The author(s) name (s) to be presented on this page, starting with the principal investigator and the designation and the institutional affiliation (if any). The other authors of the study should then be listed in the order of their importance to the contribution. Their designations and affiliations can be given suitably.

c. **Acknowledgements:** This can be third part of the title. Any form of support in terms of finances or institutional help, comments, secretarial / editorial assistance etc, can be mentioned here.

(ii) **Abstract:** This is a synopsis of the research in full detail, but in an abridged form. The ideal length would be about 150 words. It should include the problem studied, the methods used, the research design adopted (any apparatus, if used), the statistical treatment, the findings and the conclusions, as well as its applications / implications.

(iii) **Introduction:** This is the beginning of the report and thereafter the entire report follows a sequence. An introduction is not written simply as an introduction. It has three divisions in the report writing.

a. The first part is the statement of the problem that was studied. A part of this section needs to include the motivation for studying the particular problem and the theoretical or practical orientations underlying the study undertaken.

b. The second part of this section should contain a description of the review of the previous literature available in the field. This should refer to the studies related to the field of the present inquiry. A good rationale and a logical relationship between the earlier studies and the present investigation needs to be developed.

c. The third part of this section is devoted to the formulation of the hypothesis, for the study. This should include the operational definition of the hypothesis that is proposed. The hypothesis is to be clearly stated and the scientific procedures involved in testing it need to be specified. The independent and dependent variables should be identified and the design is shown in a diagram, if possible.

(iv) Method: This constitutes the main body of the report to be written. It includes a very precise account of how the study was conducted. This is the

NOTES

procedural part of the report. It is expected that anyone who reads the report and wishes to undertake a similar study should be able to do so, by following the method's section closely. There are several sub-sections that make up this section. These help in bringing to focus different parts of the method:

a. *The sample:* This requires that the sample be clearly specified, along with the type of sampling procedure used. Special characteristics like age, sex, etc., that are relevant to the study needs to be identified precisely. The total number of subjects included for the study is very significant. The different treatments should be clearly described. The control conditions should be explained carefully.
b. *Equipment used:* If any special apparatus was used for the study, the model, the manufacturing company and other details need to be given. If a photograph of the equipment can be given, it would be helpful.
c. *The design:* Every study has a blueprint on the basis of which any study is proposed to be carried out. This is the design of the study. Here, the entire procedure of the conduction of the study, along with the experimental and control conditions, need to be elaborately presented.

The steps in the introduction of the independent variable have to be dealt with at length and clearly. So, also the controls that have been used in the study, the nature and type of measurement of the dependent variable should be clearly mentioned. In short, the entire procedural information has to be given, so that its replication is possible by another interested researchers.

(v) **Results:** This refers to providing the data obtained from the study. If the raw data is significant for being presented, it is to be given in a meaningful tabulated way. Otherwise, the results are to be reported in terms of the hypothesis studied. The statistical significance of the results needs to be given. Whether the hypothesis is supported or rejected should be clearly mentioned. Tables, charts, diagrams, figures, etc. are to be presented in the most suitable manner.

The major findings can be summarized and presented. Wherever the texts can be supported by visual material, it can be done, to make for greater emphasis. The statistical treatment carried about and the levels of significance for the data should be clearly stated.

(vi) **Discussion of results:** The purpose of this section is to interpret the results of the study. It should have a detailed account of the study related to other studies in the field. The discussion should state whether the hypothesis is supported or rejected. The supported hypothesis, as well as the rejected one, both need to be explained on the basis of some theoretical assumptions. Any new hypothesis can be suggested when the results are not along the predicted lines. Any faults in the formulated hypothesis can be modified on the basis of the results obtained. These can be discussed suitably. Detailed discussion about the findings can be presented. Finally, any suggestions regarding the ways in which the problem that has been studied, need to be resolved, can be offered. Any implications of the findings can be indicated. Also suggestions for future research should be mentioned. A small paragraph could be devoted to the limitations of the study and details about overcoming some of these can also be mentioned.

(vii) References: This includes all books, journals, articles, reports, internet links for references, etc., have to be listed in alphabetical order. The citations should be presented in accordance with the manual of the American Psychological Association.

(viii) Appendix: This can be included where tests, questionnaires, any statistical treatment, computer program, etc, were used for the study but do not fit into the report due to length and other considerations.

Style of writing report

Reports have to be clear, concise, complete and easy to read. The object of reporting a study is to help others who read it to be able to follow the procedure and carry out a similar activity. The report should be clear in description and explanation. The reference works must mention authors by last name, only. Abbreviations are to be avoided, as a rule. Only the very well known abbreviations such as IQ can be used. Detailed formulae and computations can be eliminated.

The report has to be written in the past tense. All sources must be acknowledged, correctly and completely. Certain permissible and well known acronyms alone are allowed to be used in report writing, e.g., STM (Short Term Memory). This can also figure in the midst of a sentence, but not at the beginning. Ideas must be paraphrased and written. As a rule, whenever quotes are used, the relevant study should be cited.

As for as possible, use only the accepted terminology. Numbers that come in the report which are from zero to nine have to be written in words, while those from 10 or larger can be written in digits.

5.6.7 Types of Reports

Reports vary in length, format and purpose. The research report presented here is one of a scientific study. Other forms of reports are:

(i) Business reports are largely in the form of a letter. They are short, precise and factual in detail.

(ii) Banks and other financial institutions present their reports in the form of a balance sheet. This is largely a statement of accounts for their customers and shareholders.

(iii) Chemists present their reports in terms of formulae and other symbols to explain the preparation, etc.

(iv) Manufacturing companies prepare reports in the form of manuals of operation, servicing, spares, etc.

(v) Students of literature, linguistics, philosophy and others write reports that critically analyse a choice subject matter. Here quotations and descriptions are common.

(vi) Newspapers, magazines and other media personnel present items or articles in the form of reports. These include first hand account of events or opinions on a topic. Sometimes interviews of experts in a given field are reported on a theme. This kind of report is largely fact-based, with some details to explain the happenings. Some of these reports have eye-witness accounts also. The implications of these events form the later parts of the report writing.

(vii) Book-review reports analyse the content, style, format, language and other details of a book, so as to inform the readers about the book. This is often a short report.

(viii) Government reports / reports of commissions are very elaborate reports that are comprehensive in nature. These reports are the result of a detailed inquiry carried out on a topic of interest, nationwide or about a specific concern / matter.

(ix) Case history reports submitted by doctors and other healthcare personnel involve in-depth studies of cases. They are based on diagnostic findings and include plans of treatment.

Research result presentation styles

There are two ways in which the findings of a research study can be presented: (i) A technical report and (ii) A popular report.

(i) **Technical report:** This is a very precise and detailed report of the problem studied, the methods used and the results obtained, along with some limitations. It is 'how it was done' report. The findings of the study are crucial in such a report. Therefore, it begins by stating the results of the study. The format is as follows.

a. *Summary of the results:* A statement of the main findings. This can be in some detail as the readers are eager to know the results.

b. *Nature of the study:* This has the general description of the study, the operational definitions involved the hypothesis formulated, the data to be collected and the type of analysis to follow.

c. *Methods:* This section includes steps like sampling design, sample selection sample size and other similar details. Any study design if applicable.

d. *Data:* How the data was collected. The sources from where the data was obtained and other related information are included under this head.

e. *Data analysis and findings:* The statistical treatments to which the data were subjected to are to be described. The findings are to be presented in tables, clarets, diagrams etc to be given here, for easy understanding.

f. *Conclusions:* The detailed summary of the findings and the implications of the results, are to be presented.

g. *References:* The different sources which are cited and consulted for the study, to be mentioned in complete, detail.

h. *Appendix:* Questionnaires, lists, formulae and other such information to be provided in a handy manner.

The entire report must be precise, readable and simple in style. Charts, diagrams and other supporting material should be presented wherever suitable.

(ii) **Popular report:** The aim of such a report is to familiarize the average reader with some important scientific studies and their findings. So, the language should be simple, direct and familiar to the reader. So avoiding technical jargon, as far as possible, is suggested here. Practical applications of the findings of the study will have important place in such a report.

The following is the suggested outline, for such a report

a. *The findings results and their implications:* Since the focus of this report is on its usefulness; the priority lies in the sub-heading taking the lead.

b. *Findings that can be used:* Here the suggestion is how the study conclusions can be used readily. This is an action-oriented detail.

c. *The purpose of the study*: The reader can be oriented to the original reason for undertaking the study and the specific details involved.

d. *Methods used*: This should be a non-technical description of the methods used for the study as well as some details about how the sample was chosen and the data collected.

e. *Results:* This is of main interest to the reader. So the complete details of the findings along with the best possible illustration of the results in charts, diagrams, etc., are to be presented here. The language must be as non-technical as possible.

f. *Appendix:* All types of tests, questionnaires, and other details that are of interest to the public is to given here.

Make the report simple, easy to read and understand apart from an appealing presentation.

5.6.8 Oral Presentation of Reports

Oral presentations are dynamic and interactive ways of communicating research findings. At times, oral presentations are to be made to those who wish to take

NOTES

decisions on the basis of the findings of a study. Policy makers are such audiences. They wish to see and hear the results of a study for a quick understanding and easy decision making. It is always necessary for a written report to be circulated, before the oral presentation is made.

Key to effective oral presentations:

- Have liberal use of visual aids
- Use colour generously to highlight features
- Have pointed effective captions and titles with the slides.
- Illustrate with graphics wherever possible.
- Have bold fonts to show conclusions, implications, etc.

How to retain listener/viewer attention in an oral presentation:

(i) Begin with an anecdote, experience or an illustration

(ii) Then present the key aspects of the study

(iii) Identify 3 to 5 major points to be made during the talk

(iv) Connect ideas, logically and proceed

(v) Give examples as you proceed

(vi) Make eye-contact with the audience, always

(vii) Modulate your voice through the talk

(viii) Use body language to good effect

(ix) Be mindful of the time allotted

(x) End on a strong note (leave the audience with a thought or an application of the results)

(xi) Always, be available for taking questions

(xii) Never forget to add a dash of humour in your presentation

(xiii) Always have a speaking outline in front of you and the rest can be on the projection screen.

5.7 SUMMARY

- Computer is a programmable machine that receives input, stores and manipulates data and provides output in a useful format.
- Computers are not only a necessity but they have become an indispensable part of research process.
- The four primary components of a computer system are: (i) Input devices, e.g., keypad, which send data to the central processing unit (CPU); (ii) Output devices, like monitor, which provide the processed information (or results); (iii) Processor, e.g., CPU, which executes computer instructions;

NOTES

and 4. Secondary storage like hard disk, CD, DVD and pen-drive where the information is stored for further use.

- Computers are characterized by high speed, diligence, accuracy, automation, storage capability, reliability and versatility.
- Computer calculations are based on the binary system unlike the decimal system that we use. The binary system uses only two digits, 1s and 0s. Decimal numbers can be converted to binary and vice versa. Broadly, all computer computations are performed using addition in different forms.
- A computer assists the research throughout different phases of research process. It plays a different role at every stage—conceptual phase, design and planning phase, empirical phase, analytic phase and dissemination phase.
- The computer is after all a man-made machine and will only work according to the software that has been fed in the system.
- Interpretation means drawing inferences from the collected and processed research data. It establishes continuity in research studies. It acts as a guide for further research studies and explains the concepts that have been observed. It is a skill that can be best learned through practice and experience.
- A researcher must explain the use of concepts with the help some theory if possible. He should keep out extraneous information but include relevant factors while interpreting research studies. The researchers must try and make sure that the source of the collected data is trustworthy and accurate and that the data reflects good homogeneity. It is important to use proper statistical methods for analysis. It is important to be aware of and take precautions against all possible sources of error.
- The researchers also must keep a watch over the tendency to over-generalize or be biased about the results.
- Basically, a research report can be divided into three broad sections: (i) Preliminary pages—this section carries the title, date, acknowledgements, preface, contents, and a list of tables and illustrations used throughout the report; (ii) Main text—the main text includes the introduction, statement of findings and recommendations, results, implications of the results and a summary; and (iii) Closing pages—here the appendices are listed along with the bibliography of all sources consulted or relevant to the research study.
- The lay-out is the structure or format of a research report. It is based on the universally accepted format provided by the American Psychological Association manual. They have suggested eight sections under which every research report has to be written. These are (i) Title, (ii) Abstract, (iii) Introduction, (iv) Method, (v) Result, (vi) Discussion, (vii) References and (viii) Appendix (if suitable)

- Research reports vary in length and format. Research reports can be in the form of a technical report, popular report, an article, a manuscript or even an oral presentation.
- An oral presentation is made for policy makers who wish to have easy and quick access to results of the policy.

5.8 KEY TERMS

- **Binary system:** A system used by a computer for all its internal communications.
- **CPU:** The heart of a computer system which has arithmetical logic system, a cache—that stores temporary memory and control unit.
- **Technical report:** A detailed report required for record keeping or public dissemination.
- **Popular report:** A report that uses few mathematical or technical terms and more charts and diagrams, large fonts, subheadings and visuals.

5.9 ANSWERS TO 'CHECK YOUR PROGRESS'

1. A computer can be used for numerical calculations, assisting in teaching and learning, financial forecasts, weather forecasts, making clinical diagnosis, providing entertainment like playing games, watching movies and listening to music, etc.
2. Let us convert the binary number 110101 to its decimal equivalent.

   Doubling the leftmost bit we get 2
   Adding to it the bit on its right we get $2 + 1 = 3$

   Doubling again the number obtained we get 6
   Adding to it the next bit we get $6 + 0 = 6$

   Again doubling we get 12
   Adding the last bit we get $12 + 1 = 13$

   Doubling the no. obtained we get 26
   Adding the next bit we get $26 + 0 = 26$

   Doubling the no. obtained we get 52
   Adding the next bit we get $52 + 1 = 53$

   110101 (binary) = 53 (decimal)
3. Let us convert the decimal number 65 to binary number.

   Solution: table for conversion of 65 into its binary equivalent.

Number to be divided	Quotient	Remainder
65	32	1
32	16	0
16	8	0
8	4	0
4	2	0
2	1	0
1	0	1

1000001 (binary) = 65 (decimal)

4. The computer is capable of completing calculations at a very high speed.
5. The central processing unit is responsible for running programs and applications on the computer.

5.10 QUESTIONS AND EXERCISES

Short-Answer Questions

1. Define computer hardware and software.
2. Define the binary number system.
3. What is the central processing unit?
4. Name the different types of reports.

Long-Answer Questions

1. What is a computer? Elaborate on its characteristics and various uses.
2. Why is interpretation a fundamental component of research process? Explain.
3. Explain the techniques of interpretation and report writing.
4. Describe the layout of a research report, covering all important points.

5.11 FURTHER READING

Kothari, C.R. 1985. *Research Methodology: Methods & Techniques*. New Delhi: Wishwa Prakashan.

Singh, A.K. 2008. *Tests, Measurements and Research Methods in Behavioural Sciences*. New Delhi: Bharati Bhawan.

NOTES

NOTES

NOTES