Semester VI

		L	T	P	C
15CMU603B	**CORE ELECTIVE II**	**5**	**-**	**-**	**5**

RESEARCH METHODOLOGY

Learning Outcome

Research Methodology emphasis on the methodology of research and its application in managerial decision making. It encompasses various sampling techniques , source and analysis of data and techniques of report writing.

Course Objective

- Research Methods is emphasizing on the methodology of research and its application in managerial decision making,
- Understood the scope and significance of research in business decisions.
- Studied and understood sampling techniques along with hypothesis testing.

UNIT I

Research : Meaning –Objectives –Types of research –Significance of Research – Research Process –Criteria of Good Research –Identification of Research Problem – Research Design.

UNIT II

Sampling : Meaning –Steps in Sample Design – Characteristics of a Good Sample Design-Determination of Sample Size-Sampling Techniques-Probability and Non – Probability Sampling- Sampling Error.

UNIT III

Sources of Data: Methods of Data Collection –Primary Data-Interview Method-Observation Method-Questionnaire-Schedule-Secondary Data-Processing of Data-Editing-Coding-Classification –Tabulation.

UNIT IV

Analysis of Data: Hypothesis – Characteristics-Concepts of Hypothesis-Null Hypothesis-Alternative Hypothesis-Level of Significance-Test of Hypothesis-Type I and Type II error-Chi-square test-t test-F test-ANOVA-Scaling Techniques

UNIT V

Interpretation and Report Writing: Interpretation –Meaning-Technique of Interpretation , Precautions-Report Writing-Steps in Writing Report-Types of Reports-Technical and Popular Report-Oral Presentation –Precaution for Writing Research Reports.

Note: The question Paper shall cover 80% theory and 20% problem.

TEXT BOOK

T1. Kothari .**C**.**R**,(2014) Research Methodology-Methods and Techniques , New Age International (P) Limited, Publishers, New Delhi.

REFERENCES

R1:Gupta.S.P.(2014)**,** Statistical Methods and Techniques, S.Chand and Co., New Delhi.

R2: Anil Kumar Gupta (2011),Research Methodology:Methods and Techniques, Vayu Education of India, New Delhi.

R3: Krishnaswami and M.Ranganathan (2014), Methodology of research in social science , Himalaya Publishing House Pvt.Ltd., Mumbai.

KARPAGAM ACADEMY OF HIGHER EDUCATION

(Deemed to be University Established Under Section 3 of UGC Act 1956)

Coimbatore – 641 021.

LECTURE PLAN

DEPARTMENT OF COMMERCE

STAFF NAME: N.SUMATHI

SUBJECT NAME: RESEARCH METHODOLOGY SUB.CODE :15CMU603B

SEMESTER: VI CLASS : III B.COM

UNIT I

Sl.no.	Lecture Duration Period	Topics to be covered	Support Materials
1	1	Research Meaning, Definition Difference between Research and Research methodology	T: P.No:1-2
2	1	Scope of Research Objectives of Research	R1: P.No 1-8 T: P.No 2-3
3	1	Significance of research	R1: P.No 8-10
4	1	Types of research	T: P.No 2-6 R1: P.No 49-55
5	1	Characteristics of Research Qualities of good research	T: P.No 20-21
6	1	Process of research	T: P.No 10-19
7	1	Problems faced by the researchers in India	T: P.No 21-23
8	1	Scientific method of Research	R1: P.No 31-35
9	1	Research problem Identification and selection of research problem	T: P.No 24-26

10	1	Steps in research problem	T: P.No 27-28
11	1	Research Design- meaning Need and importance of research design	T: P. No: 31-32
12	1	Features of good research design	T: P. No: 32-33
13	1	Types of research design	T: P. No: 35-38 W1
14	1	Recapitulation and discussion of important questions	
		Total no. of hours planned for unit - I	**14**

UNIT-II

S.No	**Lecture Duration Period**	**Topics to be covered**	**Support Material**
1	1	Sampling – Introduction, meaning	T:55-56
2	1	Important concept relating to sampling, sources of sampling	R1:21.2
3	1	Steps in sample Design	T:57-58
4	1	Characteristics of good sample design	
5	1	determination of sample size	R1:21.10
6	1	Factors determining sample size	
7	1	Sample technique-Introduction, Meaning	T.P:58-60
8	1	Probability Sampling-Simple random sampling	
9	1	Stratified random Sampling	
10	1	Systematic sampling	
11	1	Cluster Sampling	
12	1	Non probability sampling-Judgment Sampling	T:60-63

13	1	Convenience Sampling	
14	1	Quota Sampling	
15	1	Sampling error –Introduction, Types	R1:21.12
16	1	Recapitulation and Discussion of important questions	
Total No of Hours planned for Unit- II			**16**

Unit III

S.No	Lecture Duration Period	Topics to be covered	Support Material
1	1	Source of data-Introduction	R1:39
2	1	Methods of data Collection	R1:40-42
3	1	Primary data-observation-Types	T:153-155
4	1	Observation-Merits and demerits	
5	1	Interview method – types, merits and demerits	T:100-102
6	1	Questionnaire method of data collection – merits and demerits,	R1:11.4 T:100-102
7	1	Mailed questionnaire	
8	1	Schedule Method of Data collection - merits and demerits	R1:44-47
9	1	Secondary Data-Introduction	R1:11-24
10	1	Processing of data	R1:52
11	1	Editing of data	R1:57
12	1	Coding of data	R1:58
13	1	Classification of data	R1:59

14	1	Tabulation of classified data	R1:91-94
15	1	Recapitulation and Discussion of important questions	
Total No of Hours planned for Unit- III			**15**

Unit IV

S.No	Lecture Duration Period	Topics to be covered	Support Material
1	1	Hypothesis Meaning and Types of hypothesis	T: 184 R1: 420 - 30
2	1	Sources of Hypothesis	T:185-192
3	1	Characteristics of hypothesis	
4	1	Basic concepts relating to hypothesis Null and Alternative	T:185-192
5	1	Procedures of hypothesis testing	T :31-33
6	1	Formulation of hypothesis	T : 35-38
7	1	Type I and Type II Error	T : 69-75 R1: P.No: 215-220
8	1	Chi-square Test	T:237-250
9	1	t-Test	T: P.No :198-204
10	1	F-Test	R1:953-972
11	1	ANOVA	R1:1011-1038
12	1	Scaling – meaning Importance of scaling	T: 76-78
11	1	scale classification bases	T: 77
13	1	Important scaling techniques	T: 82-85
14	1	Scale construction techniques	R1:226-246

15	1	Recapitulation and discussion of important questions	
		Total no. of hours planned for unit – IV	**15**

UNIT-V

Sl.no.	Lecture Duration Period	Topics to be covered	Support Materials
1	1	Interpretation -Meaning Need for Interpretation	T1: P.No :244-245
2	1	Techniques of Interpretation	T1: P.No :345
3	1	Precautions of Interpretation	T1: P.No :346-347
4	1	Report Writing Meaning Need for interpretation	T1: P.No :347-348
5	1	Significance of Report Writing Contents, Formats of research report	R1: P.No :436-439
6	1	Different steps in Report Writing Qualities of good research report	R1: P.No :439-445
7	1	Layout of research report	T1: P.No :349-350
8	1	Types of report: popular report, Technical report	T1: P.No :350-353
9	1	Mechanics of writing a research report	T1: P.No :353-359
10	1	Oral Presentation	
11	1	Precautions of writing research report	T1: P.No :400-410
12	1	Recapitulation and discussion of important questions	
13	1	Total no. of hours planned for unit - V	12
14	1	Discussion o previous ESE question papers	
15	1	Discussion o previous ESE question papers	
		Total no. of hours planned for unit – V & Previous ESE question paper discussion	**15**

TEXT BOOK

T1. Kothari .**C**.**R**,(2014) Research Methodology-Methods and Techniques , New Age International (P) Limited, Publishers, New Delhi.

REFERENCES

R1:Gupta.S.P.(2014)**,** Statistical Methods and Techniques, S.Chand and Co., New Delhi.

R2: Anil Kumar Gupta (2011),Research Methodology:Methods and Techniques, Vayu Education of India, New Delhi.

R3: Krishnaswami and M.Ranganathan (2014), Methodology of research in social science , Himalaya Publishing House Pvt.Ltd., Mumbai.

WEBSITE

W1:www.sciencebuddies.org

UNIT I

SYLLABUS

Research: Meaning –Objectives –Types of research –Significance of Research – Research Process –Criteria of Good Research –Identification of Research Problem – Research Design.

* * * * *

MEANING OF RESEARCH

Research is an endeavour to discover, develop and verify knowledge. It is an intellectual act that begins with the asking of questions and progressiveness through the critical examination of evidence that is both relevant and reliable to the revelation of truth. Research can be defined as the search for knowledge, or as any systematic investigation, with an open mind, to establish novel facts, solve new or existing problems, prove new ideas, or develop new theories, usually using a scientific method. The primary purpose of research is discovering, interpreting, and the development of methods and systems for the advancement of human knowledge on a wide variety of scientific matters of our world and the universe.

DEFINITIONS OF RESEARCH

Webster's New International Dictionary: "Research is careful critical enquiry or examination in seeking facts or principles, diligent investigation in order to ascertain something.

John W. Best: Research may be defined as the systematic and objective analysis and recording to controlled observations that may lead to the development of generalization, principles of theories resulting in prediction and possible ultimate control of events

Robert Ross: Research is essentially an investigation, a recording and an analysis of evidence too the purpose of gaining knowledge

Clifford Woody: Research comprises of defining and redefining problems, formulating hypothesis or suggested solutions, collecting, organizing and evaluating data making deduction and reaching conclusion and at last carefully testing conclusions to determine whether they fit in formulating hypothesis.

John Dewey: Research is considered to be the formal, systematic, intensive process of carrying on the scientific method of analysis. It involves a more systematic structure of investigation, usually in some sort of formal record of procedures and a report of result or conclusions.

Fred Kerlinger: Research is an organized enquiry designed and carried out to provide information for solving a problem.

Redman and Mory: Systematized effort to gain new knowledge

OBJECTIVES OF RESEARCH

- The purpose of research is to discover answers to questions through the application of scientific procedures. The main aim of research is to find out the truth which is hidden and which has not been discovered as yet.

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

TYPES OF RESEARCH

1) Descriptive Research

A descriptive study may be simple or complex. It determines who, what, where and how of a topic. It is concerned with describing the characteristics (e.g., the extent to which libraries are used) estimating the proportion of the people in a specified population who hold certain views or attitudes (e.g., how many favour the abolition of capital punishment?) predicting specifically (e.g., how may will cash their government bonds during a given period?) and discovering or testing whether certain variables are associated (e.g., people who spend a good deal of time for reading, go to movies often with each other)

Descriptive study may employ any of or all the methods of data collection such as interview, questionnaire, observation, tests and cumulative record cards. In the descriptive study the researcher must be careful to make a note of the bias and extravagance that may creep in at every stage of the study – formulating the objectives of the study;

designing the methods of data collection; selecting the sample; collecting, processing and analyzing the data; and reporting the findings.

2) Analytical Research

Analytical study makes use of available information by analyzing and doing critical evaluation. Analytical study makes use of higher level statistical tools which are not commonly used.

3) Applied Research

Applied research aims at finding a solution for an immediate problem faced by any business organization. This research deals with real life situations. Example: "Why have sales decreased during the last quarter"? Market research is an example of applied research. Applied research has a practical problem-solving emphasis. It brings out many new facts.

Examples:

1. Use of fibre glass body for cars instead of metal.
2. To develop a new market for the product.

4) Fundamental Research

This is otherwise known as basic research or fundamental research. Gathering knowledge for knowledge's sake is known as fundamental research. It does not have any commercial potential. It is not connected to any practical problem. e.g. Theory of Relativity. It is only for the enrichment of the knowledge.

5) Quantitative Research

Quantitative researches are based on the measurements of quantity or amounts. It means that these type of researches deals with items which are expressed in numbers.

6) Qualitative Research

Qualitative researches deals with the qualitative phenomena. i.e. anything which cannot be expressed in numerical terms. Motivation research is an example of qualitative research.

7) Conceptual Research

Conceptual research is that related to some abstract idea(s) or theory. It is generally used by philosophers and thinkers to develop new concepts or to reinterpret existing ones.

8) Empirical Research

Empirical research relies on experience or observation alone, often without due regard for system and theory. It is data-based research, coming up with conclusions which are capable of being verified by observation or experiment. We can also call it as experimental type of research. In such a research it is necessary to get at facts firsthand, at their source, and actively to go about doing certain things to stimulate the production of desired information. In such a research, the researcher must first provide himself with a working hypothesis or guess as to the probable results. He then works to get enough facts (data) to prove or disprove his hypothesis. He then sets up experimental designs which he thinks will manipulate the persons or the materials concerned so as to bring forth the desired information. Such research is thus characterized by the experimenter's control over the variables under study and his deliberate manipulation of one of them to study its effects. Empirical

research is appropriate when proof is sought that certain variables affect other variables in some way. Evidence gathered through experiments or empirical studies is today considered to be the most powerful support possible for a given hypothesis.

9) One-time research or Longitudinal Research

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

10) Field Method

Field study is a scientific enquiry aimed at discovering the relations and interactions among sociological, physiological and educational variables in real social structures and life situations like communities, schools, factories, organizations and institutions. Hence, it is called field study.

11) Exploratory Research

Explanatory research is carried, when the reason for a problem is not clear. In exploratory research, all possible reasons which are very obvious are eliminated, thereby directing the research to proceed further with limited options.

Example for Exploratory Research

Sales decline in a company may be due to:

- Inefficient service
- Improper price
- Inefficient sales force
- Ineffective promotion

- Improper quality

12) Formalized Research

Formalized research studies are those with substantial structure and with specific hypotheses to be tested.

13) Historical Research

This research is the induction of principles through research into the past and social forces which have shaped the present. Its aim is to apply reflective thinking to unsolved social problems by discovering past trends of events, facts and attitudes, and by tracing lines of development in human thought and action.

14) Decision Oriented Research

Decision-oriented research is always for the need of a decision maker and the researcher in this case is not free to embark upon research according to his own inclination.

15) Individual and Group Research

The research undertaken by an individual is called individual research. The bulk of research activities in universities, and colleges are made by the individual. The individual research is done on the basis of one's own judgement, interest and capacity.

Group research is undertaken by several researchers. Their activities are coordinated by a director, Research conducted by a firm, trade association and government agency is performed by a team of researchers under a project director. Research in colleges and universities financed by grants is done on a group basis.

16) Operations Research

This method of research has been done for solving problems by using scientific methods and quantitative techniques. While the researchers care to study the development of methods, the industrial operations researcher evinces interest in the applications of methods to solve the pressing or critical problems of their firm.

Research can also be classified as conclusion-oriented and decision-oriented. While doing conclusion-oriented research a researcher is free to pick up a problem, redesign the enquiry and is free to conceptualize as he wishes. Decision-oriented research always implies taking a rational decision. Operational research is an example of decision-oriented research.

SIGNIFICANCE OF RESEARCH

"All progress is born of inquiry. Doubt is often better than overconfidence, for it leads to inquiry, and inquiry leads to invention" is a famous Hudson Maxim in context of which the significance of research can well be understood. Increased amounts of research make progress possible. Research inculcates scientific and inductive thinking and it promotes the development of logical habits of thinking and organisation.

The role of research in several fields of applied economics, whether related to business or to the economy as a whole, has greatly increased in modern times. The increasingly complex nature of business and government has focused attention on the use of research in solving operational problems. Research, as an aid to economic policy, has gained added importance, both for government and business.

1) Research provides the basis for nearly all Government Policies in our Economic System

For instance, government's budgets rest in part on an analysis of the needs and desires of the people and on the availability of revenues to meet these needs. The cost of needs has to be equated to probable revenues and this is a field where research is most needed. Through research we can devise alternative policies and can as well examine the consequences of each of these alternatives. Decision-making may not be a part of research, but research certainly facilitates the decisions of the policy maker. Government has also to chalk out programmes for dealing with all facets of the country's existence and most of these will be related directly or indirectly to economic conditions. The plight of cultivators, the problems of big and small business and industry, working conditions, trade union activities, the problems of distribution, even the size and nature of defense services are matters requiring research. Thus, research is considered necessary with regard to the allocation of nation's resources. Another area in government, where research is necessary, is collecting information on the economic and social structure of the nation. Such information indicates what is happening in the economy and what changes are taking place. Collecting such statistical information is by no means a routine task, but it involves a variety of research problems. These days nearly all governments maintain large staff of research technicians or experts to carry on this work. Thus, in the context of government, research as a tool to economic policy has three distinct phases of operation, viz., (i) investigation of economic structure through continual compilation of facts; (ii) diagnosis of events that are taking place and the analysis of the forces underlying them; and (iii) the prognosis, i.e., the prediction of future developments.

2) Research has its Special Significance in Solving various Operational and Planning Problems of Business and Industry

Operations research and market research, along with motivational research, are considered crucial and their results assist, in more than one way, in taking business decisions. Market research is the investigation of the structure and development of a market for the purpose of formulating efficient policies for purchasing, production and sales. Operations research refers to the application of mathematical, logical and analytical techniques to the solution of business problems of cost minimization or of profit maximization or what can be termed as optimization problems. Motivational research of determining why people behave as they do is mainly concerned with market characteristics. In other words, it is concerned with the determination of motivations underlying the consumer (market) behaviour. All these are of great help to people in business and industry who are responsible for taking business decisions. Research with regard to demand and market factors has great utility in business. Given knowledge of future demand, it is generally not difficult for a firm, or for an industry to adjust its supply schedule within the limits of its projected capacity. Market analysis has become an integral tool of business policy these days. Business budgeting, which ultimately results in a projected profit and loss account, is based mainly on sales estimates which in turn depend on business research. Once sales forecasting is done, efficient production and investment programmes can be set up around which are grouped the purchasing and financing plans. Research, thus, replaces intuitive business decisions by more logical and scientific decisions.

3) Research is equally important for social scientists in studying social relationships and in seeking answers to various social problems.

It provides the intellectual satisfaction of knowing a few things just for the sake of knowledge and also has practical utility for the social scientist to know for the sake of being able to do something better or in a more efficient manner. Research in social sciences is concerned both with knowledge for its own sake and with knowledge for what it can contribute to practical concerns. "This double emphasis is perhaps especially appropriate in the case of social science. On the one hand, its responsibility as a science is to develop a body of principles that make possible the understanding and prediction of the whole range of human interactions. On the other hand, because of its social orientation, it is increasingly being looked to for practical guidance

In addition to what has been stated above, the significance of research can also be understood keeping in view the following points:

a) To those students who are to write a master's or Ph.D. thesis, research may mean careerism or a way to attain a high position in the social structure;
b) To professionals in research methodology, research may mean a source of livelihood;
c) To philosophers and thinkers, research may mean the outlet for new ideas and insights;
d) To literary men and women, research may mean the development of new styles and creative work;
e) To analysts and intellectuals, research may mean the generalizations of new theories.

Thus, research is the fountain of knowledge for the sake of knowledge and an important source of providing guidelines for solving different business, governmental and social problems. It is a sort of formal training which enables one to understand the new developments in one's field in a better way.

RESEARCH PROCESS

Research process consists of series of actions or steps necessary to effectively carry out research and the desired sequencing of these steps.

1. Formulating the Research Problem;
2. Extensive literature survey;
3. Developing the Hypothesis;
4. Preparing the Research Design;
5. Determining Sample Design;
6. Collecting the Data;
7. Execution of the Project;
8. Analysis of Data;
9. Hypothesis testing;
10. Generalizations and Interpretation, and
11. Preparation of the Report or Presentation of the Results

1) Formulating the Research Problem

There are two types of research problems, viz., those which relate to states of nature and those which relate to relationships between variables. At the very outset the researcher must single out the problem he wants to study, i.e., he must decide the general area of interest or aspect of a subject-matter that he would like to inquire into. Initially the problem may be stated in a broad general way and then the ambiguities,

if any, relating to the problem be resolved. Then, the feasibility of a particular solution has to be considered before a working formulation of the problem can be set up. The formulation of a general topic into a specific research problem, thus, constitutes the first step in a scientific enquiry. Essentially two steps are involved in formulating the research problem, viz., understanding the problem thoroughly, and rephrasing the same into meaningful terms from an analytical point of view.

The best way of understanding the problem is to discuss it with one's own colleagues or with those having some expertise in the matter. In an academic institution the researcher can seek the help from a guide who is usually an experienced man and has several research problems in mind. Often, the guide puts forth the problem in general terms and it is up to the researcher to narrow it down and phrase the problem in operational terms. In private business units or in governmental organisations, the problem is usually earmarked by the administrative agencies with whom the researcher can discuss as to how the problem originally came about and what considerations are involved in its possible solutions.

The researcher must at the same time examine all available literature to get himself acquainted with the selected problem. He may review two types of literature the conceptual literature concerning the concepts and theories, and the empirical literature consisting of studies made earlier which are similar to the one proposed. The basic outcome of this review will be the knowledge as to what data and other materials are available for operational purposes which will enable the researcher to specify his own research problem in a meaningful context. After this the researcher rephrases the problem into analytical or operational terms

i.e., to put the problem in as specific terms as possible. This task of formulating, or defining, a research problem is a step of greatest importance in the entire research process. The problem to be investigated must be defined unambiguously for that will help discriminating relevant data from irrelevant ones. Care must, however, be taken to verify the objectivity and validity of the background facts concerning the problem. Professor W.A. Neiswanger correctly states that the statement of the objective is of basic importance because it determines the data which are to be collected, the characteristics of the data which are relevant, relations which are to be explored, the choice of techniques to be used in these explorations and the form of the final report. If there are certain pertinent terms, the same should be clearly defined along with the task of formulating the problem. In fact, formulation of the problem often follows a sequential pattern where a number of formulations are set up, each formulation more specific than the preceding one, each one phrased in more analytical terms, and each more realistic in terms of the available data and resources.

2) Extensive Literature Survey

Once the problem is formulated, a brief summary of it should be written down. It is compulsory for a research worker writing a thesis for a Ph.D. degree to write a synopsis of the topic and submit it to the necessary Committee or the Research Board for approval. At this juncture the researcher should undertake extensive literature survey connected with the problem. For this purpose, the abstracting and indexing journals and published or unpublished bibliographies are the first place to go to. Academic journals, conference proceedings, government reports, books etc., must be tapped depending on the nature of the problem. In this process, it should be remembered that one source

will lead to another. The earlier studies, if any, which are similar to the study in hand should be carefully studied. A good library will be a great help to the researcher at this stage.

3) Developing the Hypothesis

After extensive literature survey, researcher should state in clear terms the working hypothesis or hypotheses. Working hypothesis is tentative assumption made in order to draw out and test its logical or empirical consequences. As such the manner in which research hypotheses are developed is particularly important since they provide the focal point for research. They also affect the manner in which tests must be conducted in the analysis of data and indirectly the quality of data which is required for the analysis. In most types of research, the development of working hypothesis plays an important role. Hypothesis should be very specific and limited to the piece of research in hand because it has to be tested. The role of the hypothesis is to guide the researcher by delimiting the area of research and to keep him on the right track. It sharpens his thinking and focuses attention on the more important facets of the problem. It also indicates the type of data required and the type of methods of data analysis to be used.

How does one go about developing working hypotheses? The answer is by using the following approach:

a) Discussions with colleagues and experts about the problem, its origin and the objectives in seeking a solution;
b) Examination of data and records, if available, concerning the problem for possible trends, peculiarities and other clues;
c) Review of similar studies in the area or of the studies on similar problems; and

d) Exploratory personal investigation which involves original field interviews on a limited scale with interested parties and individuals with a view to secure greater insight into the practical aspects of the problem.

Thus, working hypotheses arise as a result of a-priori thinking about the subject, examination of the available data and material including related studies and the counsel of experts and interested parties. Working hypotheses are more useful when stated in precise and clearly defined terms. It may as well be remembered that occasionally we may encounter a problem where we do not need working hypotheses, specially in the case of exploratory or formulative researches which do not aim at testing the hypothesis. But as a general rule, specification of working hypotheses in another basic step of the research process in most research problems.

4) Preparing the Research Design

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

the research study is that of exploration. But when the purpose happens to be an accurate description of a situation or of an association between variables, the suitable design will be one that minimises bias and maximises the reliability of the data collected and analysed.

There are several research designs, such as, experimental and non-experimental hypothesis testing. Experimental designs can be either informal designs (such as before-and-after without control, after-only with control, before-and-after with control) or formal designs (such as completely randomized design, randomized block design, Latin square design, simple and complex factorial designs), out of which the researcher must select one for his own project.

The preparation of the research design, appropriate for a particular research problem, involves usually the consideration of the following:

a) The means of obtaining the information;
b) The availability and skills of the researcher and his staff (if any);
c) Explanation of the way in which selected means of obtaining information will be organized and the reasoning leading to the selection;
d) The time available for research; and
e) The cost factor relating to research, i.e., the finance available for the purpose.

5) Determining Sample Design

All the items under consideration in any field of inquiry constitute a 'universe' or 'population'. A complete enumeration of all the items in the 'population' is known as a census inquiry. It can be presumed that in such an inquiry when all the items are covered no element of chance is

left and highest accuracy is obtained. But in practice this may not be true. Even the slightest element of bias in such an inquiry will get larger and larger as the number of observations increases. Moreover, there is no way of checking the element of bias or its extent except through a resurvey or use of sample checks. Besides, this type of inquiry involves a great deal of time, money and energy. Not only this, census inquiry is not possible in practice under many circumstances. For instance, blood testing is done only on sample basis. Hence, quite often we select only a few items from the universe for our study purposes. The items so selected constitute what is technically called a sample.

The researcher must decide the way of selecting a sample or what is popularly known as the sample design. In other words, a sample design is a definite plan determined before any data are actually collected for obtaining a sample from a given population. Thus, the plan to select 12 of a city's 200 drugstores in a certain way constitutes a sample design. Samples can be either probability samples or non-probability samples. With probability samples each element has a known probability of being included in the sample but the non-probability samples do not allow the researcher to determine this probability. Probability samples are those based on simple random sampling, systematic sampling, stratified sampling, cluster/area sampling whereas non-probability samples are those based on convenience sampling, judgement sampling and quota sampling techniques. A brief mention of the important sample designs is as follows:

a) Deliberate Sampling

Deliberate sampling is also known as purposive or non-probability sampling. This sampling method involves purposive or deliberate selection of particular units of the universe for constituting a sample which represents the universe. When population elements are selected for inclusion in the sample based on the ease of access, it can be called convenience sampling. If a researcher wishes to secure data from, say, gasoline buyers, he may select a fixed number of petrol stations and may conduct interviews at these stations. This would be an example of convenience sample of gasoline buyers. At times such a procedure may give very biased results particularly when the population is not homogeneous. On the other hand, in judgement sampling the researcher's judgement is used for selecting items which he considers as representative of the population. For example, a judgement sample of college students might be taken to secure reactions to a new method of teaching. Judgement sampling is used quite frequently in qualitative research where the desire happens to be to develop hypotheses rather than to generalise to larger populations.

b) Simple Random Sampling

This type of sampling is also known as chance sampling or probability sampling where each and every item in the population has an equal chance of inclusion in the sample and each one of the possible samples, in case of finite universe, has the same probability of being selected. For example, if we have to select a sample of 300 items from a universe of 15,000 items, then we can put the names or numbers of all the 15,000 items on slips of paper and conduct a lottery. Using the random number tables is another method of random sampling. To select the sample, each item is assigned a number from 1 to 15,000. Then, 300 five digit random numbers are selected from the table. To do this we

select some random starting point and then a systematic pattern is used in proceeding through the table. We might start in the 4th row, second column and proceed down the column to the bottom of the table and then move to the top of the next column to the right. When a number exceeds the limit of the numbers in the frame, in our case over 15,000, it is simply passed over and the next number selected that does fall within the relevant range. Since the numbers were placed in the table in a completely random fashion, the resulting sample is random. This procedure gives each item an equal probability of being selected. In case of infinite population, the selection of each item in a random sample is controlled by the same probability and that successive selections are independent of one another.

c) Systematic Sampling

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

d) Stratified Sampling

If the population from which a sample is to be drawn does not constitute a homogeneous group, then stratified sampling technique is applied so as to obtain a representative sample. In this technique, the population is stratified into a number of non-overlapping subpopulations

or strata and sample items are selected from each stratum. If the items selected from each stratum is based on simple random sampling the entire procedure, first stratification and then simple random sampling, is known as stratified random sampling.

e) Quota Sampling

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

f) Cluster Sampling and Area Sampling

Cluster sampling involves grouping the population and then selecting the groups or the clusters rather than individual elements for inclusion in the sample. Suppose some departmental store wishes to sample its credit card holders. It has issued its cards to 15,000 customers. The sample size is to be kept say 450. For cluster sampling this list of 15,000 card holders could be formed into 100 clusters of 150 card holders each. Three clusters might then be selected for the sample randomly. The sample size must often be larger than the simple random sample to ensure the same level of accuracy because is cluster sampling procedural potential for order bias and other sources of error is usually accentuated. The clustering approach can, however, make the sampling

procedure relatively easier and increase the efficiency of field work, specially in the case of personal interviews.

Area sampling is quite close to cluster sampling and is often talked about when the total geographical area of interest happens to be big one. Under area sampling we first divide the total area into a number of smaller non-overlapping areas, generally called geographical clusters, then a number of these smaller areas are randomly selected, and all units in these small areas are included in the sample. Area sampling is specially helpful where we do not have the list of the population concerned. It also makes the field interviewing more efficient since interviewer can do many interviews at each location.

g) Multi-stage Sampling

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

h) Sequential Sampling

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

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

In practice, several of the methods of sampling described above may well be used in the same study in which case it can be called mixed sampling. It may be pointed out here that normally one should resort to random sampling so that bias can be eliminated and sampling error can be estimated. But purposive sampling is considered desirable when the universe happens to be small and a known characteristic of it is to be studied intensively. Also, there are conditions under which sample designs other than random sampling may be considered better for reasons like convenience and low costs. The sample design to be used must be decided by the researcher taking into consideration the nature of the inquiry and other related factors.

6) Collecting the Data

In dealing with any real life problem it is often found that data at hand are inadequate, and hence, it becomes necessary to collect data that are appropriate. There are several ways of collecting the appropriate data which differ considerably in context of money costs, time and other resources at the disposal of the researcher.

Primary data can be collected either through experiment or through survey. If the researcher conducts an experiment, he observes some quantitative measurements, or the data, with the help of which he examines the truth contained in his hypothesis. But in the case of a survey, data can be collected by any one or more of the following ways:

a) By Observation

This method implies the collection of information by way of investigator's own observation, without interviewing the respondents. The information obtained relates to what is currently happening and is not complicated by either the past behaviour or future intentions or attitudes of respondents. This method is no doubt an expensive method and the information provided by this method is also very limited. As such this method is not suitable in inquiries where large samples are concerned.

b) Through Personal Interview

The investigator follows a rigid procedure and seeks answers to a set of pre-conceived questions through personal interviews. This method of collecting data is usually carried out in a structured way where output depends upon the ability of the interviewer to a large extent.

c) Through Telephone Interview

This method of collecting information involves contacting the respondents on telephone itself. This is not a very widely used method but it plays an important role in industrial surveys in developed regions, particularly, when the survey has to be accomplished in a very limited time.

d) By Mailing of Questionnaire

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

the questionnaire. Questionnaire to be used must be prepared very carefully so that it may prove to be effective in collecting the relevant information.

e) Through Schedules

Under this method the enumerators are appointed and given training. They are provided with schedules containing relevant questions. These enumerators go to respondents with these schedules. Data are collected by filling up the schedules by enumerators on the basis of replies given by respondents. Much depends upon the capability of enumerators so far as this method is concerned. Some occasional field checks on the work of the enumerators may ensure sincere work.

The researcher should select one of these methods of collecting the data taking into consideration the nature of investigation, objective and scope of the inquiry, finanical resources, available time and the desired degree of accuracy. Though he should pay attention to all these factors but much depends upon the ability and experience of the researcher. In this context Dr A.L Bowley very aptly remarks that in collection of statistical data commonsense is the chief requisite and experience the chief teacher.

7) Execution of the Project

Execution of the project is a very important step in the research process. If the execution of the project proceeds on correct lines, the data to be collected would be adequate and dependable. The researcher should see that the project is executed in a systematic manner and in time. If the survey is to be conducted by means of structured questionnaires, data can be readily machine-processed. In such a

situation, questions as well as the possible answers may be coded. If the data are to be collected through interviewers, arrangements should be made for proper selection and training of the interviewers. The training may be given with the help of instruction manuals which explain clearly the job of the interviewers at each step. Occasional field checks should be made to ensure that the interviewers are doing their assigned job sincerely and efficiently. A careful watch should be kept for unanticipated factors in order to keep the survey as much realistic as possible. This, in other words, means that steps should be taken to ensure that the survey is under statistical control so that the collected information is in accordance with the pre-defined standard of accuracy. If some of the respondents do not cooperate, some suitable methods should be designed to tackle this problem. One method of dealing with the non-response problem is to make a list of the non-respondents and take a small sub-sample of them, and then with the help of experts vigorous efforts can be made for securing response.

8) Analysis of Data

After the data have been collected, the researcher turns to the task of analysing them. The analysis of data requires a number of closely related operations such as establishment of categories, the application of these categories to raw data through coding, tabulation and then drawing statistical inferences. The unwieldy data should necessarily be condensed into a few manageable groups and tables for further analysis. Thus, researcher should classify the raw data into some purposeful and usable categories. Coding operation is usually done at this stage through which the categories of data are transformed into symbols that may be

tabulated and counted. Editing is the procedure that improves the quality of the data for coding. With coding the stage is ready for tabulation. Tabulation is a part of the technical procedure wherein the classified data are put in the form of tables. The mechanical devices can be made use of at this juncture. A great deal of data, specially in large inquiries, is tabulated by computers. Computers not only save time but also make it possible to study large number of variables affecting a problem simultaneously.

Analysis work after tabulation is generally based on the computation of various percentages, coefficients, etc., by applying various well defined statistical formulae. In the process of analysis, relationships or differences supporting or conflicting with original or new hypotheses should be subjected to tests of significance to determine with what validity data can be said to indicate any conclusion(s). For instance, if there are two samples of weekly wages, each sample being drawn from factories in different parts of the same city, giving two different mean values, then our problem may be whether the two mean values are significantly different or the difference is just a matter of chance. Through the use of statistical tests we can establish whether such a difference is a real one or is the result of random fluctuations. If the difference happens to be real, the inference will be that the two samples come from different universes and if the difference is due to chance, the conclusion would be that the two samples belong to the same universe. Similarly, the technique of analysis of variance can help us in analysing whether three or more varieties of seeds grown on certain fields yield significantly different results or not. In brief, the researcher can analyse the collected data with the help of various statistical measures.

9) Hypothesis Testing

After analysing the data as stated above, the researcher is in a position to test the hypotheses, if any, he had formulated earlier. Do the facts support the hypotheses or they happen to be contrary? This is the usual question which should be answered while testing hypotheses. Various tests, such as Chi square test, t-test, F-test, have been developed by statisticians for the purpose. The hypotheses may be tested through the use of one or more of such tests, depending upon the nature and object of research inquiry. Hypothesis-testing will result in either accepting the hypothesis or in rejecting it. If the researcher had no hypotheses to start with, generalisations established on the basis of data may be stated as hypotheses to be tested by subsequent researches in times to come.

10) Generalizations and Interpretation

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

11) Preparation of the Report or Presentation of the Results

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

1) The layout of the report should be as follows: (i) the preliminary pages; (ii) the main text, and (iii) the end matter.

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

The main text of the report should have the following parts:

a) **Introduction:** It should contain a clear statement of the objective of the research and an explanation of the methodology adopted in accomplishing the research. The scope of the study along with various limitations should as well be stated in this part.

b) **Summary of Findings**: After introduction there would appear a statement of findings and recommendations in non-technical language. If the findings are extensive, they should be summarised.

c) **Main Report**: The main body of the report should be presented in logical sequence and broken-down into readily identifiable sections.

d) **Conclusion**: Towards the end of the main text, researcher should again put down the results of his research clearly and precisely. In fact, it is the final summing up.

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

2) Report should be written in a concise and objective style in simple language avoiding vague expressions such as 'it seems,' 'there may be', and the like.

3) Charts and illustrations in the main report should be used only if they present the information more clearly and forcibly.

4) Calculated 'confidence limits' must be mentioned and the various constraints experienced in conducting research operations may as well be stated.

CRITERIA OF GOOD RESEARCH

- Whatever may be the types of research works and studies, one thing that is important is that they all meet on the common ground of scientific method employed by them. One expects scientific research to satisfy the following criteria:
  - The purpose of the research should be clearly defined and common concepts be used.
  - The research procedure used should be described in sufficient detail to permit another researcher to repeat the research for further advancement, keeping the continuity of what has already been attained.
  - The procedural design of the research should be carefully planned to yield results that are as objective as possible.
  - The researcher should report with complete frankness, flaws in procedural design and estimate their effects upon the findings.
  - The analysis of data should be sufficiently adequate to reveal its significance and the methods of analysis used should be

appropriate. The validity and reliability of the data should be checked carefully.

- Conclusions should be confined to those justified by the data of the research and limited to those for which the data provide an adequate basis.
- Greater confidence in research is warranted if the researcher is experienced, has a good reputation in research and is a person of integrity.

QUALITIES OF GOOD RESEARCH

1) Good research is Systematic

It means that research is structured with specified steps to be taken in a specified sequence in accordance with the well defined set of rules. Systematic characteristic of the research does not rule out creative thinking but it certainly does reject the use of guessing and intuition in arriving at conclusions.

2) Good research is Logical

This implies that research is guided by the rules of logical reasoning and the logical process of induction and deduction are of great value in carrying out research. Induction is the process of reasoning from a part to the whole whereas deduction is the process of reasoning from some premise to a conclusion which follows from that very premise.

3) Good research is Empirical

It implies that research is related basically to one or more aspects of a real situation and deals with concrete data that provides a basis for external validity to research results.

4) Good research is Replicable

This characteristic allows research results to be verified by replicating the study and thereby building a sound basis for decisions.

QUALITIES OF A GOOD RESEARCHER

A) General Qualities

1) Scientific Attitude

The first essential Quality of a successful research worker is that he must possess a scientific (systematic) frame (structure) of mind. He must have the determination (willpower / strength of mind) and ability to get the naked (hidden) facts and not to be influenced by one's own wishes.

As human beings he has certain praises (admiration) and prejudices (bias). He has also certain precarceived notions (ideas) about the problems being researched. He should keep all these things with him.

2) Imagination and Insight

Researcher must possess high degree of imagination. He should be able to go deeper and deeper into the realm (area) of abstract social phenomena (fact / event) and visualize the intangible aspects (features) of the society

3) Perseverance

Work of scientific research requires steady of mind. Researcher should not get easily discouraged. It is equally possible that he might subsequently feel that the choice of the problem was wrong. Inspite of all this he must have more courage to face the difficulties and work patiently and continuously over long periods

4) Quick Grasping Power

The researcher should possess the power to grasp the significance of things quickly

5) Clarity of Thinking

A good researcher should have clear idea about the terminology that he is going to use.

B) Specific Qualities

1) Knowledge of the Subject

The researcher should be enough knowledge in his area of research. Such knowledge helps him in preparing questionnaire and schedule to get proper information. He can enter into face to face discussion and remove any doubts arising the minds of the people regarding the study

2) Knowledge of the technique of Research

Researcher should have basic idea on tools used in his research

3) Personal Taste in the Study

A personal taste in the study will inspire him and keep his morale (confidence) in times of difficulties. A forced work is often monotonous and very tiresome

4) Familiarity about the Information

The researcher should be familiar with the people whom he is studying. Familiarity will help him to get intimate (close) information

5) Unbiased Attitude

The researcher should have no preconceptions (idea / bias) about the subject under study. He should go to his research with absolutely a clean state. He should maintain an open mind and look for data which would substantiate (validate / verify) and give his theory a new meaning.

RESEARCH DESIGN

The formidable problem that follows the task of defining the research problem is the preparation of the design of the research project, popularly known as the "research design". Decisions regarding what, where, when, how much, by what means concerning an inquiry or a research study constitute a research design. "A research design is the arrangement of conditions for collection and analysis of data in a manner that aims to combine relevance to the research purpose with economy in procedure."1 In fact, the research design is the conceptual structure within which research is conducted; it constitutes the blueprint for the collection, measurement and analysis of data. As such the design includes an outline of what the researcher will do from writing the hypothesis and its operational implications to the final analysis of data. More explicitly, the desing decisions happen to be in respect of:

1. What is the study about?
2. Why is the study being made?
3. Where will the study be carried out?
4. What type of data is required?
5. Where can the required data be found?
6. What periods of time will the study include?
7. What will be the sample design?
8. What techniques of data collection will be used?
9. How will the data be analyzed?

10. In what style will the report be prepared?

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

a) the sampling design which deals with the method of selecting items to be observed for the given study;

b) the observational design which relates to the conditions under which the observations are to be made;

c) the statistical design which concerns with the question of how many items are to be observed and how the information and data gathered are to be analysed; and

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

From what has been stated above, we can state the important features of a research design as under:

i) It is a plan that specifies the sources and types of information relevant to the research problem.

ii) It is a strategy specifying which approach will be used for gathering and analysing the data.

iii) It also includes the time and cost budgets since most studies are done under these two constraints.

In brief, research design must, at least, contain—(a) a clear statement of the research problem; (b) procedures and techniques to be used for gathering information; (c) the population to be studied; and (d) methods to be used in processing and analyzing data.

COMPONENTS OF RESEARCH DESIGN

1) Title of the Study

Enough information should be given in the title, to identify the study. The researcher should consider the following while selecting a title

- The title should be specific to the area of the study
- The title should indicate the topic of the study
- The language of the title should be professional in nature but not pedantic (dull)
- The title should be as brief as possible

2) Introduction

Under this heading a brief explanation of the genesis of the problem should be given.

3) Statement of the Problem

After a brief introduction explaining the genesis of the problem, the researcher should state the problem. While stating the problem use of clear, simple and concise statement is preferable

4) Review of Previous Studies

Under this head the researcher presents what is so far known about the problem under consideration. All related studies need not be discussed. The researcher may describe the most important ones. A

review of the previous studies enables the researcher to know the different areas covered by various studies, to concentrate on the areas where little research has been carried out, to look into various merits and shortcomings of certain studies already completed and to verify the present findings with that of the previous ones

5) Scope of the Study

This heading gives an idea about the extend of the study. The scope of the study is dependent on several factors such as the time and money available with the investigator, availability of the sample, co-operation of the respondents and the like

6) Objectives of the Study

The task of the researcher is to lay down the objectives precisely. The objectives enlighten the researcher's own mind and lead to more efficient enquiry. Once the objectives are selected, the study can be undertaken with required accuracy and within the given resources. The objectives mentioned should be well within the scope of the study

7) Hypothesis to be Tested

Hypothesis is a proposition, condition or principle which is assumed, perhaps without belief in order to draw logical conclusions. Hypotheses are formulated to explain observed facts, conditions, or behaviours and to serve as a guide in the research process. Each hypothesis is individually tested to determine whether it is tenable (reasonable) or not. Hypothesis should be stated in clear, concise and understandable language

8) Operational Definition of Concepts

All terms that might be ambiguous should be clarified. A clear understanding of the terms used in the study is important. It is necessary to identify and label the variables. The variables can be labeled as independent variable and dependent variable. An independent variable is the factor which is measured, manipulated or selected by the experiments. A dependent variable is that factor which is measured to determine the effect of independent variable

9) Geographical Area to be Covered

Under this head the area to be covered by the study is mentioned

10) Reference Period

The period of study can be mentioned under this heading

11) Methodology

The researcher should first determine the kind of information needed to answer the research questions. Secondly he must know the sources of data and finally he must know the means by which he will gather information which is known as methodology

12) Sampling

Sampling involves taking a portion of population, making observation on this smaller group and then generalizing the findings to be applied to a large population. The small group that is observed is called the sample and the large group is called population. The sample is the portion of the population and it must be representative of the population. If the sample is biased, the findings of the study cannot be generalized

13) Tools for Collection of Data

The choice of method for collecting the data is governed by the subject matter, the unit of enquiry and the scale of the study. A study of the behavior of a group would call for observational techniques, for a simple enquiry among the cross section of population, a questionnaire is adequate. A survey of general population entailing many complicated questions would call for personal interviewing

14) Plan of Analysis

Once the data have been collected, they must be reduced to meaningful results by statistical analysis so that the conclusions for generalization can be drawn from them. The researcher should describe how he plans to organize the data. He should decide the statistical treatment. The appropriateness of the technique should be discussed. He must discuss the procedure for treating the data

15) Research Report

The results should be communicated. The format consists of three parts

- Part I Preliminary Pages, which contain title page, approval sheet, preface (if any), table of contents, list of tables (if any) and list of figures (if any)
- Part II Body of the report, which covers content chapters
- Part III Supplementary pages which included bibliography appendix (if any) and index (if any)

16) Time Schedule

The researcher has to work out a time schedule for his research work. The time required includes the following:

- Time to be used for preparing the theoretical background
- Time to be used for preparing the data gathering devices such as questionnaire, interview schedule, record sheet, interviewer's manual and time and expenses sheets
- Time to be used for data collection
- Time to be used for processing the data
- Time to be used for writing the report
- Time to be used for submitting the thesis

17) Financial Budget

It is desirable to work out the budget which gives an idea about the money needed to complete the project. The cost estimates of the project will include stationery, printing, sample selection, field work, mailing, processing, tabulating, preparation of report and overheads.

The research design will differ depending on the research purpose. The research purposes may purport to gain familiarity with a phenomenon, portray accurately the characteristics of a particular individual, situation or group, determine the frequency with which something occurs and test a hypothesis of casual relationship between variables. In practice, these different types of studies are not always sharply distinguishable. The research design will be more or less the same for any type of study

* * * * *

POSSIBLE QUESTIONS

UNIT-I

PART A
One Mark
Online Examination

PART B
Eight Marks

1. Explain the meaning and significance of a research design.
2. Write the important concepts relating to research design.
3. Explain the characteristics and qualities of good research?
4. Discuss the importance of research in modern times?
5. Describe the types of research with examples?
6. Describe the process of research with diagrams.
7. Describe fully the techniques of defining a research problem.
8. Define research and explain the significance of research in detail?
9. Describe the different methods of research design.
10. Discuss the steps followed in research process with suitable chart.

KARPAGAM ACADEMY OF HIGHER EDUCATION

(DEEMED TO BE UNIVERSITY)

(ESTABILISHED UNDER SEC 3 OF UGC ACT 1956)

DEPARTMENT OF COMMERCE

III -B.COM RESEARCH METHODOLOGY-15CMU603B

UNIT-INTRODUCTION ABOUT RESEARCH

S.NO	QUESTIONS	OPTION I	OPTION II	OPTION III	OPTION IV	ANSWER
1	____ is an endeavour to discover, develop and verify knowledge	Hypothesis	Research	Character	Sample	Research
2	_____study may be simple or complex	Descriptive	Analytical	Applied	Fundamental	Descriptive
3	___ available information by analyzing and doing critical evaluation	Qualitative	Quantitative	Analytical	Applied	Analytical
4	____research aims at finding a solution for an immediate problem	Applied	Analytical	Longitudinal	Case Study	Applied
5	_____research is an example of applied research	Production	Market	Operations	Business	Market
6	___research is otherwise known as pure research	Fundamental	Operations	Individual	Decision Oriented	Fundamental
7	__research is applicable to phenomena that can be expressed in terms of quantity	Quantitative	Qualitative	Conceptual	Empirical	Quantitative

S.NO	QUESTIONS	OPTION I	OPTION II	OPTION III	OPTION IV	ANSWER
8	___research is an important type of qualitative research.	Basic	Motivation	Pure	Decision Oriented	Motivation
9	__research is important in the behavioural sciences	Qualitative	Quantitative	Applied	Empirical	Qualitative
10	by philosophers and thinkers to develop new concepts or to reinterpret existing ones	Conceptual	Field Method	Exploratory	Quantitative	Conceptual
11	research, coming up with conclusions which are capable of being verified by observation	Conceptual	Field Method	Exploratory	Empirical	Empirical
12	___research is confined to a single time-period	One time	Longitudinal	Descriptive	Analytical	One time
13	___research is carried on over several time-periods	One time	Longitudinal	Descriptive	Analytical	Longitudinal
14	__research depends upon the environment in which it is to be carried out.	Field Setting	Exploratory	Formalized	Historical	Field Setting
15	with substantial structure and with specific hypotheses to be tested	Exploratory	Formalized	Historical	Decision Oriented	Formalized
16	___research is carried, when the reason for a problem is not clear	Exploratory	Formalized	Historical	Decision Oriented	Exploratory
17	reflective thinking to unsolved social problems by discovering past trends of events, facts and	One time	Longitudinal	Descriptive	Historical	Historical

S.NO	QUESTIONS	OPTION I	OPTION II	OPTION III	OPTION IV	ANSWER
18	__research is always for the need of a decision maker	Decision oriented	Qualitative	Historical	Conceptual	Decision oriented
19	___research is done on the basis of one's own judgement, interest and capacity	Descriptive	Analytical	Group	Individual	Individual
20	___research is undertaken by several researchers	Group	Descriptive	Analytical	Applied	Group
21	__research is an example of decision-oriented research	Quantitative	Operations	Applied	Descriptive	Operations
22	__inductive thinking and it promotes the development of logical habits of thinking and	Research	Investigation	Enquiry	Study of Human Beings	Research
23	__the structure and development of a market for the purpose of formulating efficient policies for	Market	Motivation	Operations	Production	Market
24	__the determination of motivations underlying the consumer behaviour	Operations	Production	Market	Motivation	Motivation
25	__steps necessary to effectively carry out research and the desired sequencing of these	Planning	Research Process	Research Process	Transcription	Research Process
26	Steps involved in carrying out a research work is known as ___	Research Process	Research Process	Planning	Tabulation	Research Process
27	___literature concerning the concepts and theories	Conceptual	Empirical	Historical	Theoretical	Conceptual

S.NO	QUESTIONS	OPTION I	OPTION II	OPTION III	OPTION IV	ANSWER
28	__ determines the data which are to be collected	Objective of Study	Research Design	Review of Literature	Hypothesis	Objective of Study
29	__is the focal point for research	Observation	Questionnaire	Objectives of the Study	Hypothesis	Hypothesis
30	___facilitates research to be as efficient as possible yielding maximal information	Research Design	Research Process	Objectives of the Study	Hypothesis	Research Design
31	considered appropriate if the purpose of the research study is that of exploration	Flexible	Rigid	Experimental	Non-experimental	Flexible
32	be an accurate description of a situation or of an association between variables, the suitable	Latin	Rigid	Complex	Factorial	Rigid
33	A complete enumeration of all the items in the 'population' is known as___	Probability Sample	Non-Probability Sample	Census	Cluster	Census
34	___is a very important step in the research process	Review of Literature	Data Collection	Analysis work	Execution of the Project	Execution of the Project
35	related operations such as establishment of categories, the application of these categories	Analysis of Data	Data Verification	Transcription	Classification	Analysis of Data
36	categories of data are transformed into symbols that may be tabulated and counted	Editing	Coding	Tabulation	Analysis of Data	Coding
37	__is the procedure that improves the quality of the data for coding	Editing	Coding	Tabulation	Analysis of Data	Editing

S.NO	QUESTIONS	OPTION I	OPTION II	OPTION III	OPTION IV	ANSWER
38	procedure wherein the classified data are put in the form of tables	Analysis of Data	Editing	Coding	Tabulation	Tabulation
39	often trigger off new questions which in turn may lead to further researches	Generalization	Interpretation	Hypothesis Testing	Analysis of Data	Interpretation
40	The scope of the study along with various limitations should as well be stated in __chapter	Introduction Chapter	Review of Literature Chapter	Analysis and Interpretation Chapter	Summary of Findings Chapter	Introduction Chapter
41	justified by the data of the research and limited to those for which the data provide an	Introduction	Review of Literature	Conclusion	Analysis and Interpretation	Conclusion
42	__is the process of induction and deduction are of great value in carrying out research	Systematic	Logical	Empirical	Replicable	Logical
43	specified steps to be taken in a specified sequence in accordance with the well defined	Systematic	Logical	Empirical	Replicable	Systematic
44	is related basically to one or more aspects of a real situation and deals with concrete data	Systematic	Logical	Replicable	Empirical	Empirical
45	research results to be verified by replicating the study and thereby building a sound basis	Systematic	Logical	Replicable	Empirical	Replicable
46	or examination in seeking facts or principles, diligent investigation in order to	Enquiry	Inquiry	Research	Investigation	Research
47	__ research discovers or tests whether certain variables are associated	Historical	One time	Conceptual	Descriptive	Descriptive

S.NO	QUESTIONS	OPTION I	OPTION II	OPTION III	OPTION IV	ANSWER
48	concerned with generalisations and with the formulation of a theory.	Fundamental	Operations	Individual	Decision Oriented	Fundamental
49	____research are based on the measurements of quantity or amounts	Qualitative	Conceptual	Empirical	Quantitative	Quantitative
50	___ research is concerned with phenomena relating to or involving quality or kind	Conceputal	Empirical	Field Method	Quantitative	Qualitative
51	___ research is that related to some abstract idea or theory	Empirical	Conceptual	Field Method	Exploratory	Conceptual
52	when proof is sought that certain variables affect other variables in some way	Formalized	Historical	Empirical	Individual	Empirical
53	application of mathematical, logical and analytical techniques for cost	Production	Market	Motivation	Operations	Operations
54	people behave as they do is mainly concerned with market characteristics	Production	Market	Motivation	Operations	Motivation
55	This task of formulating, or defining, a research problem is a step of __	Research Process	Research Design	Hypothesis	Data Validation	Research Process
56	delimiting the area of research and to keep him on the right track	Transcription	Classification	Hypothesis	Sample Design	Hypothesis
57	__is a blue print of research	Research Process	Objectives of the Study	Research Design	Hypothesis	Research Design

S.NO	QUESTIONS	OPTION I	OPTION II	OPTION III	OPTION IV	ANSWER
58	of relevant evidence with minimal expenditure of effort, time and money	Hypothesis	Research Process	Objectives of the Study	Research Design	Research Design
59	and the development of methods and systems for the advancement of human	Interpretation	Investigation	Enquiry	Research	Research
60	available information by analyzing and doing critical evaluation	Empirical	Descriptive	Applied	Analytical	Analytical

UNIT – II

SYLLABUS

Sampling – Meaning – Steps in Sample Design – Characteristics of a Good Sample Design – Determination of Sample Size – Sampling Techniques – Probability and Non-Probability Sampling – Sampling Error

* * * * *

SAMPLING

All items in any field of inquiry constitute a 'Universe' or 'Population. A complete enumeration of all items in the 'population' is known as a census inquiry. It can be presumed that in such an inquiry, when all items are covered, no element of chance is left and highest accuracy is obtained. But in practice this may not be true. Even the slightest element of bias in such an inquiry will get larger and larger as the number of observation increases. Moreover, there is no way of checking the element of bias or its extent except through a resurvey or use of sample checks. Besides, this type of inquiry involves a great deal of time, money and energy. Therefore, when the field of inquiry is large, this method becomes difficult to adopt because of the resources involved. At times, this method is practically beyond the reach of ordinary researchers. Perhaps, government is the only institution which can get the complete enumeration carried out. Even the government adopts this in very rare cases such as population census conducted once in a decade.

Further, many a time it is not possible to examine every item in the population, and sometimes it is possible to obtain sufficiently accurate results by studying only a part of total population. In such cases there is

no utility of census surveys. Then, the researcher may make use of Sampling.

- Sampling is nothing but a proportion of Population
- However, it needs to be emphasised that when the universe is a small one, it is no use resorting to a sample survey. When field studies are undertaken in practical life, considerations of time and cost almost invariably lead to a selection of respondents i.e., selection of only a few items. The respondents selected should be as representative of the total population as possible in order to produce a miniature cross-section. The selected respondents constitute what is technically called a 'sample' and the selection process is called 'sampling technique.' The survey so conducted is known as 'sample survey'.
- Algebraically, let the population size be N and if a part of size n (which is < N) of this population is selected according to some rule for studying some characteristic of the population, the group consisting of these n units is known as 'sample'.
- Researcher must prepare a sample design for his study i.e., he must plan how a sample should be selected and of what size such a sample would be.

STEPS IN SAMPLE DESIGN

1) Type of Universe

The first step in developing any sample design is to clearly define the set of objects, technically called the universe, to be studied. The universe can be finite or infinite. In finite universe the number of items is certain, but in case of an infinite universe the number of items is infinite, i.e. we cannot have any idea about the total number of items. The population of a city, the number of workers in a factory and the like are examples of finite universes, whereas the number of stars in the sky,

listeners of a specific radio programme, throwing a dice etc. are examples of infinite universes.

2) Sampling Unit

A decision has to be taken concerning a sampling unit before selecting sample. Sampling unit may be a geographical one such as state, district, village, etc., or a construction unit such as a house, flat, etc., or it may be a social unit such as family, club, school, etc., or it may be an individual. The researcher will have to decide one or more of such units that he has to select for his study.

3) Source List

It is also known as 'sampling frame' from which sample is to be drawn. It contains the names of all items of universe. If source list is not available, researcher has to prepare it. Such a list should be comprehensive, correct, reliable and appropriate. It is extremely important for the source list to be as representative of the population as possible.

4) Size of Sample

This refers to the number of items to be selected from the universe to constitute a sample. This is a major problem before a researcher. The size of sample should neither be excessively large, nor too small. It should be optimum. An optimum sample is one which fulfills the requirements of efficiency, representativeness, reliability and flexibility. While deciding the size of sample, researcher must determine the desired precision as also an acceptable confidence level for the estimate. The size of population variance needs to be considered as in case of larger variance usually a bigger sample is needed. The size of population must be kept in view for this also limits the sample size. The parameters of

interest in a research study must be kept in view, while deciding the size of sample. Costs too dictate the size of sample that we can draw. As such, budgetary constraint must invariably be taken into consideration when we decide the sample size.

5) Parameters of Interest

In determining the sample design, one must consider the question of the specific population parameters which are of interest. For instance, we may be interested in estimating the proportion of persons with some characteristics in the population or we may be interested in knowing some average or the other measure concerning the population. There may be important sub-groups in the population about whom we would like to make estimates. All this has a strong impact upon the sample design we would accept.

6) Budgetary Constraint

Cost considerations, from practical point of view, have a major impact upon decisions relating to not only the size of the sample but also to the type of sample. This fact can even lead to the use of a non-probability sample.

7) Sampling Procedure

Finally, the researcher must decide the type of sample he will use i.e. he must decide about the technique to be used in selecting the items for the sample. In fact, this technique or procedure stands for the sample design itself. There are several sample designs out of which the researcher must choose one for his study. Obviously, he must select that design which, for a given sample size and for a given cost, has a smaller sampling error.

CHARACTERISTICS OF A GOOD SAMPLE DESIGN

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

DETERMINATION OF SAMPLE SIZE

1) Nature of Universe

Universe may be either homogenous or heterogeneous in nature. If the items of the universe are homogenous, a small sample can serve the purpose. But if the items are heterogeneous, a large sample would be required. Technically, this can be termed as the dispersion factor.

2) Number of Classes Proposed

If many class-groups are to be formed, a large sample would be required because a small sample might not be able to give a reasonable number of items in each class-group.

3) Nature of Study

If items are to be intensively and continuously studied, the sample should be small. For a general survey the size of the sample should be large, but a small sample is considered appropriate in technical surveys.

4) Type of Sampling

Sampling technique plays an important part in determining the size of the sample. A small random sample is apt to be much superior to a larger but badly selected sample.

5) Standard of Accuracy and Acceptable Confidence Level

If the standard of accuracy or the level of precision is to be kept high, we shall require relatively larger sample. For doubling the accuracy for a fixed significance level, the sample size has to be increased fourfold.

6) Availability of Finance

In practice, size of the sample depends upon the amount of money available for the study purposes. This factor should be kept in view while determining the size of sample for large samples result in increasing the cost of sampling estimates.

7) Other Considerations

Nature of units, size of the population, size of questionnaire, availability of trained investigators, the conditions under which the sample is being conducted, the time available for completion of the study are a few other considerations to which a researcher must pay attention while selecting the size of sample.

SAMPLING TECHNIQUES

1) Probability Sampling

In probability sample, every unit in the population has equal chances for being selected as a sample unit.

2) Non–probability Sampling

In non probability sampling, units in the population has unequal or zero chances for being selected as a sample unit.

RANDOM SAMPLING PROCEDURES

The importance of randomness in sampling needs no emphasis, It is a means for securing a representative sample. How can a random sample be drawn? The layman tends to think that random sampling means picking out units "at random", i.e., in a haphazard or hit-and-miss way. Experience shows that the human being is an extremely poor instrument for the conduct of a random selection. To ensure true randomness the method of selection must be independent of human judgement. There are basic procedures.

1) Lottery Method

This is the simplest and most familiar procedure of random sampling. If a sample of 10 students is to be drawn out of a list of 50 students in a section, take 50 equal size chips or slips of paper; number them from 1 to 50 each bearing only one number. Roll each slip. Put the rolled slips in a global container and thoroughly shuffle or mix them. Take 10 chips from the container one after another. Each time before drawing a chip, mix the chips in the container thoroughly. The units bearing the numbers of chips drawn constitute the random sample.

(i) Sampling with Replacement

After a number is selected by draw, it may be replaced and consequently it has a chance of being selected again. Such method is known as sampling with replacement. This is usually referred to as unrestricted random sampling.

(ii) Sampling without Replacement

Selected numbers is set aside, and so in subsequent draws, it does not get a chance of being selected again. This type of sampling is known as sampling without replacement. This is a form of restricted sampling.

2) Use of Table o Random Numbers

This is a less cumbersome, but equally valid procedure of sample selection. Tables of random numbers have been developed by Kendall and Smith (1939), Fisher and Yates (1963) and Tippett (1927). To select a random sample out of a given frame, one should simply start to read numbers from a Table of Random Numbers at any randomly selected point and pick out numbers within the range of the frame. Let us suppose that random sample of 50 is to be selected from a College populations of 500 Commerce Students. We can use any table of random numbers.

10	09	73	25	33	76	52	01
37	54	20	48	05	64	89	47
08	42	26	89	53	19	64	50
09	01	90	25	29	09	37	67
12	80	79	99	70	80	15	73
66	06	57	47	17	34	07	27
31	06	01	08	05	45	57	18
85	26	97	76	02	02	05	16
63	57	33	21	35	05	32	54
73	79	64	57	53	03	52	96

Let us suppose, we start at the top of the left hand second column. As the population consists of a three-digit figure, read three-digit columns, i.e., read 097, 542, 422, 019 and so on. All the numbers within the range of 1 to 500 may be picked out. Then the sample will consist of:

097, 422, 019, 065. 060, 269 and so on. In the above reading, 542, 807, 573, etc., are rejected because they are over 500.

When the researcher reaches the bottom of a column, he can simply move one digit to the right and start at the top of the column again, and read numbers in three-digits: 973, 420, 226, 190, 079 and so on.

The main advantage of the use of a Table of Random Numbers are:

Easy to use and ready accessibility

The Table of Random numbers is ideal for obtaining a random sample from relatively small populations. When populations are quite large say lakhs or crores, drawing numbers from the table becomes tedious.

3) Use of Computer

If the population is very large and if computer facilities are available, a computer may be used for drawing a random sample. The computer can be programmed to print out a series of random numbers as the researcher desires.

TYPES OF SAMPLING

A) Probability Sampling

1) Simple Random Sampling

This sampling technique gives each element an equal and independent chance of being selected. An equal chance means equal probability of selection, e.g., in a population of 300, each element theoretically has $1/300^{th}$ chance of being selected. In a population of

1000, each element has 1/1000th chance of being selected. Equal probability selection method is described as Epsem sampling. An independent choice means that the draw of one element will not affect the chances of other elements being selected.

Where some elements are purposely excluded from the sample, the resulting sample is not a random one, Hence, all elements should be included in the sample frame to draw a random sample.

Merits

1. All elements in the population have an equal chance of being selected
2. Of all the probability sampling techniques, simple random sampling is the easiest to apply
3. It is the most simple type of probability sampling to understand
4. It does not required a prior knowledge of the true composition of the population
5. The amount of sampling error associated with any sample drawn can easily be computed

Demerits

1. It is often impractical, because of non-availability of population list, or of difficulty in enumerating the population. For example, it is difficult to get a current accurate list of households in a city of a list of landless rural agricultural labourers who migrate from area to area in search of employment or a list of households of a nomadic tribe
2. The use of simple random sampling may be wasteful because we fail to use all of the known information about the population
3. This technique does not ensure proportionate representation to various groups constituting the population

4. The sampling error in this sampling is greater than that in other probability samples of the same size, because it is less precise than other methods
5. The size of the sample required to ensure its representativeness is equally larger under this type of sampling than under other random sampling techniques
6. A simple random design may be expensive in time and money

2) Stratified Random Sampling

This is an improved type of random sampling. In this method, the population is sub-divided into homogenous groups or strata, and from each stratum, random sample is drawn. For example university students may be divided on the basis of discipline, and each discipline group may again be divided into juniors and seniors; and the employees of a business undertaking may be divided into managers and non-managers and each of those two groups may be sub-divided into salary-grade wise strata.

a) Proportionate Stratified Sampling

This sampling involves drawing a sample from each stratum in proportion to the latter's share in the total population.

Specialization	No. of Students	Proportion of Each Stream
Production	40	0.4
Finance	20	0.2
Marketing	30	0.3
Rural Development	10	0.1
Total	100	1.0

Merits

1. It enhances the representativeness of the sample by giving proper representation to all sub-groups in the population

2. It gives higher statistical efficiency that the given by simple random sampling for a given sample size
3. It is easy to carry out this sample method
4. This method gives a self-weighing sample, the population mean can be estimated simply by calculating the sample mean

Demerits

1. A prior knowledge of the composition of the population and the distribution of the population characteristics are required to adopt this method
2. This method is very expensive in time and money. Of course its greater efficiency may offset the additional cost
3. The identification of the strata might lead to classification errors. Some elements maybe included into the wrong strata. This may vitiate the interpretation of survey results.

b) Disproportionate Stratified Sampling

This method does not give proportionate representation to strata (group). It necessarily involves giving over representation to some strata and under representation to others. There may be several disproportionate schemes. All strata may be given equal weight, even though their shares in the total population vary. Alternatively some substrata may be given greater weight and others lesser weight. When is such disproportionate weighing preferable? **Example :** Drawing one per cent as sample irrespective on the numbers of members in the sample

Merits

1. It is less time consuming compared with proportionate sampling, because the researcher is not necessarily concerned about the

proportionate representativeness of his resulting sample as in the latter method

2. It facilitates giving appropriate weighting to particular groups, which are small but more important

Demerits

1. This method does not give each stratum proportionate representation. Hence, the resulting sample may be less representative
2. This method requires a prior knowledge of the composition of the population, which is not always possible
3. This method is also subject to classification errors. It is possible that the researcher may misclassify certain elements
4. Though disproportionate sampling is a means for developing an optimal stratification scheme, its practical feasibility is doubtful because one generally does not know the relative variability in the strata nor the relative costs

3) Systematic Sampling

This method of sampling is an alternative to random sampling. It consists of taking every K^{th} item in the population after a random start with an item form 1 to K. **For example**, suppose it is desired to select a sample of 20 students, from a list of 300 students, divide the population total of 300 by 20, the quotient is 15. Select a number at random between 1 and 15, using lottery method or a table of random numbers. Suppose the selected number is 9. Then the students numbered 9, 24, 39 are selected as the sample.

As the interval between sample units is fixed, this method is also known as fixed interval method.

Merits

1. It is much simpler than random sampling. It is easy to use
2. It is easy to instruct to field investigators to use this method
3. This method may require less time. A researcher operating on a limited time schedule will prefer this method
4. This method is cheaper than simple random sampling
5. It is easier to check whether every 'k'th has been included in the sample
6. Sample is spread evenly over the population
7. It is statistically more efficient than a simple random sample when population elements are ordered chronologically, by size, class, etc., Then systematic sampling gives a better representative sample

Demerits

1. This method ignores all elements between two 'k'th elements selected. Further, except the first element, other selected elements are not chosen at random. Hence, this sampling cannot be considered to be a probably sampling in the strict sense of the term
2. As each element does not have an equal chance of being selected, the resulting sample is not a random one. For studies aiming at estimation or generalizations, this disadvantage would be serious one
3. This method may sometimes give a biased sample. If by chance, several 'k' th elements chosen represent a particular group, that group would be over-represented in the sample

4) Cluster Sampling

Where the population elements are scattered over a wider area and a list of population elements is not readily available, the use of simple or stratified random sampling method would be too expensive and time consuming. In such cases cluster sampling is usually adopted.

Cluster sampling means random selection of sampling units consisting of population elements. Each such sampling unit is a cluster of population elements. Then from each selected sampling unit, a sample of population elements is drawn by either simple random selection or stratified random selection.

Example: Suppose a researcher wants to select a random sample of 1000 households out of 40000 estimated households in a city for a survey. A direct sample of individual households would be difficult to select, because a list of households does not exist and would be too costly to prepare. Instead, he can select a random sample of a few blocks / wards. The number of blocks to be selected depends upon the average number of estimated households per block. Suppose the average number of households per block is 200, then 5 blocks comprise the sample. Since the number of households per block varies, the actual sample size depends on the block which happen to be selected. Alternatively, he can draw a sample of more blocks and from each blocks a certain number of households may be selected by systematic sampling.

Merits

1. This method is much easier and more convenient to apply when large populations are studied or large geographical areas are covered. Even a ready list of population elements is not necessary. A researcher can simply draw a random sample of geographical sections and adopt single or multistage sampling depending on the vastness of the area covered by the study
2. The cost of this method is much less when compared with other sampling methods

3. This method promotes the convenience of field work at it could be done in compact places
4. Sampling under this method does not require more time
5. Units of study can be readily substituted for other units within the same random section
6. This method is flexible. Where it involves multistage sampling, it is possible to employ different types of sampling in successive stages

Demerits

1. The cluster size may vary and this variation could increase the bias of the resulting sample. For example, if the researcher were to interview all adults in households in each selected street the number of adults would vary from house to house. There would be certain bias resulting from the large coverage of big families
2. The sampling error in this method of sampling is greater. Thus, this method is statistically less efficient than other probability sampling methods
3. Adjacent units of study (e.g. households) tend to have more similar characteristics than do units distantly apart. This affects the 'representativeness' of the sample and this effect is reflected in a greater sampling error.

5) Area Sampling

This is an important form of cluster sampling. In larger field surveys, clusters consisting of specific geographical areas like districts, taluks, villages or blocks in a city are randomly drawn. As the geographical areas are selected as sampling units in such cases, their sampling is called area sampling. It is not a separate method of sampling, but forms a part of cluster sampling.

In a country like India where a state (previously known as province) is divided into districts, districts into talukas and talukas into towns and villages, area sampling is done on the basis of these administrative units in multi-stages.

6) Multi-stage Sampling

In this method, sampling is carried out in two or more stages. The population is regarded as being composed of a number of first stage sampling units. Each of them is made up of a number of second stage units and so forth. That is, at each stage, a sampling unit is a cluster of the sampling units of the subsequent stage. First, a sample of the first stage sampling units is drawn, then from each of the selected first stage sampling unit, a sample of the second stage sampling units is drawn. The procedure continues down to the final sampling units or population elements. Appropriate random sampling method is adopted at each stage.

Merits

It results in concentration of fieldwork in compact small areas and consequently in a saving of time, labour and money

1. It is more convenient, efficient and flexible than single-stage sampling
2. It obviates the necessity of having a sampling frame covered the entire population

Demerits

The major disadvantage of the multi-stage sampling is that the procedure of estimating sampling error and cost advantage is

complicated. It is difficult for a non-statistician to follow estimation procedure.

B. Non-Probability Sampling

1) Convenience Sampling

This is non-probability sampling. It means selecting sample units in a just 'hit an miss' fashion. E.g. Interviewing people whom we happen to meet. This sampling also means selecting whatever sampling units are conveniently available e.g. a teacher may select students in his class. This method is also known as accidental sampling because the respondents whom the researcher meets accidentally are included in the sample.

Merits

1. Cheapest and simplest
2. It does not require a list of population
3. It does not require any statistical expertise

Demerits

1. Convenience sampling is highly biased, because of the researcher's subjectivity, and so it does not yield a representative sample
2. It is the least reliable sampling method. There is no way of estimating the representativeness of the sample
3. The findings cannot be generalized

2) Purposed or Judgement Sampling

This method means deliberate selection of sample units that conform to some pre-determined criteria. This is known as judgement sampling. This involves selection of cases which we judge as the most appropriate

ones for the given study. It is based on the judgement of the researcher or some expert. It does not aim at securing a cross section of a population.

The chance that a particular case be selected for the sample depends on the subjective judgement of the researcher. For example, A researcher may deliberately choose industrial undertakings in which quality circles are believed to be functioning successfully and undertakings in which quality circles are believed to be a total failure

Merits

1. It is less costly and more convenient
2. It guarantees inclusion of relevant elements in the sample. Probability sampling plans cannot give such guarantee

Demerits

1. This does not ensure the representativeness of the sample
2. This is less efficient for generalizing when compared with random sampling
3. This method requires more prior extensive information about the population one studies. Without such information, it is not possible to adjudge the suitability of the sample items to be selected
4. The method does not lend itself for using inferential statistics, because, this sampling does not satisfy the underlying assumption of randomness.

3) Quota Sampling

This is a form of convenient sampling involving selection of quota groups of accessible sampling units by traits such as sex, age, social class etc., when the population is known to consist of various categories by sex, age, religion, social class, etc., in specific proportions, each

investigator may be given an assignment of quota groups specified by the pre-determined traits in specific proportions. He can then select accessible persons, belonging to those quota groups in the area assigned to him.

Example

Sex	**Numbers**	**Age**		**Social Class**	
Male	11	20-40	5	Higher	3
Female	9	41-50	8	Middle	10
	20	51-60	4	Lower	7
		Above 60	3		**20**
			20		

Merits

1. It is considerably less costly than probability sampling
2. It takes less time
3. There is no need for a list of population. Thus, quota sampling is a suitable method of sampling a population for which no suitable frame is available
4. Field work can easily be organized. Strict supervision need not be required

Demerits

1. It may not yield a precise representative sample, and it is impossible to estimate sampling error. The findings, therefore, are not generalizable to any significant extent
2. Interviewers may tent to choose the most accessible persons; they may ignore slums or areas difficult to reach. Thus, they may fail to secure a representative sample within their quota groups
3. Strict control of field work is difficult
4. The quota of sampling is subject to higher degree of classification error, because the investigators are likely to base their classification of

respondents' social status and economic status mostly on their impressions about them

5. It is difficult for sampling on more than three variable dimensions. This is because the number of categories to be selected is a multiplication of the number of values in each variable. For instances, if we want to sample proportionate number of persons by sex, social status and age and these variables consist of two, three and three categories respectively.

4) Snowball Sampling

This is the colourful name for a technique of building up a list or a sample of a special population by using an initial set of its members as informants. For example, if a researcher wants to study the problem faced by Indians through some source like Indian Embassy. Then he can ask each one of them to supply names of other Indians known to them, and continue this procedure until he gets an exhaustive list from which he can draw a sample or make a census survey.

This sampling technique may also be used in socio-metric studies. For example, the members of a social group may be asked to name the persons with whom they have social contacts, each one of the persons so named may also be asked to do so, and so on. The researcher may thus get a constellation of associates and analyse it.

Merits

1. It is very useful in studying social groups, informal group in a formal organization, and diffusion of information among professionals of various kinds
2. It is useful for smaller populations for which no frames are readily available

Demerits

1. The major disadvantages of snowball sampling is that it does not allow the use of probability statistical methods. Elements included are dependent on the subjective choice of the original selected respondents
2. It is difficult to apply this method when the population is large
3. It does not ensure the inclusion of all elements in the lists

SAMPLING ERROR

A survey aims at estimating or inferring selected population characteristics or parameters by studying either population or a sample of the population. The research results may either from the 'true values' of the parameters under study. Such differences are known as Errors and Biases. The errors of a survey may be classified into (a) Sampling Errors (b) Sampling Biases (c) Non-sampling errors and (d) Non-sampling biases.

1) Sampling Errors

The errors which arise because of studying only a part of the total population are called sampling errors. These may arise due to non-representativeness of the samples and the inadequacy of sample size. When several samples are drawn from a population, their results would not be identical. The degree of variation of sample results is measured by standard deviation and it is known as the standard error of the concerned statistic. As sample size increases the magnitude of the error decreases. Sample size and sampling error are thus negatively correlated.

2) Sampling Biases

The average of the estimates of a population parameter derived from an infinite number of samples is called the expected value of the

estimator. The difference between this value and 'true value' of the parameter is the bias. Bias may arise (1) if the sampling is done by a non-random method. (2) if the sampling frame is incomplete or inaccurate and (3) if some sections of the population are not available / refuse to cooperate. Any of these factors will cause non-compensating errors which cannot be reduced by an increase in sample size. The only sure way of avoiding bias arising through the sampling method is to use a random method. Randomness is an essential part of the protection against selection bias.

3) Non-Sampling Errors

These are errors which arise from sources other than sampling. They include errors of observation, errors of measurement and errors of responses. Data are collected through the methods of observation or interviewing. The physical procedures of observation or interviewing are subject to imperfection which cause errors. Measurement errors consist of errors in processing and analysis. Errors of response include incorrect responses of the respondents, mistakes in noting their response etc.,

4) Non-Sampling Biases

These biases pose problems for scientific measurement. They affect both the population sample value and account for the difference between the population value and the true value. They consist of biases of observation and non-observation, response biases and process biases. Biases of observation are caused by obtaining and recording observations incorrectly. Non-observation biases arise from failure to obtain observations on some segments of the population due to either non-coverage or non-response. The latter may be due to refusals, non-at-homes, lost forms, etc., response biases consist of biases arising from imperfections in field

observation or interviewing. Processing biases are produced during coding, tabulating and computing.

5) Total Error

In sampling theory, a popular model combines sampling and non-sampling errors and biases into the Total error. This total error is the square root of the sums of squares of variable errors and squares of bias. It is often called the root means square error. The variable error are caused only by sampling errors, and VE equals the standard error of sampling. Bias is mostly caused by measurement biases. The total error depends on the length of both the legs. The sampling error / standard error leg can be shortened by improving the sample design and by increasing the sample size. The length of biases leg may be reduced by improving the tool of data collection, the precision of method of data collection, filed work, coding, processing and analysis. The measurement of sampling error does not pose much problem but the measurement of non-sampling errors require special procedures and it is a costly effect. Hence, the reduction of non-sampling error is a challenge to the researchers.

* * * * *

POSSIBLE QUESTIONS

UNIT-II

PART A

One Mark

Online Examination

PART B

Eight Marks

1. Describe the steps in sampling design?
2. Explain the various types of probability sample designs?
3. Explain the criteria for selection of good sampling design.
4. Write the merits and demerits of sampling.
5. Describe the steps in sampling design?
6. Explain the technique and usefulness of multistage sampling.
7. Examine the factors which determine the sample size.
8. Define random sampling and explain the usefulness.
9. Explain the different methods of sampling.
10. Explain the characteristics of a good sample design.

KARPAGAM ACADEMY OF HIGHER EDUCATION

(DEEMED TO BE UNIVERSITY)

(ESTABILISHED UNDER SEC 3 OF UGC ACT 1956)

DEPARTMENT OF COMMERCE

III -B.COM RESEARCH METHODOLOGY-15CMU603B

UNIT-II SAMPLING

S.NO	QUESTIONS	OPTION I	OPTION II	OPTION III	OPTION IV	ANSWER
1	All items in any field of inquiry constitute a ___	Universe	Sampling	Area Sampling	Probability Sampling	Universe
2	In __ method all items are covered, no element of chance is left and highest accuracy is	Census	Sample	Probability	Non-Probability	Census
3	____is the only institution which can adopt census method of data collection	Private	Public	Government	Educational	Government
4	In India population census is carried out ___	Once in a decade	Once in Five Years	Once in Fifteeen Years	Yearly Once	Once in a decade
5	Studying only a part of total population in known as ___	Sampling	Census	Field Enquiry	Observation	Sampling
6	___contains the names of all items of universe	Source List	Sampling Unit	Sample Design	Census Data	Source List
7	Source list otherwise known as ___	Sampling Unit	Sample Design	Census Data	Sampling Frame	Sampling Frame

S.NO	QUESTIONS	OPTION I	OPTION II	OPTION III	OPTION IV	ANSWER
8	___refers to the number of items to be selected from the universe to constitute a sample	Size of Sample	Parameters of Interest	Source List	Sampling Unit	Size of Sample
9	Sub-group of population is known as ___	Sample	Field Enquiry	Observation	Census	Sample
10	In ___sample, every unit in the population has equal chances for being selected as a sample	Probability	Non-Probability	Area	Judgement	Probability
11	In __sampling, units in the population has unequal or zero chances for being selected as a	Probability	Non-Probability	Area	Judgement	Non-Probability
12	____error is the square root of the sums of squares of variable errors and squares of bias	Sampling	Total	Non-Sampling	Bias	Total
13	___is often called the root means square error	Sampling error	Total error	Non-Sampling error	Bias	Total error
14	The errors which arise because of studying only a part of the total population are called	Non-Sampling	Sampling	Bias	Total	Sampling
15	Degree of variation of sample results is measured by __	Standard Deviation	Mean	Median	Chisquare	Standard Deviation
16	As sample size increases the magnitude of the ____error decreases	Non-Sampling	Sampling	Total	Probability	Sampling
17	___may arise if some sections of the population are not available / refuse to cooperate	Bias	Sampling error	Non-Sampling error	Non-Sampling bias	Bias

S.NO	QUESTIONS	OPTION I	OPTION II	OPTION III	OPTION IV	ANSWER
18	___errors which arise from sources other than sampling	Sampling	Non-Sampling	Sampling Bias	Non-Sampling bias	Sampling
19	Errors of observation, Errors of measurement and Errors of responses leads to ___	Bias	Sampling error	Non-Sampling error	Non-Sampling bias	Non-Sampling error
20	__is a result biases of observation and non-observation, response biases	Non-Sampling error	Non-Sampling Bias	Sampling Error	Sampling Bias	Non-Sampling Bias
21	Collecting data by reference through reference is known as ___	Quota Sampling	Snowball Sampling	Convenient Sampling	Judgement Sampling	Snowball Sampling
22	__ is a form of convenient sampling involving selection of quota groups of accessible	Judgement Sampling	Quota Sampling	Snowball Sampling	Multistage Sampling	Quota Sampling
23	___method means deliberate selection of sample units that conform to some pre-	Judgement Sampling	Quota Sampling	Snowball Sampling	Multistage Sampling	Judgement Sampling
24	Judgment Sampling otherwise known as __	Purposive Sampling	Quota Sampling	Snowball Sampling	Multistage Sampling	Purposive Sampling
25	__ means selecting sample units in a just 'hit an miss' fashion	Quota Sampling	Snowball Sampling	Convenient Sampling	Judgement Sampling	Convenient Sampling
26	In ___sampling, sampling is carried out in two or more stages	Multi-stage	Cluster	Proportionate	Random	Multi-stage
27	__sampling is an important form of cluster sampling	Area	Multi-stage	Proportionate	Systematic	Area

S.NO	QUESTIONS	OPTION I	OPTION II	OPTION III	OPTION IV	ANSWER
28	Grouping of population based on some common characteristics is known as __	Cluster	Area	Multi-stage	Random	Cluster
29	___method is also known as fixed interval method	Systematic Sampling	Cluster Sampling	Snowball Sampling	Disproportionate Sampling	Systematic Sampling
30	Taking every K^{th} item in the population after a random start with an item form 1 to K is known as ___	Systematic Sampling	Cluster Sampling	Snowball Sampling	Disproportionate Sampling	Systematic Sampling
31	___method does not give proportionate representation to strata	Disproportionate Stratified	Stratified Sampling	Sampling with Replcaement	Sampling without Replacement	Disproportionate Stratified Sampling
32	__method enhances the representativeness of the sample by giving proper	Disproportionate Stratified	Proportionate Stratified Sampling	Snowball Sampling	Quota Sampling	Proportionate Stratified Sampling
33	In ___method, the population is sub-divided into homogenous groups or strata, and from	Random Sampling	Stratified Random Sampling	Multi-stage Sampling	Purposive Sampling Method	Stratified Random Sampling
34	Lottery method is an example of ___ method of sampling	Simple Random	Stratified Random	Systematic	Cluster	Simple Random
35	When the field of inquiry is large, ___method becomes difficult to adopt because of the	Sample	Probability	Non-Probability	Census	Census
36	Studying a proportion of population is known as ____	Census	Sampling	Field Enquiry	Observation	Sampling
37	When every unit in the population has equal chances for being selected as a sample	Probability	Non-Probability	Area	Judgement	Probability

S.NO	QUESTIONS	OPTION I	OPTION II	OPTION III	OPTION IV	ANSWER
38	When respondents are selected according to the desire of enumerator it is known as ___	Probability	Non-Probability	Area	Judgement	Non-Probability
39	In sampling theory, a popular model combines sampling and non-sampling errors and	Sampling error	Total error	Non-Sampling error	Bias	Total error
40	Degree of variation of sample results is measured by __	Standard Deviation	Mean	Median	Chisquare	Standard Deviation
41	Degree of variation of sample results is known as the ____error of the concerned	Standard Deviation	Total	Non-Probability	Bias	Standard
42	___may arise if the sampling is done by a non-random method	Bias	Sampling error	Non-Sampling error	Non-Sampling bias	Bias
43	The physical procedures of observation or interviewing are subject to imperfection which	Bias	Sampling error	Non-Sampling error	Non-Sampling bias	Non-Sampling error
44	___pose problems for scientific measurement	Non-Sampling error	Non-Sampling Bias	Sampling Error	Sampling Bias	Non-Sampling Bias
45	___method means deliberate selection of sample units that conform to some pre-	Purposive Sampling	Quota Sampling	Snowball Sampling	Multistage Sampling	Purposive Sampling
46	When investigator has been restricted to collect fixed number of responses from a	Judgement Sampling	Quota Sampling	Snowball Sampling	Multistage Sampling	Quota Sampling
47	Interviewing people whom we happen to meet is known as ___sampling method of data	Probability	Convenient	Snowball Sampling	Judgement	Convenient

S.NO	QUESTIONS	OPTION I	OPTION II	OPTION III	OPTION IV	ANSWER
48	__involves selection of cases which we judge as the most appropriate ones for the given	Judgement Sampling	Quota Sampling	Snowball Sampling	Multistage Sampling	Judgement Sampling
49	___is not a separate method of sampling, but forms a part of cluster sampling.	Systematic Sampling	Area Sampling	Cluster Sampling	Disproportionate Sampling	Area Sampling
50	Where the population elements are scattered over a wider area and a list of population	Cluster	Area	Multi-stage	Random	Cluster
51	__method of sampling is an alternative to random sampling	Systematic Sampling	Cluster Sampling	Snowball Sampling	Disproportionate Sampling	Systematic Sampling
52	__ method of sampling necessarily involves giving over representation to some strata	Stratified Sampling	Sampling with Replcaement	Sampling without Replacement	Disproportionate Stratified	Disproportionate Stratified Sampling
53	___sampling involves drawing a sample from each stratum in proportion to the latter's share	Proportionate	Disproportionate	Sampling with Replcaement	Sampling without Replacement	Proportionate
54	__sampling is the most simple type of probability sampling to understand	Systematic	Cluster	Stratified Random	Simple Random	Simple Random
55	____type of inquiry involves a great deal of time, money and energy	Sample	Probability	Non-Probability	Census	Census
56	Measurement errors is an example of ___	Non-Sampling error	Non-Sampling Bias	Sampling Error	Sampling Bias	Non-Sampling error
57	___sampling also means selecting whatever sampling units are conveniently available		Convenient	Snowball Sampling	Judgement	Convenient

S.NO	QUESTIONS	OPTION I	OPTION II	OPTION III	OPTION IV	ANSWER
58	As the geographical areas are selected as sampling units in such cases, their sampling is	Multi-stage	Area	Systematic	Proportionate	Area
59	__sampling means random selection of sampling units consisting of population	Cluster	Area	Multi-stage	Random	Cluster
60	___method does not give proportionate representation to a group	Stratified Sampling	Sampling with Replcaement	Sampling without Replacement	Disproportionate Stratified	Disproportionate Stratified Sampling

UNIT III

SYLLABUS

Sources of Data - Methods of Data Collection – Primary Data – Interview Method – Observation Method – Questionnaire – Schedule – Secondary Data – Processing of Data – Editing – Coding – Classification – Tabulation.

* * * * *

METHODS OF DATA COLLECTION

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

1) Primary Data

Data directly collected by the researcher, with respect to problem under study, is known as primary data. Primary data is also the first

hand data collected by the researcher for the immediate purpose of the study.

2) Secondary Data

Secondary data are statistics that already exists. They have been gathered not for immediate use. This may be described as "Those data that have been compiled by some agency other than the user".

INTERVIEW METHOD

The interview method of collecting data involves presentation of oral-verbal stimuli and reply in terms of oral-verbal responses. This method can be used through personal interviews and, if possible, through telephone interviews. Interview is one of the popular methods of data collection. The term interview can be dissected into two terms as, 'inter' and 'view.'. The essence of interview is that one mind tries to read the other. The interviewer tries to assess the interviewed in terms of the aspects studied or issues analysed.

Personal interview method requires a person known as the interviewer asking questions generally in a face-to-face contact to the other person or persons. (At times the interviewee may also ask certain questions and the interviewer responds to these, but usually the interviewer initiates the interview and collects the information.) This sort of interview may be in the form of direct personal investigation or it may be indirect oral investigation. In the case of direct personal investigation the interviewer has to collect the information personally from the sources concerned. He has to be on the spot and has to meet people from whom data have to be collected. This method is particularly suitable for intensive investigations.

But in certain cases it may not be possible or worthwhile to contact directly the persons concerned or on account of the extensive scope of enquiry, the direct personal investigation technique may not be used. In such cases an indirect oral examination can be conducted under which the interviewer has to cross-examine other persons who are supposed to have knowledge about the problem under investigation and the information, obtained is recorded. Most of the commissions and committees appointed by government to carry on investigations make use of this method.

Types of Interview

1) Structured Interview

The method of collecting information through personal interviews is usually carried out in a structured way. As such we call the interviews as structured interviews. Such interviews involve the use of a set of predetermined questions and of highly standardised techniques of recording. Thus, the interviewer in a structured interview follows a rigid procedure laid down, asking questions in a form and order prescribed. But in case of descriptive studies, we quite often use the technique of structured interview because of its being more economical, providing a safe basis for generalisation and requiring relatively lesser skill on the part of the interviewer.

2) Unstructured Interview

As against it, the unstructured interviews are characterised by a flexibility of approach to questioning. Unstructured interviews do not follow a system of pre-determined questions and standardised techniques of recording information. In a non-structured interview, the interviewer is allowed much greater freedom to ask, in case of need,

supplementary questions or at times he may omit certain questions if the situation so requires. He may even change the sequence of questions. He has relatively greater freedom while recording the responses to include some aspects and exclude others. But this sort of flexibility results in lack of comparability of one interview with another and the analysis of unstructured responses becomes much more difficult and time-consuming than that of the structured responses obtained in case of structured interviews. Unstructured interviews also demand deep knowledge and greater skill on the part of the interviewer. Unstructured interview, however, happens to be the central technique of collecting information in case of exploratory or formulative research studies.

3) Focused Interview

Focused interview is meant to focus attention on the given experience of the respondent and its effects. Under it the interviewer has the freedom to decide the manner and sequence in which the questions would be asked and has also the freedom to explore reasons and motives. The main task of the interviewer in case of a focussed interview is to confine the respondent to a discussion of issues with which he seeks conversance. Such interviews are used generally in the development of hypotheses and constitute a major type of unstructured interviews.

4) Clinical Interview

The clinical interview is concerned with broad underlying feelings or motivations or with the course of individual's life experience. The method of eliciting information under it is generally left to the interviewer's discretion.

5) Non-directive Interview

In case of non-directive interview, the interviewer's function is simply to encourage the respondent to talk about the given topic with a bare minimum of direct questioning. The interviewer often acts as a catalyst to a comprehensive expression of the respondents' feelings and beliefs and of the frame of reference within which such feelings and beliefs take on personal significance.

Merits of Interview

1. More information and that too in greater depth can be obtained.
2. Interviewer by his own skill can overcome the resistance, if any, of the respondents; the interview method can be made to yield an almost perfect sample of the general population.
3. There is greater flexibility under this method as the opportunity to restructure questions is always there, specially in case of unstructured interviews.
4. Observation method can as well be applied to recording verbal answers to various questions.
5. Personal information can as well be obtained easily under this method.
6. Samples can be controlled more effectively as there arises no difficulty of the missing returns; non-response generally remains very low.
7. The interviewer can usually control which person(s) will answer the questions. This is not possible in mailed questionnaire approach. If so desired, group discussions may also be held.
8. The interviewer may catch the informant off-guard and thus may secure the most spontaneous reactions than would be the case if mailed questionnaire is used.
9. The language of the interview can be adopted to the ability or educational level of the person interviewed and as such misinterpretations concerning questions can be avoided.

10. The interviewer can collect supplementary information about the respondent's personal characteristics and environment which is often of great value in interpreting results.

Demerits or Weakness of Interview

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

Pre-requisites and Basic Tenets of Interviewing

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

OBSERVATION METHOD

The observation method is the most commonly used method specially in studies relating to behavioural sciences. In a way we all

observe things around us, but this sort of observation is not scientific observation. Observation becomes a scientific tool and the method of data collection for the researcher, when it serves a formulated research purpose, is systematically planned and recorded and is subjected to checks and controls on validity and reliability. Under the observation method, the information is sought by way of investigator's own direct observation without asking from the respondent. For instance, in a study relating to consumer behaviour, the investigator instead of asking the brand of wrist watch used by the respondent, may himself look at the watch.

Advantages of Observation

1. The main advantage of this method is that subjective bias is eliminated, if observation is done accurately.
2. Secondly, the information obtained under this method relates to what is currently happening; it is not complicated by either the past behaviour or future intentions or attitudes.
3. Thirdly, this method is independent of respondents' willingness to respond and as such is relatively less demanding of active cooperation on the part of respondents as happens to be the case in the interview or the questionnaire method.
4. This method is particularly suitable in studies which deal with subjects (i.e., respondents) who are not capable of giving verbal reports of their feelings for one reason or the other

Limitations of Observation

1. Firstly, it is an expensive method.
2. Secondly, the information provided by this method is very limited.

3. Thirdly, sometimes unforeseen factors may interfere with the observational task.
4. At times, the fact that some people are rarely accessible to direct observation creates obstacle for this method to collect data effectively.

Precautions before Adopting Observation Method of Data Collection

- While using this method, the researcher should keep in mind things like:
- What should be observed?
- How the observations should be recorded? Or
- How the accuracy of observation can be ensured?

Types of Observation

1) Structured Observation

In case the observation is characterised by a careful definition of the units to be observed, the style of recording the observed information, standardised conditions of observation and the selection of pertinent data of observation, then the observation is called as structured observation. Structured observation is considered appropriate in descriptive studies

2) Unstructured Observation

But when observation is to take place without these characteristics to be thought of in advance, the same is termed as unstructured observation. Whereas in an exploratory study the observational procedure is most likely to be relatively unstructured.

3) Participant Observation

We often talk about participant and non-participant types of observation in the context of studies, particularly of social sciences. This distinction depends upon the observer's sharing or not sharing the life of the group he is observing. If the observer observes by making himself, more or less, a member of the group he is observing so that he can experience what the members of the group experience, the observation is called as the participant observation.

Merits of Participant Observation

1. The researcher is enabled to record the natural behaviour of the group.
2. The researcher can even gather information which could not easily be obtained if he observes in a disinterested fashion.
3. The researcher can even verify the truth of statements made by informants in the context of a questionnaire or a schedule.

Demerits of Participant Observation

- The observer may lose the objectivity to the extent he participates emotionally; the problem of observation-control is not solved; and it may narrow-down the researcher's range of experience.

4) Non-Participant Observation

But when the observer observes as a detached emissary without any attempt on his part to experience through participation what others feel, the observation of this type is often termed as non-participant observation. (When the observer is observing in such a manner that his presence may be unknown to the people he is observing, such an observation is described as disguised observation.)

5) Uncontrolled Observation

If the observation takes place in the natural setting, it may be termed as uncontrolled observation, In non-controlled observation, no attempt is made to use precision instruments. The major aim of this type of observation is to get a spontaneous picture of life and persons. It has a tendency to supply naturalness and completeness of behaviour, allowing sufficient time for observing it. The main pitfall of non-controlled observation is that of subjective interpretation. There is also the danger of having the feeling that we know more about the observed phenomena than we actually do. Uncontrolled observation is resorted to in case of exploratory researches.

6) Controlled Observation

When observation takes place according to definite pre-arranged plans, involving experimental procedure, the same is then termed controlled observation. But in controlled observation, we use mechanical (or precision) instruments as aids to accuracy and standardisation. Such observation has a tendency to supply formalised data upon which generalisations can be built with some degree of assurance. Generally, controlled observation takes place in various experiments that are carried out in a laboratory or under controlled conditions

QUESTIONNAIRE

This method of data collection is quite popular, particularly in case of big enquiries. It is being adopted by private individuals, research workers, private and public organisations and even by governments. In this method a questionnaire is sent (usually by post) to the persons concerned with a request to answer the questions and return the questionnaire. A questionnaire consists of a number of questions printed or typed in a definite order on a form or set of forms. The questionnaire is

mailed to respondents who are expected to read and understand the questions and write down the reply in the space meant for the purpose in the questionnaire itself. The respondents have to answer the questions on their own.

The method of collecting data by mailing the questionnaires to respondents is most extensively employed in various economic and business surveys

Merits of Questionnaire

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

Demerits of Questionnaire

1. Low rate of return of the duly filled in questionnaires; bias due to no-response is often indeterminate.
2. It can be used only when respondents are educated and cooperating.
3. The control over questionnaire may be lost once it is sent.
4. There is inbuilt inflexibility because of the difficulty of amending the approach once questionnaires have been despatched.

5. There is also the possibility of ambiguous replies or omission of replies altogether to certain questions; interpretation of omissions is difficult.
6. It is difficult to know whether willing respondents are truly representative.
7. This method is likely to be the slowest of all.

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

Main Aspects of a Questionnaire

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

1) General form

So far as the general form of a questionnaire is concerned, it can either be structured or unstructured questionnaire. Structured questionnaires are those questionnaires in which there are definite,

concrete and pre-determined questions. The questions are presented with exactly the same wording and in the same order to all respondents. Resort is taken to this sort of standardisation to ensure that all respondents reply to the same set of questions. The form of the question may be either closed (i.e., of the type 'yes' or 'no') or open (i.e., inviting free response) but should be stated in advance and not constructed during questioning. Structured questionnaires may also have fixed alternative questions in which responses of the informants are limited to the stated alternatives. Thus a highly structured questionnaire is one in which all questions and answers are specified and comments in the respondent's own words are held to the minimum. When these characteristics are not present in a questionnaire, it can be termed as unstructured or non-structured questionnaire. More specifically, we can say that in an unstructured questionnaire, the interviewer is provided with a general guide on the type of information to be obtained, but the exact question formulation is largely his own responsibility and the replies are to be taken down in the respondent's own words to the extent possible; in some situations tape recorders may be used to achieve this goal.

Structured questionnaires are simple to administer and relatively inexpensive to analyse. The provision of alternative replies, at times, helps to understand the meaning of the question clearly. But such questionnaires have limitations too. For instance, wide range of data and that too in respondent's own words cannot be obtained with structured questionnaires. They are usually considered inappropriate in investigations where the aim happens to be to probe for attitudes and reasons for certain actions or feelings. They are equally not suitable when a problem is being first explored and working hypotheses sought. In such situations, unstructured questionnaires may be used effectively.

Then on the basis of the results obtained in pretest (testing before final use) operations from the use of unstructured questionnaires, one can construct a structured questionnaire for use in the main study.

2) Question Sequence

In order to make the questionnaire effective and to ensure quality to the replies received, a researcher should pay attention to the question-sequence in preparing the questionnaire. A proper sequence of questions reduces considerably the chances of individual questions being misunderstood. The question-sequence must be clear and smoothly-moving, meaning thereby that the relation of one question to another should be readily apparent to the respondent, with questions that are easiest to answer being put in the beginning. The first few questions are particularly important because they are likely to influence the attitude of the respondent and in seeking his desired cooperation. The opening questions should be such as to arouse human interest. The following type of questions should generally be avoided as opening questions in a questionnaire:

- Questions that put too great a strain on the memory or intellect of the respondent;
- Questions of a personal character
- Questions related to personal wealth, etc.

Following the opening questions, we should have questions that are really vital to the research problem and a connecting thread should run through successive questions. Ideally, the question-sequence should conform to the respondent's way of thinking. Knowing what information is desired, the researcher can rearrange the order of the questions (this is possible in case of unstructured questionnaire) to fit

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

3) Question Formulation and Wording

With regard to this aspect of questionnaire, the researcher should note that each question must be very clear for any sort of misunderstanding can do irreparable harm to a survey. Question should also be impartial in order not to give a biased picture of the true state of affairs. Questions should be constructed with a view to their forming a logical part of a well thought out tabulation plan. In general, all questions should meet the following standards—(a) should be easily understood; (b) should be simple i.e., should convey only one thought at a time; (c) should be concrete and should conform as much as possible to the respondent's way of thinking. (For instance, instead of asking. "How many razor blades do you use annually?" The more realistic question would be to ask, "How many razor blades did you use last week?"

Multiple Choice Question and the Open-End Questions. In the former the respondent selects one of the alternative possible answers put to him, whereas in the latter he has to supply the answer in his own words. The question with only two possible answers (usually 'Yes' or 'No') can be taken as a special case of the multiple choice question, or can be named as a 'closed question.' There are some advantages and disadvantages of each possible form of question. Multiple choice or closed questions have the advantages of easy handling, simple to answer, quick and relatively inexpensive to analyse. They are most amenable to statistical analysis. Sometimes, the provision of alternative replies helps to make clear the meaning of the question. But the main drawback of fixed alternative questions is that of "putting answers in people's mouths" i.e., they may force a statement of opinion on an issue about which the respondent does not infact have any opinion. They are not appropriate when the issue under consideration happens to be a complex one and also when the interest of the researcher is in the exploration of a process. In such situations, open-ended questions which are designed to permit a free response from the respondent rather than one limited to certain stated alternatives are considered appropriate. Such questions give the respondent considerable latitude in phrasing a reply. Getting the replies in respondent's own words is, thus, the major advantage of open-ended questions. But one should not forget that, from an analytical point of view, open-ended questions are more difficult to handle, raising problems of interpretation, comparability and interviewer bias.

In practice, one rarely comes across a case when one questionnaire relies on one form of questions alone. The various forms complement each other. As such questions of different forms are included in one single questionnaire. For instance, multiple-choice questions constitute the basis of a structured questionnaire, particularly

in a mail survey. But even there, various open-ended questions are generally inserted to provide a more complete picture of the respondent's feelings and attitudes.

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

Essentials of Good Questionnaire

1. To be successful, questionnaire should be comparatively short and simple i.e., the size of the questionnaire should be kept to the minimum
2. Questions should proceed in logical sequence moving from easy to more difficult questions
3. Personal and intimate questions should be left to the end
4. Technical terms and vague expressions capable of different interpretations should be avoided in a questionnaire
5. Questions may be dichotomous (yes or no answers), multiple choice (alternative answers listed) or open-ended. The latter type of questions are often difficult to analyse and hence should be avoided in a questionnaire to the extent possible
6. There should be some control questions in the questionnaire which indicate the reliability of the respondent. For instance, a question

designed to determine the consumption of particular material may be asked first in terms of financial expenditure and later in terms of weight. The control questions, thus, introduce a cross-check to see whether the information collected is correct or not

7. Questions affecting the sentiments of respondents should be avoided
8. Adequate space for answers should be provided in the questionnaire to help editing and tabulation
9. There should always be provision for indications of uncertainty, e.g., "do not know," "no preference" and so on
10. Brief directions with regard to filling up the questionnaire should invariably be given in the questionnaire itself
11. Finally, the physical appearance of the questionnaire affects the cooperation the researcher receives from the recipients and as such an attractive looking questionnaire, particularly in mail surveys, is a plus point for enlisting cooperation.
12. The quality of the paper, along with its colour, must be good so that it may attract the attention of recipients

Types of Questionnaire

1) Structured non disguised Questionnaire

Here, questions are structured so as to get the facts. The interviewer will ask the questions strictly as per the pre arranged order. Structured, non disguised is widely used in market research. Questions are presented with exactly the same wording and same order to all the respondents. The reason for standardizing question is, to ensure that all respondents reply the same question. The purpose of the question is clear. The researcher wants the respondent to choose one of the five options given above. This type of questionnaire is easy to administer. The respondents have no difficulty in answering. Because it is structured, the

frame of reference is obvious. In a non-disguised type, the purpose of the questionnaire is known to the respondent.

2) Structured disguised Questionnaire

This type of questionnaire is least used in Marketing research. This type of Questionnaire is used to find, peoples' attitude, when a direct undisguised question produces a bias. In this type of questionnaire what comes out is "What does the respondent know rather than what he feels". Therefore attempt in this method is to find the respondent's attitude.

3) Non-Structured and disguised Questionnaire

The main objective is to conceal the topic of enquiry by using a disguised stimulus. Though the stimulus is standardized by researcher, respondent is allowed to answer in an unstructured manner. The assumption made here is that individuals reaction is an indication of respondent's basic perception. Projective techniques are examples of Non structured disguised technique. The techniques involve the use of a vague stimulus, that an individual is asked to expand or describe or build a story, three common types under this category are (a) Word association (b) Sentence completion (c) Story telling.

4) Non structured - Non disguised Questionnaire

Here the purpose of the study is clear, but the responses to the question is open ended. Example: "How do you feel about the cyber law currently in practice and its need for further modification"? The initial part of the question is constant. After presenting the initial question, the interview becomes very unstructured as the interviewer probes more deeply. Respondents subsequent answer determines the direction the interviewer takes next. The question asked by interviewer varies from

person to person. This method is called "Depth interview". The major advantage of this method is freedom permitted to the interviewer. By not restricting the respondents for a set of replies, the experienced interviewers will be above to get the information from the respondent fairly and accurately. The main disadvantage of this method of interviewing is that, it takes time, and respondents may not co-operate. Another disadvantage is that coding of open ended question may pose a challenge. E.g.: When a researcher asked the respondent "Tell me something about your experience in this hospital". The answer may be "Well, the nurses are "slow" to attend and Doctor is "rude". 'Slow' and 'rude' are different qualities needing separate coding. This type of interviewing is extremely helpful in exploratory studies.

SCHEDULES

This method of data collection is very much like the collection of data through questionnaire, with little difference which lies in the fact that schedules (proforma containing a set of questions) are being filled in by the enumerators who are specially appointed for the purpose. These enumerators along with schedules, go to respondents, put to them the questions from the proforma in the order the questions are listed and record the replies in the space meant for the same in the proforma.

In certain situations, schedules may be handed over to respondents and enumerators may help them in recording their answers to various questions in the said schedules. Enumerators explain the aims and objects of the investigation and also remove the difficulties which any respondent may feel in understanding the implications of a particular question or the definition or concept of difficult terms. This method requires the selection of enumerators for filling up schedules or

assisting respondents to fill up schedules and as such enumerators should be very carefully selected

The enumerators should be trained to perform their job well and the nature and scope of the investigation should be explained to them thoroughly so that they may well understand the implications of different questions put in the schedule. Enumerators should be intelligent and must possess the capacity of cross-examination in order to find out the truth. Above all, they should be honest, sincere, hardworking and should have patience and perseverance. This method of data collection is very useful in extensive enquiries and can lead to fairly reliable results. It is, however, very expensive and is usually adopted in investigations conducted by governmental agencies or by some big organisations. Population census all over the world is conducted through this method.

DIFFERENCE BETWEEN QUESTIONNAIRE AND INTERVIEW SCHEDULE

1. The questionnaire is generally sent through mail to informants to be answered as specified in a covering letter, but otherwise without further assistance from the sender. The schedule is generally filled out by the research worker or the enumerator, who can interpret questions when necessary.
2. To collect data through questionnaire is relatively cheap and economical since we have to spend money only in preparing the questionnaire and in mailing the same to respondents. Here no field staff required. To collect data through schedules is relatively more expensive since considerable amount of money has to be spent in appointing enumerators and in importing training to them. Money is also spent in preparing schedules.

3. Non-response is usually high in case of questionnaire as many people do not respond and many return the questionnaire without answering all questions. Bias due to non-response often remains indeterminate. As against this, non-response is generally very low in case of schedules because these are filled by enumerators who are able to get answers to all questions. But there remains the danger of interviewer bias and cheating.
4. In case of questionnaire, it is not always clear as to who replies, but in case of schedule the identity of respondent is known.
5. The questionnaire method is likely to be very slow since many respondents do not return the questionnaire in time despite several reminders, but in case of schedules the information is collected well in time as they are filled in by enumerators.
6. Personal contact is generally not possible in case of the questionnaire method as questionnaires are sent to respondents by post who also in turn return the same by post. But in case of schedules direct personal contact is established with respondents.
7. Questionnaire method can be used only when respondents are literate and cooperative, but in case of schedules the information can be gathered even when the respondents happen to be illiterate.
8. Wider and more representative distribution of sample is possible under the questionnaire method, but in respect of schedules there usually remains the difficulty in sending enumerators over a relatively wider area.
9. Risk of collecting incomplete and wrong information is relatively more under the questionnaire method, particularly when people are unable to understand questions properly. But in case of schedules, the information collected is generally complete and accurate as enumerators can remove the difficulties, if any, faced by respondents

in correctly understanding the questions. As a result, the information collected through schedules is relatively more accurate than that obtained through questionnaires.

10. The success of questionnaire method lies more on the quality of the questionnaire itself, but in the case of schedules much depends upon the honesty and competence of enumerators.

11. In order to attract the attention of respondents, the physical appearance of questionnaire must be quite attractive, but this may not be so in case of schedules as they are to be filled in by enumerators and not by respondents.

12. Along with schedules, observation method can also be used but such a thing is not possible while collecting data through questionnaires.

SECONDARY DATA

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

Secondary data may either be published data or unpublished data. Usually published data are available in: (a) various publications of the central, state are local governments; (b) various publications of foreign governments or of international bodies and their subsidiary organisations; (c) technical and trade journals; (d) books, magazines and newspapers; (e) reports and publications of various associations connected with business and industry, banks, stock exchanges, etc.; (f) reports prepared by research scholars, universities, economists, etc. in

different fields; and (g) public records and statistics, historical documents, and other sources of published information.

The sources of unpublished data are many; they may be found in diaries, letters, unpublished biographies and autobiographies and also may be available with scholars and research workers, trade associations, labour bureaus and other public/ private individuals and organisations. Researcher must be very careful in using secondary data. He must make a minute scrutiny because it is just possible that the secondary data may be unsuitable or may be inadequate in the context of the problem which the researcher wants to study. In this connection Dr. A.L. Bowley very aptly observes that it is never safe to take published statistics at their face value without knowing their meaning and limitations and it is always necessary to criticise arguments that can be based on them.

CHARACTERISTICS OF SECONDARY DATA

1) Reliability of Data

The reliability can be tested by finding out such things about the said data:
(a) Who collected the data? (b) What were the sources of data? (c) Were they colleted by using proper methods? (d) At what time were they collected? (e) Was there any bias of the compiler? (f) What level of accuracy was desired? Was it achieved?

2) Suitability of Data

The data that are suitable for one enquiry may not necessarily be found suitable in another enquiry. Hence, if the available data are found to be unsuitable, they should not be used by the researcher. In this context, the researcher must very carefully scrutinize the definition of various terms and units of collection used at the time of collecting the data from the

primary source originally. Similarly, the object, scope and nature of the original enquiry must also be studied. If the researcher finds differences in these, the data will remain unsuitable for the present enquiry and should not be used

3) Adequacy of Data

If the level of accuracy achieved in data is found inadequate for the purpose of the present enquiry, they will be considered as inadequate and should not be used by the researcher. The data will also be considered inadequate, if they are related to an area which may be either narrower or wider than the area of the present enquiry

PROCESSING OF DATA

The data, after collection, has to be processed and analyzed in accordance with the outline laid down for the purpose at the time of developing the research plan. This is essential for a scientific study and for ensuring that we have all relevant data for making contemplated comparisons and analysis. Technically speaking, processing implies editing, coding, classification and tabulation of collected data so that they are amenable to analysis. The term analysis refers to the computation of certain measures along with searching for patterns of relationship that exist among data-groups. Thus, "in the process of analysis, relationships or differences supporting or conflicting with original or new hypotheses should be subjected to statistical tests of significance to determine with what validity data can be said to indicate any conclusions". But there are persons (Selltiz, Jahoda and others) who do not like to make difference between processing and analysis. They opine that analysis of data in a general way involves a number of closely related operations which are performed with the purpose of summarising the collected data and organising these in such a manner that they

answer the research question(s). We, however, shall prefer to observe the difference between the two terms as stated here in order to understand their implications more clearly.

EDITING

Editing is a process of examining the collected data to detect errors and omissions and to correct these when possible. As a matter of fact, editing involves a careful scrutiny of the completed questionnaire / schedules. Editing is done to assure that the data are accurate, consistent (constant), uniformly entered (Male=1, Female=2), as completed as possible and have been well arranged to facilitate coding and tabulation

Types of Editing

A) Field Editing

1. Field Editing – is the review of the reporting forms by the investigator for completing (translating or rewriting) what the latter has written in abbreviated or illegible (unreadable) form at the time of recording the respondent's responses
2. This type of editing is necessary in view of the fact that individual writing styles often can be difficult for others to decipher (read)
3. This sort of editing should be done as soon as possible after the interview, preferably on the very day or on the next day
4. While doing field editing, the investigator must restrain (control) himself and must not correct error of omission by simply guessing what the informant would have said if the question had been asked

B) Central Editing

1. Central Editing – should take place when all forms of schedules have been completed and returned to the office
2. This type of editing implies that all forms should get a through editing by a single editor in a small study and by a team of editors in case of large enquiry
3. Editor may correct the obvious (noticeable) errors (ex.) entry in the wrong place, entry recorded in the months when it should have been recorded in weeks,
4. In case of inappropriate (Wrong) on missing replies, the editor can sometimes determine the proper answer by reviewing the other information in the schedule
5. At times, the respondent can be contacted for clarification
6. The editor must strike out the answer if the same is inappropriate and he has no basis for determining the correct answer or the response. In such a case an editing entry of 'No answer' is called for
7. All wrong replies, which are quite obvious (unclear) must be dropped from the final results

Points to be Considered by Editor While Editing of Data

1. They should be familiar with instructions given to the interviewers and coders as well as with the editing instructions supplied to them for the purpose
2. They must make entries on the form in some distinctive (unique) and that too in a standardized form
3. They should initial all answers which they change or supply
4. Editor's initials and the date of editing should be placed on each completed form or schedule
5. While crossing-out (removing or deleting) an original entry for one reason or another, they should just draw a single line on it so that the same may remain legible

CODING

Coding refers to the process of assigning numerals or other symbols to answers so that responses can be put into a limited number of categories or classes. Such classes should be appropriate to the research problem under consideration. They must also possess the characteristics of exhaustiveness and also that of mutual exclusively which means that a specific answers can be placed in one and only one cell in a given category set. Another rule to be observed is that of unidimensionality by which is meant that every class is defined in terms of only one concept

Coding is necessary for efficient analysis and through it the several replies may be reduced to a small number of classes which contain the critical information required for analysis. Coding decisions should usually be taken at the designing stage or the questionnaire. This makes it possible to precode the questionnaire choices and which in turn is helpful for computer tabulation as one can straight forward key punch from the original questionnaire. But in case of hand coding some standard method may be used. One such standard method is to code in the margin with a colored pencil. The other method can be to transcribe the data from the questionnaire to a coding sheet. Whatever method is adopted, one should see that coding errors are altogether eliminated or reduced to the minimum level.

CLASSIFICATION

Most research studies result in a large volume of raw data which must be reduced into homogeneous groups if we are to get meaningful relationships. This fact necessitates classification of data which happens to be the process of arranging data in groups or classes on the basis of

common characteristics. Data having a common characteristic are placed in one class and in this way the entire data get divided into a number of groups or classes. Classification can be one of the following two types, depending upon the nature of the phenomenon involved:

1) Classification according to Attributes

As stated above, data are classified on the basis of common characteristics which can either be descriptive (such as literacy, sex, honesty, etc.) or numerical (such as weight, height, income, etc.). Descriptive characteristics refer to qualitative phenomenon which cannot be measured quantitatively; only their presence or absence in an individual item can be noticed. Data obtained this way on the basis of certain attributes are known as statistics of attributes and their classification is said to be classification according to attributes. Such classification can be simple classification or manifold classification. In simple classification we consider only one attribute and divide the universe into two classes—one class consisting of items possessing the given attribute and the other class consisting of items which do not possess the given attribute. But in manifold classification we consider two or more attributes simultaneously, and divide that data into a number of classes (total number of classes of final order is given by 2n, where n = number of attributes considered).* Whenever data are classified according to attributes, the researcher must see that the attributes are defined in such a manner that there is least possibility of any doubt/ambiguity concerning the said attributes.

2) Classification according to Class-Intervals

Unlike descriptive characteristics, the numerical characteristics refer to quantitative phenomenon which can be measured through some statistical units. Data relating to income, production, age, weight, etc.

come under this category. Such data are known as statistics of variables and are classified on the basis of class intervals. For instance, persons whose incomes, say, are within Rs 201 to Rs 400 can form one group, those whose incomes are within Rs 401 to Rs 600 can form another group and so on. In this way the entire data may be divided into a number of groups or classes or what are usually called, 'class-intervals.' Each group of class-interval, thus, has an upper limit as well as a lower limit which are known as class limits. The difference between the two class limits is known as class magnitude. We may have classes with equal class magnitudes or with unequal class magnitudes. The number of items which fall in a given class is known as the frequency of the given class. All the classes or groups, with their respective frequencies taken together and put in the form of a table, are described as group frequency distribution or simply frequency distribution. Classification according to class intervals usually involves the following three main problems:

i) How may classes should be there? What should be their magnitudes?

There can be no specific answer with regard to the number of classes. The decision about this calls for skill and experience of the researcher. However, the objective should be to display the data in such a way as to make it meaningful for the analyst. Typically, we may have 5 to 15 classes. With regard to the second part of the question, we can say that, to the extent possible, class-intervals should be of equal magnitudes, but in some cases unequal magnitudes may result in better classification. Hence the researcher's objective judgement plays an important part in this connection. Multiples of 2, 5 and 10 are generally preferred while determining class magnitudes. Some statisticians adopt the following formula, suggested by H.A. Sturges, determining the size of class interval:

$$i = R/(1 + 3.3$$

log N) where

i = size of class interval;

R = Range (i.e., difference between the values of the largest item and smallest item among the given items);

N = Number of items to be grouped.

It should also be kept in mind that in case one or two or very few items have very high or very low values, one may use what are known as open-ended intervals in the overall frequency distribution. Such intervals may be expressed like under Rs 500 or Rs 10001 and over. Such intervals are generally not desirable, but often cannot be avoided. The researcher must always remain conscious of this fact while deciding the issue of the total number of class intervals in which the data are to be classified.

ii) How to Choose Class Limits?

While choosing class limits, the researcher must take into consideration the criterion that the mid-point (generally worked out first by taking the sum of the upper limit and lower limit of a class and then divide this sum by 2) of a class-interval and the actual average of items of that class interval should remain as close to each other as possible. Consistent with this, the class limits should be located at multiples of 2, 5, 10, 20, 100 and such other figures. Class limits may generally be stated in any of the following forms:

Exclusive type class intervals: They are usually stated as follows:

10–20

20–30

30–40

40–50

The above intervals should be read as under:

10 and under 20

20 and under 30

30 and under 40

40 and under 50

Thus, under the exclusive type class intervals, the items whose values are equal to the upper limit of a class are grouped in the next higher class. For example, an item whose value is exactly 30 would be put in 30–40 class interval and not in 20–30 class interval. In simple words, we can say that under exclusive type class intervals, the upper limit of a class interval is excluded and items with values less than the upper limit (but not less than the lower limit) are put in the given class interval.

Inclusive type class intervals: They are usually stated as follows:

11–20

21–30

31–40

41–50

In inclusive type class intervals the upper limit of a class interval is also included in the concerning class interval. Thus, an item whose value is 20 will be put in 11–20 class interval. The stated upper limit of the class interval 11–20 is 20 but the real limit is 20.99999 and as such 11–20 class interval really means 11 and under 21.

When the phenomenon under consideration happens to be a discrete one (i.e., can be measured and stated only in integers), then we

should adopt inclusive type classification. But when the phenomenon happens to be a continuous one capable of being measured in fractions as well, we can use exclusive type class intervals.

iii) How to determine the Frequency of Each Class?

This can be done either by tally sheets or by mechanical aids. Under the technique of tally sheet, the class-groups are written on a sheet of paper (commonly known as the tally sheet) and for each item a stroke (usually a small vertical line) is marked against the class group in which it falls. The general practice is that after every four small vertical lines in a class group, the fifth line for the item falling in the same group, is indicated as horizontal line through the said four lines and the resulting flower (IIII) represents five items. All this facilitates the counting of items in each one of the class groups.

Alternatively, class frequencies can be determined, specially in case of large inquires and surveys, by mechanical aids i.e., with the help of machines viz., sorting machines that are available for the purpose. Some machines are hand operated, whereas other work with electricity. There are machines which can sort out cards at a speed of something like 25000 cards per hour. This method is fast but expensive.

TABULATION

When a mass of data has been assembled, it becomes necessary for the researcher to arrange the same in some kind of concise (brief / short) and logical order. This is known as tabulation. Thus, tabulation is the process of summarizing raw data and displaying the same in compact form. In a broader sense, tabulation is an orderly arrangement of data in column and rows

Need for Tabulation

1. It conserves (save) space and reduce explanatory and descriptive statement to a minimum
2. It facilitates the process of comparison
3. It facilitates the summation (abstract) of items and the detection of errors and omissions
4. It provides a basis for various statistical computations (calculation)

Principles of Tabulation

1. Every table should have a clear, concise (short) and adequate title so as to make the table intelligible without reference to the text and this title should always be placed just above the body of the table
2. Every table should be given a distinct number to facilitate easy reference
3. The column and the row headings of the table should be clear and brief
4. The units of measurements under each heading or sub-heading must always be indicated (i.e.) in Rs.
5. Explanatory footnotes, if any, concerning the table should be placed directly beneath (under) the table, along with the reference symbols used in the table
6. Source or sources from where the data in the table have been obtained must be indicated juts below the table
7. Usually the columns are separated from one another by lines which make the table more readable and attractive. Lines are always drawn at the top and bottom of the table and below the captions
8. There should be thick lines to separate the data under one class from the data under another class and the lines separating the sub-divisions of the classes should be comparatively thin lines

9. The columns may be numbered to facilitate reference
10. Those columns whose data are to be compared should be kept side by side. Similarly, percentages and / or averages must also be kept close to the data
11. It is generally considered better to approximate (rough / estimated) figures before tabulation as the same would reduce unnecessary details in the table itself
12. In order to emphasize the relative significance of certain categories, different kinds of type, spacing and indentations (marks) may be used
13. It is important that all column figures be properly aligned. Decimal points and (+) or (-) signs should be in perfect alignment
14. Abbreviations should be avoided to the extent possible and ditto marks should not be used in the table
15. Miscellaneous and exceptional items, if any, should be usually placed in the last row of the table
16. Table should be made as logical, clear, accurate and simple as possible. If the data happen to be very large, they should not be crowded in a single table for that would make the table unwieldy and inconvenient
17. Total of rows should normally be placed in the extreme right column and that of column should be placed at the bottom
18. The arrangement of the categories in a table may be chronological, geographical, alphabetical or according to magnitude to facilitate comparison. Above all, the table must suit the needs and requirements of an investigation

* * * * *

POSSIBLE QUESTIONS

UNIT-III

PART A

One Mark

Online Examination

PART B

Eight Marks

1. Define questionnaire method of data collection and explain its types.
2.Discuss the various tools used to collect the data?
3..Discuss the scale construction techniques in detail
4.Analyze the interview method of data collection with pros and cons?
5.Discuss the primary methods of data collection with examples?
6.Explain the techniques involved in data processing?
7. Distinguish between questionnaire and schedule.
8.What are the different types of editing? Explain each of these types with examples.
9. Discuss the different types of observation.
10.What are the merits and limitations of interview as a method of data collection?

KARPAGAM ACADEMY OF HIGHER EDUCATION

(DEEMED TO BE UNIVERSITY)

(ESTABILISHED UNDER SEC 3 OF UGC ACT 1956)

DEPARTMENT OF COMMERCE

III -B.COM RESEARCH METHODOLOGY-15CMU603B

UNIT-III METHODS OF DATA COLLECTION

S.NO	QUESTIONS	OPTION I	OPTION II	OPTION III	OPTION IV	ANSWER
1	___begins after a research design chalked out	Analysis	Interpretation	Report Writing	Data Collection	Data Collection
2	Data collected for the first time is known as ___data	Primary	Secondary	Regional	Teritarial	Primary
3	Data that have been compiled by some agency other than the user is	Teritarial	Regional	Secondary	Primary	Secondary
4	___method is particularly suitable for intensive investigations	Interview	Questionnaire	Secondary data	Observation	Interview
5	__interview method of collecting information through personal interviews	Focused	Structured	Unstructured	Clinical	Structured
6	In a ___interview, the interviewer is allowed much greater freedom to raise	Directive	Non-directive	Structured	Unstructured	Unstructured
7	___interviews are used generally in the development of hypotheses	Structured	Non-directive	Clinical	Focused	Focused
8	___interview is concerned with broad underlying feelings or motivations of an	Structured	Clinical	Focused	Non-directive	Clinical

9	In ____interview, the interviewer's function is simply to encourage the	Structured	Clinical	Focused	Non-directive	Non-directive
10	_____method is the most commonly used method specially in studies relating	Observation	Questionnaire	Interview	Secondary data	Observation
11	If the observer observes by making himself a member of the group, the observation is	Structured	Participant	Clinical	Focused	Participant
12	When the observer observes as a detached emissary, the observation of this type is	Participant	Structured	Clinical	Non-Participant	Non-Participant
13	If the observation takes place in the natural setting, it may be termed as	Focused	Structured	Non-Participant	Uncontrolled	Uncontrolled
14	When observation takes place according to definite pre-arranged plans,	Controlled	Uncontrolled	Focused	Clinical	Controlled
15	___method of data collection is quite popular, particularly in case of big enquiries	Interview Schedule	Observation	Questionnaire	Secondary data	Questionnaire
16	___ questionnaire is one in which all questions and answers are specified and	Closed ended	Unstructured	Open ended	Structured	Structured
17	____questionnaire is widely used in market research	Structured non disguised	Strucured disguised	Non-structured and disguised	Non-structured non-disguised	Structured non disguised
18	____ Questionnaire is used to find, peoples' attitude	Strucured disguised	Non-structured and disguised	Non-structured non-disguised	Structured non disguised	Strucured disguised
19	Projective techniques are examples of__	Non-structured non-disguised	Strucured disguised	Non-structured and disguised	Structured non disguised	Non-structured and disguised
20	___ explain the aims and objects of the investigation and also remove the	Agent	Surveyor	Investigator	Enumerator	Enumerator

21	__method of collecting data involves presentation of oral-verbal stimuli and reply in	Observation	Interview	Questionnaire	Secondary data	Interview
22	For carrying out descriptive study___ method of data collection is more suitable	Focused	Structured	Unstructured	Clinical	Structured
23	___interviews are characterized by a flexibility of approach to questioning	Focused	Structured	Unstructured	Clinical	Unstructured
24	____interview happens to be the central technique of collecting information in	Directive	Structured	Focused	Unstructured	Unstructured
25	___interview is meant to focus attention on the given experience of the	Structured	Unstructured	Clinical	Focused	Focused
26	___interview study the individual's life experience	Clinical	Structured	Directive	Focused	Clinical
27	___implies editing, coding, classification and tabulation of collected data	Processing	Classification	Transcription	Analysis of Data	Processing
28	___refers to the computation of certain measures along with searching for patterns	Processing	Classification	Transcription	Analysis	Analysis
29	____involves a careful scrutiny of the completed questionnaires and/or	Coding	Classification	Tabulation	Editing	Editing
30	___editing will be done as soon as possible after the interview, preferably on the	Chronological	Alphabetical	Field	Central	Field
31	____editing should take place when all forms of schedules have been	Chronological	Alphabetical	Field	Central	Central
32	Assignment of numbers and symbols for collected data is known as ___	Coding	Editing	Tabulation	Transcription	Coding

33	Reduction of data into homogenous group is known as ____	Coding	Editing	Classification	Tabulation	Classification
34	Classification of a person based on his gender is known as classification	Trait	Class Interval	Attributes	Element	Attributes
35	___characteristics refer to qualitative phenomenon which cannot be measured	Descriptive	Analytical	Uncomplicated	Basic	Descriptive
36	Classification of an individual based on single criteria is known as ___	Manifold	Simple	Double	Triple	Simple
37	Classification of an individual based on multiple criteria is known as ___	Manifold	Simple	Double	Triple	Manifold
38	In __classification we consider only one attribute and divide the universe into	Double	Triple	Manifold		Simple
39	Numerical characteristics are classified based on___	Class Interval	Trait	Element	Attributes	Class Interval
40	The difference between the two class limits is known as __	Interval	Frequency	Attributes	Class Magnitude	Class Magnitude
41	The number of items which fall in a given class is known as the __of the given class	Interval	Frequency	Attributes	Class Magnitude	Frequency
42	All the classes or groups, with their respective frequencies taken together	Simple Frequency Distribution	Binominal distribution	Polynominal distribution	Class Magnitude	Simple Frequency Distribution
43	___process helps in the presentation of all responses and observations on data	Editing	Coding	Transcription	Tabulation	Transcription
44	___conserves space and reduce explanatory and descriptive statement to a	Transcription	Tabulation	Coding	Editing	Tabulation

45	___method is independent of respondents' willingness to respond and as such is	Interview	Questionnaire	Observation	Secondary data	Observation
46	___ interviews demand deep knowledge and greater skill on the part of the	Unstructured	Directive	Non-directive	Clinical	Unstructured
47	__ data are those which have already been collected by someone else	Primary	Secondary	Regional	Teritarial	Secondary
48	___is done to assure that the data are accurate, consistent with other facts	Tabulation	Coding	Editing	Transcription	Editing
49	____editing is the review of the reporting forms by the investigator for completing	Field	Central	Chronological	Alphabetical	Field
50	___is necessary for efficient analysis and through it the several replies may be	Coding	Tabulation	Editing	Transcription	Coding
51	The process of arranging data in groups or classes are known as ___	Coding	Editing	Classification	Tabulation	Classification
52	Classification of a person based on his literacy level known as classification	Class Interval	Trait	Element	Attributes	Attributes
53	Data relating to income are classified based on _____	Class Interval	Trait	Element	Attributes	Class Interval
54	The process of presenting data in row and columns is known as ___	Tabulation	Transcription	Editing	Coding	Tabulation
55	___refers to the process of assigning numbers	Editing	Coding	Tabulation	Transcription	Coding
56	Grouping of data on the basis of common characteristics is known as___	Classification	Coding	Editing	Tabulation	Classification

57	Classification of a person based on honesty is known as classification based on____	Attributes	Class Interval	Trait	Element	Attributes
58	Data relating to production are classified based on _____	Class Interval	Trait	Element	Attributes	Class Interval
59	___is an intermediary process between data coding and data tabulation	Editing	Coding	Transcription	Tabulation	Transcription
60	__ is a combination of rows and columns	Editing	Coding	Transcription	Tabulation	Tabulation

UNIT IV

SYLLABUS

Analysis of Data: Hypothesis – Characteristics-Concepts of Hypothesis-Null Hypothesis-Alternative Hypothesis-Level of Significance-Test of Hypothesis-Type I and Type II error-Chi-square test-t test-F test-ANOVA-Scaling Techniques

HYPOTHESIS

Hypothesis is usually considered as the principal instrument in research. Its main function is to suggest new experiments and observations. In fact, many experiments are carried out with the deliberate object of testing hypotheses. Decision-makers often face situations wherein they are interested in testing hypotheses on the basis of available information and then take decisions on the basis of such testing. In social science, where direct knowledge of population parameter(s) is rare, hypothesis testing is the often used strategy for deciding whether a sample data offer such support for a hypothesis that generalisation can be made. Thus hypothesis testing enables us to make probability statements about population parameter(s). The hypothesis may not be proved absolutely, but in practice it is accepted if it has withstood a critical testing. Before we explain how hypotheses are tested through different tests meant for the purpose, it will be appropriate to explain clearly the meaning of a hypothesis and the related concepts for better understanding of the hypothesis testing techniques.

What is Hypothesis?

Ordinarily, when one talks about hypothesis, one simply means a mere assumption or some supposition to be proved or disproved. But for a researcher hypothesis is a formal question that he intends to resolve. Thus a hypothesis may be defined as a proposition or a set of proposition set forth as an explanation for the occurrence of some specified group of phenomena either asserted merely as a provisional conjecture to guide some investigation or accepted as highly probable in the light of established facts. Quite often a research hypothesis is a predictive statement, capable of being tested by scientific methods, that relates an independent variable to some dependent variable. For example, consider statements like the following ones:

"Students who receive counselling will show a greater increase in creativity than students not receiving counselling" Or

"the automobile A is performing as well as automobile B."

These are hypotheses capable of being objectively verified and tested. Thus, we may conclude that a hypothesis states what we are looking for and it is a proposition which can be put to a test to determine its validity.

CHARACTERISTICS OF HYPOTHESIS

1. Hypothesis should be clear and precise. If the hypothesis is not clear and precise, the inferences drawn on its basis cannot be taken as reliable.
2. Hypothesis should be capable of being tested. In a swamp of untestable hypotheses, many a time the research programmes have bogged down. Some prior study may be done by researcher in order to make hypothesis a testable one. A hypothesis "is testable if other

deductions can be made from it which, in turn, can be confirmed or disproved by observation."

3. Hypothesis should state relationship between variables, if it happens to be a relational hypothesis.
4. Hypothesis should be limited in scope and must be specific. A researcher must remember that narrower hypotheses are generally more testable and he should develop such hypotheses.
5. Hypothesis should be stated as far as possible in most simple terms so that the same is easily understandable by all concerned. But one must remember that simplicity of hypothesis has nothing to do with its significance.
6. Hypothesis should be consistent with most known facts i.e., it must be consistent with a substantial body of established facts. In other words, it should be one which judges accept as being the most likely.
7. Hypothesis should be amenable to testing within a reasonable time. One should not use even an excellent hypothesis, if the same cannot be tested in reasonable time for one cannot spend a life-time collecting data to test it.
8. Hypothesis must explain the facts that gave rise to the need for explanation. This means that by using the hypothesis plus other known and accepted generalizations, one should be able to deduce the original problem condition. Thus hypothesis must actually explain what it claims to explain; it should have empirical reference.

FORMULATION OF HYPOTHESIS

1. A hypothesis may originate in different ways. A cultural environment may give rise to it. In India, for example, religion and custom dominate the way of life. This has had it reaction on economic values and individual initiative in various walks of life. Such a situation

could give rise to any number of hypothesis; sociological, cultural, political and economic

2. A second source of hypothesis is folk wisdom or current popular beliefs and practices suggesting both the problems and the hypothesis
3. Analogies are often a spring of valuable hypothesis. Students of sociology in the course of their studies would have come across analogies wherein a society is compared to a biological organism, the natural law to the social law, thermo-dynamics to social dynamics etc.,
4. The history of science provides an eloquent testimony to the fact that personal and idiosyncratic experiences of the scientist contributes a great deal to the type and form of questions he may ask as also to the kinds of tentative answers to these questions that he can provide
5. Hypotheses may also rest on the findings of other studies. The researcher on the basis of the findings of other studies may hypothesise that similar relationship between specified variables will hold good in the present study, too. This is a very common way of researchers who design their study with a view to replicating another study conducted in a different concrete context
6. Another source of hypothesis formulation in cases which are expectations to accepted theory
7. Personal experience and individual reaction may give rise to hypotheses
8. A hypotheses may turn from a body of theory which, by way of logical deduction, may lead to the production that if certain conditions are present, certain results will follow. Theory is indeed an extremely fertile seed-bed of hypotheses.

PROCEDURE FOR HYPOTHESIS TESTING

To test a hypothesis means to tell (on the basis of the data the researcher has collected) whether or not the hypothesis seems to be valid. In hypothesis testing the main question is: whether to accept the null hypothesis or not to accept the null hypothesis? Procedure for hypothesis testing refers to all those steps that we undertake for making a choice between the two actions i.e., rejection and acceptance of a null hypothesis. The various steps involved in hypothesis testing are stated below:

1) Making a Formal Statement

The step consists in making a formal statement of the null hypothesis (H0) and also of the alternative Hypothesis (Ha). This means that hypothesis should be clearly stated, considering the nature of the research problem. For instance, Mr. Mohan of the Civil Engineering Department wants to test the load bearing capacity of an old bridge which must be more than 10 tons, in that case he can state his hypotheses as under:

Null Hypothesis H_0 : μ =10 tons

Alternative Hypothesis H_a: μ > 10 tons

Take another example. The average score in an aptitude test administered at the national level is 80. To evaluate a state's education system, the average score of 100 of the state's students selected on random basis was 75. The state wants to know if there is a significant difference between the local scores and the national scores. In such a situation the hypotheses may be stated as under:

Null Hypothesis H_0 : μ = 80

Alternative Hypothesis H_a: $\mu \neq 80$

The formulation of hypotheses is an important step which must be accomplished with due care in accordance with the object and nature of the problem under consideration. It also indicates whether we should use a one-tailed test or a two-tailed test. If Ha is of the type greater than (or of the type lesser than), we use a one-tailed test, but when Ha is of the type "whether greater or smaller" then we use a two-tailed test.

2) Selecting a Significance Level

The hypotheses are tested on a pre-determined level of significance and as such the same should be specified. Generally, in practice, either 5% level or 1% level is adopted for the purpose. The factors that affect the level of significance are: (a) the magnitude of the difference between sample means; (b) the size of the samples; (c) the variability of measurements within samples; and (d) whether the hypothesis is directional or non-directional (A directional hypothesis is one which predicts the direction of the difference between, say, means). In brief, the level of significance must be adequate in the context of the purpose and nature of enquiry.

3) Deciding the Distribution to Use

After deciding the level of significance, the next step in hypothesis testing is to determine the appropriate sampling distribution. The choice generally remains between normal distribution and the t-distribution. The rules for selecting the correct distribution are similar to those which we have stated earlier in the context of estimation.

4) Selecting a Random Sample and Computing an Appropriate Value

Another step is to select a random sample(s) and compute an appropriate value from the sample data concerning the test statistic

utilizing the relevant distribution. In other words, draw a sample to furnish empirical data.

5) Calculation of the Probability

One has then to calculate the probability that the sample result would diverge as widely as it has from expectations, if the null hypothesis were in fact true.

6) Comparing the Probability

Yet another step consists in comparing the probability thus calculated with the specified value for α□, the significance level. If the calculated probability is equal to or smaller than the α value in case of one-tailed test (and α □/2 in case of two-tailed test), then reject the null hypothesis (i.e., accept the alternative hypothesis), but if the calculated probability is greater, then accept the null hypothesis. In case we reject H_0, we run a risk of (at most the level of significance) committing an error of Type I, but if we accept H_0, then we run some risk (the size of which cannot be specified as long as the H_0 happens to be vague rather than specific) of committing an error of Type II.

STATISTICS

't' Test

When the size of sample is small (less than 30). In particular, it will no longer be possible for us to assume (a) that the random sampling distribution of a statistic is approximately normal and (b) that values given by the sample data are sufficiently close to the population values and can be used in their place for the calculation of the standard error of the estimate.

The removal of these assumptions makes it necessary to use entirely new techniques to deal with the problems of small samples. The division between the theories of large and small samples is, therefore, a very real one, though it is not always easy to draw a precise line of demarcation. It should be noted that as a rule, the methods and the theory of small samples are applicable to large samples, though the reverse is not true.

While dealing with small samples our main interest is not to estimate the population values as in true in large samples; rather our interest lies in testing a given hypothesis, i.e., in ascertaining whether observed values could have arisen by sampling fluctuations from some value given in advance. For example, if a sample of 15 gives a correlation coefficient of +0.4, we shall be interested not so much in the value of the correlation in the parent population, but more generally whether this value could have been arisen from an uncorrelated population. i.e. whether it is significant of correlation in the parent population.

It should be noted that the investigator who works with very small samples must know that his estimates will vary widely from sample to sample. Moreover, he must be satisfied with relatively wide confidence intervals. Precision of statement is less, of course, the wider the intervals employed. Each inference drawn from large sample results in far more precise in the limits it sets up than is an inference based on a much smaller sample.

In statistical test theory, the notion of statistical error is an integral part of hypothesis testing. The test requires an unambiguous statement of a null hypothesis, which usually corresponds to a default "state of nature", for example "this person is healthy", "this accused is not guilty" or "this product is not broken". An alternative hypothesis is the negation of null

hypothesis, for example, "this person is not healthy", "this accused is guilty" or "this product is broken". The result of the test may be negative, relative to the null hypothesis (not healthy, guilty, broken) or positive (healthy, not guilty, not broken). If the result of the test corresponds with reality, then a correct decision has been made. However, if the result of the test does not correspond with reality, then an error has occurred. Due to the statistical nature of a test, the result is never, except in very rare cases, free of error. Two types of error are distinguished: type I error and type II error.

Type I error

A type I error occurs when the null hypothesis (H_0) is true, but is rejected. It is asserting something that is absent, a false hit. A type I error may be likened to a so-called false positive (a result that indicates that a given condition is present when it actually is not present).

In terms of Folk tales, an investigator may see the wolf when there is none ("raising a false alarm"). Where the null hypothesis, H_0, is: no wolf.

The type I error rate or significance level is the probability of rejecting the null hypothesis given that it is true. It is denoted by the Greek letter α (alpha) and is also called the alpha level. Often, the significance level is set to 0.05 (5%), implying that it is acceptable to have a 5% probability of incorrectly rejecting the null hypothesis.

Type II error

A type II error occurs when the null hypothesis is false, but erroneously fails to be rejected. It is failing to assert what is present, a miss. A type II error may be compared with a so-called false negative (where an actual 'hit' was disregarded by the test and seen as a 'miss') in a test checking for a single condition with a definitive result of true or false. A Type II error is committed when we fail to believe a true alternative hypothesis.

In terms of folk tales, an investigator may fail to see the wolf ("failing to raise an alarm"). Again, H_0: no wolf.

The rate of the type II error is denoted by the Greek letter β (beta) and related to the power of a test (which equals 1−β).

Table of error types

Tabularised relations between truth/falseness of the null hypothesis and out comes of the test:

Table of error types		Null hypothesis (H_0) is	
		True	False
Decision About Null Hypothesis (H_0)	Reject	Type I error (False Positive)	Correct inference (True Positive)
	Accept (not rejected)	Correct inference (True Negative)	Type II error (False Negative)

Student's 't' Distribution

Theoretical work on t-distribution was done by W.S. Gosset (1876-1937) in the early 1900. Gosset was employed by the Guinness and Son, a Dulbin bravery, Ireland, which did not permit employees to publish research findings under their own names. So Gosset adopted the pen name "Student" and published his findings under their name. Thereafter, the t-distribution is commonly called Student's t-distribution or simply student's distribution.

The t-distribution is used when sample size is 30 or less and the population standard deviation is unknown.

The 't' statistic is defined as

$t = \bar{X} - \mu / S_{\bar{X}} \sqrt{n}$

Where $S = \sqrt{\sum (X - \bar{X})^2 / n-1}$

Sum 1: The manufacturer of a certain make of electric bulbs claims that his bulbs have a mean life of 25 months with a standard deviation of 5 months. A random sample of 6 such bulbs gave the following values

Life of Months	24	26	30	20	20	18

Can you regard the producer's claim to be valid at 1% level of significance? (Given that the table values of the appropriate test statistics at the said level are 4.032, 3.707 and 3.499 for 5, 6 and 7 degrees of freedom respectively)

Solution

Let us taken the hypothesis that there is no significant difference in the mean life of bulbs in the sample and that of the population, Applying t-test:

X	**$(X-\bar{X})$ X**	**X^2**
24	+1	1
26	+3	9
30	+7	49
20	-3	9
20	-3	9
18	-5	25
$\sum X=138$		$\sum X^2=102$

$$t= \bar{X}-\mu/SD \sqrt{X} n$$

Average of X= $\sum X/n$ = 138/6 = 23

$$SD = \sqrt{\sum X^2/n-1} = \sqrt{102/5} \sqrt{} = 20.4 = 4.517$$

$$= \frac{(23-25)}{4.517} \sqrt{6} = \frac{2 \text{ X } 2.449}{4.517} = 1.084$$

$\nu = n-1 = 6-1 = 5$

For $\nu=5$ $t_{0.01}=4.032$

The calculated value of t is less than the table value. The hypothesis is accepted. Hence, the producer's claim is not valid at 1% level of significance.

Sum 2: A random sample of size 16 has 53 as mean. The sum of the squares of the deviation taken from mean is 135. Can this sample be regarded as taken from the population having 56 as mean? Obtain 95% and 99% confidence limits of the mean of the population. (For v=15, $t_{0.05}$=2.13 for v=15, $t_{0.01}$=2.95)

Solution

Let us take the hypothesis that there is no significant difference between the sample mean and hypothetical mean, Applying t-test.

$t = \bar{X}-\mu / SD \sqrt{X\ n}$

$SD = \sqrt{\sum X^2/n-1} = \sqrt{135/15} = 3$

$= (53-56) / 3 \sqrt{\ } X\ 16$

$= 3 X 4 / 3$

$= 4$

v=16-1=15. For v=16, $t_{0.05}$=2.13.

The calculated value is more than the table value. The hypothesis is rejected. Hence, the sample has not come from a population having 56 as mean.

Sum: 3 The life time of electric bulbs for a random sample of 10 from a large consignment gave the following data:

Item	1	2	3	4	5	6	7	8	9	10
Life in '000	4.2	4.6	3.9	4.1	5.2	3.8	3.9	4.3	4.4	5.6

hours										

Can we accept the hypothesis that the average life time of bulbs is 4000 hours.

Solution

Let us take the hypothesis that there is no significant difference in the sample mean and the hypothetical population mean

X	**$(X-\bar{X})$**	**$(X-\bar{X})^2$**
4.2	-0.2	0.04
4.6	+0.2	0.04
3.9	-0.5	0.25
4.1	-0.3	0.09
5.2	+0.8	0.64
3.8	-0.6	0.36
3.9	-0.5	0.25
4.3	-0.1	0.01
4.4	0.0	0.00
5.6	+1.2	1.44
$\sum X=44$		$\sum(X-X)^2=3.12$

$$t= \bar{X}-\mu/S \; X \quad n$$

$$\text{Where } S=\sqrt{\sum(X-X)^2/\ n-1}$$

$$t= \frac{4.4-4}{0.589} \quad \sqrt{10} = \frac{0.4 \text{ X } 3.162}{0.589} = 0.2148$$

$\nu=n-1 = 10-1 = 9$

For $\nu=9$ $t_{0.05}=2.262$

The calculated value of t is less than the table value. The hypothesis is accepted. The average life time of the bulbs could be 4000 hours.

Sum 4: Two types of drugs were used on 5 and 7 patients for reducing their weight.

Drug A was imported and drug B indigenous. The decrease in the weight after using the drugs for six months was as follows:

Drug A	10	12	13	11	14		
Drug B	8	9	12	14	15	10	9

Is there a significant difference in the efficacy of the two drugs? If not, which drug should you buy (For $\nu=10$, $t_{0.05}=2.228$)

Solution

Let us take the hypothesis that there is no significant difference in the efficacy of the two drugs.

X_1	$(X_1-\bar{X}_1)$	$(X_1-\bar{X}_1)^2$	X_2	$(X_2-\bar{X}_2)$	$(X_2-\bar{X}_2)^2$
10	-2	4	8	-3	9
12	0	0	9	-2	4
13	+1	1	12	+1	1
11	-1	1	14	+3	9
14	+2	4	15	+4	16
			10	-1	1
			9	-2	4
$\sum X_1=60$		$\sum (X_1-\bar{X}_1)^2=10$	$\sum X_2=77$		$\sum (X_2-\bar{X}_2)^2=44$

$$t = \frac{\bar{X}_1+\bar{X}_2}{S}\sqrt{\frac{n_1n_2}{n_1+n_2}}$$

$$S=\sqrt{\frac{\sum(X_1-\bar{X}_1)^2+\sum(X_2-\bar{X}_2)^2}{n_1+n_2-2}}$$

$$S = \sqrt{\frac{10 + 44}{5+7-2}} = \sqrt{\frac{54}{10}} = 2.324$$

$$= \frac{12-11}{2.324} \quad \frac{5 \times 7}{5+7} = \frac{1.708}{2.324} = 0.735$$

$\nu = n_1 + n_2 - 2 = 5 + 7 - 2 = 10$

$\nu = 10$, $t_{0.05} = 2.228$

The calculated value of t is less than the table value, the hypothesis is accepted. Hence, there is no significance in the efficacy of two drugs. Since drug B is indigenous and there is no difference in the efficacy of imported and indigenous drug, we should buy indigenous drug, i.e., B.

Sum 5 : For a random sample of 10 persons, fed on diet A, the increased weight in pounds in a certain period were:

10	6	16	17	13	12	8	14	15	9

For another random sample of 12 persons, fed on diet B, the increase in the same period were:

7	13	22	15	12	14	18	8	21	23	10	17

Test whether the diets A and diet B differ significantly as regards their effect on increase in weight, Given the following.

Degrees of Freedom	19	20	21	22	23
Value at t at 5% level	2.09	2.09	2.08	2.07	2.07

Solution: Let us take the null hypothesis that A and B do not differ significantly weight regard to their effect on increase in weight.

X_1	$(X_1 - \bar{X}_1)$	$(X_1 - \bar{X}_1)^2$	X_2	$(X_2 - \bar{X}_2)$	$(X_2 - \bar{X}_2)^2$
10	-2	4	7	-8	64
6	-6	36	13	-2	4
16	+4	16	22	+7	49
17	+5	25	15	0	0
13	+1	1	12	-3	9
12	0	0	14	-1	1
8	-4	16	18	+3	9
14	+2	4	8	-7	49
15	+3	9	21	+6	36
9	-3	9	23	+8	64
			10	-5	25
			17	+2	4
$\sum X_1=120$		$\sum (X_1 - \bar{X}_1)^2=120$	$\sum X_2=180$		$\sum (X_2 - \bar{X}_2)^2=314$

Mean increase in weight of 10 persons fed on diet A

Mean increase in weight of 10 persons fed on diet B

$$t = \frac{\bar{X}_1 + \bar{X}_2}{S} \times \sqrt{\frac{n_1 n_2}{n_1 + n_2}}$$

$$S = \sqrt{\frac{\sum (X_1 - \bar{X}_1)^2 + \sum (X_2 - \bar{X}_2)^2}{n_1 + n_2 - 2}}$$

$$SD = \frac{120 + 314}{10 + 12 - 2} = \frac{434}{20} = 4.66$$

X_1=12, X_2=15, n_1=10, n_2 = 12, S=4.66. Substituting the values in the above formula:

$$t = \frac{12 - 15}{4.66} \times \frac{10 \times 12}{10 + 12} = \frac{3}{4.66} \times 2.34 = 1.51$$

For v=20, the table value of t at 5 per cent level is 2.09. The calculated value is less than the table value and hence the experiment provides no evidence against the hypothesis. We, therefore, conclude that diets A and B do not differ significantly as regards their effect on increase in weight is concerned.

Sum 6: In a test given to two groups of students, the marks obtained are as follows:

I Group	18	20	36	50	49	36	34	49	41
II Group	29	28	26	35	30	44	46		

Examine the significance of difference between the arithmetic mean of the marks secured by the students of the above two groups. (The value of t at 5% level of significance for v=14 is 2.14)

Solution

Let us take the hypothesis that there is no significant difference in the arithmetic mean of the marks secured by the students of the two groups.

Group I X_1	**$(X_1 - \bar{X}_1)$**	**$(X_1 - \bar{X}_1)^2$**	**Group II X_2**	**$(X_2 - \bar{X}_2)$**	**$(X_2 - \bar{X}_2)^2$**
18	-19	361	29	-5	25
20	-17	289	28	-6	36
36	-1	1	26	-8	64
50	+13	169	35	+1	1
49	+12	144	30	-4	16
36	-1	1	44	+10	100
34	-3	9	46	+12	144

49	+12	144			
41	+4	16			
$\sum X_1=333$		$\sum$ **$(X_1 - X_1)^2=1134$**	$\sum X_2=238$		$\sum$ **$(X_2 - X_2)^2=386$**

$$t = \frac{\bar{X}_1 - \bar{X}_2}{S} \times \sqrt{\frac{n_1 n_2}{n_1+n_2}}$$

$$S = \sqrt{\frac{\sum(X_1-\bar{X}_1)^2 + \sum(X_2-\bar{X}_2)^2}{n_1+n_2-2}}$$

$$SD = \frac{1134 + 386}{9 + 7 - 2} = 10.42$$

$$t = \frac{37-34}{10.42} = \frac{9 \times 7}{9+7} = \frac{3}{10.42} \times 1.984 = 0.571$$

$\nu = n_1+n_2-2 = 9 + 7 - 2 = 14$; For $\nu = 14$, $t_{0.05} = 2.14$

The calculated value of t is less than the table value and hence the hypothesis hold true. We, therefore, conclude that the mean marks of the students of the two groups do not differ significantly.

F-Test

The F-test is named in honor of the great Statistician R.A. Fisher. The object of F –test is to find out whether the two independent estimates of population variance differ significantly, or whether the two samples may be regarded as drawn from the normal populations having the same variance. For carrying out the test of significance, we calculate the ratio F. F is defined as

$F=S_1^2/S_2^2$, Where $S_1^2 = \sum(X_1-\bar{X}_1)^2/n_1-1$ and

$$S_2^2 = \sum(X_2-\overline{X_2})^2/n_2-1$$

It should be noted that S_1^2 is always the larger estimate of variance, i.e., $S_1^2 > S_2^2$

$$F = \frac{\text{Larger estimate of Variance}}{\text{Smaller estimate of Variance}}$$

The calculated value of F is compared with the table value for v1 and v2 at 5% or 1% level of significance. If calculated value of F is greater than the table value then the F ratio is considered significant and the null hypothesis is rejected. On the other hand, if the calculated value of F is less than the table value the null hypothesis is accepted and it is inferred that both the samples have come from the population having same variance.

Since F Test is based on the ratio of two variances, it is also known as the Variance Ratio Test. The ratio of two variances follows a distribution called the F distribution named after the famous statistician R.A. Fisher.

Sum 1: Two random samples were drawn from two normal populations and their values are :

A	66	67	75	76	82	84	88	90	92		
B	64	66	74	78	82	85	87	92	93	95	97

Test whether the two populations have the same variance at the 5% level of significance. (F=3.36) at 5% level of significance level v1=10 and v2=8.

Solution

Let us take the hypothesis that the two populations have the same variance

A X_1	**$(\bar{X_1}-X_1)$; X1**	**X_1^2**	**B X_2**	**$(\bar{X_2}-X_2)$; X_2**	**c X_2^2**
66	-14	196	64	-19	361
67	-13	169	66	-17	289
75	-5	25	74	-9	81
76	-4	16	78	-5	25
82	2	4	82	-1	1
84	4	16	85	2	4
88	8	64	87	4	16
90	10	100	92	9	81
92	12	144	93	10	100
			95	12	144
			97	14	196
$\sum X_1=720$	**$\sum X_1=0$**	**$\sum X_1^2=734$**	**$\sum X_2=913$**	**$\sum X_2=0$**	**$\sum X_2^2=1298$**

Average of $X_1 = \sum X_1/n1 = 720 / 9 = 80$;

Average of $X_2 = \sum X_2/n2 = 913/ 11 = 83$

$S_1^2 = \sum(X_1-X_1)^2/n_1-1$

$= 734 / 9-1 = 91.75$

$S_2^2 = \sum(X_2-X_2)^2/n_2-1$

$= 1298 / 11-1 = 129.80$

$F = S_1^2/S_2^2$

$= 129.8 / 191.75$

$= 1.415$

The calculated value of F is less than the table value. The hypothesis is accepted. Hence, it may be calculated that the two populations have the same variance.

Sum 2: In a sample of 8 observations, the sum of squared deviations of items from the mean was 84.4. In another sample of 10 observations, the value was found to be 102.60. Test whether the difference is significant at 5% level.

You are given that at 5% level, critical value of F for v1=7 and v2=9 degrees of freedom is 3.29 and for v1=8 and v2=10 degrees of freedom, its value is 3.07.

Solution

Let us take hypothesis that the difference in the variance of the two samples is not significant. We are given

$S_1^2 = \sum(X_1-X_1)^2/n_1-1 = 84.4/7 = 12.06$

$S_2^2 = \sum(X_2-X_2)^2/n_2-1 = 102.3/9 = 11.40$

$F = S_1^2/S_2^2 = 12.06/11.40 = 1.06$

The calculated value of F is less than the table value. Hence, we accept the hypothesis and conclude that the difference in the variance of two samples is not significant at 5% level.

Sum 3: Two samples are drawn from two normal populations. From the following data test whether the two samples have the same variance at 5% level.

Sample 1	60	65	71	74	76	82	85	87		
Sample 2	61	66	67	85	78	63	85	86	88	91

Solution: Let us take the hypothesis that the two populations have the same variance.

F $=S_1^2/S_2^2$

Sample 1 X_1	**$(X_1-\bar{X}_1)$; X1**	**X_1^2**	**Sample 2 X_2**	**$(X_2-\bar{X}_2)$; X_2**	**X_2^2**
60	-15	225	61	-16	256
65	-10	100	66	-11	121
71	-4	16	67	-10	100
74	-1	1	85	8	64
76	1	1	78	1	1
82	7	49	63	-14	196
85	10	100	85	8	64
87	12	144	86	9	81
			88	11	121
			91	14	196
$\sum X_1=600$	**$\sum X_1=0$**	**$\sum X_1^2=636$**	**$\sum X_2=770$**	**$\sum X_2=0$**	**$\sum X_2^2=1200$**

Average of X_1 = 600/8 = 75

Average of X_2 = 770/10 = 77

$S_1^2 = \sum(X_1-X_1)^2/n_1-1$ = 636 / 8-1 = 90.857

$S_2^2 = \sum(X_2-X_2)^2/n_2-1$ = 1200/10-1 = 133.33

F $= S_1^2/S_2^2$

= 133.33/90.857

= 1.467

For v1=9 and v2=7, $F_{0.05}$ = 3.68. The calculated value of F is less than the table value. The hypothesis holds good and hence we conclude that the two populations have the same variance.

Sum 4: The following data present the yields in Quintals of common ten subdivisions of equal area of two agricultural plots.

Plot 1	6.2	5.7	6.5	6.0	6.3	5.8	5.7	6.0	6.0	5.8
Plot 2	5.6	5.9	5.6	5.7	5.8	5.7	6.0	5.5	5.7	5.5

Test whether two samples taken from two random populations have the same variance. (5% point of F for v1=9 and v2=9 is 3.18)

Solution: Let us take the null hypothesis that the samples come from populations having the same variance

Plot 1 X_1	**$(X_1-\bar{X}_1)$; X1**	**X_1^2**	**Plot 2 X_2**	**$(X_2-\bar{X}_2)$; X_2**	**c X_2^2**
6.2	0.2	0.04	5.6	-0.1	0.01
5.7	-0.3	0.09	5.9	0.2	0.04
6.5	0.5	0.25	5.6	-0.1	0.01
6.0	0	0	5.7	0	0
6.3	0.3	0.09	5.8	0.1	0.01
5.8	-0.2	0.04	5.7	0	0
5.7	-0.3	0.09	6.0	0.3	0.09
6.0	0	0	5.5	-0.2	0.04
6.0	0	0	5.7	0	0
5.8	-0.2	0.04	5.5	-0.2	0.04
$\sum X_1=60$	**$\sum X_1=0$**	**$\sum X_1^2=0.64$**	**$\sum X_2=57$**	**$\sum X_2=0$**	**$\sum X_2^2=0.24$**

$F = S_1^2/S_2^2$

$S_1^2 = \sum(X_1-X_1)^2/n_1-1 = 0.64 / 9 = 0.071$

$S_2^2 = \sum(X_2-X_2)^2/n_2-1 = 0.24 / 9 = 0.027$

$F = S_1^2/S_2^2 = 0.071/0.027 = 2.63$

The value of F for 9 and 6 at 5% level of significance is 3.18. The calculated value is less than the table value. The hypothesis holds true. Hence, the samples come from population having the same variance.

CHI-SQUARE

The chi-square test is an important test amongst the several tests of significance developed by statisticians. Chi-square, symbolically written as χ^2 (Pronounced as Ki-square), is a statistical measure used in the context of sampling analysis for comparing a variance to a theoretical

variance. As a non-parametric test, it "can be used to determine if categorical data shows dependency or the two classifications are independent. It can also be used to make comparisons between theoretical populations and actual data when categories are used." Thus, the chi-square test is applicable in large number of problems. The test is, in fact, a technique through the use of which it is possible for all researchers to (i) test the goodness of fit; (ii) test the significance of association between two attributes, and (iii) test the homogeneity or the significance of population variance.

Chi-square is an important non-parametric test and as such no rigid assumptions are necessary in respect of the type of population. We require only the degrees of freedom (implicitly of course the size of the sample) for using this test. As a non-parametric test, chi-square can be used (i) as a test of goodness of fit and (ii) as a test of independence.

As a test of goodness of fit, χ^2 test enables us to see how well does the assumed theoretical distribution (such as Binomial distribution, Poisson distribution or Normal distribution) fit to the observed data. When some theoretical distribution is fitted to the given data, we are always interested in knowing as to how well this distribution fits with the observed data. The chi-square test can give answer to this. If the calculated value of χ^2 is less than the table value at a certain level of significance, the fit is considered to be a good one which means that the divergence between the observed and expected frequencies is attributable to fluctuations of sampling. But if the calculated value of χ^2 is greater than its table value, the fit is not considered to be a good one.

As a test of independence, χ^2 test enables us to explain whether or not two attributes are associated. For instance, we may be interested in knowing whether a new medicine is effective in controlling fever or not, χ^2

test will helps us in deciding this issue. In such a situation, we proceed with the null hypothesis that the two attributes (viz., new medicine and control of fever) are independent which means that new medicine is not effective in controlling fever. On this basis we first calculate the expected frequencies and then work out the value of χ^2 . If the calculated value of χ^2 is less than the table value at a certain level of significance for given degrees of freedom, we conclude that null hypothesis stands which means that the two attributes are independent or not associated (i.e., the new medicine is not effective in controlling the fever). But if the calculated value of χ^2 is greater than its table value, our inference then would be that null hypothesis does not hold good which means the two attributes are associated and the association is not because of some chance factor but it exists in reality (i.e., the new medicine is effective in controlling the fever and as such may be prescribed). It may, however, be stated here that χ^2 is not a measure of the degree of relationship or the form of relationship between two attributes, but is simply a technique of judging the significance of such association or relationship between two attributes.

In order that we may apply the chi-square test either as a test of goodness of fit or as a test to judge the significance of association between attributes, it is necessary that the observed as well as theoretical or expected frequencies must be grouped in the same way and the theoretical distribution must be adjusted to give the same total frequency as we find in case of observed distribution. χ^2 is then calculated as follows:

$$\chi^2 = \sum \frac{\left(O_{ij} - E_{ij}\right)^2}{E_{ij}}$$

If two distributions (observed and theoretical) are exactly alike, $\chi 2 = 0$; but generally due to 2 sampling errors, $\chi 2$ is not equal to zero and as such we must know the sampling distribution of $\chi 2$ so that we may find the probability of an observed $\chi 2$ being given by a random sample from the hypothetical universe. Instead of working out the probabilities, we can use ready table which gives probabilities for given values of $\chi 2$. Whether or not a calculated value of $\chi 2$ is significant can be ascertained by looking at the tabulated values of $\chi 2$ for given degrees of freedom at a certain level of significance. If the calculated value of $\chi 2$ is equal to or exceeds the table value, the difference between the observed and expected frequencies is taken as significant, but if the table value is more than the calculated value of $\chi 2$, then the difference is considered as insignificant i.e., considered to have arisen as a result of chance and as such can be ignored.

As already stated, degrees of freedom play an important part in using the chi-square distribution and the test based on it, one must correctly determine the degrees of freedom. If there are 10 frequency classes and there is one independent constraint, then there are (10 – 1) = 9 degrees of freedom. Thus, if 'n' is the number of groups and one constraint is placed by making the totals of observed and expected frequencies equal, the d.f. would be equal to (n – 1). In the case of a contingency table (i.e., a table with 2 columns and 2 rows or a table with two columns and more than two rows or a table with two rows but more than two columns or a table with more than two rows and more than two columns), the d.f. is worked out as follows: d.f. = (c – 1) (r – 1), where 'c' means the number of columns and 'r' means the number of rows.

CONDITIONS FOR THE APPLICATION OF $\chi 2$ TEST

1. Observations recorded and used are collected on a random basis.
2. All the items in the sample must be independent.
3. No group should contain very few items, say less than 10. In case where the frequencies are less than 10, regrouping is done by combining the frequencies of adjoining groups so that the new frequencies become greater than 10. Some statisticians take this number as 5, but 10 is regarded as better by most of the statisticians.
4. The overall number of items must also be reasonably large. It should normally be at least 50, howsoever small the number of groups may be.
5. The constraints must be linear. Constraints which involve linear equations in the cell frequencies of a contingency table (i.e., equations containing no squares or higher powers of the frequencies) are known are know as linear constraints.

Sum 1: A die is thrown 132 times with following results:

Number turned up	1	2	3	4	5	6
Frequency	16	20	25	14	29	28

Is the die unbiased?

Solution

Let us take the hypothesis that the die is unbiased. If that is so, the probability of obtaining any one of the six numbers is 1/6 and as such the expected frequency of any one number coming upward is 132 ×1/6 = 22. Now we can write the observed frequencies along with expected frequencies and work out the value of $\chi 2$ as follows:

No. Turned UP	Observed Frequency	Expected Frequency	(Oi – Ei)	(Oi – Ei)	(Oi – Ei) /Ei
1	16	22	–6	36	36/22
2	20	22	–2	4	4/22
3	25	22	3	9	9/22
4	14	22	–8	64	64/22
5	29	22	7	49	49/22
6	28	22	6	36	36/22

$\sum (O-E)^2/E = 9$

Hence, the calculated $\chi^2=9$

DF = (n-1) (6-1)=5

The table value of $\chi 2$ for 5 degrees of freedom at 5 per cent level of significance is 11.071. Comparing calculated and table values of $\chi 2$, we find that calculated value is less than the table value and as such could have arisen due to fluctuations of sampling. The result, thus, supports the hypothesis and it can be concluded that the die is unbiased.

Sum 2: Find the value of □2 for the following information:

Class	*A*	*B*	*C*	*D*	*E*
Observed frequency	8	29	44	15	4
Theoretical (or expected) frequency	7	24	38	24	7

Solution : Since some of the frequencies less than 10, we shall first re-group the given data as follows and then will work out the value of $\chi 2$

Class	Observed Frequency	Expected Frequency	O-E	$(O-E)^2/E$
A and *B*	(8 + 29) = 37	(7 + 24) = 31	6	36/31
C	44	38	6	36/38
D and *E*	(15 + 4) = 19	(24 + 7) = 31	–12	144/31

$\sum (O-E)^2/E = 6.76$ (Approximate)

Sum 3: Genetic theory states that children having one parent of blood type A and the other of blood type B will always be of one of three types, A, AB, B and that the proportion of three types will on an average be as 1 : 2 : 1. A report states that out of 300 children having one A parent and B parent, 30 per cent were found to be types A, 45 per cent per cent type AB and remainder type B. Test the hypothesis by $\chi 2$ test.

Solution: The observed frequencies of type A, AB and B is given in the question are 90, 135 and 75 respectively.

The expected frequencies of type A, AB and B (as per the genetic theory) should have been 75, 150 and 75 respectively.

We now calculate the value of $\chi 2$ as follows:

Type	Observed Frequency	Expected Frequency	O-E	$(O-E)^2$	$(O-E)^2/E$
A	90	75	15	225	225/75 = 3
AB	135	150	–15	225	225/150 = 1.5
B	75	75	0	0	0/75 = 0

$\sum (O-E)^2/E = 3 + 1.5 \quad = 4.5$

Table value of $\chi 2$ for 2 d.f. at 5 per cent level of significance is 5.991.

The calculated value of $\chi 2$ is 4.5 which is less than the table value and hence can be ascribed to have taken place because of chance. This supports the theoretical hypothesis of the genetic theory that on an average type A, AB and B stand in the proportion of 1 : 2 : 1.

Sum 4: The table given below shows the data obtained during outbreak of smallpox:

Particulars	Attacked	Not attacked	Total
Vaccinated	31	469	500
Not vaccinated	185	1315	1500
Total	**216**	**1784**	**2000**

Test the effectiveness of vaccination in preventing the attack from smallpox. Test your result with the help of $\chi 2$ at 5 per cent level of significance.

Solution: Let us take the hypothesis that vaccination is not effective in preventing the attack from smallpox i.e., vaccination and attack are independent.

Group	Observed Frequency	Expected Frequency	O-E	$(O-E)^2$	$(O-E)^2/E$
AB	31	54	–23	529	529/54 = 9.796
Ab	469	446	+23	529	529/44 = 1.186
aB	158	162	+23	529	529/162 = 3.265
ab	1315	1338	–23	529	529/1338 = 0.395

$\sum (O-E)^2/E = 14.462$

Degrees of freedom in this case = $(r - 1)(c - 1) = (2 - 1)(2 - 1) = 1$.

The table value of χ^2 for 1 degree of freedom at 5 per cent level of significance is 3.841. The calculated value of χ^2 is much higher than this table value and hence the result of the experiment does not support the hypothesis. We can, thus, conclude that vaccination is effective in preventing the attack from smallpox.

Sum 5: Two research workers classified some people in income groups on the basis of sampling studies. Their results are as follows:

Investigators	Income groups			Total
	Poor	Middle	Rich	
A	160	30	10	200
B	140	120	40	300
Total	**300**	**150**	**50**	**500**

Show that the sampling technique of at least one research worker is defective.

Solution

Let us take the hypothesis that the sampling techniques adopted by research workers are similar (i.e., there is no difference between the techniques adopted by research workers). This being so, the expectation of A investigator classifying the people in

OF	EF	(O-E)	$(O-E)^2/E$
160	120	40	1600/120 = 13.33
30	60	–30	900/60 = 15.00
10	20	–10	100/20 = 5.00
140	180	–40	1600/180 = 8.88
120	90	30	900/90 = 10.00
40	30	10	100/30 = 3.33

$\sum (O-E)^2/E = 55.54$

Degrees of freedom = $(c - 1)(r - 1)$ = = (3 – 1) (2 – 1) = 2.

The table value of $\chi 2$ for two degrees of freedom at 5 per cent level of significance is 5.991. The calculated value of $\chi 2$ is much higher than this table value which means that the calculated value cannot be said to have arisen just because of chance. It is significant. Hence, the hypothesis does not hold good. This means that the sampling techniques

adopted by two investigators differ and are not similar. Naturally, then the technique of one must be superior than that of the other.

Sum 6: In an anti a malarial campaign in a certain area, quinine was administered to 812 persons out of a total population of 3248. The number of fever cases is shown below:

Treatment	Fever	No Fever	Total
Quinine	20	792	812
No Quinine	220	2216	2436
Total	**240**	**3008**	**3248**

Discuss the usefulness of Quinine in checking malaria.

Solution

Let us take the hypothesis that quinine is not effective in checking malaria.

Observed Frequency	Expected Frequency	$(O-E)^2$	$(O-E)^2/E$
20	60	1600	26.667
220	180	1600	8.889
792	752	1600	2.128
2216	2256	1600	0.709
			$(\sum((O-E)^2/E)=38.393$

$$\chi^2 = \Sigma \frac{(O_{ij} - E_{ij})^2}{E_{ij}}$$

= 38.393

v= (r-1)(c-1) = (2-1) (2-1)=1

v=1, $\chi^2_{0.05}$ = 3.84

The calculated value of χ^2 is greater than the table value. The hypothesis is rejected. Hence, quinine is useful in checking malaria.

Sum 7: Based on information on 1000 randomly selected fields about the tenancy status of the cultivation of these fields and use of fertilizers,

collected in an agro-economic survey, the following classification was noted:

Treatment	Owned	Rented	Total
Using Fertilizer	416	184	600
Not using Fertilizer	64	336	400
Total	**480**	**520**	**1000**

Would you conclude that owner cultivators are more inclined towards the use of fertilizer at 5% level? Carry out chi-square test as per testing procedure.

Solution

Let us take the hypothesis that ownership of fields and the use of fertilizers are independent attributes.

Observed Frequency	Expected Frequency	$(O-E)^2$	$(O-E)^2/E$
416	288	16384	56.889
64	192	16384	85.333
184	312	16384	52.513
336	208	16384	78.769
			$(\sum((O-E)^2/E)=273.504$

$$\chi^2 = \Sigma \frac{\left(O_{ij} - E_{ij}\right)^2}{E_{ij}}$$

= 273.504

v= (r-1)(c-1) = (2-1) (2-1)=1

v=1, $\chi^2_{0.05}$ = 3.84

The calculated value of χ^2 is much more than the table value. The hypothesis is rejected. Hence, it can be concluded that owner's cultivators are more inclined towards the use of fertilizers.

Sum 8: In an experiment on immunization of cattle from tuberculosis, the following results were obtained.

Particulars	Affected	Not Affected
Inoculated	12	26
Not Inoculated	16	6

Calculate χ^2 and discuss the effect of vaccine in controlling suspectability to tuberculosis. (5% value of χ^2 for one degree of freedom =3.84)

Solution:

Let us take the hypothesis that the vaccine is not effective in controlling susceptibility to tuberculosis.

Observed Frequency	Expected Frequency	$(O-E)^2$	$(O-E)^2/E$
12.5	17.7	27.04	1.528
15.5	10.3	27.04	2.625
25.5	20.3	27.04	1.332
6.5	11.7	27.04	2.311
			$(\sum((O-E)^2/E)=7.796$

$$\chi^2 = \sum \frac{\left(O_{ij} - E_{ij}\right)^2}{E_{ij}}$$

= 7.796

v= (r-1)(c-1) = (2-1) (2-1)=1

v=1, $\chi^2_{0.05}$ = 3.84

Since the calculated value of χ^2 is greater than the table value, the hypothesis is not true. We, therefore, conclude that vaccine is effective in controlling susceptibility to tuberculosis.

Sum 9: From the data given below about the treatment of 250 patients suffering from a disease, state whether the new treatment is superior to the conventional treatment.

Treatment	**Favourable**	**Unfavourable**	**Total**
New	140	30	170
Conventional	60	20	80
Total	**200**	**50**	**250**

Solution

Observed Frequency	**Expected Frequency**	**(O-E)²**	**(O-E)²/E**
140	136	16	0.118
60	64	16	0.250
30	34	16	0.471
20	16	16	1.000
			(∑((O-E)²/E)=1.839

$$\chi^2 = \Sigma \frac{\left(O_{ij} - E_{ij}\right)^2}{E_{ij}}$$

= 1.839

v= (r-1)(c-1) = (2-1) (2-1)=1

v=1, $\chi^2_{0.05}$ = 3.84

The calculated value of χ^2 is less than the table value. The hypothesis is accepted. Hence, there is no significant difference between the new and conventional treatment.

Sum 10: 1000 students at College level are graded according to IQ and their economic conditions. Use Chi-square test to find out whether there is any association between economic conditions and the level of IQ.

Economic Condition	**Intelligent Quotient**			**Total**
	High	**Medium**	**Low**	
Rich	160	300	140	600
Poor	140	100	160	400
Total	**300**	**400**	**300**	**1000**

Solution

Let us take the hypothesis that there is no association between economic conditions and the level of IQ. On the basis of this hypothesis the expected frequencies corresponding to (a) and (b) are:

Observed Frequency	**Expected Frequency**	**(O-E)²**	**(O-E)²/E**
160	180	400	2.222
140	120	400	3.333
300	240	3600	15.000
100	160	3600	22.500
140	180	1600	8.889
160	120	1600	13.333
			(∑((O-E)²/E)=65.277

$$\chi^2 = \Sigma \frac{\left(O_{ij} - E_{ij}\right)^2}{E_{ij}}$$

= 65.277

v= (r-1)(c-1) = (2-1) (3-1)=2

v=2, $\chi^2_{0.05}$ = 5.99

The calculated value of χ^2 is much greater than the table value. The hypothesis is rejected. Hence, there is association between economic condition and the level of IQ.

ANALYSIS OF VARIANCE (ANOVA)

Analysis of variance (abbreviated as ANOVA) is an extremely useful technique concerning researches in the fields of economics, biology, education, psychology, sociology, and business/industry and in researches of several other disciplines. This technique is used when multiple sample cases are involved. As stated earlier, the significance of the difference between the means of two samples can be judged through either z-test or the t-test, but the difficulty arises when we happen to examine the significance of the difference amongst more than two sample means at the same time. The ANOVA technique enables us to perform this simultaneous test and as such is considered to be an important tool of analysis in the hands of a researcher. Using this technique, one can draw inferences about whether the samples have been drawn from populations having the same mean.

The ANOVA technique is important in the context of all those situations where we want to compare more than two populations such as in comparing the yield of crop from several varieties of seeds, the gasoline mileage of four automobiles, the smoking habits of five groups of university students and so on. In such circumstances one generally does not want to consider all possible combinations of two populations at a time for that would require a great number of tests before we would be able to arrive at a decision. This would also consume lot of time and money, and even then certain relationships may be left unidentified (particularly the interaction effects). Therefore, one quite often utilizes

the ANOVA technique and through it investigates the differences among the means of all the populations simultaneously.

Professor R.A. Fisher was the first man to use the term 'Variance'* and, in fact, it was he who developed a very elaborate theory concerning ANOVA, explaining its usefulness in practical field. Later on Professor Snedecor and many others contributed to the development of this technique. ANOVA is essentially a procedure for testing the difference among different groups of data for homogeneity. "The essence of ANOVA is that the total amount of variation in a set of data is broken down into two types, that amount which can be attributed to chance and that amount which can be attributed to specified causes."1 There may be variation between samples and also within sample items. ANOVA consists in splitting the variance for analytical purposes. Hence, it is a method of analyzing the variance to which a response is subject into its various components corresponding to various sources of variation. Through this technique one can explain whether various varieties of seeds or fertilizers or soils differ significantly so that a policy decision could be taken accordingly, concerning a particular variety in the context of agriculture researches. Similarly, the differences in various types of feed prepared for a particular class of animal or various types of drugs manufactured for curing a specific disease may be studied and judged to be significant or not through the application of ANOVA technique. Likewise, a manager of a big concern can analyze the performance of various salesmen of his concern in order to know whether their performances differ significantly.

Thus, through ANOVA technique one can, in general, investigate any number of factors which are hypothesized or said to influence the dependent variable. One may as well investigate the differences amongst

various categories within each of these factors which may have a large number of possible values. If we take only one factor and investigate the differences amongst its various categories having numerous possible values, we are said to use one-way ANOVA and in case we investigate two factors at the same time, then we use two-way ANOVA. In a two or more way ANOVA, the interaction (i.e., inter-relation between two independent variables/factors), if any, between two independent variables affecting a dependent variable can as well be studied for better decisions.

BASIC PRINCIPLES OF ANOVA

The basic principle of ANOVA is to test for differences among the means of the populations by examining the amount of variation within each of these samples, relative to the amount of variation between the samples. In terms of variation within the given population, it is assumed that the values of (X_{ij}) differ from the mean of this population only because of random effects i.e., there are influences on (X_{ij}) which are unexplainable, whereas in examining differences between populations we assume that the difference between the mean of the jth population and the grand mean is attributable to what is called a 'specific factor' or what is technically described as treatment effect. Thus while using ANOVA, we assume that each of the samples is drawn from a normal population and that each of these populations has the same variance. We also assume that all factors other than the one or more being tested are effectively controlled. This, in other words, means that we assume the absence of many factors that might affect our conclusions concerning the factor(s) to be studied.

In short, we have to make two estimates of population variance viz., one based on between samples variance and the other based on

within samples variance. Then the said two estimates of population variance are compared with F-test, wherein we work out.

Estimate of population variance based on between sample variance

F= ----------

Estimate of population variance based on within sample variance

Sum 1: Set up an analysis of variance table for the following per acre production data for three varieties of wheat, each grown on 4 plots and state if the variety differences are significant.

Plot of land	Per acre production data Variety of wheat		
	A	B	C
1	6	5	5
2	7	5	4
3	3	3	3
4	8	7	4

Solution: We can solve the problem by the direct method or by short-cut method, but in each case we shall get the same result. We try below both the methods.

Solution through direct method: First we calculate the mean of each of these samples:

Average of X_1 = 6+7+3+8 / 4 = 6

Average of X_2 = 5+5+3+7 / 4 = 5

Average of X_3 = 5+4+3+4 / 4 = 4

Mean of the sample means or $\overline{\overline{X}} = \frac{\overline{X}_1 + \overline{X}_2 + \overline{X}_3}{k}$

= 6 + 5 + 3 /3

= 5

Now we work out SS between and SS within Samples

$$SS \text{ between} = n_1\left(\overline{X}_1 - \overline{\overline{X}}\right)^2 + n_2\left(\overline{X}_2 - \overline{\overline{X}}\right)^2 + n_3\left(\overline{X}_3 - \overline{\overline{X}}\right)^2$$

= $4(6 – 5)^2 + 4(5 – 5)^2 + 4(4 – 5)^25$
= 4 + 0 + 4
= 8

$$SS \text{ within} = \Sigma\left(X_{1i} - \overline{X}_1\right)^2 + \Sigma\left(X_{2i} - \overline{X}_2\right)^2 + \Sigma\left(X_{3i} - \overline{X}_3\right)^2, \qquad i = 1, 2, 3, 4$$

= $(6 – 6)^2 + (7 – 6)^2 + (3 – 6)^2 + (8 – 6)^2$ } +
$(5 – 5)^2 + (5 – 5)^2 + (3 – 5)^2 + (7 – 5)^2$ } +
$(5 – 4)^2 + (4 – 4)^2 + (3 – 4)^2 + (4 – 4)^2$ }
{0 + 1 + 9 + 4} + {0 + 0 + 4 + 4} + {1 + 0 + 1 + 0}
14 + 8 + 2
24

$$SS \text{ for total variance} = \Sigma\left(X_{ij} - \overline{\overline{X}}\right)^2 \qquad i = 1, 2, 3\ldots$$

$$j = 1, 2, 3\ldots$$

= $(6 – 5)^2 + (7 – 5)^2 + (3 – 5)^2 + (8 – 5)^2$
+ $(5 – 5)^2 + (5 – 5)^2 + (3 – 5)^2$
+ $(7 – 5)^2 + (5 – 5)^2 + (4 – 5)^2$
+ $(3 – 5)^2 + (4 – 5)^2$
= 1 + 4 + 4 + 9 + 0 + 0 + 4 + 4 + 0 + 1 + 4 + 1
= 32

Source of Variation	SS	d.f.	MS	F-ratio	5% F-limit(from the F-table)
Between sample	8	(3 – 1) = 2	8/2 = 4.00	4.00/2.67 = 1.5	$F(2, 9) = 4.26$
Within sample	24	(12 – 3) = 9	24/9 =		

			2.67		
Total	32	(12 – 1) = 11			

The above table shows that the calculated value of F is 1.5 which is less than the table value of 4.26 at 5% level with d.f. being v1 = 2 and v2= 9 and hence could have arisen due to chance. This analysis supports the null-hypothesis of no difference is sample means. We may, therefore, conclude that the difference in wheat output due to varieties is insignificant and is just a matter of chance.

Sum 2: To assess the significance of possible variation in performance in a certain test between the grammar schools of a city, a common test was given to a number of students taken at random from the senior fifth class of each of the four schools concerned. The results are given below. Make an analysis of variance of data.

Schools			
A	**B**	**C**	**D**
8	12	18	13
10	11	12	9
12	9	16	12
8	14	14	16
7	4	4	15

Solution

Sample 1 X_1		**Sample 2 X_2**	**Sample 3 X_3**	**Sample 4 X_4**
8		12	18	13
10		11	12	9
12		9	16	12
8		14	6	16
7		4	8	15
Total	**45**	**50**	**60**	**65**
Avg. of X	**9**	**10**	**12**	**13**

Grand Mean of $\bar{\bar{X}}$ = $\bar{X}_1 + \bar{X}_2 + \bar{X}_3 + \bar{X}_4 / N$

= 9 + 10 + 12 + 13 /4

= 44/4

= 11

Grand Mean of all Samples of $\bar{\bar{X}}$ = 45 + 50 + 60 + 65 / 20

= 11

Variance between Samples

To obtain variance between samples, calculate square of the deviation of the various samples from the grand average. The mean of sample I is 9 but the grand mean is 11. So we will take the difference between 9 and 11 and square it. Similarly for sample II the mean is 10 but the grand average is 11 and so will take the difference between 10 and 11 and square it. Thus we will have the following table.

Sample 1 $(\bar{X}_1 - \bar{\bar{X}}_1)^2$	**Sample 2** $(\bar{X}_2 - \bar{\bar{X}}_2)^2$	**Sample 3** $(\bar{X}_3 - \bar{\bar{X}}_3)^2$	**Sample 4** $(\bar{X}_4 - \bar{\bar{X}}_4)^2$
4	1	1	4
4	1	1	4
4	1	1	4
4	1	1	4
4	1	1	4
20	**5**	**5**	**20**

Sum of the squares between the samples = 20 + 5 + 5 + 20 = 50

Mean sum of squares between the samples is 50 / 4-1 = 16.7 (because there are four samples and the degrees of freedom are 4-1=3)

Variance within the Samples

Here we find the sum of the squares is the deviations of various items in a sample from the mean values of respective samples. Thus for the first sample, then mean is 9 and so we will take deviations from 10, and so on. The squared deviations are given in the following table.

Sample 1		**Sample 2**		**Sample 3**		**Sample 4**	
X_1	$(X_1 - \bar{X}_1)^2$	X_2	$(X_2 - \bar{X}_2)^2$	X_3	$(X_3 - \bar{X}_3)^2$	X_4	$(X_4 - \bar{X}_4)^2$
8	1	12	4	18	36	13	0
10	1	11	1	12	0	9	16
12	9	9	1	16	16	12	1
8	1	14	16	6	36	16	9
7	4	4	36	8	16	15	4
$\sum(X_1 - \bar{X}_1)^2=16$		**$\sum(X_2 - \bar{X}_2)^2=58$**		**$\sum(X_3 - \bar{X}_3)^2=104$**		**$\sum(X_4 - \bar{X}_4)^2=30$**	

Total sum of squares within the samples = 16 + 58 + 104 + 30 = 208

Mean Sum of Squares within the Samples

= 208 / 20-4 = 208 / 16 = 13

It is advisable to check up the calculations by finding out total variation. Total variation is calculated by taking the squares of the deviation of each item from the grand average.

Sample 1		**Sample 2**		**Sample 3**		**Sample 4**	
X_1	$(X_1 - \bar{X}_1)^2$	X_2	$(X_2 - \bar{X}_2)^2$	X_3	$(X_3 - \bar{X}_3)^2$	X_4	$(X_4 - \bar{X}_4)^2$
8	9	12	1	18	49	13	4
10	1	11	0	12	1	9	4
12	1	9	4	16	25	12	1
8	9	14	9	6	25	16	25
7	16	4	49	8	9	15	16
$\sum(X_1 - \bar{X}_1)^2=36$		**$\sum(X_2 - \bar{X}_2)^2=63$**		**$\sum(X_3 - \bar{X}_3)^2=109$**		**$\sum(X_4 - \bar{X}_4)^2=50$**	

Total sum of squares = 36 + 63 + 109 + 50 = 258

Degrees of Freedom = 20 – 1 = 19

When we add the sum of squares between samples and sum of squares within the samples, we get the same total, i.e. 50 + 208 = 258. Hence, our calculations are correct.

Source of Variation	Sum of Squares	Degrees of Freedom	Mean Square
Between Samples	50	3	16.7
Within Samples	208	16	13.0
Total	**258**	**19**	

$$F = \frac{\text{Variance between Samples}}{\text{Variance within Samples}} = \frac{16.7}{13} = 1.285$$

The table value of F for v1=3 and v2=16 at 5% level of significance =3.24. The calculated value of F is less than the table value and hence the difference in the mean values of the sample is not significant. i.e., the samples could have come from the same variance.

Sum 3: The three samples below have been obtained from normal populations with equal variances. Test the hypothesis that the sample means are equal.

8	7	12
10	5	9
7	10	13
14	9	12
11	9	14

The table value of F at 5% level of significance for v1=1 and v2=12 is 3.88.

Solution

Let us taken the null hypothesis that there is no significant difference in the means of three samples.

X_1	X_2	X_3
8	7	12
10	5	9
7	10	13
14	9	12
11	9	14
Total 50	**40**	**60**
$\bar{X}$ **10**	**8**	**12**

$$\bar{\bar{X}} = 10 + 8 + 12 / 3 \quad = 10$$

Variance between Samples

$(\bar{X}_1 - \bar{\bar{X}})^2$	$(\bar{X}_2 - \bar{\bar{X}})^2$	$(\bar{X}_3 - \bar{\bar{X}})^2$
0	4	4
0	4	4
0	4	4
0	4	4
0	4	4
0	**20**	**20**

Sum of Squares between samples = 0 + 20 + 20 =40

Variance within Samples

X_1	$(X_1 - \bar{X}_1)^2$	X_2	$(X_2 - \bar{X}_2)^2$	X_3	$(X_3 - \bar{X}_3)^2$
8	4	7	1	12	0
10	0	5	9	9	9
7	9	10	4	13	1
14	16	9	1	12	0
11	1	9	1	14	4
	30		**16**		**14**

Sum of Squares within Samples = 30 + 16 + 14 = 60

ANOVA TABLE

Source of Variation	Sum of Squares	V	Mean Square
Between	40	2	22
Within	60	12	5
Total	**100**	**14**	

F=20/5 = 4

For v1=2 and v2=12, $F_{0.05}$ = 3.88

The calculated value of F is greater than the table value. The hypothesis is rejected. Hence, there is significant difference in the sample means.

SCALING TECHNIQUES

In research we quite often face measurement problem (since we want a valid measurement but may not obtain it), specially when the concepts to be measured are complex and abstract and we do not possess the standardised measurement tools. Alternatively, we can say that while measuring attitudes and opinions, we face the problem of their valid measurement. Similar problem may be faced by a researcher, of course in a lesser degree, while measuring physical or institutional concepts

As such we should study some procedures which may enable us to measure abstract concepts more accurately. This brings us to the study of scaling techniques.

MEANING OF SCALE

Scaling describes the procedures of assigning numbers to various degrees of opinion, attitude and other concepts. This can be done in two

ways viz., (i) making a judgement about some characteristic of an individual and then placing him directly on a scale that has been defined in terms of that characteristic and (ii) constructing questionnaires in such a way that the score of individual's responses assigns him a place on a scale. It may be stated here that a scale is a continuum, consisting of the highest point (in terms of some characteristic e.g., preference, favourableness, etc.) and the lowest point along with several intermediate points between these two extreme points. These scale-point positions are so related to each other that when the first point happens to be the highest point, the second point indicates a higher degree in terms of a given characteristic as compared to the third point and the third point indicates a higher degree as compared to the fourth and so on

Numbers for measuring the distinctions of degree in the attitudes/opinions are, thus, assigned to individuals corresponding to their scale-positions. All this is better understood when we talk about scaling technique(s). Hence the term 'scaling' is applied to the procedures for attempting to determine quantitative measures of subjective abstract concepts. Scaling has been defined as a "procedure for the assignment of numbers (or other symbols) to a property of objects in order to impart some of the characteristics of numbers to the properties in question."

MEASUREMENT OF SCALES

1) Nominal scale

Nominal scale is simply a system of assigning number symbols to events in order to label them. The usual example of this is the assignment of numbers of basketball players in order to identify them. Such numbers cannot be considered to be associated with an ordered scale for their order is of no consequence; the numbers are just convenient labels for the particular class of events and as such have no

quantitative value. Nominal scales provide convenient ways of keeping track of people, objects and events. One cannot do much with the numbers involved. For example, one cannot usefully average the numbers on the back of a group of football players and come up with a meaningful value. Neither can one usefully compare the numbers assigned to one group with the numbers assigned to another. The counting of members in each group is the only possible arithmetic operation when a nominal scale is employed. Accordingly, we are restricted to use mode as the measure of central tendency. There is no generally used measure of dispersion for nominal scales. Chi-square test is the most common test of statistical significance that can be utilized, and for the measures of correlation, the contingency coefficient can be worked out.

Nominal scale is the least powerful level of measurement. It indicates no order or distance relationship and has no arithmetic origin. A nominal scale simply describes differences between things by assigning them to categories. Nominal data are, thus, counted data. The scale wastes any information that we may have about varying degrees of attitude, skills, understandings, etc. In spite of all this, nominal scales are still very useful and are widely used in surveys and other ex-post-facto research when data are being classified by major sub-groups of the population.

2) Ordinal Scale

The lowest level of the ordered scale that is commonly used is the ordinal scale. The ordinal scale places events in order, but there is no attempt to make the intervals of the scale equal in terms of some rule.

Rank orders represent ordinal scales and are frequently used in research relating to qualitative phenomena. A student's rank in his graduation class involves the use of an ordinal scale. One has to be very careful in making statement about scores based on ordinal scales. For instance, if Ram's position in his class is 10 and Mohan's position is 40, it cannot be said that Ram's position is four times as good as that of Mohan. The statement would make no sense at all. Ordinal scales only permit the ranking of items from highest to lowest. Ordinal measures have no absolute values, and the real differences between adjacent ranks may not be equal. All that can be said is that one person is higher or lower on the scale than another, but more precise comparisons cannot be made.

Thus, the use of an ordinal scale implies a statement of 'greater than' or 'less than' (an equality statement is also acceptable) without our being able to state how much greater or less. The real difference between ranks 1 and 2 may be more or less than the difference between ranks 5 and 6. Since the numbers of this scale have only a rank meaning, the appropriate measure of central tendency is the median. A percentile or quartile measure is used for measuring dispersion. Correlations are restricted to various rank order methods. Measures of statistical significance are restricted to the non-parametric methods.

3) Interval Scale

In the case of interval scale, the intervals are adjusted in terms of some rule that has been established as a basis for making the units equal. The units are equal only in so far as one accepts the assumptions on which the rule is based. Interval scales can have an arbitrary zero, but it is not possible to determine for them what may be called an absolute zero or the unique origin. The primary limitation of the interval scale is the lack of a true zero; it does not have the capacity to measure

the complete absence of a trait or characteristic. The Fahrenheit scale is an example of an interval scale and shows similarities in what one can and cannot do with it. One can say that an increase in temperature from 30° to 40° involves the same increase in temperature as an increase from 60° to 70°, but one cannot say that the temperature of 60° is twice as warm as the temperature of 30° because both numbers are dependent on the fact that the zero on the scale is set arbitrarily at the temperature of the freezing point of water. The ratio of the two temperatures, 30° and 60°, means nothing because zero is an arbitrary point.

Interval scales provide more powerful measurement than ordinal scales for interval scale also incorporates the concept of equality of interval. As such more powerful statistical measures can be used with interval scales. Mean is the appropriate measure of central tendency, while standard deviation is the most widely used measure of dispersion. Product moment correlation techniques are appropriate and the generally used tests for statistical significance are the 't' test and 'F' test.

4) Ratio Scale

Ratio scales have an absolute or true zero of measurement. The term 'absolute zero' is not as precise as it was once believed to be. We can conceive of an absolute zero of length and similarly we can conceive of an absolute zero of time. For example, the zero point on a centimeter scale indicates the complete absence of length or height. But an absolute zero of temperature is theoretically unobtainable and it remains a concept existing only in the scientist's mind. The number of minor traffic-rule violations and the number of incorrect letters in a page of type script represent scores on ratio scales. Both these scales have absolute zeros and as such all minor traffic violations and all typing errors can be assumed to be equal in significance. With ratio scales

involved one can make statements like "Jyoti's" typing performance was twice as good as that of "Reetu." The ratio involved does have significance and facilitates a kind of comparison which is not possible in case of an interval scale.

Ratio scale represents the actual amounts of variables. Measures of physical dimensions such as weight, height, distance, etc. are examples. Generally, all statistical techniques are usable with ratio scales and all manipulations that one can carry out with real numbers can also be carried out with ratio scale values. Multiplication and division can be used with this scale but not with other scales mentioned above. Geometric and harmonic means can be used as measures of central tendency and coefficients of variation may also be calculated.

Thus, proceeding from the nominal scale (the least precise type of scale) to ratio scale (the most precise), relevant information is obtained increasingly. If the nature of the variables permits, the researcher should use the scale that provides the most precise description. Researchers in physical sciences have the advantage to describe variables in ratio scale form but the behavioural sciences are generally limited to describe variables in interval scale form, a less precise type of measurement.

IMPORTANT SCALING TECHNIQUES

A) Rating Scale

In rating scale, the rater makes judgment about some characteristic of a subject and places him directly on some point on the scale. Rating scale may be either a graphic rating scale or an itemized random scale.

1) Graphic Rating Scale

It is quite simple and various points are usually put along the line to form a continuum and the rater indicates his rating by making a tick mark at the appropriate point on a line that runs from one extreme to other.

2) The Itemized Rating Scale

It is also known as numerical scale, presents a series of statements from which a respondent select one as best reflecting his evaluation. These statements are ordered progressively in terms of more or less of some property. It provide more information and meaning to the rater and increases reliability. This form is difficult to develop and the statement may not say exactly what the respondent would like to express.

The advantage of rating scale is that the results attained from these scales are comparable favorably with alternative methods. They require less time, interesting to us and have wide range of applications. One of the disadvantages of rating scales is that it suffers from lack of reliability and validity.

B) Attitude Scales

The attitude scales are constructed with sets of rating scales designed to measures one or more aspects of an individual's or group's attitude towards some objects. The individual responses to various scales may be aggregated to provide a single attitude for the individual. There are different attitude scales. They are

1) Likert's Summated Scale

It was developed by Likert, which is frequently used in the measurement of social attitude. It uses only the definitely favorable and unfavorable statement and does not take into account the intermediate

position, and the respondent is asked to react. The respondent indicates his agreements or disagreement with each statement. Each response is given a numerical score and the score are totaled to measure the respondent's attitude. The overall scale represents the respondent's position on the continuum of favorable – unfavorable towards an issue.

Half of the statements (usually 15) included in the questionnaire are favorable and the rest ones are unfavorable.

Strongly Agree	-5	or	+2
Agree	-4	or	+1
Undecided	-3	or	0
Disagree	-2	or	-1
Strongly Disagree	-1	or	-2

Procedure

i. At the out set, subjects are dividing in to a couple of arbitrarily defined groups. For instance, those subjects with top 25% of all total scores and those with the lowest 25% of all total scores are constructed to be in possession of the most favorable and least favorable attitudes

ii. Thereafter the researcher calculates the mean score for each statement separately

iii. The difference between the two mean scores, in respect of each statement is calculated

iv. Finally, all statements are ranked according to their difference in mean scores. Those with mean differences near zero are considered poor and therefore eliminated.

Advantages

i. Item analysis increases the degree of homogeneity or internal consistency in the set of statements

ii. This method is less difficult

iii. Since a wide range of answers are given to the subject, they don't find it difficult to respond and express the intensity of their feeling

iv. Since there is no involvement of the outside group of judges in selection of the statements, it does not suffer from the problem of subjectivity

Disadvantages

i. Ties in rank are likely to occur quite frequently due to equality in total score values

ii. The response pattern of two persons having exactly identical scores may be significantly different

iii. It suffers from the problem of interpretation which does not arise in Thurstone's scale

iv. The subject is required to respond to all statement, whereas in Thurstone scale, he is required to check only those statement with which he agrees

2) Thurstone's Equal Appearing Interval Scale

L.L. Thurstone is the inventor of this scale. This scale consists of 15 to 20 statements which form a continuum of attitudes towards a subject ranging from the most favorable to the least favorable.

Procedure

i. The researcher gather a large number of statements, usually 20 or more, that express various points of view towards a group, idea or practice

ii. These statements are then submitted to a panel of judges, each of whom arranges them in eleven group or piles ranging from one extreme to another in position. Each of the judges is requested to place generally in the first pile the statements which he thinks are unfavorable issue, in the second pile to place those statement which he thinks all next most unfavorable and he goes on doing so in this manner till in the 11th pile he put the statement which he considers to be most favorable

iii. This sorting by each judge a composite position for each of the items

iv. For items that are retained, each is given its median scale value between one and 11th as established by the panel. In other words the scale value of any one statement is computed as the median position to which it is assigned by the group of judges

v. A final selection of statement is then made. For this purpose a sample of statement, whose median scores is spread evenly from one extreme to the other is taken. The statement so selected constitutes the final scale to be administered to respondents

vi. After developing the scale, the respondent is asked during the administration of scale to check the statements with which they agree. Respondents score is equal to the average of scale values attached to the items he endorses. This average is either median or mean.

Disadvantages

i. The procedure involved in the construction is very costly and time consuming

ii. The scale does not allow subjects to express the intensity of their feelings

iii. The scale values assigned to statements are influenced by the attitudes, background and intelligence of judges who see things differently from actual respondents

iv. This method is not completely objective, it involves subjective decision

3) Bogardus's Scale of Social Distance

Bogardus used cumulative scale containing a number of questions regarding a particular issue. The respondent is required to express his agreement or disagreement over that issue. The respondents who answer favorably have higher aggregate score than those who answer unfavorably. The score is computed by counting the number of items which is responded favorably by the respondent. The respondent is placed on a particular position on the scale on the basis of his scores. Bogardus used this scale to know social distance by measuring the attitude of individual towards a particular social group. In this scale a number of suggested relationships are listed to which member of an ethic group are admitted. The respondent is asked to indicate to which social group is admissible to him for each specified relationship in terms of his willingness to accept social distance.

The respondent is asked to tick off against each of 7 categories of relationship, he is willing to accept a average number of a particular ethic, social or nationality group, because the respondents first feeling reactions are known by this. In this scale the respondent has to express his reaction to each race as a group without having any regard to any individual members of a group, whether he likes of dislikes.

In order to calculate social distance mathematically weights are assigned to different categories of relationship.

Procedure

i. The weight and percentage response in respect of each category are to be placed in rows
ii. The percentage responses are to be multiplied by its weight
iii. The product is to be summed up so as to indicate social distance

Limitations

i. Although it is expected that the respondent has to express his reaction to each race as a group without having any regard to any individual members of that group, in reality the influence of any individual members may not be wiped out from mind of respondent while giving preferences
ii. The score is this scale does not indicates the actual extent or exact degree of preference of a group
iii. It is not always possible on the part of the respondent to be acquainted completely with a group, difficulty to express his attitude towards that group

4) Guttman's Cumulative Scale

Cumulative scale or Louis Guttmann's scalograms analysis, consist of a series of statements to which a respondent expresses his agreement or disagreement. The special feature is that they are cumulative in nature. A respondent replies favorably to item no.3 and also to item no. 2 and 1. The score of the individual is worked out by counting the number of point concerning the number of statement he answers favorably. If the investigator knows this total score, he can estimate as to how a respondent has answered individual statements consulting cumulative scales. The major scale of this type of cumulative scales is the Guttman's scalogram.

Scalogram analysis refers to the procedure for determining whether a set of items forms an unidimensional scale. A scale is said to be

unidimensional if the responses fall into a pattern in which endorsement of item reflecting extreme position results. The perfect scale in terms of Guttman's technique implies that an informant who responds to given question will have a high total score than informants who responded to it negatively.

Procedure

i. The area of concept in first defined with reference to problem in hand
ii. Ten or twelve statement are selected which are assumed to be representative of selected area
iii. The statements are arranged in form of 3 or 5 point scale so that subjects can indicates the intensity of their attitude for each item
iv. 10 or 12 items are submitted to a sample of 100 or more respondents who will check the items. This step is designed to determine the scalability of item
v. After the total score is attained for each person by adding up the weights of categories checked, the questionnaires are arranged in rank order from high to low according to total score
vi. A table of scalogram is presented from data as the questionnaire by recording separately the response of each person to each category by providing a column for each person and a row for each category
vii. The reproducibility of each item is determined on the basis that none with a tower score ranks higher in any item than any person with a higher total score. By this method the error of reproducibility is minimized.

Limitation

i. This method is not frequently used for simple reason that its development procedure is tedious and complex

ii. As regards the measurement of attitude towards some objects or prediction of behaviour relating to such objects, the unidimensional scale may not always prove to be most effective basis.

Measurement is a systematic way of assigning number or names to object and their attitudes. And scaling is the procedure for determining the quantitative measure of abstract concepts. A scale consists of a set of statements logically related referring to same attitude. Thus a scale may be used to measure the characteristic of respondent or to evaluate object presented to him.

POSSIBLE QUESTIONS

UNIT-IV

PART A

One Mark

Online Examination

PART B

Eight Marks

1. Explain the basic concept relating to hypothesis?

2 . Discuss the types of measurement scales with suitable examples.

3.Define hypothesis and discuss the importance of hypothesis in social research?

4. From the data given below about the treatment of 250 patients suffering from a disease, state whether the new treatment is superior to the conventional treatment.

	No.of patients	
	Favourable	**Not favourable**
Total		

New	140	30	170
Conventional	60	20	80
Total	**200**	**50**	**250**

Use Chi-square test at 5per cent level of significance.

Note: (The chi square value at 5 per cent level of significance for 1 degrees of freedom is 3.84)

(The chi square value at 1 per cent level of significance for 1 degrees of freedom is 6.635)

5. Explain the assumptions of Chi-square test.

6.What do you understand by hypothesis? How is it formulated and tested.

7. Raju Resraurant near the railway station at falna has been having average sales 500 tea Cups per day. Because of the development of bus stand nearby, it expects to increase in Sales. During the first 12 days after the start of the bus Stand, the daily sales were as under:

550,570,490,615,505,580,570,460,600,580,530,526.

On the basis of this sample information, can one conclude that raju restaurants sales have increased? Use t test at 5% level of significance

Note: (The t value at 5 per cent level of significance for 11 degrees of freedom is 1.796)

(The t value at 1 per cent level of significance for 11 degrees of freedom is 2.718)

8.Explain type I and type II error with suitable example.

9. Briefly explain the methods of scaling techniques with suitable examples .

10.The table given below shows the data obtained during outbreak of smallpox

	Attacked	**Not Attacked**	**Total**
Vaccinated	31	469	500
Not Vaccinated	185	1315	1500
Total	216	1784	2000

Test the effectiveness of vaccination in preventing the attack from smallpox. Test your

result with the help of chi-square at 5 per cent level of significance.

Note: (The chi square value at 5 per cent level of significance for 1 degrees of freedom is 3.84)

(The chi square value at 1 per cent level of significance for 1 degrees of freedom is 6.635)

KARPAGAM ACADEMY OF HIGHER EDUCATION

(DEEMED TO BE UNIVERSITY)

(ESTABILISHED UNDER SEC 3 OF UGC ACT 1956)

DEPARTMENT OF COMMERCE

III -B.COM RESEARCH METHODOLOGY-15CMU603B

UNIT-IV ANALYSIS OF DATA

S.NO	QUESTIONS	OPTION I	OPTION II	OPTION III	OPTION IV	ANSWER
1	___is usually considered as the principal instrument in research	Hypothesis	Report	Table of Contents	List of Charts	Hypothesis
2	relationship between__, if it happens to be a relational hypothesis	Variables	Data	Report	Objectives	Variables
3	which predicts the direction of the difference between two items	Directional	Vertical	Horizontal	alternative	Directional
4	Level of significance measures __ percentage	Data	Error	Inaccuracy	Fault	Error
5	__ denotes the chances of occurrence	Chance	Likelihood	Odds	Probability	Probability
6	probability is equal to or smaller than the α value in case of one-tailed test	Valid	Applicable	Null	Legtimate	Null
7	method A with method B about its superiority and if we proceed on the assumption that	Null Hypothesis	Alternative Hypothesis	Substitute Hypothesis	Option Hypothesis	Null Hypothesis
8	the method A is superior or the method B is inferior, we are then stating what is termed as __	Null Hypothesis	Alternative Hypothesis	Substitute Hypothesis	Option Hypothesis	Alternative Hypothesis

S.NO	QUESTIONS	OPTION I	OPTION II	OPTION III	OPTION IV	ANSWER
9	___hypothesis is also known as statistical hypothesis	Alternative	Null	Convinving	Authoratative	Null
10	___level is the maximum value of the probability of rejecting H_0	Implication	Connotation	Significance	Substance	Significance
11	We may reject H_0 when H_0 is true, then it is known as __	Type II Error	Type I Error	Type III Error	Type IV Error	Type I Error
12	We may accept H_0 when in fact H_0 is not true, then the error is known as ___	Type II Error	Type I Error	Type III Error	Type IV Error	Type II Error
13	make probability statements about population parameter	Concept	Hypothesis	Number	Data	Hypothesis
14	____hypothesis is usually the one which one wishes to prove	Alternative	Null	Convinving	Authoratative	Alternative
15	__ hypothesis is the one which one wishes to disprove	Alternative	Null	Convinving	Authoratative	Null
16	__error is denoted by β (beta) known as b error	Type I	Type II	Type III	Type IV	Type II
17	is usually determined in advance and is understood as the level of	Type I	Type II	Type III	Type IV	Type I
18	____main function is to suggest new experiments and observations	Report	Table of Contents	List of Charts	Hypothesis	Hypothesis
19	__ hypothesis represents the hypothesis we are trying to reject	Alternative	Null	Convinving	Authoratative	Null

S.NO	QUESTIONS	OPTION I	OPTION II	OPTION III	OPTION IV	ANSWER
20	of hypothesis which should have been accepted	Type I	Type II	Type III	Type IV	Type I
21	__error means accepting the hypothesis which should have been rejected	Type I	Type II	Type III	Type IV	Type II
22	___ is nothing but an assumption	Report	Table of Contents	Hypothesis	List of Charts	Hypothesis
23	't' test is employed when the sample size is less than ___	30	40	50	60	30
24	__ test should be employed, when the sample size is less than 30	T	ANOVA	Chisquare	Correlation	T
25	Student test is also known as ___	t' Test	Vertical			t' Test
26	The variable ____ranges from minus infinity to plus infinity	ANOVA	Chisquare	t' distribution	Regression	t' distribution
27	distribution, the t-distribution is ____ and has a mean zero	Symmetrical	Irregular	Asymmetrical	Uneven	Symmetrical
28	When the sample size is greater than __, it is known as large sample	30	40	50	60	30
29	among more than two groups____ test is employed	Analysis of Variance	Chisquare	Correlation	F Test	Analysis of Variance
30	To find mean difference between two groups ___ is employed	't' Test	F Test	Chisquare	Analysis of Variance	't' Test

S.NO	QUESTIONS	OPTION I	OPTION II	OPTION III	OPTION IV	ANSWER
31	exists any association between variables ___test is employed	't' Test	F Test	Chisquare	Analysis of Variance	Chisquare
32	technique specially designed to test whether the means of more than	Analysis of Variance	Chisquare	Correlation	F Test	Analysis of Variance
33	Chi-square is an example of _____ test	Parametric	Non-Parametric	Psychometric	Frequency	Non-Parametric
34	't' test is an example of _____ test	Parametric	Non-Parametric	Psychometric	Frequency	Parametric
35	observed frequency and expected frequency is known as __	't' Test	F Test	Chisquare	Analysis of Variance	Chisquare
36	freedom, if the Chi-square table has 3 rows and 3 columns ___	2	3	4	5	4
37	freedom, if the Chi-square table has 4 rows and 3 columns ___	3	4	5	6	6
38	freedom, if the Chi-square table has 3 rows and 2 columns ___	1	2	3	4	2
39	"procedure for the assignment of numbers to a property of objects	Scaling	Measurement	Numbering	Calculation	Scaling
40	of assigning number symbols to events in order to label them	Nominal	Ordinal	Interval	Ratio	Nominal
41	statements from which a respondent select one as best reflecting his	Graphic Rating Scale	Itemized Rating Scale	Attitude Scales	Likert's Summated Scale	Itemized Rating Scale

S.NO	QUESTIONS	OPTION I	OPTION II	OPTION III	OPTION IV	ANSWER
42	sets of rating scales designed to measures one or more aspects of an	Graphic Rating Scale	Itemized Rating Scale	Attitude Scales	Likert's Summated Scale	Attitude Scales
43	statements which form a continuum of attitudes towards a subject ranging	Thurstone's Equal Appearing Interval Scale	Bogardus's Scale of Social Distance	Guttman's Cumulative Scale	Likert's Summated Scale	Thurstone's Equal Appearing Interval Scale
44	statements to which a respondent expresses his agreement or	Thurstone's Equal Appearing Interval Scale	Bogardus's Scale of Social Distance	Guttman's Cumulative Scale	Likert's Summated Scale	Guttman's Cumulative Scale
45	convenient ways of keeping track of people, objects and events	Ordinal	Interval	Ratio	Nominal	Nominal
46	ordered scale that is commonly used is the __scale	Ordinal	Interval	Ratio	Nominal	Ordinal
47	__scales and are frequently used in research relating to	Interval	Ratio	Nominal	Ordinal	Ordinal
48	arbitrary zero, but it is not possible to determine for them what may be called	Interval	Ratio	Nominal	Ordinal	Interval
49	__scales have an absolute or true zero of measurement	Interval	Ratio	Nominal	Ordinal	Ratio
50	procedures of assigning numbers to various degrees of opinion,	Scaling	Measurement	Numbering	Calculation	Scaling
51	Assignment of numbers just for identification is an example of __ scaling	Ordinal	Interval	Ratio	Nominal	Nominal
52	__scale represents the actual amounts of variables	Interval	Ratio	Nominal	Ordinal	Ratio

S.NO	QUESTIONS	OPTION I	OPTION II	OPTION III	OPTION IV	ANSWER
53	makes judgment about some characteristic of a subject and places him	Rating	Itemized	Attitude	Likerts	Rating
54	___ test can be applied, while using nominal scaling	Chisquare	Mean	Median	Mode	Chisquare
55	Chisquare test can be employed only when the sample size is above ___	30	40	50	60	30
56	___ test is employed to find association between variables	Correlation	Regression	Chisquare	t' Test	Chisquare
57	___ test is employed when the sample size is less than 30	Analysis of Variance	t'	Chisquare	Correlation	t'
58	Questions with three or five point scale answer is known as	Graphic Rating Scale	Itemized Rating Scale	Attitude Scales	Likert's Summated Scale	Likert's Summated Scale
59	Assigning number for just identification purpose is known as ___	Nominal	Ordinal	Interval	Ratio	Nominal
60	When asnweres offered by respondents are ranked, it is known as ___ scaling	Nominal	Ordinal	Interval	Ratio	Ordinal

UNIT – V

SYLLABUS

Interpretation and Report Writing – Interpretation – Meaning – Technique of Interpretation, Precautions – Report Writing – Steps in Writing Report – Types of Reports – Technical and Popular Report – Oral Presentation – Precautions for Writing Research Reports

INTERPRETATION

After collecting and analyzing the data, the researcher has to accomplish the task of drawing inferences followed by report writing. This has to be done very carefully, otherwise misleading conclusions may be drawn and the whole purpose of doing research may get vitiated. It is only through interpretation that the researcher can expose relations and processes that underlie his findings. In case of hypotheses testing studies, if hypotheses are tested and upheld several times, the researcher may arrive at generalizations. But in case the researcher had no hypothesis to start with, he would try to explain his findings on the basis of some theory. This may at times result in new questions, leading to further researches. All this analytical information and consequential inference(s) may well be communicated, preferably through research report, to the consumers of research results who may be either an individual or a group of individuals or some public/private organisation.

MEANING OF INTERPRETATION

Interpretation refers to the task of drawing inferences from the collected facts after an analytical and/or experimental study. In fact, it is

a search for broader meaning of research findings. The task of interpretation has two major aspects viz., (i) the effort to establish continuity in research through linking the results of a given study with those of another and (ii) the establishment of some explanatory concepts. "In one sense, interpretation is concerned with relationships within the collected data, partially overlapping analysis. Interpretation also extends beyond the data of the study to include the results of other research, theory and hypotheses.". Thus, interpretation is the device through which the factors that seem to explain what has been observed by researcher in the course of the study can be better understood and it also provides a theoretical conception which can serve as a guide for further researches.

Why Interpretation?

Interpretation is essential for the simple reason that the usefulness and utility of research findings lie in proper interpretation. It is being considered a basic component of research process because of the following reasons:

1. It is through interpretation that the researcher can well understand the abstract principle that works beneath his findings. Through this he can link up his findings with those of other studies, having the same abstract principle, and thereby can predict about the concrete world of events. Fresh inquiries can test these predictions later on. This way the continuity in research can be maintained.
2. Interpretation leads to the establishment of explanatory concepts that can serve as a guide for future research studies; it opens new avenues of intellectual adventure and stimulates the quest for more knowledge.

3. Researcher can better appreciate only through interpretation why his findings are what they are and can make others to understand the real significance of his research findings.
4. The interpretation of the findings of exploratory research study often results into hypotheses for experimental research and as such interpretation is involved in the transition from exploratory to experimental research. Since an exploratory study does not have a hypothesis to start with, the findings of such a study have to be interpreted on a post-factum basis in which case the interpretation is technically described as 'post factum' interpretation.

TECHNIQUES OF INTERPRETATION

- Interpretation is not an easy job and it requires a great skill on the part of the investigator. The investigator gets the required expertise to apply the techniques. The techniques of interpretation are given below:

1) Relationship between Variables

The basic object of every analytical research is to find out the relationship between any two variables. There may be three types of relationship

- Symmetrical Relationship
- Reciprocal Relationship
- Asymmetrical Relationship

The interpretation of data can be made with the help of these relationships

2) Percentages

Percentages are used in making comparison between two or more series of data. They are also used to describe the relationships.

3) Averages

There are three forms of averages such as arithmetic mean, median, mode. Though there are other measures of central tendency, the above three measures are commonly used. Instead of using long statistical tables, the use of average makes the interpretation very simple.

4) Dispersion

Dispersion refers to the amount or the magnitude of the spread. Measures of dispersion include range, inter quartile range, average deviation and standard deviation. These measures help to interpret the data more scientifically

PRECAUTIONS IN INTERPRETATION

1. At the outset, researcher must invariably satisfy himself that (a) the data are appropriate, trustworthy and adequate for drawing inferences; (b) the data reflect good homogeneity; and that (c) proper analysis has been done through statistical methods
2. The researcher must remain cautious about the errors that can possibly arise in the process of interpreting results. Errors can arise due to false generalization and/or due to wrong interpretation of statistical measures, such as the application of findings beyond the range of observations, identification of correlation with causation and the like. Another major pitfall is the tendency to affirm that definite relationships exist on the basis of confirmation of particular hypotheses. In fact, the positive test results accepting the hypothesis must be interpreted as “being in accord” with the hypothesis, rather than as “confirming the validity of the hypothesis”. The researcher

must remain vigilant about all such things so that false generalization may not take place. He should be well equipped with and must know the correct use of statistical measures for drawing inferences concerning his study.

3. He must always keep in view that the task of interpretation is very much intertwined with analysis and cannot be distinctly separated. As such he must take the task of interpretation as a special aspect of analysis and accordingly must take all those precautions that one usually observes while going through the process of analysis viz., precautions concerning the reliability of data, computational checks, validation and comparison of results.
4. He must never lose sight of the fact that his task is not only to make sensitive observations of relevant occurrences, but also to identify and disengage the factors that are initially hidden to the eye. This will enable him to do his job of interpretation on proper lines. Broad generalisation should be avoided as most research is not amenable to it because the coverage may be restricted to a particular time, a particular area and particular conditions. Such restrictions, if any, must invariably be specified and the results must be framed within their limits.
5. The researcher must remember that "ideally in the course of a research study, there should be constant interaction between initial hypothesis, empirical observation and theoretical conceptions. It is exactly in this area of interaction between theoretical orientation and empirical observation that opportunities for originality and creativity lie." He must pay special attention to this aspect while engaged in the task of interpretation.

REPORT WRITING

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

STEPS IN WRITING REPORTS

Research reports are the product of slow, painstaking (careful / meticulous), accurate inductive (logical / reasonable) work

1) Logical Analysis of the Subject Matter

It is the first step which is primarily concerned with the development of a subject. There are two ways in which to develop a subject – (a) logically and (b) Chronologically (arranging data as per time of occurrence). The logical development is made on the basis of mental connections and associations between the one thing and another by means of analysis. Logical treatment often consists in developing the

material from the simple possible to the most complex structures. Chronological development is based on a connection or sequence in time or occurrence. The directions for doing or making something usually follow the chronological order

2) Preparation of the Final Outline

It is the next step in writing the research report. Outlines (hints) are the framework upon which long written works are constructed. They are in aid (help) to the logical organization of the material and a reminder of the points to be stressed in the report.

3) Preparation of the Rough Draft

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

4) Rewriting and Polishing of the Rough Draft

This step happens to be most difficult part of all formal writing. Usually this step requires more time than the writing of the rough report. The careful revision makes the difference between a mediocre (ordinary) and a good piece of writing. While rewriting and polishing, one should check the report for weaknesses in logical development or presentation. The researcher should also 'see whether or not the material, as it is presented, has unity and cohesion (organization); does the report stand upright and firm and exhibit a definite pattern, like a marble arch? Or

does it resemble an old wall of moldering cement and loose bricks. In addition, the researcher should give due attention to the fact that in his rough draft he has been consistent (reliably / steady) or not. He should check the mechanics of writing – grammar, spelling and usage

5) Preparation of the Final Bibliography

Next in order comes the task of the preparation of the final bibliography. The bibliography, which is generally appended (add on) to the research report, is a list of books in some way pertinent to the research which has been done. It should contain all those works which the researcher has consulted. The bibliography should be arranged alphabetically and may be divided into two parts; the first part may contain the names of books and pamphlets, and the second part may contain the name of magazines and newspaper articles. Generally, this pattern of bibliography is considered convenient and satisfactory from the point of view of reader, thought it is not the only way of presenting bibliography

6) Writing the Final Draft

This consists the last step. The final draft should be written in a concise (brief) and objective style and in simple language, avoiding vague expressions such as "it seems". While writing the final draft, the researcher must avoid abstract (theoretical) terminology and technical jargon. Illustrations and examples based on common experiences must be incorporated in the final draft as they happen to most effective in communicating the research findings to others. A research report should not be dull (boring), but must enthuse (motivate) people and maintain interest and must show originality. It must be remembered that every report should be an attempt to solve some intellectual problem and must

contribute to the solution of a problem and must add to the knowledge of both the researcher and the reader

TYPES OF REPORTS

1. Research reports vary greatly in length and type. In each individual case, both the length and the form are largely dictated (determined / ordered) by the problems at hand
2. For instance, business firms prefer reports in the letter form, just one or two pages in length.
3. Banks, Insurance organizations and financial institutions are generally fond of (having a liking for) the short balance-sheet type of tabulation for their annual reports to their customers and shareholders
4. Mathematicians prefer to write the results of their investigations in the form of algebraic (numerical) notations (Symbols + / -)
5. Chemists (Scientist trained in Chemistry) report their results in symbols and formulae (method)
6. Students of literature usually write long reports presenting the critical analysis of some writer or period or the like with a liberal use of quotations from the works of the author under discussion
7. In the field of education and psychology, the favorite form in the report on the results of experimentation accompanied by the detailed statistical tabulations
8. Clinical psychologists and social pathologists (diagnosis of disease) frequently find it necessary to make use of the case-history (all the relevant information previous gatherer) form
9. News items in the daily papers are also forms of report writing. They represent firsthand on-the-scene accounts of the events described or compilation (collection) of interviews with persons who were on the scene

(area). In such report the first paragraph usually contains the important information in detail and the succeeding paragraphs contain material which is progressively less and less important

10. Book – reviews which analyze the content of the book and report on the author's intentions, his success or failure in achieving his aims, his language, his style, scholarship (learning / research / study), bias or his point of view, such reviews also happen to be a kind of short report
11. The reports prepared by governmental bureaus (agency) special commissions, and similar other organizations are generally very comprehensive (full / complete) reports on the issues involved. Such reports are usually considered as important research products
12. Similarly, Ph.D.theses and dissertation are also a form of report-writing, usually completed by students in academic institutions
13. The above narration throws light on the fact that the results of a research investigation can be presented in a number of ways viz., a technical report, popular report,
14. Which method of presentation to be used in a particular study depends on the circumstances under which the study arose (take place) and the nature of the results
15. A technical report is used whenever a full written report of the study is required whether for record-keeping or for public dissimilation
16. A popular report is used if the research results have policy implications.

A) TECHNICAL REPORT

In technical report the main emphasis is on (i) the methods employed

(ii) assumptions made in the course of the study (iii) the detailed presentation of the findings including their limitations and supporting data

1) Summary of Results

A brief review of the main findings just in two or three pages

2) Nature of the Study

Description of the general objectives of study, formulation of the problem in operational terms, the working hypothesis, the type of analysis and data required

3) Methods Employed

Specific methods used in the study and their limitations. For instance, in sampling studies we should give details of sample design viz. sample size, sample selection etc.,

4) Data

Discussion of data collected, their sources, characteristics and limitations. If secondary data are used, their suitability of the problem at hand be fully assessed.

5) Analysis of Data and Presentation of Findings

The analysis of data and presentation of the findings of the study with supporting data in the form of tables and charts be fully narrated (explained). This, in fact, happens to be the main body of the report usually extending over several chapters

6) Conclusions

A detailed summary of the findings and the policy implications drawn from the results be explained

7) Bibliography

Bibliography of various sources consulted be prepared and attached

8) Technical Appendices

Appendices be given for all technical matters relating to questionnaire, mathematical derivations, elaboration on particular technique of analysis and the like ones

9) Index

Index must be prepared and be given invariably in the report at the end. Even in technical report, simple presentation and ready availability of the findings remain an important consideration and as such the liberal use of charts and diagrams is considered desirable

B) POPULAR REPORT

The popular report is one which emphasis on simplicity and attractiveness. The simplification should be sought (required) through clear writing, minimization of technical, particularly mathematical, details and liberal use of charts and diagrams. Attractive layout along with large print, many subheadings, even an occasional cartoon now and then is another characteristic feature of the popular report

1) Findings and their Implications

Emphasis (importance) in the report is given on the findings of most practical interest and on the implication of these findings

2) Recommendation for Action

Recommendations for action on the basis of the findings of the study is made in this section of the report

3) Objectives of the Study

A general review of how the problem arise is presented along with the specific objectives of the project under study

4) Methods Employed

A brief and non-technical description of the methods and techniques used, including a short review of the data on which the study is based, is given in this part of the report

5) Results

This section constitutes the main body of the report wherein the results of the study are presented in clear and non-technical terms with the liberal use of all sorts of illustrations such as charts, diagrams and the like ones

6) Technical Appendices

More detailed information on methods used, forms etc, is presented in the form of appendices. But the appendices are often not detailed if the report is entirely meant for general public

ORAL PRESENTATION

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

4. In order to overcome this difficulty, a written report may be circulated before the oral presentation and referred to frequently during the discussion.
5. Oral presentation is effective when supplemented by various visual devices. Use of slides, wall charts and blackboards is quite helpful in contributing to clarity and in reducing the boredom, if any. Distributing a board outline, with a few important tables and charts concerning the research results, makes the listeners attentive who have a ready outline on which to focus their thinking. This very often happens in academic institutions where the researcher discusses his research findings and policy implications with others either in a seminar or in a group discussion.
6. Thus, research results can be reported in more than one ways, but the usual practice adopted, in academic institutions particularly, is that of writing the Technical Report and then preparing several research papers to be discussed at various forums in one form or the other. But in practical field and with problems having policy implications, the technique followed is that of writing a popular report. Researches done on governmental account or on behalf of some major public or private organisations are usually presented in the form of technical reports.

PRECAUTIONS FOR WRITING RESEARCH REPORTS

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

1. While determining the length of the report (since research reports vary greatly in length), one should keep in view the fact that it should be

long enough to cover the subject but short enough to maintain interest. In fact, report-writing should not be a means to learning more and more about less and less.

2. A research report should not, if this can be avoided, be dull; it should be such as to sustain reader's interest.
3. Abstract terminology and technical jargon should be avoided in a research report. The report should be able to convey the matter as simply as possible. This, in other words, means that report should be written in an objective style in simple language, avoiding expressions such as "it seems," "there may be" and the like.
4. Readers are often interested in acquiring a quick knowledge of the main findings and as such the report must provide a ready availability of the findings. For this purpose, charts, graphs and the statistical tables may be used for the various results in the main report in addition to the summary of important findings.
5. The layout of the report should be well thought out and must be appropriate and in accordance with the objective of the research problem.
6. The reports should be free from grammatical mistakes and must be prepared strictly in accordance with the techniques of composition of report-writing such as the use of quotations, footnotes, documentation, proper punctuation and use of abbreviations in footnotes and the like.
7. The report must present the logical analysis of the subject matter. It must reflect a structure wherein the different pieces of analysis relating to the research problem fit well.
8. A research report should show originality and should necessarily be an attempt to solve some intellectual problem. It must contribute to the solution of a problem and must add to the store of knowledge.

9. Towards the end, the report must also state the policy implications relating to the problem under consideration. It is usually considered desirable if the report makes a forecast of the probable future of the subject concerned and indicates the kinds of research still needs to be done in that particular field.
10. Appendices should be enlisted in respect of all the technical data in the report.
11. Bibliography of sources consulted is a must for a good report and must necessarily be given.
12. Index is also considered an essential part of a good report and as such must be prepared and appended at the end.
13. Report must be attractive in appearance, neat and clean, whether typed or printed.
14. Calculated confidence limits must be mentioned and the various constraints experienced in conducting the research study may also be stated in the report.
15. Objective of the study, the nature of the problem, the methods employed and the analysis techniques adopted must all be clearly stated in the beginning of the report in the form of introduction.

* * * * *

POSSIBLE QUESTIONS

UNIT-V

PART A

One Mark

Online Examination

PART B

Eight Marks

1. Explain the mechanics of writing research report in detail
2. Elaborate the stages involved in preparing a research report?
3. Explain the techniques of interpretation in detail.
4. Discuss the precautions should be followed while writing a research report.
5. Explain the need and importance of Interpretation in research?
6. Explain the significance of research report in modern times?
7. Explain the techniques of interpretation in detail.
8. Discuss the precautions should be followed while writing a research report.
9. Elucidate the layout of research report covering all relevant points?
10. Explain the mechanics of writing a research report?

KARPAGAM ACADEMY OF HIGHER EDUCATION

(DEEMED TO BE UNIVERSITY)

(ESTABILISHED UNDER SEC 3 OF UGC ACT 1956)

DEPARTMENT OF COMMERCE

III -B.COM RESEARCH METHODOLOGY-15CMU603B

UNIT-V REPORT WRITING

S.NO	QUESTIONS	OPTION I	OPTION II	OPTION III	OPTION IV	ANSWER
1	continuity in research through linking the results of a given study with those	Interpretation	Analysis	Editing	Coding	Interpretation
2	__are used in making comparison between two or more series of data	Symetrical Relationship	Percentage	Reciprocal Relationship	Asymetrical Relationship	Percentage
3	Mean, Median and Mode is an example of ___	Average	Deviation	Dispersion	Relation	Average
4	__refers to the amount or the magnitude of the spread	Dispersion	Central Tendency	Time Series	Non-Parametric Test	Dispersion
5	component of the research study for the research task remains incomplete	Data Collection	Research Report	Hypothesis	Objectives	Research Report
6	Writing of ___is the last step in a research study	Summary	Report	Conclusion	Suggestion	Report
7	the basis of mental connections and associations between the	Illogical	Irrational	Unscientific	Logical	Logical
8	__development is based on a connection or sequence in time or occurrence	Chronological	Logical	Irrational	Unscientific	Chronological

S.NO	QUESTIONS	OPTION I	OPTION II	OPTION III	OPTION IV	ANSWER
9	___are the framework upon which long written works are constructed	Logical Analysis of the Subject Matter	Preparation of the Final Outline	Preparation of the Rough Draft	Rewriting and Polishing of the Rough Draft	Preparation of the Final Outline
10	The books, journals referred by the researcher are mentioned under __	Bibliography	Webliography	Citation	Quotation	Bibliography
11	References collected from website are mentioned under __	Bibliography	Webliography	Citation	Quotation	Webliography
12	___prefer reports in the letter form, just one or two pages in length	Business firms	Banks	Mathematicians	Chemists	Business firms
13	short balance-sheet type of tabulation for their annual reports to their	Business firms	Banks	Mathematicians	Chemists	Banks
14	short balance-sheet type of tabulation for their annual reports to their	Business firms	Insurance	Mathematicians	Chemists	Insurance
15	short balance-sheet type of tabulation for their annual reports to their	Business firms	Financial Institutions	Mathematicians	Chemists	Financial Institutions
16	results of their investigations in the form of algebraic notations	Mathematicians	Chemists	Students of Literature	Education	Mathematicians
17	___ report their results in symbols and formulae	Mathematicians	Chemists	Students of Literature	Education	Chemists
18	reports presenting the critical analysis of some writer or period or the like	Mathematicians	Chemists	Students of Literature	Education	Students of Literature
19	favorite form in the report on the results of experimentation	Mathematicians	Chemists	Students of Literature	Education	Education

S.NO	QUESTIONS	OPTION I	OPTION II	OPTION III	OPTION IV	ANSWER
20	favorite form in the report on the results of experimentation	Mathematicians	Chemists	Students of Literature	Psychology	Psychology
21	__frequently find it necessary to make use of the case-history form	Psychology	Clinical Psychologists	News Items	Book Reviews	Clinical Psychologists
22	__frequently find it necessary to make use of the case-history form	Psychology	Social Pathologists	News Items	Book Reviews	Social Pathologists
23	content of the book and report on the author's intentions	News Items	Book Reviews	Internet	Debates	Book Reviews
24	___reports are generally very comprehensive	Business firms	Banks	Government	Financial Institutions	Government
25	informations are presented in exhaustive form	Popular	Technical	General	Common	Technical
26	are presented by liberal use of graph, charts, diagrams etc.,	Popular	Technical	General	Common	Popular
27	summary of the findings and the policy implications drawn from the results be	Methods Employed	Data	Analysis of Data	Conclusion	Conclusion
28	A brief review of the main findings is presented in __	Nature of Study	Methods Employed	Summary	Analysis	Summary
29	In __ where various sources consulted be prepared and attached	Bibliography	Webliography	Citation	Quotation	Bibliography
30	technical matters relating to questionnaire, mathematical derivations,	Bibliography	Webliography	Citation	Appendices	Appendices

S.NO	QUESTIONS	OPTION I	OPTION II	OPTION III	OPTION IV	ANSWER
31	__must be prepared and be given invariably in the report at the end	Bibliography	Webliography	Index	Appendices	Index
32	Presenting author names in an alphabetical order is known as ___	Bibliography	Webliography	Indexing	Appendices	Indexing
33	for data collection is affixed in the report under the heading___	Appendices	Bibliography	Webliography	Indexing	Appendices
34	the study through oral verbal stimulus is known as __	Oral Presentation	Written Presentation	Printed	Typed	Oral Presentation
35	the identification of materials used in quotations in the report	Footnotes	Quotations	Proverbs	Maxim	Footnotes
36	For business firms report should not exceed __ pages	One	Two	Three	Four	One
37	sheet or financial statements are presented to __	Banks	Hospitals	Government	University	Banks
38	Report prepared with algebraic notations is an example of __ report	Mathematics	Chemistry	Physics	Language	Mathematics
39	making use of symbols and formulae is an example of __ report	Mathematics	Chemistry	Physics	Language	Chemistry
40	a full written report of the study is required whether for record-keeping or for	Popular	Technical	General	Common	Technical
41	__report is used if the research results have policy implications	Popular	Technical	General	Common	Popular

S.NO	QUESTIONS	OPTION I	OPTION II	OPTION III	OPTION IV	ANSWER
42	For shareholders information is presented in ___form	Tabulation	Textual	Manuscript	Transcript	Tabulation
43	Liberal use of quotations is found in ___ report	Language related	Mathematics	Chemistry	Psychology	Language related
44	informations are presented in first page of ___	Books	Magazines	Journals	Newspaper	Newspaper
45	Using technical jargons is an example of __ report	Popular	Technical	General	Common	Technical
46	number of graphs, charts and pictures is an example of _ report	Popular	Technical	General	Common	Popular
47	Website reference may be mentioned as ___	Bibliography	Webliography	Citation	Quotation	Webliography
48	Drawing inferences for the analysis contents is known as ___	Interpretation	Analysis	Editing	Coding	Interpretation
49	quartile range, average deviation and standard deviation	Dispersion	Central Tendency	Time Series	Non-Parametric Test	Dispersion
50	consists in developing the material from the simple possible to the most	Logical	Unsound	Specious	Unreasonable	Logical
51	organization of the material and a reminder of the points to be stressed	Logical Analysis of the Subject Matter	Preparation of the Final Outline	Preparation of the Rough Draft	Rewriting and Polishing of the Rough Draft	Preparation of the Final Outline
52	informations are presented in technical language	Popular	Technical	General	Common	Technical

S.NO	QUESTIONS	OPTION I	OPTION II	OPTION III	OPTION IV	ANSWER
53	In __ report where informations are presented in simple form	Popular	Technical	General	Common	Popular
54	__Report is one which emphasis on simplicity and attractiveness	Popular	Technical	General	Common	Popular
55	data collection is affixed in the report under the heading___	Bibliography	Webliography	Indexing	Appendices	Appendices
56	of the page on which the reference or quotation which they identify or	Footnotes	Quotations	Proverbs	Maxim	Footnotes
57	___is a search for broader meaning of research findings	Analysis	Editing	Coding	Interpretation	Interpretation
58	analysis of the subject and the preparation of the final outline	Logical Analysis of the Subject Matter	Preparation of the Final Outline	Preparation of the Rough Draft	Rewriting and Polishing of the Rough Draft	Preparation of the Rough Draft
59	drawing inferences from the collected facts after an analytical and/or	Analysis	Editing	Coding	Interpretation	Interpretation
60	Drawing inferences for the statistical analysis is known as __	Interpretation	Analysis	Editing	Coding	Interpretation

Reg No …….

[15CMU603B]

KARPAGAM ACADEMY OF HIGHER EDUCATION

(Deemed University Established Under Section 3 of UCG Act 1956)

COIMBATORE- 641 021

(For the candidates admitted from 2015 onwards)

I INTERNAL EXAMINATION, JANUARY -2018

III B.COM –SIXTH SEMESTER

RESEARCH METHODOLOGY

Time: 2 Hours **Maximum: 50 Marks**

Date:

PART - A (20X 1 = 20 Marks)

CHOOSE THE CORRECT ANSWER

1. A study to gain familiarity with a phenomenon to achieve new insights is_____

a) Exploratory research b) Descriptive Research

c) Diagnostic Research d) Qualitative Research

2. Critical evaluation made by the researcher with the facts and information already available is called _____

a) Exploratory research b) Analytical Research

c) Diagnostic Research d) Historical Research

3. To develop new concepts or to reinterpret existing ones, philosophers and thinkers use____

a) Empirical Research b) Conceptual Research

c) Pure Research d) Basic Research

4. The methods or techniques used by researchers in performing research operations is called___

a) Research Design b) Research Methodology

c) Research techniques d) Research process

5. Empirical literature comprises of _____

a) Concepts and Theories b) Earlier studies

c) Scientific enquiry d) Specific Details

6. Deliberate sampling is also known as _____

a) Purposive sampling b) Probability Sampling

c) Random Sampling d) System sampling

7. The population elements are selected for inclusion in the sample based on the ease of access_____

a) Convenience sampling b) Judgment sampling

c) Random Sampling d) Probability Sampling

8. Data Collected by filling up the Schedules by the enumerators on the basis of replies given by Respondents are _____

a) Questionnaire b) Schedule c) Interview d) Observation

9. Method of selecting items to be observed for the given study is called_____

a) Sampling design b) Statistical Design

c) Operational Design d) Observational Design

10. Variables that can be expressed only in integer values are _____

a) Discrete variables b) Extraneous Variables

c) Continuous Variables d) discontinuous variable

11. Primary data is _____ source of collection

a)Second b)Indirect c)No direct d)Direct

12. Good research is _____

a) Progressive b) Systematic c) Informative d) Non informative

13.The main objective of research design is _____

a) Budgeting on time b) Budgeting on money

c) Number of respondents to be selected d)Prepare a structure

14. The formal, systematic and intensive process of carrying on a scientific method of analysis is ___ a) Research Design b) Research c) Interpretation d)Research analysis

15. The device which would retain the actual wording of the respondents is _____

a) Recording b) Writing c) Copying d) Editing

16. Research depends upon _____

a)Random method b)Systematic method c) Horizontal method d)Vertical method

17.---------- constitutes the blue print collection measurement and analysis of data.

a) Applied thinking b)Operational research c) Subjective assessment

d) Research design

18._____ research is based on the measurement of quantity or amount.

a)Quantitiative b) Qualititative c)Experimental d)Non experimental

19. _____ research is concerned with qualitative phenomena

a)Quantitative b)Qualitative c)Non experimental d)Experimental

20. _____ research is related to some abstract idea or theory

a) Conceptual b) Empirical c) Conceptual & Empirical d) Applied

PART – C (3X 10 = 30Marks)

ANSWER ALL THE QUESTIONS

21.(a)Describe the scope and types of research with examples?

(or)

(b)Explain the significance of research in detail.

22.(a)Describe the process of research with diagrams.

(or)

(b)Explain the characteristics and qualities of good research?

23.(a)Describe the steps in sampling design?

(or)

(b)Explain the types of sampling design.