Research Methodology

Methods & Techniques

C. R. KOTHARI
Former Principal, College of Commerce
University of Rajasthan, Jaipur (India)

PUBLISHING FOR ONE WORLD

NEW AGE INTERNATIONAL (P) LIMITED, PUBLISHERS

New Delhi • Bangalore • Chennai • Cochin • Guwahati • Hyderabad
Jalandhar • Kolkata • Lucknow • Mumbai • Ranchi

Visit us at **www.newagepublishers.com**

1

Research Methodology: An Introduction

MEANING OF RESEARCH

Research in common parlance refers to a search for knowledge. One can also define research as a scientific and systematic search for pertinent information on a specific topic. In fact, research is an art of scientific investigation. The Advanced Learner's Dictionary of Current English lays down the meaning of research as "a careful investigation or inquiry specially through search for new facts in any branch of knowledge."[1] Redman and Mory define research as a "systematized effort to gain new knowledge."[2] Some people consider research as a movement, a movement from the known to the unknown. It is actually a voyage of discovery. We all possess the vital instinct of inquisitiveness for, when the unknown confronts us, we wonder and our inquisitiveness makes us probe and attain full and fuller understanding of the unknown. This inquisitiveness is the mother of all knowledge and the method, which man employs for obtaining the knowledge of whatever the unknown, can be termed as research.

Research is an academic activity and as such the term should be used in a technical sense. According to Clifford Woody research comprises defining and redefining problems, formulating hypothesis or suggested solutions; collecting, organising and evaluating data; making deductions and reaching conclusions; and at last carefully testing the conclusions to determine whether they fit the formulating hypothesis. D. Slesinger and M. Stephenson in the Encyclopaedia of Social Sciences define research as "the manipulation of things, concepts or symbols for the purpose of generalising to extend, correct or verify knowledge, whether that knowledge aids in construction of theory or in the practice of an art."[3] Research is, thus, an original contribution to the existing stock of knowledge making for its advancement. It is the persuit of truth with the help of study, observation, comparison and experiment. In short, the search for knowledge through objective and systematic method of finding solution to a problem is research. The systematic approach concerning generalisation and the formulation of a theory is also research. As such the term 'research' refers to the systematic method

[1] *The Advanced Learner's Dictionary of Current English*, Oxford, 1952, p. 1069.

[2] L.V. Redman and A.V.H. Mory, *The Romance of Research*, 1923, p.10.

[3] *The Encyclopaedia of Social Sciences*, Vol. IX, MacMillan, 1930.

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 defence 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 day 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.

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

Research is equally important for social scientists in studying social relationships and in seeking answers to various social problems. It provides the intellectual satisfaction of knowing a few things just for the sake of knowledge and also has practical utility for the social scientist to know for the sake of being able to do something better or in a more efficient manner. Research in social sciences is concerned both with knowledge for its own sake and with knowledge for what it can contribute to practical concerns. "This double emphasis is perhaps especially appropriate in the case of social science. On the one hand, its responsibility as a science is to develop a body of principles that make possible the understanding and prediction of the whole range of human interactions. On the other hand, because of its social orientation, it is increasingly being looked to for practical guidance in solving immediate problems of human relations."[6]

[6] Marie Jahoda, Morton Deutsch and Stuart W. Cook, *Research Methods in Social Relations*, p. 4.

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 a 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 generalisations of new theories.

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

Research Methods versus Methodology

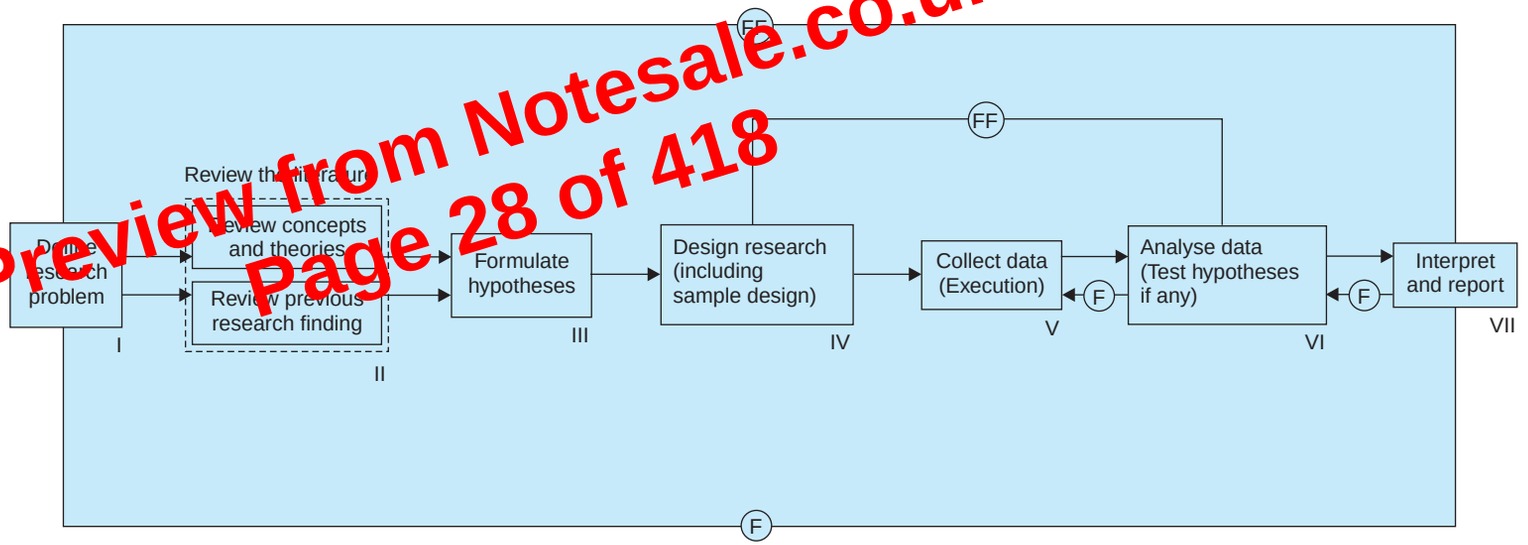
It seems appropriate at this juncture to explain the difference between research methods and research methodology. *Research methods* may be understood as all those methods/techniques that are used for conduction of research. *Research methods or techniques*, thus refer to the methods the researchers*

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

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

From what has been stated above, we can say that methods are more general. It is the methods that generate techniques. However, in practice, the two terms are taken as interchangeable and when we talk of research methods we do, by implication, include research techniques within their compass.

RESEARCH PROCESS IN FLOW CHART



Where (F) = feed back (Helps in controlling the sub-system to which it is transmitted)
(FF) = feed forward (Serves the vital function of providing criteria for evaluation)

Fig. 1.1

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:

(i) *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.

(ii) *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.

(iii) *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 *n*th element is selected until the desired number is secured.

(iv) *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*.

(v) *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.

(vi) *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.

(vii) *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.

(viii) *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

2

Defining the Research Problem

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

WHAT IS A RESEARCH PROBLEM?

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

(i) There must be an individual (or a group or an organisation), let us call it '*I*,' to whom the problem can be attributed. The individual or the organisation, as the case may be, occupies an environment, say '*N*', which is defined by values of the uncontrolled variables, Y_j.

(ii) There must be at least two courses of action, say C_1 and C_2, to be pursued. A course of action is defined by one or more values of the controlled variables. For example, the number of items purchased at a specified time is said to be one course of action.

(iii) There must be at least two possible outcomes, say O_1 and O_2, of the course of action, of which one should be preferable to the other. In other words, this means that there must be at least one outcome that the researcher wants, i.e., an objective.

(iv) The courses of action available must provides some chance of obtaining the objective, but they cannot provide the same chance, otherwise the choice would not matter. Thus, if $P(O_j \mid I, C_j, N)$ represents the probability that an outcome O_j will occur, if I select C_j in N, then $P(O_1 \mid I, C_1, N) \neq P(O_1 \mid I, C_2, N)$. In simple words, we can say that the choices must have unequal efficiencies for the desired outcomes.

* We talk of a research problem or hypothesis in case of descriptive or hypothesis testing research studies. Exploratory or formulative research studies do not start with a problem or hypothesis, their problem is to find a problem or the hypothesis to be tested. One should make a clear statement to this effect. This aspect has been dealt with in chapter entitled "Research Design".

one in terms of the available data and resources and is also analytically meaningful. All this results in a well defined research problem that is not only meaningful from an operational point of view, but is equally capable of paving the way for the development of working hypotheses and for means of solving the problem itself.

Questions

1. Describe fully the techniques of defining a research problem.
2. What is research problem? Define the main issues which should receive the attention of the researcher in formulating the research problem. Give suitable examples to elucidate your points.

   (*Raj. Uni. EAFM, M. Phil. Exam. 1979*)
3. How do you define a research problem? Give three examples to illustrate your answer.

   (*Raj. Uni. EAFM, M. Phil. Exam. 1978*)
4. What is the necessity of defining a research problem? Explain.
5. Write short notes on:
   (a) Experience survey;
   (b) Pilot survey;
   (c) Components of a research problem;
   (d) Rephrasing the research problem.
6. "The task of defining the research problem often follows a sequential pattern". Explain.
7. "Knowing what data are available often serves to narrow down the problem itself as well as the technique that might be used." Explain the underlying idea in this statement in the context of defining a research problem.
8. Write a comprehensive note on the "task of defining a research problem".

(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 analysing data.

NEED FOR RESEARCH DESIGN

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

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

bias and unreliability must be ensured. Whichever method is selected, questions must be well examined and be made unambiguous; interviewers must be instructed not to express their own opinion; observers must be trained so that they uniformly record a given item of behaviour. It is always desirable to pre-test the data collection instruments before they are finally used for the study purposes. In other words, we can say that "*structured instruments*" are used in such studies.

In most of the descriptive/diagnostic studies the researcher takes out sample(s) and then wishes to make statements about the population on the basis of the sample analysis or analyses. More often than not, sample has to be designed. Different sample designs have been discussed in detail in a separate chapter in this book. Here we may only mention that the problem of designing samples should be tackled in such a fashion that the samples may yield accurate information with a minimum amount of research effort. Usually one or more forms of probability sampling, or what is often described as random sampling, are used.

To obtain data free from errors introduced by those responsible for collecting them, it is necessary to supervise closely the staff of field workers as they collect and record information. Checks may be set up to ensure that the data collecting staff perform their duty honestly and without prejudice. "As data are collected, they should be examined for completeness, comprehensibility, consistency and reliability."[2]

The data collected must be processed and analysed. This includes steps like coding the interview replies, observations, etc.; tabulating the data; and performing several statistical computations. To the extent possible, the processing and analysing procedure should be planned in detail before actual work is started. This will prove economical in the sense that the researcher may avoid unnecessary labour such as preparing tables for which he later finds he has no use or on the other hand, re-doing some tables because he failed to include relevant data. Coding should be done carefully to avoid error in coding and for this purpose the reliability of coders needs to be checked. Similarly, the accuracy of tabulation may be checked by having a sample of the tables re-done. In case of mechanical tabulation the material (i.e., the collected data or information) must be entered on appropriate cards which is usually done by punching holes corresponding to a given code. The accuracy of punching is to be checked and ensured. Finally, statistical computations are needed and as such averages, percentages and various coefficients must be worked out. Probability and sampling analysis may as well be used. The appropriate statistical operations, along with the use of appropriate tests of significance should be carried out to safeguard the drawing of conclusions concerning the study.

Last of all comes the question of reporting the findings. This is the task of communicating the findings to others and the researcher must do it in an efficient manner. The layout of the report needs to be well planned so that all things relating to the research study may be well presented in simple and effective style.

Thus, the research design in case of descriptive/diagnostic studies is a comparative design throwing light on all points narrated above and must be prepared keeping in view the objective(s) of the study and the resources available. However, it must ensure the minimisation of bias and maximisation of reliability of the evidence collected. The said design can be appropriately referred to as a *survey design* since it takes into account all the steps involved in a survey concerning a phenomenon to be studied.

[2] Claire Selltiz *et al., op. cit.*, p. 74.

The difference between research designs in respect of the above two types of research studies can be conveniently summarised in tabular form as under:

Table 3.1

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

3. Research design in case of hypothesis-testing research studies: Hypothesis-testing research studies (generally known as experimental studies) are those where the researcher tests the hypotheses of causal relationships between variables. Such studies require procedures that will not only reduce bias and increase reliability, but will also permit drawing inferences about causality. Usually experiments meet this requirement. Hence, when we talk of research design in such studies, we often mean the design of experiments.

Professor R.A. Fisher's name is associated with experimental designs. Beginning of such designs was made by him when he was working at Rothamsted Experimental Station (Centre for Agricultural Research in England). As such the study of experimental designs has its origin in agricultural research. Professor Fisher found that by dividing agricultural fields or plots into different blocks and then by conducting experiments in each of these blocks, whatever information is collected and inferences drawn from them, happens to be more reliable. This fact inspired him to develop certain experimental designs for testing hypotheses concerning scientific investigations. Today, the experimental designs are being used in researches relating to phenomena of several disciplines. Since experimental designs originated in the context of agricultural operations, we still use, though in a technical sense, several terms of agriculture (such as treatment, yield, plot, block etc.) in experimental designs.

BASIC PRINCIPLES OF EXPERIMENTAL DESIGNS

Professor Fisher has enumerated three principles of experimental designs: (1) the Principle of Replication; (2) the Principle of Randomization; and the (3) Principle of Local Control.

6. Latin square design (L.S. design) is an experimental design very frequently used in agricultural research. The conditions under which agricultural investigations are carried out are different from those in other studies for nature plays an important role in agriculture. For instance, an experiment has to be made through which the effects of five different varieties of fertilizers on the yield of a certain crop, say wheat, it to be judged. In such a case the varying fertility of the soil in different blocks in which the experiment has to be performed must be taken into consideration; otherwise the results obtained may not be very dependable because the output happens to be the effect not only of fertilizers, but it may also be the effect of fertility of soil. Similarly, there may be impact of varying seeds on the yield. To overcome such difficulties, the L.S. design is used when there are two major extraneous factors such as the varying soil fertility and varying seeds.

The Latin-square design is one wherein each fertilizer, in our example, appears five times but is used only once in each row and in each column of the design. In other words, the treatments in a L.S. design are so allocated among the plots that no treatment occurs more than once in any one row or any one column. The two blocking factors may be represented through rows and columns (one through rows and the other through columns). The following is a diagrammatic form of such a design in respect of, say, five types of fertilizers, viz., A, B, C, D and E and the two blocking factor viz., the varying soil fertility and the varying seeds:

FERTILITY LEVEL

Seeds difference	I	II	III	IV	V
X_1	A	B	C	D	E
X_2	B	C	D	E	A
X_3	C	D	E	A	B
X_4	D	E	A	B	C
X_5	E	A	B	C	D

Fig. 3.7

The above diagram clearly shows that in a L.S. design the field is divided into as many blocks as there are varieties of fertilizers and then each block is again divided into as many parts as there are varieties of fertilizers in such a way that each of the fertilizer variety is used in each of the block (whether column-wise or row-wise) only once. The analysis of the L.S. design is very similar to the two-way ANOVA technique.

The merit of this experimental design is that it enables differences in fertility gradients in the field to be eliminated in comparison to the effects of different varieties of fertilizers on the yield of the crop. But this design suffers from one limitation, and it is that although each row and each column represents equally all fertilizer varieties, there may be considerable difference in the row and column means both up and across the field. This, in other words, means that in L.S. design we must assume that there is no interaction between treatments and blocking factors. This defect can, however, be removed by taking the means of rows and columns equal to the field mean by adjusting the results. Another limitation of this design is that it requires number of rows, columns and treatments to be

such a design the means for the columns provide the researcher with an estimate of the main effects for treatments and the means for rows provide an estimate of the main effects for the levels. Such a design also enables the researcher to determine the interaction between treatments and levels.

(ii) *Complex factorial designs:* Experiments with more than two factors at a time involve the use of complex factorial designs. A design which considers three or more independent variables simultaneously is called a complex factorial design. In case of three factors with one experimental variable having two treatments and two control variables, each one of which having two levels, the design used will be termed 2 × 2 × 2 complex factorial design which will contain a total of eight cells as shown below in Fig. 3.13.

2 × 2 × 2 COMPLEX FACTORIAL DESIGN

		Experimental Variable			
		Treatment A		Treatment B	
		Control Variable 2 Level I	Control Variable 2 Level II	Control Variable 2 Level I	Control Variable 2 Level II
Control Variable 1	Level I	Cell 1	Cell 3	Cell 5	Cell 7
	Level II	Cell 2	Cell 4	Cell 6	Cell 8

Fig. 3.13

In Fig. 3.14 a pictorial presentation is given of the design shown below.

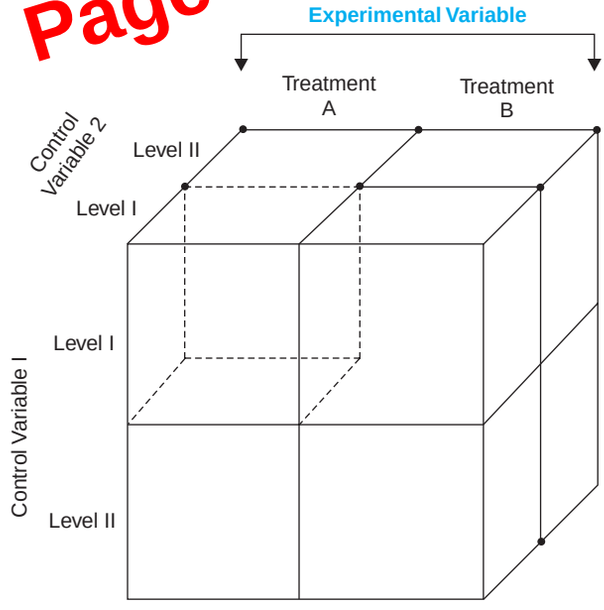


Fig. 3.14

Table 4.1

City number	*No. of departmental stores*	*Cumulative total*	*Sample*	
1	35	35	10	
2	17	52		
3	10	62	60	
4	32	94		
5	70	164	110	160
6	28	192		
7	26	218	210	
8	19	237		
9	26	263	260	
10	66	329	310	
11	37	366	360	
12	44	410	410	
13	33	443		
14	29	472	460	
15	28	500		

(vii) Sequential sampling: This sampling design is somewhat complex sample design. The ultimate size of the sample under this technique is not fixed in advance, but is determined according to mathematical decision rules on the basis of information yielded as survey progresses. This is usually adopted in case of acceptance sampling plan in context of statistical quality control. When a particular lot is to be accepted or rejected on the basis of a single sample, it is known as single sampling; when the decision is to be taken on the basis of two samples, it is known as double sampling and in case the decision rests on the basis of more than two samples but the number of samples is certain and decided in advance, the sampling is known as multiple sampling. But when the number of samples is more than two but it is neither certain nor decided in advance, this type of system is often referred to as sequential sampling. Thus, in brief, we can say that in sequential sampling, one can go on taking samples one after another as long as one desires to do so.

CONCLUSION

From a brief description of the various sample designs presented above, we can say that normally one should resort to simple random sampling because under it bias is generally eliminated and the sampling error can be estimated. But purposive sampling is considered more appropriate when the universe happens to be small and a known characteristic of it is to be studied intensively. There are situations in real life under which sample designs other than simple random samples may be considered better (say easier to obtain, cheaper or more informative) and as such the same may be used. In a situation when random sampling is not possible, then we have to use necessarily a sampling design other than random sampling. At times, several methods of sampling may well be used in the same study.

Questions

1. What do you mean by 'Sample Design'? What points should be taken into consideration by a researcher in developing a sample design for this research project.
2. How would you differentiate between simple random sampling and complex random sampling designs? Explain clearly giving examples.
3. Why probability sampling is generally preferred in comparison to non-probability sampling? Explain the procedure of selecting a simple random sample.
4. Under what circumstances stratified random sampling design is considered appropriate? How would you select such sample? Explain by means of an example.
5. Distinguish between:
   (a) Restricted and unrestricted sampling;
   (b) Convenience and purposive sampling;
   (c) Systematic and stratified sampling;
   (d) Cluster and area sampling.
6. Under what circumstances would you recommend:
   (a) A probability sample?
   (b) A non-probability sample?
   (c) A stratified sample?
   (d) A cluster sample?
7. Explain and illustrate the procedure of selecting a random sample.
8. "A systematic bias results from errors in the sampling procedures." What do you mean by such a systematic bias? Describe the important causes responsible for such a bias.
9. (a) The following are the number of departmental stores in 10 cities: 35, 27, 24, 32, 42, 30, 34, 40, 29 and 38. If we want to select a sample of 15 stores using cities as clusters and selecting within clusters proportional to size, how many stores from each city should be chosen? (Use a starting point of 4).
   (b) What sampling design might be used to estimate the weight of a group of men and women?
10. A certain population is divided into five strata so that $N_1 = 2000$, $N_2 = 2000$, $N_3 = 1800$, $N_4 = 1700$, and $N_5 = 2500$. Respective standard deviations are: $\sigma_1 = 1.6$, $\sigma_2 = 2.0$, $\sigma_3 = 4.4$, $\sigma_4 = 4.8$, $\sigma_5 = 6.0$ and further the expected sampling cost in the first two strata is Rs 4 per interview and in the remaining three strata the sampling cost is Rs 6 per interview. How should a sample of size $n = 226$ be allocated to five strata if we adopt proportionate sampling design; if we adopt disproportionate sampling design considering (i) only the differences in stratum variability (ii) differences in stratum variability as well as the differences in stratum sampling costs.

research problem and the judgement of the researcher. But one can certainly consider three types of validity in this connection: (i) Content validity; (ii) Criterion-related validity and (iii) Construct validity.

(i) *Content validity* is the extent to which a measuring instrument provides adequate coverage of the topic under study. If the instrument contains a representative sample of the universe, the content validity is good. Its determination is primarily judgemental and intuitive. It can also be determined by using a panel of persons who shall judge how well the measuring instrument meets the standards, but there is no numerical way to express it.

(ii) *Criterion-related validity* relates to our ability to predict some outcome or estimate the existence of some current condition. This form of validity reflects the success of measures used for some empirical estimating purpose. The concerned criterion must possess the following qualities:

Relevance: (A criterion is relevant if it is defined in terms we judge to be the proper measure.)

Freedom from bias: (Freedom from bias is attained when the criterion gives each subject an equal opportunity to score well.)

Reliability: (A reliable criterion is stable or reproducible.)

Availability: (The information specified by the criterion must be available.)

In fact, a Criterion-related validity is a broad term that actually refers to (i) *Predictive validity* and *(ii) Concurrent validity.* The former refers to the usefulness of a test in predicting some future performance whereas the latter refers to the usefulness of a test in closely relating to other measures of known validity. Criterion-related validity is expressed as the coefficient of correlation between test scores and some measure of future performance or between test scores and scores on another measure of known validity.

(iii) *Construct validity* is the most complex and abstract. A measure is said to possess construct validity to the degree that it confirms to predicted correlations with other theoretical propositions. Construct validity is the degree to which scores on a test can be accounted for by the explanatory constructs of a sound theory. For determining construct validity, we associate a set of other propositions with the results received from using our measurement instrument. If measurements on our devised scale correlate in a predicted way with these other propositions, we can conclude that there is some construct validity.

If the above stated criteria and tests are met with, we may state that our measuring instrument is valid and will result in correct measurement; otherwise we shall have to look for more information and/or resort to exercise of judgement.

2. Test of Reliability

The test of reliability is another important test of sound measurement. A measuring instrument is reliable if it provides consistent results. Reliable measuring instrument does contribute to validity, but a reliable instrument need not be a valid instrument. For instance, a scale that consistently overweighs objects by five kgs., is a reliable scale, but it does not give a valid measure of weight. But the other way is not true i.e., a valid instrument is always reliable. Accordingly reliability is not as valuable as validity, but it is easier to assess reliability in comparison to validity. If the quality of reliability is satisfied by an instrument, then while using it we can be confident that the transient and situational factors are not interfering.

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."[3]

Scale Classification Bases

The number assigning procedures or the scaling procedures may be broadly classified on one or more of the following bases: (a) subject orientation; (b) response form; (c) degree of subjectivity; (d) scale properties; (e) number of dimensions and (f) scale construction techniques. We take up each of these separately.

(a) Subject orientation: Under it a scale may be designed to measure characteristics of the respondent who completes it or to judge the stimulus object which is presented to the respondent. In respect of the former, we presume that the stimuli presented are sufficiently homogeneous so that the between-stimuli variation is small as compared to the variation among respondents. In the latter approach, we ask the respondent to judge some specific object in terms of one or more dimensions and we presume that the between-respondent variation will be small as compared to the variation among the different stimuli presented to respondents for judging.

(b) Response form: Under this we may classify the scales as categorical and comparative. Categorical scales are also known as rating scales. These scales are used when a respondent scores some object without direct reference to other objects. Under comparative scales, which are also known as ranking scales, the respondent is asked to compare two or more objects. In this sense the respondent may state that one object is superior to the other or that three models of pen rank in order 1, 2 and 3. The essence of ranking is, in fact, a relative comparison of a certain property of two or more objects.

(c) Degree of subjectivity: With this basis the scale data may be based on whether we measure subjective personal preferences or simply make non-preference judgements. In the former case, the respondent is asked to choose which person he favours or which solution he would like to see employed, whereas in the latter case he is simply asked to judge which person is more effective in some aspect or which solution will take fewer resources without reflecting any personal preference.

(d) Scale properties: Considering scale properties, one may classify the scales as nominal, ordinal, interval and ratio scales. Nominal scales merely classify without indicating order, distance or unique origin. Ordinal scales indicate magnitude relationships of 'more than' or 'less than', but indicate no distance or unique origin. Interval scales have both order and distance values, but no unique origin. Ratio scales possess all these features.

(e) Number of dimensions: In respect of this basis, scales can be classified as 'unidimensional' and 'multidimensional' scales. Under the former we measure only one attribute of the respondent or object, whereas multidimensional scaling recognizes that an object might be described better by using the concept of an attribute space of 'n' dimensions, rather than a single-dimension continuum.

[3] Bernard S. Phillips, *Social Research Strategy and Tactics*, 2nd ed., p. 205.

minimises the dimensionality of the solution space. This approach utilises all the information in the data in obtaining a solution. The data (i.e., the metric similarities of the objects) are often obtained on a bipolar similarity scale on which pairs of objects are rated one at a time. If the data reflect exact distances between real objects in an *r*-dimensional space, their solution will reproduce the set of interpoint distances. But as the true and real data are rarely available, we require random and systematic procedures for obtaining a solution. Generally, the judged similarities among a set of objects are statistically transformed into distances by placing those objects in a multidimensional space of some dimensionality.

The *non-metric approach* first gathers the non-metric similarities by asking respondents to rank order all possible pairs that can be obtained from a set of objects. Such non-metric data is then transformed into some arbitrary metric space and then the solution is obtained by reducing the dimensionality. In other words, this non-metric approach seeks "a representation of points in a space of minimum dimensionality such that the rank order of the interpoint distances in the solution space maximally corresponds to that of the data. This is achieved by requiring only that the distances in the solution be monotone with the input data."[9] The non-metric approach has come into prominence during the sixties with the coming into existence of high speed computers to generate metric solutions for ordinal input data.

The significance of MDS lies in the fact that it enables the researcher to study "the perceptual structure of a set of stimuli and the cognitive processes underlying the development of this structure. Psychologists, for example, employ multidimensional scaling techniques in an effort to scale psychophysical stimuli and to determine appropriate labels for the dimensions along which these stimuli vary."[10] The MDS techniques, infact, do away with the need in the data collection process to specify the attribute(s) along which the several brands, say of a particular product, may be compared as ultimately the MDS analysis itself reveals such attribute(s) that presumably underlie the expressed relative similarities among objects. Thus, MDS is an important tool in attitude measurement and the techniques falling under MDS promise "a great advance from a series of unidimensional measurements (e.g., a distribution of intensities of feeling towards single attribute such as colour, taste or a preference ranking with indeterminate intervals), to a perceptual mapping in multidimensional space of objects ... company images, advertisement brands, etc."[11]

In spite of all the merits stated above, the MDS is not widely used because of the computation complications involved under it. Many of its methods are quite laborious in terms of both the collection of data and the subsequent analyses. However, some progress has been achieved (due to the pioneering efforts of Paul Green and his associates) during the last few years in the use of non-metric MDS in the context of market research problems. The techniques have been specifically applied in "finding out the perceptual dimensions, and the spacing of stimuli along these dimensions, that people, use in making judgements about the relative similarity of pairs of Stimuli."[12] But, "in the long run, the worth of MDS will be determined by the extent to which it advances the behavioral sciences."[13]

[9] Robert Ferber (ed.), *Handbook of Marketing Research,* p. 3–51.

[10] *Ibid.*, p. 3–52.

[11] G.B. Giles, *Marketing,* p. 43.

[12] Paul E. Green, *Analyzing Multivariate Data,* p. 421.

[13] Jum C. Nunnally, *Psychometric Theory*, p. 496.

(c) Likert-type scale;
(d) Arbitrary scales;
(e) Multidimensional scaling (MDS).

9. Describe the different methods of scale construction, pointing out the merits and demerits of each.
10. "Scaling describes the procedures by which numbers are assigned to various degrees of opinion, attitude and other concepts." Discuss. Also point out the bases for scale classification.

6

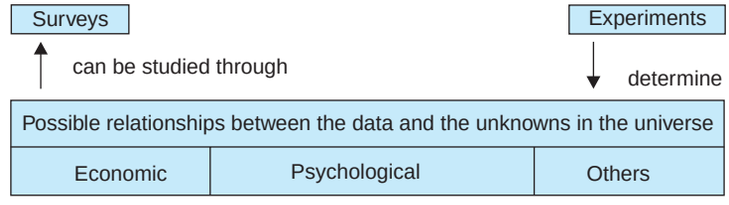
Methods of Data Collection

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. We describe the different methods of data collection, with the pros and cons of each method.

COLLECTION OF PRIMARY DATA

We collect primary data during the course of doing experiments in an experimental research but in case we do research of the descriptive type and perform surveys, whether sample surveys or census surveys, then we can obtain primary data either through observation or through direct communication with respondents in one form or another or through personal interviews.* This, in other words, means

* An experiment refers to an investigation in which a factor or variable under test is isolated and its effect(s) measured. In an experiment the investigator measures the effects of an experiment which he conducts intentionally. Survey refers to the method of securing information concerning a phenomena under study from all or a selected number of respondents of the concerned universe. In a survey, the investigator examines those phenomena which exist in the universe independent of his action. The difference between an experiment and a survey can be depicted as under:



that there are several methods of collecting primary data, particularly in surveys and descriptive researches. Important ones are: (i) observation method, (ii) interview method, (iii) through questionnaires, (iv) through schedules, and (v) other methods which include (a) warranty cards; (b) distributor audits; (c) pantry audits; (d) consumer panels; (e) using mechanical devices; (f) through projective techniques; (g) depth interviews, and (h) content analysis. We briefly take up each method separately.

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

However, observation method has various limitations. Firstly, it is an expensive method. Secondly, the information provided by this method is very limited. Thirdly, sometimes unforeseen factors may interfere with the observational task. At times, the fact that some people are rarely accessible to direct observation creates obstacle for this method to collect data effectively.

While using this method, the researcher should keep in mind things like: What should be observed? How the observations should be recorded? Or how the accuracy of observation can be ensured? In case the observation is characterised by a careful definition of the units to be observed, the style of recording the observed information, standardised conditions of observation and the selection of pertinent data of observation, then the observation is called as *structured observation.* But when observation is to take place without these characteristics to be thought of in advance, the same is termed *as unstructured observation.* Structured observation is considered appropriate in descriptive studies, whereas in an exploratory study the observational procedure is most likely to be relatively unstructured.

We often talk about participant and non-participant types of observation in the context of studies, particularly of social sciences. This distinction depends upon the observer's sharing or not sharing the life of the group he is observing. If the observer observes by making himself, more or less, a member of the group he is observing so that he can experience what the members of the group experience, the observation is called as the *participant observation.* But when the observer observes as a detached emissary without any attempt on his part to experience through participation what others feel, the observation of this type is often termed as *non-participant observation.* (When the observer is observing in such a manner that his presence may be unknown to the people he is observing, such an observation is described as *disguised observation.*)

the interviewer in a structured interview follows a rigid procedure laid down, asking questions in a form and order prescribed. As against it, the *unstructured interviews* are characterised by a flexibility of approach to questioning. Unstructured interviews do not follow a system of pre-determined questions and standardised techniques of recording information. In a non-structured interview, the interviewer is allowed much greater freedom to ask, in case of need, supplementary questions or at times he may omit certain questions if the situation so requires. He may even change the sequence of questions. He has relatively greater freedom while recording the responses to include some aspects and exclude others. But this sort of flexibility results in lack of comparability of one interview with another and the analysis of unstructured responses becomes much more difficult and time-consuming than that of the structured responses obtained in case of structured interviews. Unstructured interviews also demand deep knowledge and greater skill on the part of the interviewer. Unstructured interview, however, happens to be the central technique of collecting information in case of exploratory or formulative research studies. But in case of descriptive studies, we quite often use the technique of structured interview because of its being more economical, providing a safe basis for generalisation and requiring relatively lesser skill on the part of the interviewer.

We may as well talk about focussed interview, clinical interview and the non-directive interview. *Focussed interview* is meant to focus attention on the given experience of the respondent and its effects. Under it the interviewer has the freedom to decide the manner and sequence in which the questions would be asked and has also the freedom to explore reasons and motives. The main task of the interviewer in case of a focussed interview is to confine the respondent to a discussion of issues with which he seeks conversance. Such interviews are used generally in the development of hypotheses and constitute a major type of unstructured interviews. The *clinical interview* is concerned with broad underlying feelings or motivations or with the course of individual's life experience. The method of eliciting information under it is generally left to the interviewer's discretion. In case of *non-directive interview*, the interviewer's function is simply to encourage the respondent to talk about the given topic with a bare minimum of direct questioning. The interviewer often acts as a catalyst to a comprehensive expression of the respondents' feelings and beliefs and of the frame of reference within which such feelings and beliefs take on personal significance.

Despite the variations in interview-techniques, the major advantages and weaknesses of personal interviews can be enumerated in a general way. The chief merits of the interview method are as follows:

(i) More information and that too in greater depth can be obtained.

(ii) 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.

(iii) There is greater flexibility under this method as the opportunity to restructure questions is always there, specially in case of unstructured interviews.

(iv) Observation method can as well be applied to recording verbal answers to various questions.

(v) Personal information can as well be obtained easily under this method.

(vi) Samples can be controlled more effectively as there arises no difficulty of the missing returns; non-response generally remains very low.

(vii) 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.

(xiii) Case study techniques are indispensable for therapeutic and administrative purposes. They are also of immense value in taking decisions regarding several management problems. Case data are quite useful for diagnosis, therapy and other practical case problems.

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

(i) Case situations are seldom comparable and as such the information gathered in case studies is often not comparable. Since the subject under case study tells history in his own words, logical concepts and units of scientific classification have to be read into it or out of it by the investigator.

(ii) Read Bain does not consider the case data as significant scientific data since they do not provide knowledge of the "impersonal, universal, non-ethical, non-practical, repetitive aspects of phenomena."[8] Real information is often not collected because the subjectivity of the researcher does enter in the collection of information in a case study.

(iii) The danger of false generalisation is always there in view of the fact that no set rules are followed in collection of the information and only few units are studied.

(iv) It consumes more time and requires lot of expenditure. More time is needed under case study method since one studies the natural history cycles of social units and that too minutely.

(v) The case data are often vitiated because the subject, according to Read Bain, may write what he thinks the investigator wants; and the greater the rapport, the more subjective the whole process is.

(vi) Case study method is based on several assumptions which may not be very realistic at times, and as such the usefulness of case data is always subject to doubt.

(vii) Case study method can be used only in a limited sphere., it is not possible to use it in case of a big society. Sampling is also not possible under a case study method.

(viii) Response of the investigator is an important limitation of the case study method. He often thinks that he has full knowledge of the unit and can himself answer about it. In case the same is not true, then consequences follow. In fact, this is more the fault of the researcher rather than that of the case method.

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

Question

1. Enumerate the different methods of collecting data. Which one is the most suitable for conducting enquiry regarding family welfare programme in India? Explain its merits and demerits.

[8] Pauline V. Young, *Scientific social surveys and research*, p. 262.

(vi) Surveys are concerned with hypothesis formulation and testing the analysis of the relationship between non-manipulated variables. Experimentation provides a method of hypothesis testing. After experimenters define a problem, they propose a hypothesis. They then test the hypothesis and confirm or disconfirm it in the light of the controlled variable relationship that they have observed. The confirmation or rejection is always stated in terms of probability rather than certainty. Experimentation, thus, is the most sophisticated, exacting and powerful method for discovering and developing an organised body of knowledge. The ultimate purpose of experimentation is to generalise the variable relationships so that they may be applied outside the laboratory to a wider population of interest.*

(vii) Surveys may either be census or sample surveys. They may also be classified as social surveys, economic surveys or public opinion surveys. Whatever be their type, the method of data collection happens to be either observation, or interview or questionnaire/opinionnaire or some projective technique(s). Case study method can as well be used. But in case of experiments, data are collected from several readings of experiments.

(viii) In case of surveys, research design must be rigid, must make enough provision for protection against bias and must maximise reliability as the aim happens to be to obtain complete and accurate information. Research design in case of experimental studies, apart reducing bias and ensuring reliability, must permit drawing inferences about causality.

(ix) Possible relationships between the data and the unknowns in the universe can be studied through surveys where as experiments are meant to determine such relationships.

(x) Causal analysis is considered relatively more important in experiments where as in most social and business surveys our interest lies in understanding and controlling relationships between variables and as such correlation analysis is relatively more important in surveys.

* John W. Best and James V. Kahn, "*Research in Education*", 5th ed., Prentice-Hall of India Pvt. Ltd., New Delhi, 1986, p.111.

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

Median is a positional average and is used only in the context of qualitative phenomena, for example, in estimating intelligence, etc., which are often encountered in sociological fields. Median is not useful where items need to be assigned relative importance and weights. It is not frequently used in sampling statistics.

Mode is the most commonly or frequently occurring value in a series. The mode in a distribution is that item around which there is maximum concentration. In general, mode is the size of the item which has the maximum frequency, but at items such an item may not be mode on account of the effect of the frequencies of the neighbouring items. Like median, mode is a positional average and is not affected by the values of extreme items. it is, therefore, useful in all situations where we want to eliminate the effect of extreme variations. Mode is particularly useful in the study of popular sizes. For example, a manufacturer of shoes is usually interested in finding out the size most in demand so that he may manufacture a larger quantity of that size. In other words, he wants a modal size to be determined for median or mean size would not serve his purpose. but there are certain limitations of mode as well. For example, it is not amenable to algebraic treatment and sometimes remains indeterminate when we have two or more model values in a series. It is considered unsuitable in cases where we want to give relative importance to items under consideration.

Geometric mean is also useful under certain conditions. It is defined as the *n*th root of the product of the values of *n* times in a given series. Symbolically, we can put it thus:

$$\text{Geometric mean (or G.M.)} = \sqrt[n]{X_1 X_2 X_3 \ldots X_n}$$

$$= \sqrt[n]{X_1 \cdot X_2 \cdot X_3 \ldots X_n}$$

where

G.M. = geometric mean,

n = number of items.

X_i = ith value of the variable X

π = conventional product notation

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

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

$$= 6$$

The most frequently used application of this average is in the determination of average per cent of change i.e., it is often used in the preparation of index numbers or when we deal in ratios.

Harmonic mean is defined as the reciprocal of the average of reciprocals of the values of items of a series. Symbolically, we can express it as under:

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

$$= \text{Rec.}\frac{\text{Rec.}\, X_1 + \text{Rec.}\, X_2 + \ldots + \text{Rec.}\, X_n}{n}$$

Alternatively, we can work out the partial correlation coefficients thus:

$$r_{yx_1 \cdot x_2} = \frac{r_{yx_1} - r_{yx_2} \cdot r_{x_1x_2}}{\sqrt{1 - r^2_{yx_2}}\ \sqrt{1 - r^2_{x_1x_2}}}$$

and

$$r_{yx_2 \cdot x_1} = \frac{r_{yx_2} - r_{yx_1} \cdot r_{x_1x_2}}{\sqrt{1 - r^2_{yx_1}}\ \sqrt{1 - r^2_{x_1x_2}}}$$

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

ASSOCIATION IN CASE OF ATTRIBUTES

When data is collected on the basis of some attribute or attributes, we have statistics commonly termed as statistics of attributes. It is not necessary that the objects may process only one attribute; rather it would be found that the objects possess more than one attribute. In such a situation our interest may remain in knowing whether the attributes are associated with each other or not. For example, among a group of people we may find that some of them are inoculated against small-pox and among the inoculated we may observe that some of them suffered from small-pox after inoculation. The important question which may arise for the observation is regarding the efficiency of inoculation for its popularity will depend upon the immunity which it provides against small-pox. In other words, we may be interested in knowing whether inoculation and immunity from small-pox are associated. Technically, we say that the two attributes are associated if they appear together in a greater number of cases than is to be expected if they are independent and not simply on the basis that they are appearing together in a number of cases as is done in ordinary life.

The association may be positive or negative (negative association is also known as disassociation). If class frequency of *AB*, symbolically written as (*AB*), is greater than the expectation of *AB* being together if they are independent, then we say the two attributes are positively associated; but if the class frequency of *AB* is less than this expectation, the two attributes are said to be negatively associated. In case the class frequency of *AB* is equal to expectation, the two attributes are considered as independent i.e., are said to have no association. It can be put symbolically as shown hereunder:

If $(AB) > \frac{(A)}{N} \times \frac{(B)}{N} \times N$, then *AB* are positively related/associated.

If $(AB) < \frac{(A)}{N} \times \frac{(B)}{N} \times N$, then *AB* are negatively related/associated.

If $(AB) = \frac{(A)}{N} \times \frac{(B)}{N} \times N$, then *AB* are independent i.e., have no association.

'economic barometers measuring the economic phenomenon in all its aspects either directly by measuring the same phenomenon or indirectly by measuring something else which reflects upon the main phenomenon.

But index numbers have their own limitations with which researcher must always keep himself aware. For instance, index numbers are only approximate indicators and as such give only a fair idea of changes but cannot give an accurate idea. Chances of error also remain at one point or the other while constructing an index number but this does not diminish the utility of index numbers for they still can indicate the trend of the phenomenon being measured. However, to avoid fallacious conclusions, index numbers prepared for one purpose should not be used for other purposes or for the same purpose at other places.

2. Time series analysis: In the context of economic and business researches, we may obtain quite often data relating to some time period concerning a given phenomenon. Such data is labelled as 'Time Series'. More clearly it can be stated that series of successive observations of the given phenomenon over a period of time are referred to as time series. Such series are usually the result of the effects of one or more of the following factors:

(i) *Secular trend* or long term trend that shows the direction of the series in a long period of time. The effect of trend (whether it happens to be a growth factor or a declining factor) is gradual, but extends more or less consistently throughout the entire period of time under consideration. Sometimes, secular trend is simply stated as trend (or *T*).

(ii) *Short time oscillations* i.e., changes taking place in the short period of time only and such changes can be the effect of the following factors:

(a) *Cyclical fluctuations* (or *C*) are the fluctuations as a result of business cycles and are generally referred to as long term movements that represent consistently recurring rises and declines in an activity.

(b) *Seasonal fluctuations* (*or S*) are of short duration occurring in a regular sequence at specific intervals of time. Such fluctuations are the result of changing seasons. Usually these fluctuations involve patterns of change within a year that tend to be repeated from year to year. Cyclical fluctuations and seasonal fluctuations taken together constitute short-period regular fluctuations.

(c) *Irregular fluctuations* (*or I*), also known as Random fluctuations, are variations which take place in a completely unpredictable fashion.

All these factors stated above are termed as components of time series and when we try to analyse time series, we try to isolate and measure the effects of various types of these factors on a series. To study the effect of one type of factor, the other type of factor is eliminated from the series. The given series is, thus, left with the effects of one type of factor only.

For analysing time series, we usually have two models; (1) multiplicative model; and (2) additive model. Multiplicative model assumes that the various components interact in a multiplicative manner to produce the given values of the overall time series and can be stated as under:

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

where

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

certain level of significance is compared with the calculated value of *t* from the sample data, and if the latter is either equal to or exceeds, we infer that the null hypothesis cannot be accepted.*

4. *F distribution:* If $(\sigma_{s1})^2$ and $(\sigma_{s2})^2$ are the variances of two independent samples of size n_1 and n_2 respectively taken from two independent normal populations, having the same variance, $(\sigma_{p1})^2 = (\sigma_{p2})^2$, the ratio $F = (\sigma_{s1})^2 / (\sigma_{s2})^2$, where $(\sigma_{s1})^2 = \Sigma \left(\overline{X}_{1i} - \overline{X}_1\right)^2 / n_1 - 1$ and $(\sigma_{s2})^2 = \Sigma \left(\overline{X}_{2i} - \overline{X}_2\right)^2 / n_2 - 1$ has an *F* distribution with $n_1 - 1$ and $n_2 - 1$ degrees of freedom.

F ratio is computed in a way that the larger variance is always in the numerator. Tables have been prepared for *F* distribution that give critical values of *F* for various values of degrees of freedom for larger as well as smaller variances. The calculated value of *F* from the sample data is compared with the corresponding table value of *F* and if the former is equal to or exceeds the latter, then we infer that the null hypothesis of the variances being equal cannot be accepted. We shall make use of the *F* ratio in the context of hypothesis testing and also in the context of ANOVA technique.

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

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

Tables are there that give the value of χ^2 for given d.f. which may be used with calculated value of χ^2 for relevant d.f. at a desired level of significance for testing hypotheses. We will take it up in detail in the chapter 'Chi-square Test'.

CENTRAL LIMIT THEOREM

When sampling is from a normal population, the means of samples drawn from such a population are themselves normally distributed. But when sampling is not from a normal population, the size of the

* This aspect has been dealt with in details in the context of testing of hypotheses later in this book.

researcher usually makes these two types of estimates through sampling analysis. While making estimates of population parameters, the researcher can give only the best point estimate or else he shall have to speak in terms of intervals and probabilities for he can never estimate with certainty the exact values of population parameters. Accordingly he must know the various properties of a good estimator so that he can select appropriate estimators for his study. He must know that a good estimator possesses the following properties:

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

(ii) An estimator should have a relatively small variance. This means that the most efficient estimator, among a group of unbiased estimators, is one which has the smallest variance. This property is technically described as the *property of efficiency*.

(iii) An estimator should use as much as possible the information available from the sample. This property is known as the *property of sufficiency*.

(iv) An estimator should approach the value of population parameter as the sample size becomes larger and larger. This property is referred to as the *property of consistency*.

Keeping in view the above stated properties, the researcher must select appropriate estimator(s) for his study. We may now explain the methods which will enable us to estimate with reasonable accuracy the population mean and the population proportion, the two widely used concepts.

ESTIMATING THE POPULATION MEAN (μ)

So far as the point estimate is concerned, the sample mean $\overline{X}$ is the best estimator of the population mean, μ, and its sampling distribution, so long as the sample is sufficiently large, approximates the normal distribution. If we know the sampling distribution of $\overline{X}$, we can make statements about any estimate that we may make from the sampling information. Assume that we take a sample of 36 students and find that the sample yields an arithmetic mean of 6.2 i.e., $\overline{X} = 6.2$. Replace these student names on the population list and draw another sample of 36 randomly and let us assume that we get a mean of 7.5 this time. Similarly a third sample may yield a mean of 6.9; fourth a mean of 6.7, and so on. We go on drawing such samples till we accumulate a large number of means of samples of 36. Each such sample mean is a separate point estimate of the population mean. When such means are presented in the form of a distribution, the distribution happens to be quite close to normal. This is a characteristic of a distribution of sample means (and also of other sample statistics). Even if the population is not normal, the sample means drawn from that population are dispersed around the parameter in a distribution that is generally close to normal; the mean of the distribution of sample means is equal to the population mean.[5] This is true in case of large samples as per the dictates of the central limit theorem. This relationship between a population distribution and a distribution of sample

[5] C. William Emory, *Business Research Methods*, p.145

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

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

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

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

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

or
$$40 \pm (1.96)(0.75)$$

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

Illustration 2

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

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

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

(3) If the desired confidence level is .90, what will be the upper and lower limits of the confidence interval for the mean number of accidents per intersection per year?

Solution: The given information can be written as under:

$N = 2400$ (This means that population is finite)

$n = 64$

$\bar{X} = 3.2$

$\sigma_s = 0.8$

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

Now we can answer the given questions thus:

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

Hence,

$$\hat{\sigma}_p = \sigma_s = 0.8$$

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

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

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

$$= \frac{0.8}{\sqrt{64}} \times \sqrt{\frac{2336}{2399}}$$

$$= (0.1)\ (.97)$$

$$= .097$$

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

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

$$= 3.2 \pm (1.645)\ (.097)$$

$$= 3.2 \pm .16 \text{ accidents per intersection.}$$

When the sample size happens to be a large one or when the population standard deviation is known, we use normal distribution for detemining confidence intervals for population mean as stated above. But how to handle estimation problem when population standard deviation is not known and the sample size is small (i.e., when $n \leq 30$)? In such a situation, normal distribution is not appropriate, but we can use *t*-distribution for our purpose. While using *t*-distribution, we assume that population is normal or approximately normal. There is a different *t*-distribution for each of the possible degrees of freedom. When we use *t*-distribution for estimating a population mean, we work out the degrees of freedom as equal to $n - 1$, where n means the size of the sample and then can look for cirtical value of '*t*' in the *t*-distribution table for appropriate degrees of freedom at a given level of significance. Let us illustrate this by taking an example.

Illustration 3

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

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

$$\bar{X} = 36.8 \text{ tons per shift}$$

$$\sigma_s = 2.8 \text{ tons per shift}$$

$$n = 4$$

degrees of freedom = $n - 1 = 4 - 1 = 3$ and the critical value of '*t*' for 90 per cent confidence interval or at 10 per cent level of significance is 2.353 for 3 *d.f.* (as per the table of *t*-distribution).

9

Testing of Hypotheses I

(Parametric or Standard Tests of Hypotheses)

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 generalization 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 meaningfully, 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 A 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.

when the sampling result (i.e., observed evidence) has a less than 0.05 probability of occurring if H_0 is true. In other words, the 5 per cent level of significance means that researcher is willing to take as much as a 5 per cent risk of rejecting the null hypothesis when it (H_0) happens to be true. Thus the significance level is the maximum value of the probability of rejecting H_0 when it is true and is usually determined in advance before testing the hypothesis.

(c) *Decision rule or test of hypothesis:* Given a hypothesis H_0 and an alternative hypothesis H_a, we make a rule which is known as decision rule according to which we accept H_0 (i.e., reject H_a) or reject H_0 (i.e., accept H_a). For instance, if (H_0 is that a certain lot is good (there are very few defective items in it) against H_a) that the lot is not good (there are too many defective items in it), then we must decide the number of items to be tested and the criterion for accepting or rejecting the hypothesis. We might test 10 items in the lot and plan our decision saying that if there are none or only 1 defective item among the 10, we will accept H_0 otherwise we will reject H_0 (or accept H_a). This sort of basis is known as decision rule.

(d) *Type I and Type II errors:* In the context of testing of hypotheses, there are basically two types of errors we can make. We may reject H_0 when H_0 is true and we may accept H_0 when in fact H_0 is not true. The former is known as Type I error and the latter as Type II error. In other words, Type I error means rejection of hypothesis which should have been accepted and Type II error means accepting the hypothesis which should have been rejected. Type I error is denoted by α (alpha) known as α error, also called the level of significance of test; and Type II error is denoted by β (beta) known as β error. In a tabular form the said two errors can be presented as follows:

Table 9.2

	Decision	
	Accept H_0	Reject H_0
H_0 (true)	Correct decision	Type I error (α error)
H_0 (false)	Type II error (β error)	Correct decision

The probability of Type I error is usually determined in advance and is understood as the level of significance of testing the hypothesis. If type I error is fixed at 5 per cent, it means that there are about 5 chances in 100 that we will reject H_0 when H_0 is true. We can control Type I error just by fixing it at a lower level. For instance, if we fix it at 1 per cent, we will say that the maximum probability of committing Type I error would only be 0.01.

But with a fixed sample size, *n*, when we try to reduce Type I error, the probability of committing Type II error increases. Both types of errors cannot be reduced simultaneously. There is a trade-off between two types of errors which means that the probability of making one type of error can only be reduced if we are willing to increase the probability of making the other type of error. To deal with this trade-off in business situations, decision-makers decide the appropriate level of Type I error by examining the costs or penalties attached to both types of errors. If Type I error involves the time and trouble of reworking a batch of chemicals that should have been accepted, whereas Type II error means taking a chance that an entire group of users of this chemical compound will be poisoned, then

HYPOTHESIS TESTING OF MEANS

Mean of the population can be tested presuming different situations such as the population may be normal or other than normal, it may be finite or infinite, sample size may be large or small, variance of the population may be known or unknown and the alternative hypothesis may be two-sided or one-sided. Our testing technique will differ in different situations. We may consider some of the important situations.

1. *Population normal, population infinite, sample size may be large or small but variance of the population is known, H_a may be one-sided or two-sided*:

   In such a situation *z*-test is used for testing hypothesis of mean and the test statistic *z* is worked our as under:

$$z = \frac{\bar{X} - \mu_{H_0}}{\sigma_p / \sqrt{n}}$$

2. *Population normal, population finite, sample size may be large or small but variance of the population is known, H_a may be one-sided or two-sided*:

   In such a situation *z*-test is used and the test statistic *z* is worked out as under (using finite population multiplier):

$$z = \frac{\bar{X} - \mu_{H_0}}{\left(\sigma_p / \sqrt{n}\right) \times \sqrt{(N-n)/(N-1)}}$$

3. *Population normal, population infinite, sample size small and variance of the population unknown, H_a may be one-sided or two-sided:*

   In such a situation *t*-test is used and the test statistic *t* is worked out as under:

$$t = \frac{\bar{X} - \mu_{H_0}}{\sigma_s / \sqrt{n}} \text{ with d.f. } = (n-1)$$

and

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

4. *Population normal, population finite, sample size small and variance of the population unknown, and H_a may be one-sided or two-sided:*

   In such a situation *t*-test is used and the test statistic '*t*' is worked out as under (using finite population multiplier):

$$t = \frac{\bar{X} - \mu_{H_0}}{\left(\sigma_s / \sqrt{n}\right) \times \sqrt{(N-n)/(N-1)}} \text{ with d.f. } = (n-1)$$

Table 9.3: Names of Some Parametric Tests along with Test Situations and Test Statistics used in Context of Hypothesis Testing

Unknown parameter	*Test situation (Population characteristics and other conditions. Random sampling is assumed in all situations along with infinite population*	*Name of the test and the test statistic to be used*		
		One sample	*Two samples*	
			Independent	*Related*
1	*2*	*3*	*4*	*5*
Mean (μ)	Population(s) normal or Sample size large (i.e., $n > 30$) *or* population variance(s) known	*z*-test and the test statistic	*z*-test for difference in means and the test statistic	

$$z = \frac{X - \mu_{H_0}}{\sigma_p / \sqrt{n}}$$

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

In case σ_p is not known, we use σ_s in its place calculating

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

is used when two samples are drawn from the same population. In case σ_p is not known, we use σ_{s12} in its place calculating

$$\sigma_{s12} = \sqrt{\frac{n_1 \left(\sigma_{s1}^2 + D_1^2 \right) + n_2 \left(\sigma_{s2}^2 + D_2^2 \right)}{n_1 + n_2}}$$

where $D_1 = \left(\bar{X}_1 - \bar{X}_{12} \right)$

$D_2 = \left(\bar{X}_2 - \bar{X}_{12} \right)$

$$\bar{X}_{12} = \frac{n_1 \bar{X}_1 + n_2 \bar{X}_2}{n_1 + n_2}$$

Contd.

S. No.	X_i	Hypothesised mean $m_{H_0} = 578$ kg.	$D_i = (X_i - \mu_{H_0})$	D_i^2
5	572	578	–6	36
6	578	578	0	0
7	570	578	–8	64
8	572	578	–6	36
9	596	578	18	324
10	544	578	–34	1156
$n = 10$			$\sum D_i = -60$	$\sum D_i^2 = 1816$

$$\therefore \quad A = \sum D_i^2 / (\sum D_i)^2 = 1816/(-60)^2 = 0.5044$$

Null hypothesis $H_0 : \mu_{H_0} = 578$ kg.

Alternate hypothesis $H_a : \mu_{H_0} \neq 578$ kg.

As H_a is two-sided, the critical value of *A*-statistic from the *A*-statistic table (Table No. 10 given in appendix at the end of the book) for $(n - 1)$ i.e., $10 - 1 = 9$ d.f. at 5% level is 0.276. Computed value of *A* (0.5044), being greater than 0.276 shows that *A*-statistic is insignificant in the given case and accordingly we accept H_0 and conclude that the mean breaking strength of copper wire' lot maybe taken as578 kg. weight. Thus, the inference on the basis of *t*-statistic stands verified by *A*-statistic.

Illustration 6

Raju Restaurant near the railway station at Falna has been having average sales of 500 tea cups per day. Because of the development of bus stand nearby, it expects to increase its 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 Restaurant's sales have increased? Use 5 per cent level of significance.

Solution: Taking the null hypothesis that sales average 500 tea cups per day and they have not increased unless proved, we can write:

$H_0 : \mu = 500$ cups per day

$H_a : \mu > 500$ (as we want to conclude that sales have increased).

As the sample size is small and the population standard deviation is not known, we shall use *t*-test assuming normal population and shall work out the test statistic *t* as:

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

(To find $\bar{X}$ and σ_s we make the following computations:)

Table 9.7

Sample one				*Sample two*			
S.No.	X_{1i}	$X_{1i}-A_1$ ($A_1=10$)	$(X_{1i}-A_1)^2$	*S.No.*	X_{2i}	$X_{2i}-A_2$ ($A_2=8$)	$(X_{2i}-A_2)^2$
1.	12	2	4	1.	8	0	0
2.	15	5	25	2.	10	2	4
3.	11	1	1	3.	14	6	36
4.	16	6	36	4.	10	2	4
5.	14	4	16	5.	13	5	25
6.	14	4	16				
7.	16	6	36				
$n_1=7$;	$\Sigma(X_{1i}-A_1)=28$;		$\Sigma(X_{1i}-A_1)^2=134$	$n_2=5$;	$\Sigma(X_{2i}-A_2)=15$;		$\Sigma(X_{2i}-A_2)^2=69$

$$\therefore \quad \bar{X}_1 = A_1 + \frac{\Sigma(X_{1i}-A_1)}{n_1} = 10 + \frac{28}{7} = 14 \text{ ounces}$$

$$\bar{X}_2 = A_2 + \frac{\Sigma(X_{2i}-A_2)}{n_2} = 8 + \frac{15}{5} = 11 \text{ ounces}$$

$$\sigma^2_{s_1} = \frac{\Sigma(X_{1i}-A_1)^2 - \left[\Sigma(X_{1i}-A_1)\right]^2/n_1}{(n_1-1)}$$

$$= \frac{134-(28)^2/7}{7-1} = 3.667 \text{ ounces}$$

$$\sigma^2_{s_2} = \frac{\Sigma(X_{2i}-A_2)^2 - \left[\Sigma(X_{2i}-A_2)\right]^2/n_2}{(n_2-1)}$$

$$= \frac{69-(15)^2/5}{5-1} = 6 \text{ ounces}$$

Hence,
$$t = \frac{14-11}{\sqrt{\frac{(7-1)(3.667)+(5-1)(6)}{7+5-2}} \times \sqrt{\frac{1}{7}+\frac{1}{5}}}$$

Solution: Let the sales before campaign be represented as X and the sales after campaign as Y and then taking the null hypothesis that campaign does not bring any improvement in sales, we can write:

$H_0: \mu_1 = \mu_2$ which is equivalent to test $H_0: \bar{D} = 0$

$H_a: \mu_1 < \mu_2$ (as we want to conclude that campaign has been a success).

Because of the matched pairs we use paired t-test and work out the test statistic 't' as under:

$$t = \frac{\bar{D} - 0}{\sigma_{diff.}/\sqrt{n}}$$

To find the value of t, we first work out the mean and standard deviation of differences as under:

Table 9.9

Shops	*Sales before campaign* X_i	*Sales after campaign* Y_i	*Difference* $(D_i = X_i - Y_i)$	*Difference squared* D_i^2
A	53	58	–5	25
B	28	29	–1	1
C	31	30	1	1
D	48	55	–7	49
E	50	56	–6	36
F	47	45	2	9
$n = 6$			$\sum D_i = -21$	$\sum D_i^2 = 121$

$$\therefore \quad \bar{D} = \frac{\sum D_i}{n} = -\frac{21}{6} = -3.5$$

$$\sigma_{diff.} = \sqrt{\frac{\sum D_i^2 - (\bar{D})^2 \cdot n}{n-1}} = \sqrt{\frac{121 - (-3.5)^2 \times 6}{6-1}} = 3.08$$

Hence, $$t = \frac{-3.5 - 0}{3.08/\sqrt{6}} = \frac{-3.5}{1.257} = -2.784$$

Degrees of freedom = $(n - 1) = 6 - 1 = 5$

As H_a is one-sided, we shall apply a one-tailed test (in the left tail because H_a is of less than type) for determining the rejection region at 5 per cent level of significance which come to as under, using table of t-distribution for 5 degrees of freedom:

$$R: t < -2.015$$

The observed value of t is – 2.784 which falls in the rejection region and thus, we reject H_0 at 5 per cent level and conclude that sales promotional campaign has been a success.

Solution: *Using A-test:* Using *A*-test, we work out the test statistic for the given problem as under:

$$A = \frac{\sum D_i^2}{\left(\sum D_i\right)^2} = \frac{121}{(-21)^2} = 0.2744$$

Since H_a in the given problem is one-sided, we shall apply one-tailed test. Accordingly, at 5% level of significance the table value of *A*-statistic for $(n-1)$ or $(6-1) = 5$ d.f. in the given case is 0.372 (as per table of *A*-statistic given in appendix). The computed value of *A*, being 0.2744, is less than this table value and as such *A*-statistic is significant. This means we should reject H_0 (alternately we should accept H_a) and should infer that the sales promotional campaign has been a success.

HYPOTHESIS TESTING OF PROPORTIONS

In case of qualitative phenomena, we have data on the basis of presence or absence of an attribute(s). With such data the sampling distribution may take the form of binomial probability distribution whose mean would be equal to $n \cdot p$ and standard deviation equal to $\sqrt{n \cdot p \cdot q}$, where p represents the probability of success, q represents the probability of failure such that $p + q = 1$ and n, the size of the sample. Instead of taking mean number of successes and standard deviation of the number of successes, we may record the proportion of successes in each sample, in which case the mean and standard deviation (or the standard error) of the sampling distribution may be obtained as follows:

Mean proportion of successes $= (n \cdot p)/n = p$

and standard deviation of the proportion of successes $= \sqrt{\frac{p \cdot q}{n}}$.

If n is large, the binomial distribution tends to become normal distribution, and as such for proportion testing purposes we make use of the test statistic z as under:

$$z = \frac{\hat{p} - p}{\sqrt{\frac{p \cdot q}{n}}}$$

where $\hat{p}$ is the sample proportion.

For testing of proportion, we formulate H_0 and H_a and construct rejection region, presuming normal approximation of the binomial distribution, for a predetermined level of significance and then may judge the significance of the observed sample result. The following examples make all this quite clear.

Illustration 13

A sample survey indicates that out of 3232 births, 1705 were boys and the rest were girls. Do these figures confirm the hypothesis that the sex ratio is 50 : 50? Test at 5 per cent level of significance.

Solution: Starting from the null hypothesis that the sex ratio is 50 : 50 we may write:

Solution: Let $H_0: \hat{p} = p$ (there is no difference between sample proportion and population proportion)

and $H_a: \hat{p} \neq p$ (there is difference between the two proportions)

and on the basis of the given information, the test statistic z can be worked out as under:

$$z = \frac{\hat{p} - p}{\sqrt{p \cdot q \frac{N - n}{nN}}} = \frac{\frac{20}{100} - .05}{\sqrt{(.05)(.95)\frac{2000 - 100}{(100)(2000)}}}$$

$$= \frac{0.150}{0.021} = 7.143$$

As the H_a is two-sided, we shall determine the rejection regions applying two-tailed test at 5 per cent level and the same works out to as under, using normal curve area table:

$$R: |z| > 1.96$$

The observed value of z is 7.143 which is in the rejection region and as such we reject H_0 and conclude that there is a significant difference between the proportion of smokers in the college and university.

HYPOTHESIS TESTING FOR COMPARING A VARIANCE TO SOME HYPOTHESISED POPULATION VARIANCE

The test we use for comparing a sample variance to some theoretical or hypothesised variance of population is different than z-test or the t-test. The test we use for this purpose is known as chi-square test and the test statistic symbolised as χ^2, known as the chi-square value, is worked out. The chi-square value to test the null hypothesis viz, $H_0: \sigma_s^2 = \sigma_p^2$ worked out as under:

$$\chi^2 = \frac{\sigma_s^2}{\sigma_p^2}(n - 1)$$

where σ_s^2 = variance of the sample

σ_p^2 = variance of the population

$(n - 1)$ = degree of freedom, n being the number of items in the sample.

Then by comparing the calculated value of χ^2 with its table value for $(n - 1)$ degrees of freedom at a given level of significance, we may either accept H_0 or reject it. If the calculated value of χ^2 is equal to or less than the table value, the null hypothesis is accepted; otherwise the null hypothesis is rejected. This test is based on chi-square distribution which is not symmetrical and all

context of analysis of variance. The following examples illustrate the use of *F*-test for testing the equality of variances of two normal populations.

Illustration 19

Two random samples drawn from two normal populations are:

Sample 1	20	16	26	27	23	22	18	24	25	19		
Sample 2	27	33	42	35	32	34	38	28	41	43	30	37

Test using variance ratio at 5 per cent and 1 per cent level of significance whether the two populations have the same variances.

Solution: We take the null hypothesis that the two populations from where the samples have been drawn have the same variances i.e., $H_0 : \sigma^2_{p_1} = \sigma^2_{p_2}$. From the sample data we work out $\sigma^2_{s_1}$ and $\sigma^2_{s_2}$ as under:

Table 9.10

Sample 1			*Sample 2*		
X_{1i}	$(X_{1i} - \bar{X}_1)$	$(X_{1i} - \bar{X}_1)^2$	X_{2i}	$(X_{2i} - \bar{X}_2)$	$(X_{2i} - \bar{X}_2)^2$
20	–2	4	27	–8	64
16	–6	36	33	–2	4
26	4	16	42	7	49
27	5	25	35	0	0
23	1	1	32	–3	9
22	0	0	34	–1	1
18	–4	16	38	3	9
24	2	4	28	–7	49
25	3	9	41	6	36
19	–3	9	43	8	64
			30	–5	25
			37	2	4
$\Sigma X_{1i} = 220$		$\Sigma (X_{1i} - \bar{X}_1)^2 = 120$	$\Sigma X_{2i} = 420$		$\Sigma (X_{2i} - \bar{X}_2)^2 = 314$
$n_1 = 10$			$n_2 = 12$		

$$\bar{X}_1 = \frac{\Sigma X_{1i}}{n_1} = \frac{220}{10} = 22; \qquad \bar{X}_2 = \frac{\Sigma X_{2i}}{n_2} = \frac{420}{12} = 35$$

$$\therefore \qquad \sigma^2_{s_1} = \frac{\Sigma (X_{1i} - \bar{X}_1)^2}{n_1 - 1} = \frac{120}{10 - 1} = 13.33$$

28. (i) A random sample from 200 villages was taken from Kanpur district and the average population per village was found to be 420 with a standard deviation of 50. Another random sample of 200 villages from the same district gave an average population of 480 per village with a standard deviation of 60. Is the difference between the averages of the two samples statistically significant? Take 1% level of significance.

(ii) The means of the random samples of sizes 9 and 7 are 196.42 and 198.42 respectively. The sums of he squares of the deviations from the mean are 26.94 and 18.73 respectively. Can the samples be constituted to have been drawn from the same normal population? Use 5% level of significance.

29. A farmer grows crops on two fields *A* and *B*. On *A* he puts Rs. 10 worth of manure per acre and on *B* Rs 20 worth. The net returns per acre exclusive of the cost of manure on the two fields in the five years are:

Year	1	2	3	4	5
Field A, Rs per acre	34	28	42	37	44
Field B, Rs per acre	36	33	48	38	50

Other things being equal, discuss the question whether it is likely to pay the farmer to continue the more expensive dressing. Test at 5% level of significance.

30. *ABC* Company is considering a site for locating their another plant. The company insists that any location they choose must have an average auto traffic of more than 2000 trucks per day passing the site. They take a traffic sample of 20 days and find an average volume per day of 2140 with standard deviation equal to 100 trucks.

Answer the following:

(i) If $\alpha = .05$, should they purchase the site?

(ii) If we assume the population mean to be 2140, what is the β error?

(i) First of all calculate the expected frequencies on the basis of given hypothesis or on the basis of null hypothesis. Usually in case of a 2 × 2 or any contingency table, the expected frequency for any given cell is worked out as under:

$$\text{Expected frequency of any cell} = \left[\frac{\text{(Row total for the row of that cell)} \times \text{(Column total for the column of that cell)}}{\text{(Grand total)}}\right]$$

(ii) Obtain the difference between observed and expected frequencies and find out the squares of such differences i.e., calculate $(O_{ij} - E_{ij})^2$.

(iii) Divide the quantity $(O_{ij} - E_{ij})^2$ obtained as stated above by the corresponding expected frequency to get $(O_{ij} - E_{ij})^2/E_{ij}$ and this should be done for all the cell frequencies or the group frequencies.

(iv) Find the summation of $(O_{ij} - E_{ij})^2/E_{ij}$ values or what we call $\sum \frac{(O_{ij} - E_{ij})^2}{E_{ij}}$. This is the required χ^2 value.

The χ^2 value obtained as such should be compared with relevant table value of χ^2 and then inference be drawn as stated above.

We now give few examples to illustrate the use of χ^2 test.

Illustration 3

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 χ^2 as follows:

Table 10.2

No. turned up	*Observed frequency* O_i	*Expected frequency* E_i	$(O_i - E_i)$	$(O_i - E_i)^2$	$(O_i - E_i)^2/E_i$
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

Hence, $$\chi^2 = \Sigma \frac{(O_{ij} - E_{ij})^2}{E_{ij}} = 55.54$$

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

The table value of χ^2 for two degrees of freedom at 5 per cent level of significance is 5.991.

The calculated value of χ^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.

Illustration 8

Eight coins were tossed 256 times and the following results were obtained:

Numbers of heads	0	1	2	3	4	5	6	7	8
Frequency	2	6	30	52	67	56	32	10	1

Are the coins biased? Use χ^2 test.

Solution: Let us take the hypothesis that the coins are not biased. If that is so, the probability of any one coin falling with head upward is 1/2 and with tail upward is 1/2 and it remains the same whatever be the number of throws. In such a case, the expected values of getting 0, 1, 2, … heads in a single throw in 256 throws of eight coins will be worked out as follows*.

Table 10.7

Events or No. of heads	*Expected frequencies*
0	${}^8C_0\left(\frac{1}{2}\right)^0\left(\frac{1}{2}\right)^8 \times 256 = 1$
1	${}^8C_1\left(\frac{1}{2}\right)^1\left(\frac{1}{2}\right)^7 \times 256 = 8$
2	${}^8C_2\left(\frac{1}{2}\right)^2\left(\frac{1}{2}\right)^6 \times 256 = 28$

contd.

*The probabilities of random variable i.e., various possible events have been worked out on the binomial principle viz., through the expansion of $(p + q)^n$ where $p = 1/2$ and $q = 1/2$ and $n = 8$ in the given case. The expansion of the term ${}^nC_r\, p^r\, q^{n-r}$ has given the required probabilities which have been multiplied by 256 to obtain the expected frequencies.

CONVERSION OF CHI-SQUARE INTO COEFFICIENT OF CONTINGENCY (*C*)

Chi-square value may also be converted into coefficient of contingency, especially in case of a contingency table of higher order than 2 × 2 table to study the magnitude of the relation or the degree of association between two attributes, as shown below:

$$C = \sqrt{\frac{\chi^2}{\chi^2 + N}}$$

While finding out the value of *C* we proceed on the assumption of null hypothesis that the two attributes are independent and exhibit no association. Coefficient of contingency is also known as coefficient of Mean Square contingency. This measure also comes under the category of non-parametric measure of relationship.

IMPORTANT CHARACTERISTICS OF χ^2 TEST

(i) This test (as a non-parametric test) is based on frequencies and not on the parameters like mean and standard deviation.

(ii) The test is used for testing the hypothesis and is not useful for estimation.

(iii) This test possesses the additive property as has already been explained.

(iv) This test can also be applied to a complex contingency table with several classes and as such is a very useful test in research work.

(v) This test is an important non-parametric test as no rigid assumptions are necessary in regard to the type of population, no need of parameter values and relatively less mathematical details are involved.

CAUTION IN USING χ^2 TEST

The chi-square test is no doubt a most frequently used test, but its correct application is equally an uphill task. It should be borne in mind that the test is to be applied only when the individual observations of sample are independent which means that the occurrence of one individual observation (event) has no effect upon the occurrence of any other observation (event) in the sample under consideration. Small theoretical frequencies, if these occur in certain groups, should be dealt with under special care. The other possible reasons concerning the improper application or misuse of this test can be (i) neglect of frequencies of non-occurrence; (ii) failure to equalise the sum of observed and the sum of the expected frequencies; (iii) wrong determination of the degrees of freedom; (iv) wrong computations, and the like. The researcher while applying this test must remain careful about all these things and must thoroughly understand the rationale of this important test before using it and drawing inferences in respect of his hypothesis.

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 analysing 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 analyse 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.

THE BASIC PRINCIPLE 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 *j*th 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.

$$F = \frac{\text{Estimate of population variance based on between samples variance}}{\text{Estimate of population variance based on within samples variance}}$$

[1] Donald L. Harnett and James L. Murphy, *Introductory Statistical Analysis*, p. 376.

Source of variation	*SS*	*d.f.*	*MS*	*F-ratio*	*5% F-limit*
Interaction	29.23*	4*	$\frac{29.23}{4}$	$\frac{7.308}{0.389}$	$F(4, 9) = 3.63$
Within samples (Error)	3.50	$(18 - 9) = 9$	$\frac{3.50}{9} = 0.389$		
Total	76.28	$(18 - 1) = 17$			

* These figures are left-over figures and have been obtained by subtracting from the column total the total of all other value in the said column. Thus, interaction $SS = (76.28) - (28.77 + 14.78 + 3.50) = 29.23$ and interaction degrees of freedom $= (17) - (2 + 2 + 9) = 4$.

The above table shows that all the three *F*-ratios are significant of 5% level which means that the drugs act differently, different groups of people are affected differently and the interaction term is significant. In fact, if the interaction term happens to be significant, it is pointless to talk about the differences between various treatments i.e., differences between drugs or differences between groups of people in the given case.

Graphic method of studying interaction in a two-way design: Interaction can be studied in a two-way design with repeated measurements through graphic method also. For such a graph we shall select one of the factors to be used as the *X*-axis. Then we plot the averages for all the samples on the graph and connect the averages for each level of the other factor by a distinct mark (or a coloured line). If the connecting lines do not cross over each other, then the graph indicates that there is no interaction, but if the lines do cross, they indicate definite interaction or inter-relation between the two factors. Let us draw such a graph for the data of illustration 3 of this chapter to see whether there is any interaction between the two factors viz., the drugs and the groups of people.

Graph of the averages for amount of blood pressure reduction in millimeters of mercury for different drugs and different groups of people.*

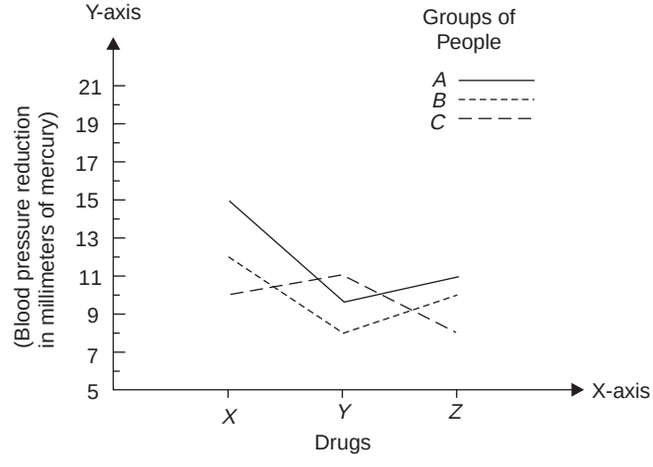


Fig. 11.1

* Alternatively, the graph can be drawn by taking different group of people on *X*-axis and drawing lines for various drugs through the averages.

Source of variation	*SS*	*d.f.*	*MS*	*F-ratio*	*5% F-limit*
Between varieties	48.50	3	$\frac{48.50}{3} = 16.17$	$\frac{16.17}{1.75} = 9.24$	$F(3, 6) = 4.76$
Residual or error	10.50	6	$\frac{10.50}{6} = 1.75$		
Total	113.00	15			

The above table shows that variance between rows and variance between varieties are significant and not due to chance factor at 5% level of significance as the calculated values of the said two variances are 8.85 and 9.24 respectively which are greater than the table value of 4.76. But variance between columns is insignificant and is due to chance because the calculated value of 1.43 is less than the table value of 4.76.

ANALYSIS OF CO-VARIANCE (ANOCOVA)

WHY ANOCOVA?

The object of experimental design in general happens to be to ensure that the results observed may be attributed to the treatment variable and to no other causal circumstances. For instance, the researcher studying one independent variable, X, may wish to control the influence of some uncontrolled variable (sometimes called the covariate or the concomitant variable), Z, which is known to be correlated with the dependent variable, Y, then he should use the technique of analysis of covariance for a valid evaluation of the outcome of the experiment. "In psychology and education primary interest in the analysis of covariance rests in its use as a procedure for the statistical control of an uncontrolled variable."[2]

ANOCOVA TECHNIQUE

While applying the ANOCOVA technique, the influence of uncontrolled variable is usually removed by simple linear regression method and the residual sums of squares are used to provide variance estimates which in turn are used to make tests of significance. In other words, covariance analysis consists in subtracting from each individual score (Y_i) that portion of it Y_i' that is predictable from uncontrolled variable (Z_i) and then computing the usual analysis of variance on the resulting ($Y - Y'$)'s, of course making the due adjustment to the degrees of freedom because of the fact that estimation using regression method required loss of degrees of freedom.*

[2] George A-Ferguson, *Statistical Analysis in Psychology and Education*, 4th ed., p. 347.

* Degrees of freedom associated with adjusted sums of squares will be as under:

Between	$k - 1$
within	$N - k - 1$
Total	$N - 2$

$$\text{Correction factor for } X = \frac{(\Sigma X)^2}{N} = 7616.27$$

$$\Sigma Y = 33 + 72 + 105 = 210$$

$$\text{Correction factor for } Y = \frac{(\Sigma Y)^2}{N} = 2940$$

$$\Sigma X^2 = 9476 \quad \Sigma Y^2 = 3734 \quad \Sigma XY = 5838$$

$$\text{Correction factor for } XY = \frac{\Sigma X \cdot \Sigma Y}{N} = 4732$$

Hence, total *SS* for $X = \Sigma X^2$ – correction factor for X

$= 9476 - 7616.27 = 1859.73$

$$SS \text{ between for } X = \left\{\frac{(49)^2}{5} + \frac{(114)^2}{5} + \frac{(175)^2}{5}\right\} - \{\text{correction factor for } X\}$$

$= (480.2 + 2599.2 + 6125) - (7616.27)$

$= 1588.13$

SS within for X = (total *SS* for X) – (*SS* between for X)

$= (1859.73) - (1588.13) = 271.60$

Similarly we work out the following values in respect of Y

total *SS* for $Y = \Sigma Y^2$ – correction factor for Y

$= 3734 - 2940 = 794$

$$SS \text{ between for } Y = \left\{\frac{(33)^2}{5} + \frac{(72)^2}{5} + \frac{(105)^2}{5}\right\} - \{\text{correction factor for } Y\}$$

$= (217.8 + 1036.8 + 2205) - (2940) = 519.6$

SS within for Y = (total *SS* for Y) – (*SS* between for Y)

$= (794) - (519.6) = 274.4$

Then, we work out the following values in respect of both X and Y

Total sum of product of $XY = \Sigma XY$ – correction factor for XY

$= 5838 - 4732 = 1106$

$$SS \text{ between for } XY = \left\{\frac{(49)(33)}{5} + \frac{(114)(72)}{5} + \frac{(175)(105)}{5}\right\} - \text{correction factor for } XY$$

$= (323.4 + 1641.6 + 3675) - (4732) = 908$

SS within for XY = (Total sum of product) – (*SS* between for XY)

$= (1106) - (908) = 198$

$$= \frac{198}{274.40} = 0.7216$$

	Deviation of initial group means from general mean (= 14) in case of Y	*Final means of groups in X (unadjusted)*
Group I	–7.40	9.80
Group II	0.40	22.80
Group III	7.00	35.00

Adjusted means of groups in X = (Final mean) – b (deviation of initial mean from general mean in case of Y)

Hence,

Adjusted mean for Group I = (9.80) – 0.7216 (–7.4) = 15.14

Adjusted mean for Group II = (22.80) – 0.7216 (0.40) = 22.51

Adjusted mean for Group III = (35.00) – 0.7216 (7.00) = 29.95

Questions

1. (a) Explain the meaning of analysis of variance. Describe briefly the technique of analysis of variance for one-way and two-way classifications.
   (b) State the basic assumptions of the analysis of variance.
2. What do you mean by the additive property of the technique of the analysis of variance? Explain how this technique is superior in comparison to sampling.
3. Write short notes on the following:
   (i) Latin-square design.
   (ii) Coding in context of analysis of variance.
   (iii) F-ratio and its interpretation.
   (iv) Significance of the analysis of variance.
4. Below are given the yields per acre of wheat for six plots entering a crop competition, there of the plots being sown with wheat of variety A and three with B.

Variety	*Yields in fields per acre*		
	1	*2*	*3*
A	30	32	22
B	20	18	16

   Set up a table of analysis of variance and calculate F. State whether the difference between the yields of two varieties is significant taking 7.71 as the table value of F at 5% level for $v_1 = 1$ and $v_2 = 4$.
   (*M.Com. II Semester EAFM Exam., Rajasthan University, 1976*)
5. A certain manure was used on four plots of land A, B, C and D. Four beds were prepared in each plot and the manure used. The output of the crop in the beds of plots A, B, C and D is given below:

		Brands of gasoline			
		W	*X*	*Y*	*Z*
Cars	*A*	13	12	12	11
		11	10	11	13
	B	12	10	11	9
		13	11	12	10
	C	14	11	13	10
		13	10	14	8

13. The following are paired observations for three experimental groups concerning an experimental involving three methods of teaching performed on a single class.

Method A to Group I		*Method B to Group II*		*Method C to Group III*	
X	*Y*	*X*	*Y*	*X*	*Y*
33	20	35	31	15	15
40	32	50	45	10	20
40	22	10	5	5	10
32	24	50	33	35	15

X represents initial measurement of achievement in a subject and *Y* the final measurement after subject has been taught. 12 pupils were assigned at random to 3 groups of 4 pupils each, one group from one method as shown in the table.

Apply the technique of analysis of covariance for analysing the experimental results and then state whether the teaching methods differ significantly at 5% level. Also calculate the adjusted means on *Y*.

[Ans. *F*-ratio is not significant and hence there is no difference due to teaching methods.

Adjusted means on *Y* will be as under:

For Group I 20.70

For Group II 24.70

For Group III 22.60]

The test statistic, utilising the McNemer test, can be worked out as under:

$$\chi^2 = \frac{(|A - D| - 1)^2}{(A + D)} = \frac{(|200 - 100| - 1)^2}{(200 + 100)}$$

$$= \frac{99 \times 99}{300} = 32.67$$

Degree of freedom = 1.

From the Chi-square distribution table, the value of χ^2 for 1 degree of freedom at 5% level of significance is 3.84. The calculated value of χ^2 is 32.67 which is greater than the table value, indicating that we should reject the null hypothesis. As such we conclude that the change in people's attitude before and after the experiment is significant.

4. *Wilcoxon Matched-pairs Test (or Signed Rank Test)*

In various research situations in the context of two-related samples (i.e., case of matched paires such as a study where husband and wife are matched or when we compare the output of two similar machines or where some subjects are studied in context of before-after experiment) when we can determine both direction and magnitude of difference between matched values, we can use an important non-parametric test viz., Wilcoxon matched-pairs test. While applying this test, we first find the differences (d_i) between each pair of values and assign rank to the differences from the smallest to the largest without regard to sign. The actual signs of each difference are then put to corresponding ranks and the test statistic T is calculated which happens to be the smaller of the two sums viz., the sum of the negative ranks and the sum of the positive ranks.

While using this test, we may come across two types of tie situations. One situation arises when the two values of some matched pair(s) are equal i.e., the difference between values is zero in which case we drop out the pair(s) from our calculations. The other situation arises when two or more pairs have the same difference value in which case we assign ranks to such pairs by averaging their rank positions. For instance, if two pairs have rank score of 5, we assign the rank of 5.5 i.e., (5 + 6)/2 = 5.5 to each pair and rank the next largest difference as 7.

When the given number of matched pairs after considering the number of dropped out pair(s), if any, as stated above is equal to or less than 25, we use the table of critical values of T (Table No. 7 given in appendix at the end of the book) for the purpose of accepting or rejecting the null hypothesis of no difference between the values of the given pairs of observations at a desired level of significance. For this test, the calculated value of T must be equal to or smaller than the table value in order to reject the null hypothesis. In case the number exceeds 25, the sampling distribution of T is taken as approximately normal with mean $U_T = n(n + 1)/4$ and standard deviation

$$\sigma_T = \sqrt{n(n + 1)\,(2n + 1)/24}\,,$$

where n = [(number of given matched pairs) – (number of dropped out pairs, if any)] and in such situation the test statistic z is worked out as under:

$$z = \frac{T - U_T}{\sigma_T}$$

We may now explain the use of this test by an example.

Illustration 4

An experiment is conducted to judge the effect of brand name on quality perception. 16 subjects are recruited for the purpose and are asked to taste and compare two samples of product on a set of scale items judged to be ordinal. The following data are obtained:

Pair	*Brand A*	*Brand B*
1	73	51
2	43	41
3	47	43
4	53	41
5	58	47
6	47	32
7	52	24
8	58	58
9	38	38
10	61	53
11	56	42
12	56	56
13	38	44
14	58	57
15	65	40
16	75	68

Test the hypothesis, using Wilcoxon matched-pairs test, that there is no difference between the perceived quality of the two samples. Use 5% level of significance.

Solution: Let us first write the null and alternative hypotheses as under:

H_0: There is no difference between the perceived quality of two samples.

H_a: There is difference between the perceived quality of the two samples.

Using Wilcoxon matched-pairs test, we work out the value of the test statistic T as under:

Table 12.4

Pair	*Brand A*	*Brand B*	*Difference* d_i	*Rank of* $\lvert d_i \rvert$	*Rank with signs* +	–
1	73	51	22	13	13	…
2	43	41	2	2.5	2.5	…

Contd.

values of the first sample (and call it R_1) and also the sum of the ranks assigned to the values of the second sample (and call it R_2). Then we work out the test statistic i.e., U, which is a measurement of the difference between the ranked observations of the two samples as under:

$$U = n_1 \cdot n_2 + \frac{n_1(n_1 + 1)}{2} - R_1$$

where n_1, and n_2 are the sample sizes and R_1 is the sum of ranks assigned to the values of the first sample. (In practice, whichever rank sum can be conveniently obtained can be taken as R_1, since it is immaterial which sample is called the first sample.)

In applying U-test we take the null hypothesis that the two samples come from identical populations. If this hypothesis is true, it seems reasonable to suppose that the means of the ranks assigned to the values of the two samples should be more or less the same. Under the alternative hypothesis, the means of the two populations are not equal and if this is so, then most of the smaller ranks will go to the values of one sample while most of the higher ranks will go to those of the other sample.

If the null hypothesis that the $n_1 + n_2$ observations came from identical populations is true, the said 'U' statistic has a sampling distribution with

$$\text{Mean} = \mu_U = \frac{n_1 \cdot n_2}{2}$$

and Standard deviation (or the standard error)

$$= \sigma_U = \sqrt{\frac{n_1 n_2 (n_1 + n_2 + 1)}{12}}$$

If n_1 and n_2 are sufficiently large (i.e., both greater than 8), the sampling distribution of U can be approximated closely with normal distribution and the limits of the acceptance region can be determined in the usual way at a given level of significance. But if either n_1 or n_2 is so small that the normal curve approximation to the sampling distribution of U cannot be used, then exact tests may be based on special tables such as one given in the, appendix,* showing selected values of Wilcoxon's (unpaired) distribution. We now can take an example to explain the operation of U test.

Illustration 5

The values in one sample are 53, 38, 69, 57, 46, 39, 73, 48, 73, 74, 60 and 78. In another sample they are 44, 40, 61, 52, 32, 44, 70, 41, 67, 72, 53 and 72. Test at the 10% level the hypothesis that they come from populations with the same mean. Apply U-test.

Solution: First of all we assign ranks to all observations, adopting low to high ranking process on the presumption that all given items belong to a single sample. By doing so we get the following:

*Table No. 6 given in appendix at the end of the book.

probability of 0.05, given the null hypothesis and the significance level. If the calculated probability happens to be greater than 0.05 (which actually is so in the given case as 0.056 > 0.05), then we should accept the null hypothesis. Hence, in the given problem, we must conclude that the two samples come from populations with the same mean.

(The same result we can get by using the value of W_l. The only difference is that the value maximum $W_l - W_l$ is required. Since for this problem, the maximum value of W_l (given $s = 5$ and $l = 4$) is the sum of 6 through 9 i.e., 6 + 7 + 8 + 9 = 30, we have Max. $W_l - W_l = 30 - 27 = 3$ which is the same value that we worked out earlier as W_s, – Minimum W_s. All other things then remain the same as we have stated above).

(b) *The Kruskal-Wallis test* (*or H test*)*:* This test is conducted in a way similar to the *U* test described above. This test is used to test the null hypothesis that '*k*' independent random samples come from identical universes against the alternative hypothesis that the means of these universes are not equal. This test is analogous to the one-way analysis of variance, but unlike the latter it does not require the assumption that the samples come from approximately normal populations or the universes having the same standard deviation.

In this test, like the *U* test, the data are ranked jointly from low to high or high to low as if they constituted a single sample. The test statistic is *H* for this test which is worked out as under:

$$H = \frac{12}{n(n+1)} \sum_{i=1}^{k} \frac{R_i^2}{n_i} - 3(n+1)$$

where $n = n_1 + n_2 + ... + n_k$ and R_i being the sum of the ranks assigned to n_i observations in the *i*th sample.

If the null hypothesis is true that there is no difference between the sample means and each sample has at least five items*, then the sampling distribution of *H* can be approximated with a chi-square distribution with $(k - 1)$ degrees of freedom. As such we can reject the null hypothesis at a given level of significance if *H* value calculated, as stated above, exceeds the concerned table value of chi-square. Let us take an example to explain the operation of this test:

Illustration 7

Use the Kruskal-Wallis test at 5% level of significance to test the null hypothesis that a professional bowler performs equally well with the four bowling balls, given the following results:

Bowling Results in Five Games

With Ball No. *A*	271	282	257	248	262
With Ball No. *B*	252	275	302	268	276
With Ball No. *C*	260	255	239	246	266
With Ball No. *D*	279	242	297	270	258

* If any of the given samples has less than five items then chi-square distribution approximation can not be used and the exact tests may be based on table meant for it given in the book "Non-parametric statistics for the behavioural sciences" by S. Siegel.

Table 12.8: Calculation of Spearman's

S. No.	*Interview score X*	*Aptitude test score Y*	*Rank X*	*Rank Y*	*Rank Difference 'd_i' (Rank X) – (Rank Y)*	*Differences squared* d_i^2
1	81	113	21	15	6	36
2	88	88	11	27	–16	256
3	55	76	35	32	3	9
4	83	129	18	9	9	81
5	78	99	24.5	21	3.5	12.25
6	93	142	6	3	3	9
7	65	93	32	25	7	49
8	87	136	13	6	7	49
9	95	82	1.5	30	–28.5	812.25
10	76	91	26	26	0	0
11	60	83	34	28.5	5.5	30.25
12	85	96	15.5	22.5	–7	49
13	93	126	6	10	–4	16
14	66	108	31	18.5	12.5	156.25
15	90	95	9.5	24	–14.5	210.25
16	69	65	29	35	–6	36
17	87	96	13	22.5	–9.5	90.25
18	68	101	30	20	10	100
19	81	111	21	16	5	25
20	84	121	17	12	5	25
21	82	89	19	28.5	–9.5	90.25
22	90	79	9.5	31	–21.5	462.25
23	63	71	33	33	0	0
24	78	108	24.5	18.5	6	36
25	73	68	27	34	–7	49
26	79	121	23	12	11	121
27	72	109	28	17	11	121
28	95	121	1.5	12	–10.5	110.25
29	81	140	21	4	17	289
30	87	132	13	8	5	25
31	93	135	6	7	–1	1
32	85	143	15.5	2	13.5	182.25
33	91	118	8	14	–6	36
34	94	147	3.5	1	2.5	6.25
35	94	138	3.5	5	–1.5	225
$n = 35$						$\sum d_i^2 = 3583$

$$\text{Spearman's } `r' = 1 - \left\{\frac{6\sum d_i^2}{n\left(n^2 - 1\right)}\right\} = 1 - \left\{\frac{6 \times 3583}{35\left(35^2 - 1\right)}\right\}$$

$$= 1 - \frac{21498}{42840} = 0.498$$

Since $n = 35$ the sampling distribution of r is approximately normal with a mean of zero and a standard deviation of $1/\sqrt{n-1}$. Hence the standard error of r is

$$\sigma_r = \frac{1}{\sqrt{n-1}} = \frac{1}{\sqrt{35-1}} = 0.1715$$

As the personnel manager wishes to test his hypothesis at 0.01 level of significance, the problem can be stated:

Null hypothesis that there is no correlation between interview score and aptitude test score i.e., $\mu_r = 0$.

Alternative hypothesis that there is positive correlation between interview score and aptitude test score i.e., $\mu_r > 0$.

As such one-tailed test is appropriate which can be indicated as under in the given case:

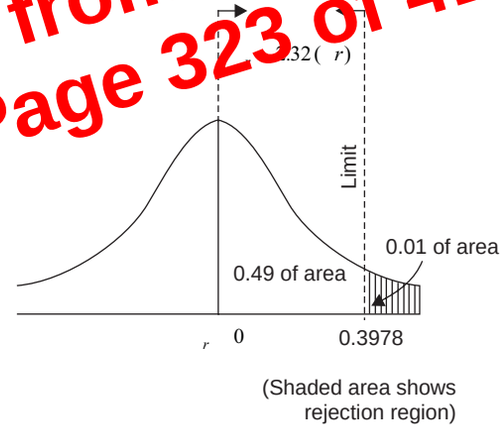


Fig. 12.5

By using the table of area under normal curve, we find the appropriate z value for 0.49 of the area under normal curve and it is 2.32. Using this we now work out the limit (on the upper side as alternative hypothesis is of > type) of the acceptance region as under:

$$\mu_r + (2.32)\,(0.1715)$$
$$= 0 + 0.3978$$
$$= 0.3978$$

(ii) If N is larger than 7, we may use χ^2 value to be worked out as: $\chi^2 = k(N-1)$. W with d.f. $= (N-1)$ for judging W's significance at a given level in the usual way of using χ^2 values.

(f) Significant value of W may be interpreted and understood as if the judges are applying essentially the same standard in ranking the N objects under consideration, but this should never mean that the orderings observed are correct for the simple reason that all judges can agree in ordering objects because they all might employ 'wrong' criterion. Kendall, therefore, suggests that the best estimate of the 'true' rankings of N objects is provided, when W is significant, by the order of the various sums of ranks, R_j. If one accepts the criterion which the various judges have agreed upon, then the best estimate of the 'true' ranking is provided by the order of the sums of ranks. The best estimate is related to the lowest value observed amongst R_j.

This can be illustrated with the help of an example.

Illustration 9

Seven individuals have been assigned ranks by four judges at a certain music competition as shown in the following matrix:

	Individuals						
	A	*B*	*C*	*D*	*E*	*F*	*G*
Judge 1	1	3	2	5	7	4	6
Judge 2	2	4	1	3	7	5	6
Judge 3	3	4	1	2	7	6	5
Judge 4	1	2	5	4	6	3	7

Is there significant agreement in ranking assigned by different judges? Test at 5% level. Also point out the best estimate of the true rankings.

Solution: As there are four sets of rankings, we can work out the coefficient of concordance (W) for judging significant agreement in ranking by different judges. For this purpose we first develop the given matrix as under:

Table 12.9

K = 4	*Individuals*							$\therefore N = 7$
	A	*B*	*C*	*D*	*E*	*F*	*G*	
Judge 1	1	3	2	5	7	4	6	
Judge 2	2	4	1	3	7	5	6	
Judge 3	3	4	1	2	7	6	5	
Judge 4	1	2	5	4	6	3	7	
Sum of ranks (R_j)	7	13	9	14	27	18	24	$\Sigma R_j = 112$
$\left(R_j - \bar{R}_j\right)^2$	81	9	49	4	121	4	64	$\therefore s = 332$

Correlation Matrix, *R*

Variables

		X_1	X_2	X_3	….	X_k
Variables	X_1	r_{11}	r_{12}	r_{13}	….	r_{1k}
	X_2	r_{21}	r_{22}	r_{23}	….	r_{3k}
	X_3	r_{31}	r_{32}	r_{33}	….	r_{3k}
	.	.	.	.		
	.	.	.	.		
	X_k	r_{k1}	r_{k2}	r_{k3}	….	r_{kk}

The main diagonal spaces include unities since such elements are self-correlations. The correlation matrix happens to be a symmetrical matrix.

(ii) Presuming the correlation matrix to be positive manifold (if this is not so, then reflections as mentioned in case of centroid method must be made), the first step is to obtain the sum of coefficients in each column, including the diagonal element. The vector of column sums is referred to as U_{a1} and when U_{a1} is normalized, we call it V_{a1}. This is done by squaring and summing the column sums in U_{a1} and then dividing each element in U_{a1} by the square root of the sum of squares (which may be termed as normalizing factor). Then elements in V_{a1} are accumulatively multiplied by the first row of R to obtain the first element in a new vector U_{a2}. For instance, in multiplying V_{a1} by the first row of R, the first element in V_{a1} would be multiplied by the r_{11} value and this would be added to the product of the second element in V_{a1} multiplied by the r_{12} value, which would be added to the product of third element in V_{a1} multiplied by the r_{13} value and so on for all the corresponding elements in V_{a1} and the first row of R. To obtain the second element of U_{a2}, the same process would be repeated i.e., the elements in V_{a1} are accumulatively multiplied by the 2nd row of R. The same process would be repeated for each row of R and the result would be a new vector U_{a2}. Then U_{a2} would be normalized to obtain V_{a2}. One would then compare V_{a1} and V_{a2}. If they are nearly identical, then convergence is said to have occurred (If convergence does not occur, one should go on using these trial vectors again and again till convergence occurs). Suppose the convergence occurs when we work out V_{a8} in which case V_{a7} will be taken as V_a (the characteristic vector) which can be converted into loadings on the first principal component when we multiply the said vector (i.e., each element of V_a) by the square root of the number we obtain for normalizing U_{a8}.

(iii) To obtain factor B, one seeks solutions for V_b, and the actual factor loadings for second component factor, B. The same procedures are used as we had adopted for finding the first factor, except that one operates off the first residual matrix, R_1 rather than the original correlation matrix R (We operate on R_1 in just the same way as we did in case of centroid method stated earlier).

(iv) This very procedure is repeated over and over again to obtain the successive PC factors (viz. C, D, etc.).

(C) Maximum Likelihood (ML) Method of Factor Analysis

The ML method consists in obtaining sets of factor loadings successively in such a way that each, in turn, explains as much as possible of the population correlation matrix as estimated from the sample correlation matrix. If R_s stands for the correlation matrix actually obtained from the data in a sample, R_p stands for the correlation matrix that would be obtained if the entire population were tested, then the ML method seeks to extrapolate what is known from R_s in the best possible way to estimate R_p (but the PC method only maximizes the variance explained in R_s). Thus, the ML method is a statistical approach in which one maximizes some relationship between the sample of data and the population from which the sample was drawn.

The arithmetic underlying the ML method is relatively difficult in comparison to that involved in the PC method and as such is understandable when one has adequate grounding in calculus, higher algebra and matrix algebra in particular. Iterative approach is employed in ML method also to find each factor, but the iterative procedures have proved much more difficult than what we find in the case of PC method. Hence the ML method is generally not used for factor analysis in practice.*

The loadings obtained on the first factor are employed in the usual way to obtain a matrix of the residual coefficients. A significance test is then applied to indicate whether it would be reasonable to extract a second factor. This goes on repeatedly in search of one factor after another. One stops factoring after the significance test fails to reject the null hypothesis for the residual matrix. The final product is a matrix of factor loadings. The ML factor loadings can be interpreted in a similar fashion as we have explained in case of the centroid or the PC method.

ROTATION IN FACTOR ANALYSIS

One often talks about the rotated solutions in the context of factor analysis. This is done (i.e., a factor matrix is subjected to rotation) to attain what is technically called "simple structure" in data. Simple structure according to L.L. Thurstone is obtained by rotating the axes** until:

(i) Each row of the factor matrix has one zero.

(ii) Each column of the factor matrix has p zeros, where p is the number of factors.

(iii) For each pair of factors, there are several variables for which the loading on one is virtually zero and the loading on the other is substantial.

(iv) If there are many factors, then for each pair of factors there are many variables for which both loadings are zero.

(v) For every pair of factors, the number of variables with non-vanishing loadings on both of them is small.

All these criteria simply imply that the factor analysis should reduce the complexity of all the variables.

* The basic mathematical derivations of the ML method are well explained in S.A. Mulaik's, *The Foundations of Factor Analysis.*

** Rotation constitutes the geometric aspects of factor analysis. Only the axes of the graph (wherein the points representing variables have been shown) are rotated keeping the location of these points relative to each other undisturbed.

4. *Data:* Discussion of data collected, their sources, characteristics and limitations. If secondary data are used, their suitability to the problem at hand be fully assessed. In case of a survey, the manner in which data were collected should be fully described.

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. 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.

The order presented above only gives a general idea of the nature of a technical report; the order of presentation may not necessarily be the same in all the technical reports. This, in other words, means that the presentation may vary in different reports; even the different sections outlined above will not always be the same, nor will all these sections appear in any particular report.

It should, however, be remembered that even in a technical report, simple presentation and ready availability of the findings remain an important consideration and as such the liberal use of charts and diagrams is considered desirable.

(B) Popular Report

The popular report is one which gives emphasis on simplicity and attractiveness. The simplification should be brought 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. Besides, in such a report emphasis is given on practical aspects and policy implications.

We give below a general outline of a popular report.

1. *The findings and their implications:* Emphasis in the report is given on the findings of most practical interest and on the implications of these findings.

2. *Recommendations for action:* Recommendations for action on the basis of the findings of the study is made in this section of the report.

3. *Objective 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 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.

a thousand words. Statistics are usually presented in the form of tables, charts, bars and line-graphs and pictograms. Such presentation should be self explanatory and complete in itself. It should be suitable and appropriate looking to the problem at hand. Finally, statistical presentation should be neat and attractive.

9. *The final draft:* Revising and rewriting the rough draft of the report should be done with great care before writing the final draft. For the purpose, the researcher should put to himself questions like: Are the sentences written in the report clear? Are they grammatically correct? Do they say what is meant'? Do the various points incorporated in the report fit together logically? "Having at least one colleague read the report just before the final revision is extremely helpful. Sentences that seem crystal-clear to the writer may prove quite confusing to other people; a connection that had seemed self evident may strike others as a *non-sequitur.* A friendly critic, by pointing out passages that seem unclear or illogical, and perhaps suggesting ways of remedying the difficulties, can be an invaluable aid in achieving the goal of adequate communication."[6]

10. *Bibliography:* Bibliography should be prepared and appended to the research report as discussed earlier.

11. *Preparation of the index:* At the end of the report, an index should invariably be given, the value of which lies in the fact that it acts as a good guide, to the reader. Index may be prepared both as subject index and as author index. The former gives the names of the subject-topics or concepts along with the number of pages on which they have appeared or discussed in the report, whereas the latter gives the similar information regarding the names of authors. The index should always be arranged alphabetically. Some people prefer to prepare only one index common for names of authors, subject-topics, concepts and the like ones.

PRECAUTIONS IN 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,

[6] Claire Selltiz and others, *Research Methods in Social Relations* rev., Methuen & Co. Ltd., London, 1959, p. 454.

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.

CONCLUSION

In spite of all that has been stated above, one should always keep in view the fact report-writing is an art which is learnt by practice and experience, rather than by mere doctrination.

Questions

1. Write a brief note on the 'task of interpretation' in the context of research methodology.
2. "Interpretation is a fundamental component of research process", Explain. Why so?
3. Describe the precautions that the researcher should take while interpreting his findings.
4. "Interpretation is an art of drawing inferences, depending upon the skill of the researcher". Elucidate the given statement explaining the technique of interpretation.

are being replaced by devices such as bubble memories and optical video discs. In brief, computer technology has become highly sophisticated and is being developed further at a very rapid speed.

THE COMPUTER SYSTEM

In general, all computer systems can be described as containing some kind of input devices, the CPU and some kind of output devices. Figure 15.1 depicts the components of a computer system and their inter-relationship:

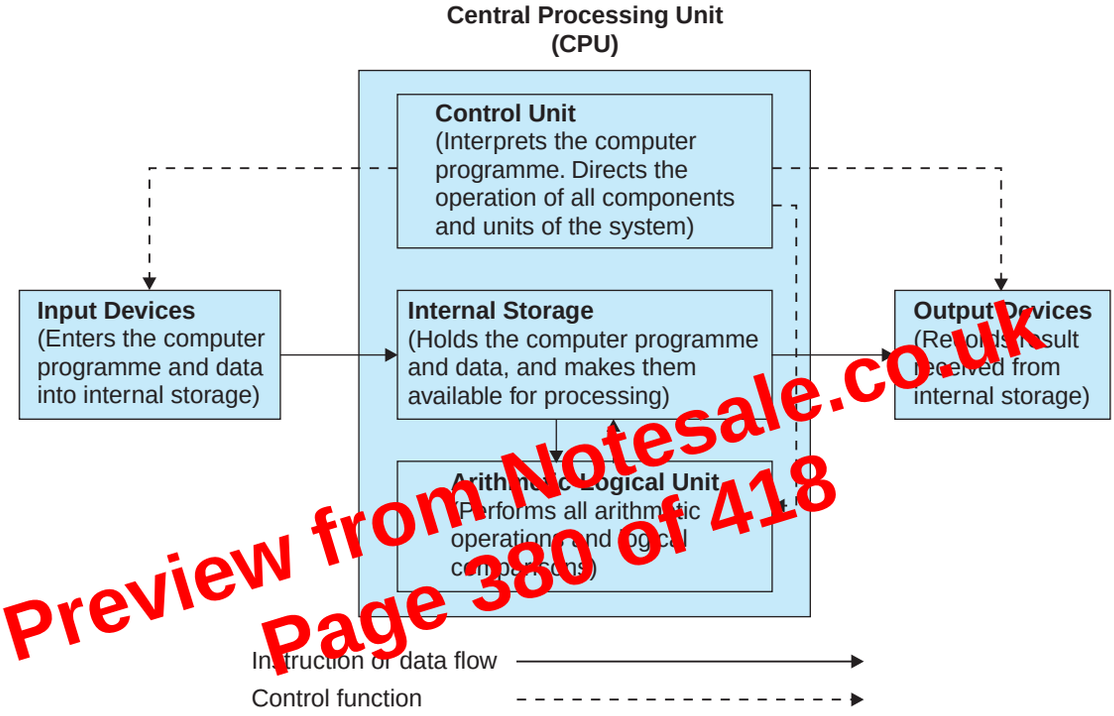


Fig. 15.1

The function of the input-output devices is to get information into, and out of, the CPU. The input devices translate the characters into binary, understandable by the CPU, and the output devices retranslate them back into the familiar character i.e., in a human readable form. In other words, the purpose of the input-output devices is to act as translating devices between our external world and the internal world of the CPU i.e., they act as an interface between man and the machine. So far as CPU is concerned, it has three segments viz. (i) internal storage, (ii) control unit, and (iii) arithmetic logical unit. When a computer program or data is input into the CPU, it is in fact input into the internal storage of the CPU. The control unit serves to direct the sequence of computer system operation. Its function extends to the input and output devices as well and does not just remain confined to the sequence of operation within the CPU. The arithmetic logical unit is concerned with performing the arithmetic operations and logical comparisons designated in the computer program.

In terms of overall sequence of events, a computer program is input into the internal storage and then transmitted to the control unit, where it becomes the basis for overall sequencing and control of computer system operations. Data that is input into the internal storage of the CPU is available for

Solution:

Decimal number	Binary number		According to complementary method	
14	0001110		0001110	
Subtract 72	1001000	Step 1.	+ 0110111	(ones complement of 1001000)
–58		Step 2.	01000101	
		Step 3.	→ 0	(add 0 as no carry)
			1000101	
		Result	– 0111010	(recomplement and attach a negative sign). Its decimal equivalent is –58.

The computer performs the division operation essentially by repeating this complementary subtraction method. For example, 45 ÷ 9 may be thought of as 45 – 9 = 36 – 9 = 27 – 9 = 18 – 9 = 9 – 9 = 0 (minus 9 five times).

Binary Fractions

Just as we use a decimal point to separate the whole and decimal fraction parts of a decimal number, we can use a binary point in binary numbers to separate the whole and fractional parts. The binary fraction can be converted into decimal fraction as shown below:

$$0.101 \text{ (binary)} = (1 \times 2^{-1}) + (0 \times 2^{-2}) + (1 \times 2^{-3})$$
$$= 0.5 + 0.0 + 0.125$$
$$= 0.625 \text{ (decimal)}$$

To convert the decimal fraction to binary fraction, the following rules are applied:

(i) Multiply the decimal fraction repeatedly by 2. The whole number part of the first multiplication gives the first 1 or 0 of the binary fraction;

(ii) The fractional part of the result is carried over and multiplied by 2;

(iii) The whole number part of the result gives the second 1 or 0 and so on.

Illustration 7

Convert 0.625 into its equivalent binary fraction.

Solution:

Applying the above rules, this can be done as under:

$$0.625 \times 2 = 1.250 \rightarrow 1$$
$$0.250 \times 2 = 0.500 \rightarrow 0$$
$$0.500 \times 2 = 1.000 \rightarrow 1$$

Hence, 0.101 is the required binary equivalent.

n	r_0	.10	.25	.40	.50
20	0	.1216	.0032	.0000	.0000
	1	.3917	.0243	.0005	.0000
	2	.6768	.0912	.0036	.0002
	3	.8669	.2251	.0159	.0013
	4	.9567	.4148	.0509	.0059
	5	.9886	.6171	.1255	.0207
	6	.9975	.7857	.2499	.0577
	7	.9995	.8981	.4158	.1316
	8	.9999	.9590	.5955	.2517
	9	1.0000	.9861	.7552	.4119
	10	1.0000	.9960	.8723	.5881
	11	1.0000	.9990	.9433	.7483
	12	1.0000	.9998	.9788	.8684
	13	1.0000	1.0000	.9934	.9423
	14	1.0000	1.0000	.9983	.9793
	15	1.0000	1.0000	.9996	.9941
	16	1.0000	1.0000	1.0000	.9987
	17	1.0000	1.0000	1.0000	.9998
	18	1.0000	1.0000	1.0000	1.0000
	19	1.0000	1.0000	1.0000	1.0000
	20	1.0000	1.0000	1.0000	1.0000

1	2	3	4	5	6
40	0.368	0.263	0.191	0.158	0.102
60	0.369	0.262	0.189	0.155	0.099
120	0.369	0.261	0.187	0.153	0.095
∞	0.370	0.260	0.185	0.151	0.092

* n = number of pairs

Source: *The Brit. J. Psychol*, Volume XLVI, 1955, p. 226.