NATIONAL OPEN UNIVERSITY OF NIGERIA

COURSE CODE: ARD 501

COURSE TITLE: STATISTICS AND RESEARCH METHODS

COURSE GUIDE

ARD 501
STATISTICS AND RESEARCH METHODS

Course Team Dr. Olufemi A. Oladunni (Course Developer/Writer)
Agricultural and Rural Management Training
Institute (ARMTI), Ilorin
Prof. J. O. Otitolaiye (Course Editor) – Kogi State
University, Anyigba

NATIONAL OPEN UNIVERSITY OF NIGERIA

National Open University of Nigeria
Headquarters
14/16 Ahmadu Bello Way

Victoria Island, Lagos

Abuja Office
5 Dar es Salaam Street
Off Aminu Kano Crescent
Wuse II, Abuja

e-mail: centralinfo@nou.edu.ng
URL: www.nou.edu.ng

Published by
National Open University of Nigeria

Printed 2014

ISBN: 978-058-028-X

CONTENTS **PAGE**

INTRODUCTION

Statistics and Research Methods in Extension is a three-credit unit course available to all students offering Agricultural Extension and other related courses.

Research has been seen as the continuous effort of mankind to uncover the unknown as well as advance known areas. Since creation, human beings have made efforts to deal with their environment so that they can exist peacefully in it. Hence, they developed the quest to expand the knowledge of how to remove the obstacles hindering their peaceful existence in their environment.

Human beings and their quest for advancing knowledge, which is referred to as research are inseparable. Over the years, people have sharpened their methods of advancing knowledge through vigorous development in research and statistics. This development has led to rapid progress in making discoveries and inventions for the benefit of man. Research is therefore the process of discovering new knowledge, invention of new objects, and correcting errors in the existing knowledge and inventions.

In peoples struggles for improved nutrition, agro-industrial raw materials and income, man has made efforts to develop its agriculture by teaching farmers improved farm practices. This is the major responsibility of extension workers. The process of extending improved farm practices to farmers is made more efficient through improved knowledge and understanding of the farmers, their learning process and the techniques for efficient communication by extension agents with farmers.

The concept of statistics and research methods remain an important area of advancing the development and adoption of improved technologies in agriculture. The purpose underlying the study of statistics and research methods is to develop your skills and knowledge in this area.

WHAT YOU WILL LEARN IN THIS COURSE

The course consists of units and a course guide. This guide tells you briefly what the course is about, what course materials you will be using and how you can work with these materials.

The course also gives you guidance in respect of your tutor-marked assignment which will be made available in the assignment file. There will be regular tutorial classes that are related to the course. It is advisable for you to attend these tutorial sessions.

COURSE AIMS

This course aims to provide you with an understanding of the research and statistical methods used to explain and draw inferences on the outcome of research findings. It also aims to provide you with approach to research studies for academic purpose. Administrators and other stakeholders in the sector will also find the course useful.

COURSE OBJECTIVES

To achieve the aims set out, the course has a set of objectives. Each unit has specific objectives which are included at the beginning of the unit. You should read and study these objectives which are included at the beginning of the unit. You may wish to refer to them during your study to check on your progress. Make sure you look at the unit objectives after the completion of each unit. By doing this, you would have followed the instructions as stated in the unit.

Below are the comprehensive objectives of the course as a whole. By meeting these objectives, you would have achieved the aims of the course as a whole. It is expected that after going through the course, you should be able to:

- discuss research problems
- develop hypothesis and relate it to the objectives of the research
- design research and prepare the questionnaire for data collection
- use statistical tools to analyse data collected
- present the findings in narrative, tabular and graphical forms.

WORKING THROUGH THIS COURSE

To complete this course, you are required to read each study unit, the textbooks and other materials which may be provided and or recommended by the National Open University of Nigeria.

Each unit contains self-assessment exercise and at certain points in the course, you would be required to submit assignments for assessment purposes. At the end of the course, there is a final examination. Below you will find listed all the components of the course, what you have to do and how you should allocate your time to each unit in order to complete the course on time and successfully too.

This course entails that you spend a lot of time to read and practice. I advice that you attend the tutorial sessions as organised by NOUN, by so doing; you would have the opportunity of comparing your knowledge with that of other students.

THE COURSE MATERIALS

The main components of the course are:

1. The Course Guide
2. Study Units

3. Assignments
4. Presentation Schedule 5. References/Further Reading.

STUDY UNITS

The study units in this course are as follows:

Module 1	**Research Problems and Statement of Hypothesis**
Unit 1	Research: Meaning, Importance and Characteristics
Unit 2	Selection and Formulation of a Research Problem
Unit 3	Developing Hypothesis, Objectives and Identification of Variables
Module 2	**Research Design, Questionnaire Design and Data Collection**
Unit 1	Research Design
Unit 2	Questionnaire Design
Unit 3	Measurement and Data Collection Methods
Module 3	**Statistical Theory and Different Statistical Methods for Handling Data**
Unit 1	Sampling and Statistical Tools
Module 4	**Presentation of Research Findings in Narrative, Tabular and Graphical Forms**
Unit 1	Data Presentation
Unit 2	Report Writing

In module 1, the first unit focuses on research, its meaning, importance and characteristics. The second unit deals with selection and formulation of a research problem. The third unit focuses on developing hypothesis, objectives and identification of variables. Module 2 has three units dealing with research design, questionnaire design, measurement and data collection.

Module 3 is concerned with sampling and statistical tools, while module 4 deals with data presentation and report writing.

Sampling and statistical tools and data presentation units may take between two to three weeks because of their detailed and quantitative nature of the discussion which require some calculations. All other units consist of one or two weeks work. Each of the unit includes an introduction, objectives, reading materials, exercises, conclusion, summary, tutor-marked assignment, references and other resources related to the required reading. In general, these exercises test you on the materials you have just covered or require you to apply it in some way thereby assisting you to evaluate your progress and reinforce your comprehension of the material. Together with the TMAs, these exercises will help you achieve the stated learning objectives of the individual units and of the course as a whole.

PRESENTATION SCHEDULE

Your course materials have important dates for the early and timely completion and submission of your TMAs as well as attending tutorials. You should remember that you are required to submit all your assignments by the stipulated time and date. You should guard against falling behind in your work.

ASSESSMENT

There are three aspects to the assessment of the course. The first is made up of self-assessment exercises, the second consists of the tutor-marked assignments and the third is the written examination/end-of-courseexamination.

You are advised to do the exercises. In tackling the assignments, you are expected to apply information, knowledge and techniques you gathered during the course. The assignments must be submitted to your facilitator for formal assessment in accordance with the deadlines stated in the presentation schedule and the assignment file. The work you submit to your tutor for assessment will count for 30 percent of your total course work. At the end of the course you will need to sit for a final or end of course examination of about three hour duration. This examination will count for 70 percent of your total course marks.

TUTOR-MARKED ASSIGNMENT (TMA)

The TMA is a continuous assessment component of your course. It accounts for 30 percent of the total score. You will be given four TMAs to answer. Three of these must be answered before you are allowed to sit for the end-of-course examination. The TMAs would be given to you by your facilitator and returned after you have done the assignment. Assignment questions for the units in this course are contained in the assignment file. You will be able to complete your assignment from the information and material contained in your reading, references and study units. However, it is desirable

at all degree levels of education to demonstrate that you have read and researched more into your references, which will give you a wider viewpoint and could provide you with a deeper understanding of the subject.

Make sure that each assignment reaches your facilitator on or before the deadline given in the presentation schedule and assignment file. If for any reason you cannot complete your work on time, contact your facilitator before the assignment is due to discuss the possibility for an extension. Extension will not be granted after the due date unless there are exceptional circumstances.

FINAL EXAMINATION AND GRADING

The end of course examination for *Statistics and Research Methods in Extension* will be for three hours and has a value of 70 percent of the total course work. The examination will consist of questions, which will reflect the type of self-assessment exercises and tutor-marked assignments you have previously encountered. All areas of the course will be assessed.

You are advised to use the time between finishing the last unit of the course and sitting for the final examination to revise the entire course. You may find it useful to review your self-assessment exercises, TMAs and comments on them before the examination. The end-of-courseexamination covers information from all parts of the course.

COURSE MARKING SCHEME

Assignment	**Marks**
Assignments 1 – 4	Four assignments, best three marks of the four count for 10 percent each – 30 percent of course marks
End of course examination	70 percent of overall course marks
Total	100 percent of course materials

FACILITATORS/TUTORS AND TUTORIALS

There are 24 hours of tutorials provided in support of this course. You will be notified of the dates, times and location of these tutorials as well as the name and phone number of your facilitator, as soon as you are allocated a tutorial group.

Your facilitator will mark and comment on your assignments, keep a close watch on your progress and any difficulties you might face and provide assistance to you during the course. You are expected to mail your tutor-marked assignment to your facilitator before the schedule date (at least two working days are required). They will be marked by your tutor and returned to you as soon as possible.

Do not delay to contact your facilitator by telephone or e-mail if you need assistance.

Contact your tutor or facilitator if:

- you do not understand any part of the study or the assigned readings
- you have difficulty with the self-tests
- you have a question or problem with an assignment or with the grading of an assignment

You should endeavour to attend the tutorials. This is the only chance to have face-to-face contact with your course facilitator and to ask questions which may be answered instantly. You can raise any problem encountered in the course of your study.

SUMMARY

Statistics and Research Methods in Extension (AEM 501) is a course that intends to provide basic knowledge and skills in the area of research to students in agricultural extension and/or any other area in the social sciences.
On completion of this course, you will be equipped with necessary techniques to conduct research, analyse the data and infer from the data collected by using some statistical techniques discussed in the course.

In addition you should be able to:

- mention the different approaches to research in extension
- engage in research design and make a choice of design to use
- design questionnaire for survey in extension
- analyse the results of the data collected using various statistical tools as discussed in this course
- present your research findings that will address the need of different audiences.

All the above, among others will be the necessary skills and knowledge available to you in this course. I wish you success in the course and hope you will find it interesting and rewarding.

MAIN COURSE

CONTENTS **PAGE**

MODULE 1 RESEARCH PROBLEMS AND STATEMENT OF HYPOTHESIS

Unit 1 Research: Meaning, Importance and Characteristics

Unit 2	Selection and Formulation of a Research Problem
Unit 3	Developing Hypothesis, Objectives and Identification of Variables

UNIT 1 RESEARCH: MEANING, IMPORTANCE AND CHARACTERISTICS

CONTENTS

1.0 INTRODUCTION

Research is the orderly procedure by which man increases his knowledge. It is different from accidental discovery because it follows a series of steps designed precisely for the purpose of developing information.

Research is the continuous effort of mankind to uncover the unknown. Since creation, human beings have struggled to deal with the environment they found themselves in so that it can be conducive for existence.

Research is a careful, critical and disciplined enquiry, varying in techniques and method. According to the nature and condition of the problem identified, research is directed towards clarification or resolution (or both) of the problem. The purpose of research is to discover answers to questions or problems through the application of scientific procedures.

Extension refers to an out of school education services for the members of the family and

others directly or indirectly engaged in farm production to enable them adopt improved practices in production, management, conservation and marketing.

Extension may be defined as the science of developing the capability of the people for sustainable improvement in their quality of life. The main aim of extension is human development.

2.0 OBJECTIVES

At the end of this unit, you should be able to:

- explain the concept of research
- discuss the importance of research in development context
- explain the different types of research
- assess the type of data needed in social research.

3.0 MAIN CONTENT 3.1 Concept of Research

As a follow up from the introduction, research is a scientific investigation aimed at generating dependable solutions to problems through planned and systematic collection, analysis and interpretation of data. In some cases, research may be designed to:

- solve specific research problems which are strategic
- generate new understanding (basic research)
- adjust technology to specific circumstances (adaptive research)
- describe, explain or establish cause and effect relationship (survey research).

Essentially, all forms of research involve the seeking or identification of a problem or problems to investigate. The whole attempt is to discover facts and relationship that will make knowledge and understanding more effective.

3.2 Importance of Research in the Development Context

Development is a process of change from the existing level to a planned new level. It implies that the level aimed at is better (physical, economic, social, etc.) than the existing one. It has no saturation point and therefore, is a continuous process which is positive and purposive. In this context the questions to be asked are:

- what can be done to meet the challenges posed at facilitating development and increase the likelihood that it will move in a constructive direction?
- what kind of institutional arrangements can be designed that would improve the

conditions of people?

- how can change be introduced without destroying the existing cultural patterns and values that provide necessary meaning and stability to people.

These are some of the challenges confronting the society which forms the scope for social research. In this plan, research is designed to contribute to the understanding of:

i. Ways to facilitate constructive change in the direction of meeting human needs and expanding the participation of people in the political, economic and social process.

ii. Ways to minimise the coercive, destructive and psychological disabling consequences of rapid social change. Research that addresses itself to such broad questions like, how can economic developments be facilitated in highly traditional agricultural societies? Or how can the psychological or social dislocations from rapid development change be minimised and counteracted?

By and large, the goal of social research is to improve the standard of living in society. Some people feel that social research is a waste of time, effort and money and think that more fundamental research is necessary for practical benefits. However, it must be conceded that any problem can be solved only on the basis of research evidence rather than personal opinions guided by prejudices or dogmatism.

Many development programmes could not meet with much expected success because the programmes were not guided by scientific study aimed at improvement in comparative cases. Where such investigation was conducted, the research methods were improperly utilised or inappropriately adopted to the socio-cultural setting that the obtained data could not appreciably improve the development programmes. In view of this, it is advocated that research in the social sciences should not be undertaken if it has no implications for a theory, or if it lacks even minimum needs of scientific method or explanatory power. Again, any research that cannot provide answers to social issues or problems should not be undertaken. Therefore, development research is systematic, representative and non-arbitrary information gathered from the field reality and aimed at the solution of priority problems or amelioration of unwanted conditions.

3.3 Research Characteristics and Types

Research should identify problem in the development sphere, objectives or hypothesis (tentative answer or educated guess of the possible outcome), variable (derived from the objectives), method of investigation and data management.

3.4 Types of Research

Research is generally classified as fundamental or basic research and applied or practical research.

a. **Fundamental or basic research**: According to Young (1996), its main purpose is gathering knowledge for knowledge's sake. This is more an intellectual exploration arising from insatiable intellectual curiosity. It is not concerned with solving any practical problem, but with discovering underlying, and if possible, universal laws and theories.

b. **Applied or practical research**: According to Young (1996), this is the gathering of knowledge that could aid in the betterment of human destiny. In reality, no sharp line of demarcation can be drawn between applied and basic research. Each is dependent upon the other for development and verification. It has gradually been realised that many of the results of research are not being adopted by the people. The simple reason is that the researches were conducted under controlled conditions, far away from the reality of people's living and working conditions and as such do not fit their situations. This has given rise to the concept of adaptive research, which emphasises that research be socially useful, and conducted in situations where people live and work. In agriculture, this is known as on-farm research i.e. the research carried out on farmers' farm.

c. **Action research**: This is the method of trying consciously to find out whether or not certain activities actually lead to the results that were anticipated (Khaparde, 1998). The objective is to find solutions to problems in a specific context. Action research is a form of social research, which aims at better insight into the problem by learning from the experience gained in an attempt to solve the problems. Five basic steps are involved in the process.

i. Analysing the problem with the local people and decide on the objectives.
ii. Inducing appropriate action in the situation and working on it with the local people.
iii. Monitoring and evaluation of the results with the local people
iv. Identify constraints.
v. Suggest measures to improve the situation.

Action research leads to a deeper understanding of the situation by the local people and by the research team. It avoids high theory and complex methods. The emphasis is on immediate problem of local relevance.

3.5 Types of Educational Research

Any attempt to classify types of educational research poses a problem. The classification made by Best and Kahn (1992) is scientific, wide and comprehensive, and all researches are likely to fall under one of the following three types or a combination thereof.

a. **Historical research**: This is one which investigates, records, analyses and interprets the event of the past for the purpose of discovering sound generalisations that are helpful in understanding the past and the present and, to a limited extents in, anticipating the future.

b. **Descriptive research**: A descriptive research is one which describes records,

analyses and interprets the condition that exists. In such a research an attempt is made to discover relationship between existing non-manipulated variables, apart from some comparison or contrast among these variables. Descriptive research basically describes "what is". It is also known as non-experimental research.

c. **Experimental research**: In this, certain variables are controlled or manipulated and their effects are examined upon some other variables. Thus experimental research basically describes what will be, when variable are carefully controlled or manipulated.

A comparison of the characteristics and types of experimental, nonexperimental (descriptive) research is given below:

- Despite dissimilarities in experimental and non-experimental (descriptive) research, focus is to verify the postulated relationship between independent and dependent variables.
- Researches in psychology, sociology and education are more of non-experimental. This is because the scope of experimentation and control of variables are difficult in these situations.

3.6 Non-Experimental or Descriptive Research

(i) Field studies: Any ex-post-facto scientific study which systematically discovers relations and interactions among variables in real life situations such as a school, college, factory, community etc. may be called a field study. An ex-post-facto study is one where the investigator tries to trace an effect that has already been produced, to its probable causes. In field experiment, the independent variables are manipulated and its impact upon the dependent variable is examined whereas in field study the investigator does not manipulate variables, rather he or she aims at discovering the relationship and interactions among sociological, psychological and educational variable.

According to Katz (1953), field studies have been divided into two types:

- exploratory field studies
- hypothesis testing field studies.

a. **Exploratory field studies**: These are the studies that intend to discover significant variable in the field situation and finds out relations among those variables so that the ground work for better and more systematic testing of hypothesis can be laid. Thus, exploratory field study seeks what is; it does not

rather seek to produce relations to be found later. For example, productivity of persons may be correlated with several factors like age, education, attitude, values, etc. But in this research, no hypothesis is formulated relating to the productivity and any of these variables.

b. **Hypothesis testing field studies**: Here, the investigator formulates some hypothesis and then proceeds to test them. The research provides some concrete evidence for such testing in order to predict relationships among variables. In field studies, reliability and validity are to be tested. The investigator, for example, formulates the hypothesis that an unfavourable attitude may result in lower productivity. On the basis of the results obtained in the study, one can verify the truth of hypothesis. The hypothesis testing field study is more popular than exploratory field study.

Due to large number of uncontrolled variables, precise measurement of variable is a difficult task and this adversely affects the internal and external validity of field study. A field study also suffers from lack of practicability. It is time and cost consuming activity.

(ii) Ex-post-facto-research

An ex-post-facto research is one in which the investigator attempt to trace an effect which has already occurred to its probable causes. The effect becomes the dependent variable and the probable causes the independent variables. In ex-post-facto research, the investigator has no direct control over the independent variables whose manifestations occur first (as already occurred) and then their effects became obvious. In this type of research, it is difficult for the investigator to control the independent variables either by manipulation or by randomisation. For example, the investigator wants to study the major determinants (factors) of academic achievement among primary school children. The three factors, socio-economic status (SES), motivation and intelligence are most likely to produce differences in academic achievement. After analysis of the data, it is found that higher motivation and higher intelligence are associated with higher academic achievement. In this ex-postfacto study, the dependent variable is academic achievement and the independent variables are socio-economic status, motivation and intelligence, over which the investigator has no direct control.

In ex-post-facto research, the investigator has no direct control over the independent variable which occurs prior to the effect they produce and some variables are inherently not directly or experimentally manipulable. Variables, like home background, school environment, aptitude, intelligence, parental influence are not manipulable. The main drawback of this type of research is that the prediction regarding the relationship between the independent and dependent variables becomes obvious and the investigator may not be able to provide a plausible explanation for the relationship between two types of variables, and a spurious conclusion may be reached as because two factors go together, one is the cause and the other is the effect. Despite the limitations, ex-post-facto research is a popular method for educational and sociological problems.

(iii) Survey research

This may be defined as a technique whereby researcher studies the whole

population with respect to certain sociological and psychological variables. For example, if a researcher wants to study how many people of both sexes in Nigeria adopt contraceptive device as a measure of birth control, this shall constitute a survey research. It is impossible to make an approach to each member of the population, or universe because it requires a lot of time, money and resources. So, a convenient random sample, which is considered to be representative of the whole universe, is selected and subsequently an inference regarding the entire population is drawn from the population for studying the relative incidence, distribution and relationship of psychological and sociological variable. The method is termed as a sample survey.

Survey research is dependent upon the following factors:

a. As survey research deals with characteristics, attitudes and behaviour of individuals or a group of individuals called a sample, direct contact with those persons must be established by the researcher.
b. The success of survey research depends upon the willingness and the cooperativeness of the sample selected for the study.
c. Survey research requires the researcher to be trained personnel, having social intelligence, manipulative skill and research insight.

Depending on the ways of collecting data, survey research can be classified into different categories, namely:

- personal interview
- mail questionnaire panel technique
- telephone survey.

Survey research remains at the surface and it does not penetrate into the depth of the problem being investigated. This is a time consuming, costly and demands expertise, research knowledge and sophistication on the part of the research. Despite these challenges, survey research is an important and indispensable tool for studying social attitude, beliefs, values, etc. with maximal accuracy at the economic rate.

(iv) Content/document analysis

This is a method of systematic examination of communications or of current records or documents. Instead of questioning respondents according to some scale items or observing their behaviour directly, the content analysis takes the communications or documents generated by the respondents and systematically find out the frequency or proportion of their appearances.

In document or content analysis, the primary sources of data are: letters, diaries, autobiographies, records, reports, printed forms, themes or other academic work, books, periodicals, bulletins or catalogues, syllabus, pictures, films, cartons, etc. This analysis is applicable to a wide variety of issues such as creativity, attitude,

ethnocentrisms, stereotypes, curriculum changes, values, interest, religiosity, college budget, etc. It can also be used to examine the effect of experimental manipulation upon the dependent variables. If the investigator wants to study the effect or practice upon the improvement of handwriting of children, content analysis is an important and useful research design.

(v) Case study

This is a one way method of organising social data for the purpose of viewing social reality. It tends to examine a social unit as a whole. The unit may be a person, a family, a social group, a social institution or even a community (Goode and Hatt, 1981 and Best and Kahn, 1992).

Since case study is a descriptive research, no variable is manipulated. The researcher tends to study the aspects of what and why of the social unit, which means he/she not only tries to explain the complex behavioural pattern of the social unit but also tries to locate those factors responsible for such complex behavioural pattern. In case study, the researcher gathers data usually through methods of observation, interview, questionnaire, opinionnaire, checklist and other psychological tests. Analysis of recorded data from newspapers, government agencies as well as interviewing the suspects, friends, relatives is also common.

Case study may be of two types – the individual case study and the community case study. In individual case study, since the social unit consists of one individual or person, it emphasises indepth analysis and is fruitful in developing some hypothesis to be tested, but it is not useful in making broad generalisations. The community case study is one in which the social unit is not a person, rather, a family or a social group. Such case study is a thorough observation and analysis of a group of people who are living together in a particular geographical territory. The community case study tries to deal with different elements of the community life such as location, prevailing economic activity, climate and natural resources, historical development, social structure, life values, health education, recreation, religious expression, and impact of outside world, etc.

The main advantage of case study method is that it provides sufficient basal facts for developing a suitable hypothesis regarding the social unit being studied (Goode and Hatt, 1981). This is possible because of the in-depth analysis of the concerned social unit. The opinion is that case study provides the opportunity of careful examination of all the relevant facts and data on the basis of which a questionnaire or an opinionnaire or any psychological test is to be developed. The main problem in case study is the response of the researcher himself/herself. The researcher may come to feel a false sense of certainty about the conclusions arrived at. The subjective bias of the researcher is a constant threat to objective data gathering in case study. As a consequence, the conclusion may lose its dependability and validity, and the study becomes questionable. It is also a costly method in terms of time and money, and cause and effect relationship is not established. Despite these limitations, case study is a useful method of organising research observations in the social sciences.

(vi) Ethnographic studies

Ethnographic study is a method of field observation or observation of behaviour in natural setting. It consists of participant observation, conversation and the use of informants to study the cultural and social characteristics of primitive people, whose numbers are small and who are geographically and culturally isolated. It is sometimes known as cultural anthropology or more recently as naturalistic inquiry. Nowadays, such observation and conservation have also been extended to the study of different social groups. In ethnographic studies, major emphasis is put on language analysis, marriage, child-rearing practices, religious beliefs and practices, social relations, political institutions, etc. To effectively conduct ethnographic study:

i. The researcher should personally go to the people of the tribe and live for a long period of time to become an integrated member of the social group. He or she should also learn the native language of the tribe for better communication and adjustment with the people
ii. The researcher should have the skill to interpret observation in terms of tribe's concepts, feelings and values and at the same
time supplement his or her own judgment in making objective interpretation of observation.
i. The researcher should be trained or at least should train the informants to record the field data in the tribe's own language and cultural perspective.

Ethnographic study is conducted in real-life setting and natural behaviour is observed. The external validity of ethnographic study is generally high and so its generalisation is valid and sound. But sometimes, the researchers or their informants may fail to maintain the position of neutrality and may be overwhelmed by the strong feeling and emotion of the subjects. This may defeat the basic purpose of the study and invalidate the conclusions of the study. Besides, such study requires trained personnel as well as much time and patience on the part of the researchers who have to live with the tribal community.

SELF-ASSESSMENT EXERCISE

i. Have you ever carried out a research before?
ii. List the different types of research. iii. List the factors that survey research is dependent upon.

3.7 Other Types of Research

i. **Social intervention experiments:** A social intervention experiment is initiated and conducted with a practical end in view. Its purpose is to bring about a tangible and worthwhile change in the course of a natural social process. The experiment is conducted in a real life situation and involves participation of a band of workers with varying orientations, interests, skills and purposes. The two agencies mainly responsible for the management and conduct of the experiment consist of a

research team and an action team. The research team designs the experiment, work out its details, formulates and defines the treatment variable, develops the measures for assessing the treatment effect, takes decision about the sampling method assigns units to treatment, undertakes the analysis of the data, and prepares a report. The action team delivers the treatment to the target population, and collects, records and stores the data in a manner that they are readily available for the final analysis and preparation of the report. The successful completion of the experiment principally depends upon coordination between the research team and the action team.

Social intervention experiments may be designed:

- to evaluate an on-going social intervention programme, or
- to make a preliminary assessment of a proposed intervention under controlled conditions, as a step towards the development of a full scale social innovation plan. A social intervention experiment aims to reach a set of three different though related goals such as, ultimate goal, instrumental goal, and performance goal. The ultimate goal is the realisation of the social change in the target population which the intervention is sought to bring about. The instrumental goal is the measure adopted for the achievement of the ultimate goal. The performance goal relates to the efficient delivery of the treatment to the target population. **ii Cross-cultural research:** This is one in which people belonging to different cultures interact with each other. When an educated, urban oriented researcher goes to study a group of rural or tribal people, he or she encounters a cross-cultural situation. A comparative study of two cultures is also cross-cultural research. Cross-cultural research is more common in anthropology and may be faithfully used in other disciplines. Cross-cultural research provides scope for in-depth and holistic understanding of societies. The findings may help in spreading literacy, improved agriculture, health and family planning practices, awareness about environment among the people. The finding may also help in reaching a target group of people with specific programmes. Participatory research methods may help in making cross-cultural research and come out with findings relevant to the target group.

Cross-cultural research is generally conducted with small groups or communities. Participant observation is not enough. The most important point of cross-cultural research is intuitive understating of the culture of the community. The response of the people comes out of the cultural background which is a whole complex of beliefs, customs, norms, values, tradition, and is transmitted from one generation to another. In cross-cultural research individual variation is held constant, while variation in the culture is the subject of study. In other word, the members composing a culture are treated as homogenous, because it is the capital, representative or normative aspect of behaviour of a cultural unit that is taken into consideration.

Limitations of cross-cultural research

a. Inadequacy of sampling.

b. Problem of defining and measuring the variables.
c. Reliability of data.

In spite of these inadequacies, cross-cultural research has immense value in theory testing, theory building and increasing the range of the variables which facilitate statistical treatment to which the results may be subjected to.

i **Participatory research:** This is an overarching method that can be employed in different social science research method. It essentially implies involvement of the subjects/stakeholders of research output in the process of research.

ii **Evaluation research:** This is a process of applying scientific procedures to accumulate reliable and valid evidence in the manner and extent to which specific activities produce particular effects or outcomes. A programme can be evaluated if it has clearly specified goals and the manner in which the programme is implemented is clear.

Evaluation implies making reasonable judgments about programme outcome based on objective evidence gathered through acceptable methods of research. Evaluation is determination of the results attained by some activity designed to accomplish some valued goal or objective. A programme is an organised set of activities, projects, processes or services oriented toward achievement of specific objectives. Evaluation is the process of determining systematically and objectively the relevance, efficiency, effectiveness and impact of the activity in the light of the objectives.

3.8 Types of Data Collected in Social Research

Whatever may be the type of research, generally five types of data are collected in social science research.

i. Physical environment – natural resources, constraint.
ii. Demographic characteristic of people: what kind of people inhabits a particular area of habitat?
iii. Socio-economic environment - living condition of the people, income, occupation, employment situation, land holding, agricultural activities and outcome, housing and so on. How people live? iv. Activities of the people; what people do? Mass media contact, mobility.
v. People's opinion and activities on a variety of issues of concern.

The details on each of these categories may be varied depending upon the research problem and social context in which it is done.

SELF-ASSESSMENT EXERCISE

i. What is a social research?

ii. List the types of data collected in social research. iii. List the limitations in cross-cultural research.

4.0 CONCLUSION

Research in extension mainly focuses on human science where the totality of social life is brought to play. The different roles of human life constitute a vehicle for dissemination of improved technologies in agricultural production.

5.0 SUMMARY

In this unit, we have examined the concept of research, its importance and characteristics in the developmental context. The concept of social science research, its objectives, the motivating factors for social research, the basic assumptions of social research, and the different sources of data and utility of social science research were dealt with in detail. The different approaches to research and types of research are described. We have also discussed the types of data generally collected in social research.

Key words

Research: It means quest, search, pursuit, and search for truth. The purpose of research is to discover answers to questions or problems through the application of scientific procedures.

Research methodology: It is the description, explanation and justification of various methods of conducting research. It may be understood as a science of studying how research is done scientifically. In it, we study the various steps that are generally adopted by a researcher in studying the research problem along with the logic behind them.

Social Research: It may be defined as the systematic method of discovering new facts verifying old facts, their consequence, interrelationship, casual explanations and the natural laws which govern them.

6.0 TUTOR-MARKED ASSIGNMENT

1. How do you define research?
2. What is the need for research?
3. What are the different approaches to research?
4. Indicate the different types of research.

7.0 REFERENCES/FURTHER READING

Ackroff, R. L. (1953). *The Design of Social Research.* The University of Chicago Press.

Densch, M. & Cooks, W. (1963). *Action Research and its Importance in Under-Developed Economy*. Planning, Research and Action Institute, Planning Dept.,

Government of UP.

Kerlinger, F. N. (1965). *Foundation of Behavioural Research*. New York: Rinehart & Winston.

Olowu, T. A. (2004). *Research Methods in Agricultural Extension*.

Ray, G. L. & Sagar, M. (1999*). Research Methods in Social Sciences and Extension Education.* Calcutta: Naya Prokash.

Settiz, C. *et al.* (n.d.). *Research Methods in Social Relations*. U.S.A: Methuen & Co. Ltd.

UNIT 2 SELECTION AND FORMULATION OF A RESEARCH PROBLEM

CONTENTS

1.0 INTRODUCTION

The purpose of research is to find answers to questions or solutions to problems through the application of specific procedures. This particular unit examines the essential steps in the research process, the factors to be considered in selecting a topic for research which may consequently lead to formulation and evaluation of problem.

2.0 OBJECTIVES

At the end of this unit, you should be able to:

- define a research problem
- identify factors that determine the choice of a problem
- describe sources of research problems
- formulate and evaluate research problem.

3.0 MAIN CONTENT

3.1 Research Problem

The research problem is an account of the major issues that the researcher wants to investigate and explain. It is the main question that the researcher hopes to provide answer to in his study. This is a complex phenomenon, which cannot be understood by mere feeling and observation, but requires solution through reasoning or thinking and application of basis scientific research tools and procedures. According to Adedoyin (2004), a problem is an interrogative sentence or statement that asks what relation exists between two or more variables and which contains implications for empirical testing of its stated relation or relations.

The purpose of research is to find answers to questions or solutions to problems through application of specific procedures.

Fifteen essential steps have been identified in the research process. They are:.

- selecting the field, topic or subject for research
- surveying the field to comprehend the research problem
- developing a bibliography
- formulating or defining the problem
- differentiating and outlining the elements of the problem
- clarifying the elements in the problem according to their relations to data or

evidence
- determining the data or evidence required
- ascertaining the availability of data or evidence
- testing the solvability of problem
- collecting data and information
- synthesising and arranging data preparation to their analysis
- analysing and interpreting the data
- arranging the data for presentation
- selecting and using citations, references and footnotes
- developing the form and style of the research exposition or report.

There are no standard rules that, either singly or collectively, will guarantee the suitability of a research problem. However, the following points should guide the researcher in problem identification.

i. The problem should be sufficiently original that it does not involve objectionable duplication.
ii. The problem must be researchable. iii. The problem must be significant and must be capable of adding new information to existing state of knowledge, provide solutions to problem, generate new technology or improve existing ones, establish the relevance or applicability of technology, described characteristics of variables and explain relationships.
i. Investigation into the problem must be feasible in the circumstances in which the researcher finds himself or herself.
ii. The problem must have a theoretical value.
iii. The problem should avoid moral and ethical issues.

3.2 Selecting a Topic for Research

The range of potential topic for social research is as broad as the range of social behaviour itself. In the development context, either in rural or agricultural development, there are many problems seeking solutions like low productivity in farms, poverty reduction, unemployment or access to services and facilities for the welfare of the people.

The general topics of study may be suggested by some practical concerns or by some scientific or intellectual interest or one may want to find out the impacts of a particular programme. It may also be to predict the future course of events in order to plan an appropriate action. Scientific or intellectual interest may also suggest an equally wide range of topics for research (impact of profit on rural lives). One may be interested in the phenomenon that has been already studied to some extent. If one happens to be working in a research field in which there is highly developed theoretical system, one may want to test specific predictions based on the theory.

Therefore, the scope in selecting a topic for research depends on:

- the nature of interest of the researcher

- the limitation and recommendation of previous research in the field of interest
- the focus of concern of the organisation or department where one works
- the interests of the sponsoring agency (like World Bank) of any research project.

SELF-ASSESSMENT EXERCISE

i. What is a research problem? ii. What is the purpose of research? iii. List four points that will guide a researcher in problem identification.

3.3 Problem Formulation

A problem is an interrogative sentence or statement that asks what relation exists between two or more variables (e.g. what is the relation between training and adoption level of farmers?). The answer to the question is what is being sought in the research.

The next step would be to specify the research problem. It may not always be possible for a researcher to formulate the problem simply, clearly and completely. It may take the researcher quite sometimes, thought and research before he/she can clearly say what questions he/she has been seeking answers to. The scientific enquiry should be geared to the solutions of the problems. It becomes pre-requisite to make the problems concrete and explicit as a first step in formulation of research. The selection of a research topic may be determined by other scientific considerations. The formulation of the topic into a research problem is the first step in a scientific enquiry. It should be influenced by the requirement of the scientific procedure. However, there is no fool proof rule which will guide the investigator in formulating significant questions about a given research area.

Characteristics of research problems

a. The problem must reflect felt needs (i.e. the need for increase in agricultural productivity for food security).
b. The problems are non-hypothetical. The research problem must be based on factual evidence in which case the testing of their validity is not necessary, otherwise it is relegated to the status of hypothesis if it is relevant to the case in question.
c. Problems suggest meaningful, testable hypothesis: Hypothesis are testable when information about their validity may be collected and analysed.
d. Problems are relevant and manageable: An unmanageable research yields few benefit to anyone. The most likely outcome of such research is significant treatment of part of the problem and neglect of other parts thus achieving little details.
e. Researchable problem differs from problematic situation. A problematic situation is a situation which exists as a generalised situation. A researchable problem must

be identified and defined with specificity for instance “the increasing rate of migration of people from rural to urban communities” or “an increasing rate of crime in the cities”. A problematic situation can be the source of a variety of researchable problems.

f. Is the problem one which perennially causes concern (e.g. low agricultural productivity)?

g. Can the effects of a solution to the problem be specifically enumerated?

h. is the problem one which can be really solved by research (by new information)? Or is it one which is the result of such things as lack of funds, inadequate personalities, in appropriate attitudes of the persons concerned or situation in which the research information can make no contribution? (Political considerations sometimes outweigh research findings).

i. Can the solution be obtained in some other way? Is it necessary to study in the community or in the laboratory?

Examples of problem statement

The low level of technology on poultry farms for egg production contributes to high cost of production, low average productivity and a deficient marketing system of poultry products. These factors cause low profits to the producer, price fluctuations for the consumers and deterioration in the balance of payments because of the need to import poultry product (eggs).

The problem of downy mildew disease of maize has led to low productivity in maize production. This has led to low supply of maize for the livestock feed millers and other users of maize. Consequently the millers resort to importation of maize at exorbitant prices to supplement the national output.

There are three criteria for good problem statement.

i. The problem should express a relation between two or more variables.

ii. The problem should be stated clearly and unambiguously in question form.

iii. The problem and problem statement should be such as to imply possibility of empirical testing. This criterion is often difficult to satisfy in some research.

There are certain conditions that are conducive to the formulation of significant problems. Among these conditions are:

- systematic immersion in the subject matter through first hand observation
- the study of existing literature
- discussions with persons who have accumulated much research experience in the field of study.

So the first step in the formulation is the discovery of a problem in need of a solution. It is also necessary to select a topic that would yield a task of manageable size. The task must be reduced to one that can be handled in a single study or divided into a number of sub-questions that can be dealt with in separate

studies.

3.4 Factors Determining the Selection of a Problem

According to Adedoyin (2004), two sets of factors should be considered in choosing a good research problem. These factors are those related to the researcher and those related to the environment. Such factors as epitomised by scholars are as follows:

i. **Workability**: The researcher must consider whether he or she has enough time and funds to carry out the study successfully. ii. **Methodology**: It is necessary to consider whether it is feasible to adopt what appears to be the most appropriate methodology. iii. **Coverage**: In general terms, the more restricted the study, the more thorough the work is likely to be. Coverage refers to both the breadth of the problem and the depth of analysis.

i. **Interest**: It is obvious that a problem, which is of interest to the researcher, would be studied more efficiently by him/her than another in which he/she is not interested. Interest in the problem area will generate enthusiasm in reading extensively about the subject and willingness to be thorough in the analysis.
ii. **Theoretical value**: The work must contribute to knowledge. The study should be properly placed within the context of previous studies. The problem must derive from or be linked to a theory.
iii. **Practical value**: Basic research does not emphasise practical value. But in all other forms of research, practical value is of great importance.

3.5 Problem Evaluation

A researcher needs to consider the following in problem evaluation:

i. Discovering the nature of the problem - who has the problem? Is it capable of solution through research? ii. Determining the end objectives – what is gained by solving the problem and significant to whom and for what?

i. Specifying the possible causes of action to solve the problem where the investigation can begin, how data could be collected and from what sources
ii. Investigate the setting of the problem situation for possible alterations and simplification of the problem.

The definition of the problem should include significance of the problem, definition of the terms necessary for a better understanding of the problem and assumptions underlying the study of the problem and limitation of the scope of the problem area.

There are two types of definitions.

a. Constitutive definition which defines a construct with other constructs. A construct is a concept which has the added meaning of having been deliberately and consciously invented or adapted for a specific scientific purpose (e.g.

achievement of mastery in a certain task, anxiety – subjective fear).

b. Operational definition which assigns a meaning to a construct in specifying the activities or operations necessary to measure the construct. It is a sort of manual of instructions to the investigator. A measured operational definition is one that describes how a variable will be measured (e.g. social participation, cosmopolitans).

An experimental operational definition spells out the details of how the investigator manipulates the variable. Evaluation of research problem after it has been identified and analysed is essential so as to access the characteristics of the problem. The evaluation exercise must reveal that the research problem fulfils the following criteria:

vi. Satisfy the personal goal of the researcher and the expectation of the agency or organisation.

vii. The problem must be of interest to the researcher so as to attract genuine commitment to the research endeavour.

viii. The problem must also be significant in the sense that its investigation should be capable of adding new information into the present state of knowledge.

ix. Research into the problem should as well be feasible. With regards to the organisation or agency, the problem should be researchable and research into the problem should generate solution to existing problems or open up new opportunities, approaches or methods.

x. The researcher must be genuinely interested in it without a biased mind. It is important to choose a research problem or topic in which one has an open mind. This allows for creativity, originality, objectivity, initiative, ingenuity, foresight and all other situations conducive to original thinking.

xi. The researcher must possess the necessary skill, background and knowledge needed to study the problem effectively. It is better for the researcher to focus on problem within his/her field of interest, specialisation, competence or expertise.

xii. Necessary tools, equipment, laboratories and subjects needed to conduct the research must be accessible. This emphasises the fact that research into the problem must be feasible. It is essential that the data and other necessary factors essential for in-depth study of the problem are available in the situation in which the researcher operates.

xiii. The researcher must have time and fund to complete the study. The budget in terms of fund and time to adequately pursue research into the problem must be arranged in advance. Availability of these factors must be ensured before embarking on the research work. This will facilitate good and reliable process and outcome.

xiv. The study must meet the scope, the significance and topical demand by the institution interested in the study. The research work must be well focused in terms of addressing the problem of the sponsors or the end users of the research findings. It is also necessary to ensure adequate coverage in terms of both the breadth of the problem and the depth of the analysis.

xv. There must be access to adequate data for the study. The researcher needs to ensure that there is fair access to relevant data source. (Source: AESON, 2004).

Types of data collected in social research

a. **Survey research**: This involves people's opinion and activities on varieties of concern or directly eliciting responses from subject by questioning.
b. **Direct observation**: This involves the utilisation of either human observers (participants) or mechanical observers (camera, tape recorder) to observe some physical trace of behaviour or the behaviour itself (what people do).
c. **Physical-trace evidence**: This involves natural resources, e.g. fossils and litters.
d. **Document analysis**: This involves the use of record keeping activities of government agencies, private institutions, media organisation and private citizens.

The details of the data may vary depending on the research problem and the social context. The data may also be further classified into primary data, secondary data, time series data, cross-sectional data or experimental data.

SELF-ASSESSMENT EXERCISE

i. List the criteria for good problem statement. ii. What are the steps in problem formulation? iii. List the points to consider in problem evaluation.

4.0 CONCLUSION

From the discussions above, a researcher must not be alien to the problem at hand. He/she must be familiar with the environment where the research problem has been identified and be able to do a proper formulation followed with evaluation.

5.0 SUMMARY

In this unit, we have learnt that:

- the purpose of research is to find answers to questions or solutions to problems through application of specific procedures

- the scope in selecting a topic for research depends on the nature of interest and the organisation/agency
- the proper formulation of the research problem is necessary in order to be able to arrive at a solution to identified problem
- the evaluation of a research problem is essential so as to access the characteristics of the problem.

6.0 TUTOR-MARKED ASSIGNMENT

1. List 15 essential steps in the research process.
2. What are the factors than can determine the choice of a problem by a researcher?
3. The research problems must fulfill some criteria for evaluation to be done. What are these criteria?
4. Give an example of problem statement in agricultural extension.

7.0 REFERENCES/FURTHER READING

Helm, S. G.C. (1970). *Research Concepts in Human Behaviour.* Appleton Century.

Kerlinger, F. N. (1965). *Foundation of Behavioural Research*. New York: Rinehart & Winston.

Olowu, T. A. (2004). *Research Methods in Agricultural Extension*. AESON Publication.

UNIT 3 DEVELOPING HYPOTHESIS, OBJECTIVES AND IDENTIFICATION OF VARIABLES

CONTENTS

1.0 INTRODUCTION

After the problem has been formulated as discussed in the last unit, the hypothesis is derived from the problem. It is a tentative answer to the research question or an educated guess of the research outcome. This unit examines the criteria for hypothesis statement, the development of hypothesis and the identification of variables.

2.0 OBJECTIVES

At the end of this unit, you should be able to:

- develop hypothesis for the research problem formulated
- list the criteria for good hypothesis and type of hypothesis
- describe hypothesis testing and levels of significance identify types of errors in research and
- describe the different types of variable.

3.0 MAIN CONTENT 3.1 Definition and Nature of Hypothesis

A hypothesis is a conjectural statement of the relation between two or more variables. It is a tentative answer to the research question or an educated guess of the research outcome. Hypothesis is always in a declarative sentence form and they relate either generally or specifically from variable to variable. The testable proposition is called hypothesis. It is a proposition, condition or principle which assured perhaps, without a belief in order to draw out its logical consequences. By this method, we can test its agreement with facts which are known and may be determined.

With regards to definition, Onyeogi (2000) defines hypothesis as:

- a proposition stating in a testable form the relationship between two or more variables
- a conjectural statement of the relationship between two or more variables.

Thus, hypothesis does not automatically furnish the researcher with the authentic answer to the problem under investigation until such a hypothesis is tested and confirmed in the face of relevant data.

Hypothesis constitutes part of the researcher's attempt at explaining casual relationships. It is thus an indispensable tool of the research process which enables one

to restrict and streamline one's search for the ultimate solution to the research problem under investigation. According to Young (1992), "Without a working hypothesis, the researcher would find it difficult, laborious and time consuming to make adequate discriminations in the complex interplay of factors before him/her. The hypothesis guides him/her in the selection of pertinent facts needed to explain the problem at hand. It also saves him/her from becoming lost in a welter of irrelevance."

3.2 Criteria and Sources for Good Hypothesis Statement

a. They must be formed as "if then" relationship and stated in such manner that their implications and relationship to the problem can be shown logically. The explicit use of the words "if then" is not necessarily required, the relationship however, is critical and often an explicit "if then" statement will assume an accurate relationship.
b. They should be stated as simple as possible both in terms of theoretical complexities and implication and in terms of number of variable.
c. They must be capable of verification or rejection within the limits of the research resources.
d. They must be stated in a manner which provides direction for the research.
e. Taken together they must be adequate and efficient in suggesting one or more meaningful solution to the problem. They must provide for an acceptable level of confidence in the results, but at the same time economise the use of scarce resources.

Hypothesis may be developed from various sources. A hypothesis may be based simply on hunch. It may rest on the findings of other studies and the expectation that a similar relationship between two or more variables may hold good in the present study. It may be an outcome from a body of theory that through process of logical deduction could lead to the prediction that if certain conditions are present, certain results will follow. However, the basis for correct formulation of hypothesis is the knowledge of the researcher, the knowledge being founded primarily in theory. The broader the experience of the researcher in relating theory to applied problem, the more efficient he will be in formulating appropriate hypothesis.

3.3 Functions of Hypothesis

i. It ensures the optimal use of a researcher's valuable time and other scarce resources by limiting the scope of the inquiry. ii. It transforms research questions into testable propositions.
iii. It leads to discovery of additions to knowledge by helping to confirm or disconfirm particular theories or propositions. iv. It determines the types of data needed for an inquiry and suggests the most appropriate instrument for data collection.

i. It suggests the most appropriate methods and tools for the analysis of data.
ii. It provides the framework for drawing the conclusion of a research endeavour. In fact, conclusions are direct response to the hypothesis formulated for the study as confirmed or discontinued by data analysis.

3.4 Types of Hypothesis

Hypotheses are mostly classified according to the way they are stated in the research process. Based on this method of classification, there are two types of hypothesis:

a. **Null Hypothesis**: It is a succinct way to express the testing of obtained data against chance expectations. The null hypothesis (also termed statistical hypothesis) is a proposition that stipulates that there would be no relationship or difference between the variables being studied and that any such relationship or difference if found to exist does so accidentally or as a result of chance. It is more or less the negation of the directional or alternative hypothesis. The standard error is a means of testing the null hypothesis. It expresses the null hypothesis since it is a measure of expected chance fluctuation around the mean zero.

b. **Alternative Hypothesis**: This is also known as directional hypothesis. It specifies the researcher's expectation of his/her empirical test. It indicates that a relationship or difference exists between two variables or groups and goes further to state the nature or direction of that relationship or differences. An example can be stated thus: "Farmers who have attended extension workshop in the past would produce higher crop output than those who have not attended". This same hypothesis may be stated in a non-directional manner that "there will be a significant difference between the productivity levels of farmers who have attended extension workshops and those who had not attended."

For further emphasis, regardless of the sources of the hypothesis, it performs two important functions:

- it serves as a guide to the kind of data that must be collected in order to answer the research question
- it shows the way in which data can be organised most efficiently in the analysis.

It goes without saying that formulation and verification of hypothesis is a goal of scientific enquiry. Yet there is no shortcut to this goal. In many cases of social relation or social research, significant hypothesis do not exist. Much exploratory research, therefore, must be done before hypothesis can be formulated. Such exploratory work is an inevitable step for scientific progress.

A problem cannot be solved if it cannot be reduced to hypothesis, as a problem is a question not directly testable. Therefore, hypothesis is an important and indispensable tool of scientific research. The reasons for this belief are:

- they are working instrument and can be deduced from theory or from other hypothesis
- they can be tested and shown to be probably true or false
- they are powerful tools for the advancement of knowledge because they enable the researcher to get outside himself/herself.

The problem and hypothesis direct investigation helps to deduce specific empirical manifestation implied by them. It advances scientific knowledge by helping the investigator to confirm or reject a theory. Hypothesis incorporates the theory or part of it, in testable or near testable form.

Some examples of hypothesis

Based on the problems mentioned earlier, the following hypothesis can be stated:

i. There exists in the country improved methods of poultry production which, if used by the producers would increase their profit.
ii. Farmers have not adopted new methods because they are unaware of their existence.
iii. Special credit sources are necessary if farmers are to adopt improved methods of production:
a. Farmers are unable to obtain new technologies due to financial limitations.
b. Farmers are unable to obtain credit which limits their ability to finance changes in production methods.
iv. Egg price stabilisation programme could induce farmers to adopt improved methods of production.

SELF-ASSESSMENT EXERCISE

i. Define hypothesis.
ii. List the functions of hypothesis. iii. List the different types of hypothesis.

3.4 Hypothesis Testing and Levels of Significance

In qualitative studies, the main purpose of the analysis is to test hypothesis, which form the basis of the study and to discuss these in relation to theories in the field. Hypothesis testing is often referred to as significance testing.

A test of significance is conducted by comparing the values of a statistics computed from a sample with values predicted by the sampling distribution under the assumption that the null hypothesis is true. Tests are made at essentially arbitrary levels of significance, usually the five percent or the one percent level. For a difference to be taken as statistically significant or not, the probability that the given difference could have arisen "by chance" must be ascertained. Before the investigator makes a judgment of significance or non-significance, some critical point(s) must be designated along the probability scale which will serve to separate these two judgment categories.

For convenience, researchers have chosen several arbitrary standards call "level of significance" of which the 0.05 and 0.01 levels are most often used. The confidence with which a researcher rejects or accepts a null hypothesis depends upon the level of significance adopted by him/her. Such level of significance must have been set before

she collects his/her data. It is not a good practice to shift from a higher to a lower level after data have been collected. The 0.01 level of significance is more exact than 0.05 level.

Significance is reported in terms of probability of errors. This is going to be discussed in the next section.

3.5 Types of Errors

There are few errors in the nature of problems and hypotheses. They include:

a. Scientific problems are not moral and ethical questions. Value statements that indicate cultural or personal judgments or preferences should be avoided.
b. Another common defect of problem statements often occurs in doctoral thesis: listing the methodological points or problems as sub-problems. These have two characteristics that make them easy to detect.
i. They are substantive problems but they spring from the basic problem
ii. They clearly relate to techniques or methods of sampling, measuring or analysing the data.

Hypothesis testing involves risks because answers are provided in terms of probability. Nobody is absolutely sure that the observed differences or relationships between two variables are not due to chance. The probability value (p-value) is an indication of the odds against the results of the study occurring by chance. There is the chance that the results obtained might have been influenced by forces other than the ones provided for in the study. Therefore, the null hypothesis may be rejected when it should in reality be accepted. Alternatively, the null hypothesis may not be rejected when in reality it should have been rejected.

Type I Errors

These errors are made when the researcher rejects a null hypothesis by making a difference or relationship significant, although no true difference or relationship exists. In other words, type I error is committed by rejecting null hypothesis when it is true, thereby making a non-significant difference or relationship to appear to be significant.

Type II Errors

These errors are made when a researcher accepts a null hypothesis by making a difference or relationship not significant, when a true difference or relationship actually exists. In other words, Type II error is committed by accepting null hypothesis when it is not true, thereby making a significant difference or relationship to appear to be nonsignificant.

The table below presents the decisions or inference that could lead to errors.

Table 1: Error Types in Hypothesis Testing

Decision	**Reality**	
	No Difference	**Difference**
Reject Ho	Type I Error	Correct Decision
Accept Ho	Correct decision	Type II Error

Source: Obilade (1987)

Whenever the significance is doubtful or uncertain, the best way to guide against both types of erroneous inference is to demand or seek more evidence. Additional data, repetition of the experiment and better controls will often make possible a correct judgment.

Setting a high level of significance tends to prevent Type I errors but encourage the appearance of Type II errors. The advice given is that the researcher must decide on which kind of wrong inference he/she would rather avoid, as apparently he/she can prevent one type of error only at the risk of making the other more likely.

The most generally acceptable practice is to set level of significance of at least 0.0 1 in most experimental research, that is, to risk Type II errors by preventing those of Type I. However, it has been expressed that 0.05 level of significance is often satisfactory, especially on preliminary work.

3.6 Developing the Objectives of a Research

Objectives are usually expressed in a layman terminology and are directed as much to the audience or client as to the researcher. The primary objective of research will be either:

- to suggest or recommend to the audience practical means of problem resolution
- to provide information to clarify an unknown situation.

Generally, the objectives taken as a group will:

- define the limits of the research project for the researcher
- clarify the need of conducting the research
- identify the audience
- describe the expected product of the research for the audience.

The objectives link the theoretical relationships presented in the hypothesis to the analytical and methodological orientation necessary for conducting the research. An objective specifies what the researcher intends to do or find in the project and suggests the research procedures to be used. Research objectives are neither political objective nor are they objectives of an action programme of the government, but the information to be obtained. However, the information, in turn, can be used by the audience or sponsor of the research for an action programme. Objectives of a research can be divided into two: Primary and secondary objectives. Primary objectives are broad in scope while secondary objective are more specific.

Based on the hypothesis stated earlier, the following objectives can be given:

- to determine the obstacles to the adoption and optimal use of new technology by poultry farmers
- secondary objectives listed in order of their relationship to the hypothesis
- determine if presently known modern technology is profitable to poultry farmers given their present resource situation and market outlook
- determine if the extension service is effectively providing necessary information to farmers concerning possible alternatives for production
- determine if the required changes to adopt new technology on forms are outside the financial means of the farmers
- ascertain whether present credit sources are adequate in providing for the farmer' needs related to new practices
- obtain farmers' opinion about a price stabilisation programme and their possible reaction to it with respect to changes in use of technology and concurrent production.

Research work plan

The work plan is a guide showing different activities and the time allotted to each. A good work plan shows when the project begins and ends.

It could be in tabular form, making it easy to understand and follow. A work plan assists both the researcher and the funding agency both to match activities with cost and to evaluate and monitor the progress being made on the research.

In some research activities, timing is of utmost importance. For example, research undertaking in agriculture has to be precisely timed with the work plan indicating where too much or too little time is being allocated.

Example of a work plan

For a research planned for 12 months, a work plan may look as follows:

3.7 Variables

Variables refer to any aspect of behaviour or any condition that change. More specifically, variables are those attributes of objects, events, things and beings, which vary and can be measured. In other words, variables have the characteristics or conditions that can be observed, manipulated or controlled by the researcher. Examples of variables include age, education, housing, income, social participation, risk orientation, innovation proneness, value orientation, information on seeking behaviour, level of adoption, intelligence, aptitude, etc. When a characteristic has only one value, it is a constant, not a variable.

Variables can be classified in several ways. Some commonly accepted classifications are presented below:

i. Dependent and independent variables

Dependent variable may be defined as the phenomenon or characteristic hypothesised to be the outcome, effect, consequence or output of some input variable which has preceded it in time. On the other hand, phenomenon or characteristic hypothesised to be the input or antecedent variable is called independent variable. It is presumed to cause the dependent variable and is selected or manipulated prior to measuring the outcome of dependent variable.

For example, if one wants to study the effect of religion upon attitude towards family planning, the individual may take several religious groups and study their attitude towards family planning. By this, the researcher may be able to predict which religious group has a favourable attitude or unfavourable attitude towards family planning. Here the religious groups constitute the example of independent variables and the attitude towards family planning constitutes the example of dependent variable.

It may be mentioned here that a variable can be dependent in one study and independent in another. For example, level of adoption is generally recognised as a dependent variable. It may also be treated as an independent variable if it is intended to study the contribution of adoption in enhancing income.

ii. Moderator variable and control variable

The moderator variables are special type of independent variables which are hypothesised to modify the relationship between independent variables. Age, intelligence, etc. are examples of moderator variables.

Control variable are those which may affect the relationship between the independent and dependent variables, and which are controlled (effects cancelled out) by eliminating the variable, holding the variable constant or using statistical methods. The difference between a control and a moderator variable is that the effects of the control variable are minimised, eliminated or held constant while the effects of the moderator variables are studied. Since both control and moderator variables are independent variables, it is up to the researcher to determine the independent, moderator and control variables.

iii. Intervening variables

A variable which is hypothesised to exist but cannot be observed and is presumed to occur to explain the relationship between the independent and dependent variables is called intervening or hidden variable e.g. the heredity on child intelligence against environment.

According to Kerlinger (1973), the constructs, which are nonobservable, are been called intervening variables. It can neither be seen, nor heard nor felt. It is inferred from the behaviour. Hostility is inferred from presumably aggressive acts. Motivation is an intervening variable. Motivation is a construct invented by man to account for presumably motivated behaviour.

iv. Qualitative attributes and quantitative variable

The qualitative character refers to those which cannot be manipulated after the researcher has started and which consists of categories that cannot be ordered in magnitude. It refers to quality, characteristic or attribute and hence known as attribute characters. Characteristics such as colour, race, sex, religion, etc. are of qualitative type since qualitative characters cannot be ordered in magnitude. Their precise measurements are not possible. However, we may obtain frequencies (a qualitative variable), corresponding to different categories of opinion (a qualitative character), by assigning values in order.

The quantitative variable refers to those variables which are composed of categories that can be ordered in magnitude. That means it may exist in greater or smaller amounts. Examples of quantitative variables are age, income, size of land holding, size of group, intelligence length of experience in the cultivation of a particular crop, adoption quotient and other numerous characteristics. With the quantitative variables, precise measurements are possible because they can easily be ordered in terms of increasing or decreasing magnitude.

A variable which is manipulated by the researcher is active variable. Examples of active variables are reward, punishment, methods of teaching, etc. But anxiety is one-such variable which can be categorised as attribute variable as well as active variable. Anxiety can be manipulated by giving a set of instructions to the subjects. In this case, it becomes an active variable. Anxiety can be measured with the help of a scale or test, and then it constitutes the example of an attribute variable.

v. Continuous and discrete variables

Quantitative variables can be further divided into two categories – continuous variable and discrete variables. A continuous variable is one which is capable of being measured in any arbitrary degree of finesse or exactness. Age, height, intelligence, income, level of adoption, etc. are some examples of continuous variable. Such variables can be measured in the smallest degree of fineness, e.g. height: - take any value within some range. Its value is not fixed (e.g. 1.52-1.525—1.595-1.600). The discrete variables are those variables whose values can be determined by counting. These are not capable of being measured in any arbitrary degree of fineness or exactness because the variables contain a clear gap. For example, the number of members in a group may be 10, 15 or 20, etc. A discrete variable consists only of whole numbers, fractional values such as 10½

15½ or 20½ cannot occur. The number of females in a particular local government area or state, the number of books in a library and so on is some examples of discrete variables.

vi. Stimulus variable and response variable

A stimulus variable is the condition or manipulation created by the researcher so as to evoke response in an organism. The general classes of things the researchers observe that relate to the environment, situation or condition of stimulation are referred to as stimulus variables. The stimulus variables, also known as action variables, may be items like a slide show, a field day, methods demonstration, etc.

Any kind of behaviour of the respondent is called response or behavioural variable. This refers to some action or response of an individual. It may also refer to the frequency with which a particular event occurs or it may be the scale value of a particular event. The responses of farmers for questions like, have you conducted a demonstration? Yes/No or have you attended the field day? Yes/No, are examples of response or behavioural variables?

vii. Extraneous variables

These are those independent variables that are not related to the purpose of the study, but may affect the dependent variable. It is, therefore essential that extraneous variables are controlled. Suppose, an investigator is interested in studying the efficacy of method of instruction on the achievement scores (dependent variable) of some trainees. The methods to be evaluated are lecture, seminar and discussion (independent variable). The researcher discovers that the achievement scores i.e. the dependent variable is positively correlated with intelligence (an extraneous variable) of the subject (trainees). That is, trainees with high intelligence tend to score high on the achievement test and those who are low on intelligence score are low on achievement test. Thus, the variable, intelligence (not of direct interest to the investigator) needs to be controlled because it is a source of variance, which may influence the achievement scores.

SELF-ASSESSMENT EXERCISE

i. What is a variable?
ii. List the different types of variables. iii. Distinguish between dependent and independent variables.

4.0 CONCLUSION

From the discussion in this unit, it is evident that the foundation for a valuable research is laid on the critical processes of hypothesis testing, level of significance, the type of errors and the different variables addressed in the research problem.

5.0 SUMMARY

To sum up, scientific inquiry is an understanding geared to the solution of problems. The first step in formulating the research is to make the problem concrete and explicit. A researcher should identify some aspects of the topic which can be formulated into a specific research problem which is feasible to investigate with the resources and time available. The next step is to propose a tentative solution to the problems, in the form of a testable proposition. This testable proposition is called hypothesis. Also, we have seen the identification of different types of variables which constitute the important, initial endeavour of a researcher.

6.0 TUTOR-MARKED ASSIGNMENT

1. What is meant by hypothesis?
2. What are the different kinds of hypothesis?
3. What is meant by variable in research?
4. What are the different types of variables?
5. Formulate a simple hypothesis for a research.
6. Develop three objectives related to the hypothesis formulated.
7. Do you think identification of variables is necessary to develop the research tools?

7.0 REFERENCES/FURTHER READING

Adedoyin, S. F. (2004). *Research Problems and Statements of Hypothesis*. AESON Publication.

Amali, E. (2005). *Developing a Research Work Plan, Budget and Possible Source of Finance: Basic Issue in Research Methodology*. Faculty of Business and Social Sciences.
University of Ilorin, Ilorin, Nigeria.

Chris, O. A. & Peter, E. H. (1982). *Planning and Conducting Applied Agricultural Research*. Westview Press Inc.

Helm, S. G. C. (1970). *Research Concepts in Human Behaviour*. Appleton Century.

Kerlinger, F. N. (1965). *Foundation of Behavioural Research.* New York: Rinehart & Winston.

Obilade, T. (1978). *Research Methods in Language and Communication.* Ibadan: Odusote Bookstores Ltd.

Ogolo, M. B. (1996). *Student's Guide to Writing Research and Project Proposals*. Port-Harcourt: City-Creeks Publishers, Rivers State, Nigeria.

MODULE 2 RESEARCH DESIGN, QUESTIONNAIRE DESIGN AND DATA COLLECTION

UNIT 1 RESEARCH DESIGN

CONTENTS

1.0 INTRODUCTION

Research problems are to be dealt with in different ways. Some research problems may require only library research, whereas others may require conducting a series of personal interviews at different levels. Each method has its own advantages and disadvantages. One method may be most appropriate for a particular research problem and another may be for some other problems. Then the question is how do we know which method is to be adopted for analysing the given research problem. It is for this reason research should select a research design.

2.0 OBJECTIVES

At the end of this unit, you should be able to:

- provide a definition and purpose of research design
- explain the different types of research designs.

3.0 MAIN CONTENT

3.1 Illustration of the Format of the Research Design

Research design, in simple terms, means an outline of the research project's working or form in which the research is to be set up. It comprises of prior decisions that make up the master plan for executing a research project. The master plan should at least include the following:

- statement of the study objectives, the output of the research
- statement of the data inputs, or causal data, on the basis of which the solution is to be reached
- the analytical method with which the inputs will be treated or calculated.

For instance, if the research problem is to find out the impact of New Rice for Africa (NERICA) cultivation on the input use and the yield per hectare of paddy, the hypothesis here is that NERICA cultivation increases yield per hectare and reduces the usage of water and other essential inputs. The design would have three elements.

1. The objective is to measure the impact of NERICA cultivation on input use and output levels.
2. The relevant data should be gathered with the help of a well designed schedule or a questionnaire.
3. The methodology, the tools and techniques that would be used to analyse the data.

3.2 Benefits of a Research Design

A research design serves as a bridge between what has been established (the research objectives) and what is to be done. If there is no explicit design, the researchers would have only foggy notions about what to do. It is extremely desirable that the design be put in writing. The design thus guides the conduct of the research.

SELF-ASSESSMENT EXERCISE

i. What are the research designs that you have used in the past?
ii. Examine and list their advantages

3.3 Types of Research Designs

Research designs are categorised into three types. These are exploratory, descriptive and causal.

a. **Exploratory research**: This is most commonly unstructured, informal research that is undertaken to acquire background information about the general nature of the research problem. It is used when one is seeking wise-guts into the general nature of a problem, the possible decision alternatives, and relevant variables that need to be considered.
b. **Descriptive research**: This describes phenomena without establishing association between factors. The data may be the behavioural variables of people who are under the study and the situational variables that existed or are forthcoming. A study may measure the adoption levels of hybrid maize technology by farmers, the number of training and demonstration held to increase the adoption levels, etc. It does not study the magnitude of the effect of these trainings/demonstrations in the adoption levels. But it may provide some basis for drawing inferences.
c. **Causal research**: This is carried out only when conclusive research is able to determine whether there is a causal connection between an action that the decision maker is considering and the objective being sought.

Example: A fertiliser company placed a new fertiliser based on precision farming for sale. Before launching this product, market survey was conducted on its sales potential. However, after launching the product, sales did not increase as per the anticipation. The company manager hypothesised that the low volume of sales is due to poor advertisement campaign on the benefits of the fertiliser among farming community. Now the challenge was to test the hypothesis (extensive advertisement) would bring about the objective (higher sales volume).

From the above example, to test causation a conclusive research is very important. Absolute proof of causation is difficult to prove. Causal research would need to address the question of causation. For instance, training to the farmers on new methods of cultivation was really a cause for increased production.

The following conditions are necessary to decide cause and effect statements.

i. Strong evidence of association between an action and observed outcome need to be gathered. In other words, there should be enough evidence of variation. This means that the variables occur together.

ii. There should be evidence that action preceded the result or an outcome. This means time sequence of variables (that the causal event occurs either before or simultaneously with the effect) is important.
iii. The outcome or result should be explained by only this factor not by any other factor. If the hypothesis is that maize production increases with the use of organic manures, keep all other things constant.

How are these causal relationships determined? They are determined through the use of experiment. These research designs can fall under two sub-categories (1) experimental (2) quasi-experimental.

Experimental designs

We can conduct experiments in two settings – in the laboratory or in the field. In a laboratory experiment, the researcher administers the treatment to subjects in a controlled environment. This desired setting assists the researcher in minimising unwanted effects of extraneous variables through the control and manipulation of certain variables. However, laboratory experiments provide a high level of internal validity because the respondents can be carefully controlled/or manipulated. We can conclude that the results obtained from a laboratory experiment are due to the changes in the treatment variables(s) because we are eliminating the effects of many extraneous variables. But the results of laboratory experiments may not be able to project the real world situation which means external validity of the results is questioned. Since the human factor as a variable is involved, more and more doubts of the external validity of the results coming out of laboratory experiment increases.

Field experiments

Field experiments are conducted in the real world environment. This form of experimentation provides a high level of external validity as the respondents are reacting as they would be in normal circumstances.

Experimental designs are intended to measure cause and effect relationships. They can measure concomitant variation, the degree of change in one variable (say maize production), when other variable (say fertiliser consumption) is changed. For example, in agriculture, researchers hypothesise that using a particular pesticide (the cause or treatment) would reduce or kill the pest (effect). To demonstrate this, two farmers suffering from the same problem are chosen. The treatment is administered in one of the farmer's field (the test group) while the second farmer is the control group. The difference in the yield rate between the test group and the control group is attributed to the effects of the pesticide.

Classic experimental designs: Randomisation of subjects into control and treatment groups is a classic experimental method. The two broad classes of classic experimental design are:

a. **Between subjects design**

In a between subjects design, the values of the dependent variable for one group of participants (the group of farmers who have prior experience in using a knapsack sprayer) are compared with the values for another group of participants (for example, the group of farmers who have never used the knapsack sprayer before). In between subjects design, responses from a given participant appear in only one group.

b. Factorial designs

Let us imagine a design where we have a training programme where we would like to look at a variety of programme variations to see which works best. For instance, we would like to vary the amount of time other farmers receive instruction with one group getting one hour per week and another group four hours per week. We would like to vary the setting with one group getting the training in the field and another group in the classroom. We could think about having four separate groups to do this, but when we are varying the amount of time in instruction what setting would we use in class or in the field? And when we are studying the setting, what amount of instruction time would be used- one hour, four hours or something else?

In factorial design, a factor is a major independent variable. In the above example, we have two factors: time spent in instruction (through 4 hours/week) and training setting (field and classroom). A level is a sub division of a factor. In this example, time in instruction has two levels and setting has two levels. In this particular example, we have 2 x 2 factorial designs. The number notation will tell us the number of factors and number of levels. The number of different treatment groups that we have in any factorial design can easily be determined by multiplying through the number notation. For instance, in our example we have 2 x 2 = 4 groups.

A null outcome would be indicated by the average retention score being the same for all four groups of the factorial design. Unequal mean retention scores would indicate a main effect of time factor in imparting training and training setting or an interaction effect of both.

Randomised block designs

In the randomised block designs, the researcher will identify a single extraneous factor that he thinks might affect the test units' response to the treatment. Therefore, the researcher will attempt to isolate that extraneous factor by blocking out its effects. Then he stratifies the subjects and for each stratum, a factorial design is run.

Latin-square design

This design blocks out the effect of two extraneous factors. In this design, treatment on each subject is tested one at a time in some sequence. As a result of rotating the treatments, it is expected that the extraneous factors will offset each other and wash out.

Quasi-experimental design

An experiment should be used whenever it is feasible and is costjustified in conclusive studies. The experimental design discussed earlier is not feasible in social sciences because human factor is involved whose behaviour is very random.

The quasi experimental designs do not require rigorous qualities of experimental design. However, they are also not descriptive because they do yield some quantitative indicators of association between their variables. There are five types of quasi experimental designs. They are (1) one-shot case study (2) one-group pretest-post test design (3) longitudinal design (4) static group comparison design (5) simulations.

One-shot case study

This is the simplest of all the designs. In this type, test units (those who are exposed to the causal variable) are not selected randomly but on some other basis, such as self selection. These subjects are first exposed to the causal (or treatment) variable x, and then measurements are taken afterward.

Example: The non-agriculture faculty working in an agriculture organisation is asked to undergo training voluntarily in some important aspects of agriculture. After the training is over, the performance of the faculty that has undergone training voluntarily is compared with the faculty that has not undergone training.

However, this type of design suffers poor sample selection, and inability to compare the prior performance of the faculty who has undergone training.

Non-equivalent group, post-test only

The non-equivalent, post-test only design consists of administering an outcome measure to two groups or to a programme/treatment group and a comparison. For example, one group of students might receive reading instruction using a whole language programme while the other receives a phonetics based programme. After 12 weeks, a reading comprehension test can be administered to see which programme was more effective.

A major problem with this design is that the two groups might not be necessarily the same before any instruction takes place and may differ in important ways that influence what reading progress they are able to make. For instance, if it is found that the students in the phonetics groups perform better, there is no way of determining if they are better prepared or better readers even before the programme and/or whether other factors are influential to their growth.

Non-equivalent group, pre-test-post-test

The non-equivalent group, pretest-post-test design partially eliminates a major limitation of the non-equivalent group, post-test only design. At the start of the study, the researcher empirically assesses the differences in the two groups. Therefore, if the researcher finds that one group performs better than the other on the post-test, he/she can rule out mutual differences (if the groups were in fact similar on the pretest).

This design is subject to some limitations.

i. The test units are volunteered. ii. Validity of the history because of events intervening between pretest and post-test. iii. The environment that may be socio-economic may have changed during that period.

i. Maturation (changes in the subjects). For instance the faculty which has undergone training might have improved its performance due to experience during the training period.
ii. There is no control group to compare with (faculty who did not have the training).
iii. Sometimes the pre-test data is collected at the same time as the post-test data, as when the researcher asks for recollection data of the "before" state. This is known as a proxy pre-test-post-test design and has additional validity problems since the pre-test data are usually significantly less reliable.

Ex-post facto research

Ex-post facto means "from what is done afterwards". In the context of research, it means "after the fact" or "retrospectively" and refers to those studies which investigate possible cause and effect relationship by observing an existing condition or state of affairs and searching back in time of plausible causal factors.

Steps in ex-post facto design

i. Formulate the research problem including identification of factors that may influence dependent variable(s). ii. Identify alternate hypothesis that may explain the relationships. iii. Identify and select the subject groups. iv. Collect and analyse data.

Ex-post facto studies cannot prove causation, but may provide insight into understanding of phenomenon.

Longitudinal designs or time series designs

In time series designs, several assessments (or measurements) are obtained from the treatment group as well as from the control group. This occur prior to and after the application of the treatment. The series of observations before and after can provide rich information about phenomena that we are studying. Because measures at several points in time prior and subsequent to the programme are likely to provide a more reliable picture of achievement, the time series design is sensitive to trends in performance. Even in this design, limitations and problems of non-equivalent group, pre-test-post-test design is still applicable.

This design is an extension of pre-test-post-test design. It provides repetitive measurements of same kind of events at various points in time. Changes that take place during those intervals are registered. This type of design is very useful in monitoring studies.

SELF-ASSESSMENT EXERCISE

i. Briefly explain the following:
a. Exploratory research
b. Descriptive research
c. Causal research.
ii. What is factorial design?

4.0 CONCLUSION

Research design is a detailed blueprint used to guide a research study towards its objectives.

5.0 SUMMARY

In this unit, we have learnt that research design presents the important features of the research methods to be used with the justification of the strengths and limitations of the chosen method relative to the alternatives. Exploratory research, descriptive research, causal research, experimental design, quasi-experimental design, factorial design among others was discussed.

6.0 TUTOR-MARKED ASSIGNMENT

1. Is a research design always necessary before a research study is conducted?
2. In what ways do exploratory, descriptive and causal research designs differ? How will the difference influences the relative importance of each research approach at each phase of research project?
3. Is experimental design useful in the social sciences? How?

7.0 REFERENCES/FURTHER READING

Cook, T. D. & Donald T. C. (2002). *Quasi-Experimentation Design and Analysis Issues for Field Setting*. Boston: Houghton-Mifflin, 1979.

Creswell, J. W. (2002). *Research Design: Qualitative, Quantitative, and Mixed Methods Approach.* Thousand Oaks, C.A.: Sage Publications.

Duncan, S. C. & T. E. D. (1994). 'Modeling Incomplete Longitudinal Substance Use Using Latent Growth Cure Methodology'. *Multivariate Behavioral Research*, 29:313-338.

Leedy, P. & Jeanne E. O. (2004). *Practical Research: Planning and Design* (8th ed.).

Levin, I. (1999). *Relating Statistics and Experimental Design*. Thousand Oaks, C.A: Sage Publications.

Pedhazur, E. J. & Schmelkin, L. P. (1991). *Measurement, Design, and Analysis: An Integrated Approach*. Lawrence Erlbaum Association.

Shadish, W. R, Thomas, D. C. & Donald T. C. (2002*). Experimental and Quasi-Experimental Designs for Generalised Causal Inference*. Boston: Houghton-Mifflin.

UNIT 2 QUESTIONNAIRE DESIGN

CONTENTS

1.0 INTRODUCTION

Questionnaire design is one of the major tools in social survey. Survey has been defined as a fact-finding study that covers all types of social investigations. Survey methods have attracted considerable attention in recent years as sources of information on the attitudes and behaviour of people (living in rural or urban areas). The purpose of many surveys is simply to provide someone with information. For example, an extension agent may want to know the rate of adoption of certain seeds among farmers in a given community, how families of different size, composition and social status react to new technologies. Whatever may be the case, the purpose of social survey is to explain the relationships between certain variables.

Moser and Kalton (1971) emphasise that some of the principal ways of collecting information in social survey are mail, questionnaire and personal interviews.

2.0 OBJECTIVES

At the end of this unit, you should be able to:

- develop questionnaire for social survey
- structure a questionnaire

- determine what it takes to pre-test questionnaire
- describe the process of administering questionnaire.

3.0 MAIN CONTENT

3.1 Structuring a Questionnaire

A questionnaire or recording schedule is the instrument on which the principal ways of collecting survey data depend.

These two terms (the questionnaire or recording schedule) are sometimes used to distinguish between a situation in which the respondent fills in the answers and that in which an interviewer asks the questions and records the answers. To the survey designer, the distinction is very important because in one case, he/she is producing a document to be understood by the respondent unaided, while in the other case the users will be persons specially trained to handle such document.

In questionnaire design, the convenience of the interviewer and respondent should be considered. Layout and printing should be such that editing and coding can proceed smoothly. One of the skills in designing a questionnaire is to make clear which questions are to be answered and by whom? Sometimes, whole sections concern only a subgroup of respondents, in which case this should be quite plain to whoever is recording the answers. Frequently, 'a question' consists of a main question and one or more "dependent" questions. For example, one may want to know the number of people that are growing rice in a community. If rice growers are identified, one may want to probe further into the methods used in rice growing. In this case only rice growers will answer the second question.

It is obvious that the survey planner must vigorously examine every question and exclude any that is not strictly relevant to the survey's objectives. In this, the pilot survey or pre-test is the most helpful tool. When setting the scope of a questionnaire, one important criterion to consider is that of the practicability of the question. It is not good asking persons' opinions about something they do not understand or about events too long ago for them to remember accurately or about matters which although concern them, they are unlikely to have accurate information on or that are so personal or emotional that valid answers cannot be expected by formal direct questioning.

Guidelines on questionnaire design

I. Question content

In considering the contents of questionnaires, the following must be considered:

i. The researcher should keep in mind, when writing each item, that the major purpose of the questionnaire is in two folds: to translate the research objective into specific information and to assist the interviewer in motivating the respondent to communicate, the required information.

ii. The language used must be gauged to both the level of the group to be interviewed and the precision of the data needed. In choosing the language for a questionnaire, the population being studied should be kept in mind. The aim in question wording is to communicate with respondents as nearly as possible in their own language.

iii. The social acceptability of the possible alternative answers must be considered. Most respondents will lie or refuse to respond rather than destroy their ego by giving a socially unacceptable answer. Others may attempt to attract attention to themselves by exaggeration.

iv. Leading questions: a leading question is one which by its contents, structure or wording, leads the respondent in the direction of a certain answer. For example, "do you read only daily newspapers, such as *Punch* or *Tribune*?" Respondents may seek refuge in the answers named. Hence, either all or none of the alternatives should be stated.

v. Each question should be limited to a single idea. Double-barreled and ambiguous questions such as did the extension agent communicate clearly and accurately should be avoided at all costs. If an ambiguous word creeps in, different people will understand the question differently and will, in effect, be answering different questions. The following example can be considered, "is your work made more difficult because you are ploughing manually? The question may be asked of all farmers in the survey irrespective of whether they are ploughing manually. What then, would a "No" answer mean?

vi. Subjects which people do not like to discuss in public present a problem to the questionnaire designer. Respondents are often embarrassed to discuss private matters, to give low prestige answers and to admit to socially unacceptable behaviour and attitudes. If, for instance, questions on number of children and frequency of taking a bath were asked from local farmers, many of them would probably refuse to reply and others may distort their answers.

II. Question order

It is usually best to arrange the sequence of questions from the more general to the more specific. In putting the individual questions together to form the questionnaire, the order of questions needs to be planned. The order may affect the refusal rate and there is plenty of evidence that it may also influence the answers obtained.

At the start of the interview, the respondent is unsure of himself/herself and so the opening questions should be ones to put him at ease and build up rapport between him/her and the interviewer. They should be interesting questions, which he/she will have no difficulty in answering, and they should not be sensitive topics, otherwise he/she may refuse to continue with the interview. The questions should then proceed in a logical manner moving from topic to topic in a way that indicates to the respondent the relationship between the questions.

When determining the order of questions within a topic, the conditioning effect of earlier questions should be considered. It is not good asking "can you name any insecticide" if a previous question has mentioned all known insecticides. In order words, knowledge questions must not be preceded by relevant information. Even though interest may centre on specific issues, it is a good idea to start with broad question about the subject and then to narrow down to the specific issues.

3.2 Pre-testing Questionnaire

It is exceedingly difficult to plan a survey without a good deal of knowledge of its subject matter, the population it is to cover, the way people will react to questions and, even the answers they are likely to give. One must be able to have an idea of how long the survey will take, how many interviewers will be needed, and how much it will cost.

How, without trial interviews, can one be sure that the questions will be as meaningful to the average respondent as to the survey expert? How is one to decide which questions are worth asking at all? Commonsense suggests the necessity of doing a few test interviews or sending out trial forms by way of preparing for the main survey. This is known as "Pretest". Pre-test is a standard practice widely used in research surveys.

Role of pre-test in survey

Pre-test provides guidance on the following:

i. The adequacy of sampling frame from which it is proposed to select the sample. For example, one might be planning to use the tax register of farmers as the basis for drawing a sample.
ii. The variability (with regard to the subject under investigation) within the population to be surveyed. This is important in determining an efficient sample design since the very decision on sample size requires some knowledge of the variability of the population.
iii. The non-response rate to be expected. The probable numbers of results and non-contacts can be roughly estimated from the pretest and the effectiveness of various ways of reducing nonresponse can be compared.
iv. The suitability of the method of collecting the data. Alternative methods are available, and one needs data on their relative cost, accuracy and likely response rates to make a sensible choice.
v. The adequacy of the questionnaire. The questionnaire will have been previously

tested informally on colleagues and friends, but the pre-test offers a way of trying it with the kind of interviewers and respondents to be used on the main survey.

vi. The probable cost and duration of the main survey and of its various stages. If it appears that the survey will take too long or be too expensive, the pre-test can be valuable in suggesting when economies can be made.

From the above, pre-test can help to guide the choice between alternative methods of collecting the data, ordering the questions, wording and so forth. Questionnaires should therefore be designed, so as to ensure a strict testing of these alternatives.

SELF-ASSESSMENT EXERCISE

i. List the guidelines in questionnaire design.
ii. State the role of pre-test in surveys.

3.3 Administering Questionnaire

For a meaningful and successful administration of questionnaire, special permission is needed to approach certain categories of individuals or institutions. The request usually takes the form of a letter of intent, stating what the research project is about, the methods to be used and the nature of the sample. This should be as brief as possible and in clear language. A copy of the interview schedule may be required, though this is more likely at the state or local than at national level.

Arranging for an announcement about a survey either on the local radio or through the village head before interviewing begins is a good way of spreading information about its acceptability to the authorities and gaining cooperation from the people. Announcements in churches and mosques are also useful.

Before the work in each village begins, the supervisor must have visited the chief to discuss the survey, emphasising that the replies will be confidential and would not be revealed to government officials.

The major tasks of an interviewer are to locate his/her sample members, to obtain interviews with them and to ask the questions and record the answers as instructed. The amount of time the interviewer spends on different aspects of his/her work depends, however, on such factors as the length of the questionnaire and the nature of the sample.

Having located the respondents, the interviewer has to obtain an interview. The aim of his/her introductory procedures must be to increase the respondent's motivation to co-operate. He/she can achieve this by first stating the organisation he/she represents and perhaps showing an authorisation card. In many cases, this, together with a brief statement of why the survey is being done, is enough to secure cooperation. When the survey answers are to be treated as confidential and anonymous, this should be made clear to the respondent. It is often worth explaining in simple terms, how the sample was selected and that lack of cooperation would make it less representative. The interview

should remove any suspicion that he/she is out to ask test questions, (i.e. to find out how much the respondent knows).

In most large-scale surveys, the aim is to attain uniformity in the asking of questions and recording of answers. In consequence, the training of interviewers is oriented towards efficiency in following instructions. An interviewer is expected to ask all applicable questions: to ask them in the order given and with no more elucidation and probing than is explicitly allowed, and to make no unauthorised variations in the wordings.

Characteristics of inadequate response

i. Partial response in which the respondent gives a relevant but incomplete answer. ii. Non-response, when the respondent remains silent or refuses to answer the question. iii. Irrelevant response, in which the respondent does not answer the question asked.

i. Inaccurate response, when the question is answered by a reply which is biased or distorted.
ii. Verbalised response problems, when the respondent explains why he/she cannot answer the question, perhaps because he/she fails to understand it because he/she thinks it is irrelevant or inappropriate.

Techniques for dealing with inadequate response

i. Interviewer should allow a brief expectant pause to develop, thus indicating that something more is required, but the pause must not be allowed to become a long one or it may have a negative effect.
ii. Expectant glances can also be useful, they can similarly be overdone.
iii. Give the respondent encouragement by an expression such as "that is interesting" or "I see".
i. When the interviewer needs to ask a supplementary question, it must be a neutral one, such as "how do you mean? "Can you tell me more about that?"
ii. Repeating part of the respondent's answer in a question manner can also be useful, but an attempt to summarise the respondent's answer for this can also be dangerous.

Recording

Another important aspect of questionnaire administration is recording. Recording answers may seem straight forward, but bias can be introduced here as everywhere else. An interviewer must not be so busy writing that he is not really listening; otherwise he may write what he/she expects to hear

rather than what was said. One must keep in mind that there is no "correct" answer. Some will be quite contrary to what is expected and some may seem threatening or misinformed. Nevertheless, if all answers could be predicted in advance, then there would be no need for a survey. In any case, people have a right to their opinions. Write answers down as accurately as possible with no rounding out to what was "probably possible". There is no need for complete sentences if only a phrase was used but what is written must be understandable to an outsider. If it does not make sense, probe.

Ball point pens must be used rather than pencils, because the latter will rub off and be illegible. Use a plastic bag to keep schedules dry and keep them neat and together so they do not get lost. There is much work to be done on them, after the interviewing is completed.

SELF-ASSESSMENT EXERCISE

i. What is inadequate response in questionnaire administration?
ii. List some of the symptoms of inadequate response. iii. Mention two techniques of dealing with inadequate response.

4.0 CONCLUSION

In this particular unit, we have not attempted to deal comprehensively with the subject of questionnaire design. The points selected for discussion have been those thought to be of interest to learners embarking on survey. To the problem of questionnaire design in general, there is no easy solution. Even if one follows all the accepted principles, there usually remains a choice of several questions forms, each of which seems satisfactory.

The answer to a question lies in a detailed pre-test, more than anything else. It is the essence of a good questionnaire. The interviewing technique for a pre-test tends to resemble the informal interviewing. This is different from formal interviewing of the final survey.

6.0 SUMMARY

Question designing remains a matter of common sense, experience and avoidance of known pitfalls. It is not yet, a matter of applying theoretical rules. Alternative versions of questions must be vigorously tested in pre-tests, and as such, tests of practicability must play a crucial role in questionnaire design.

6.0 TUTOR-MARKED ASSIGNMENT

1. Mention the two major guidelines in questionnaire design.
2. In each of them, describe three main points to be considered.
3. What are the roles of pre-test in survey?
4. List the symptoms of inadequate response.
5. Describe the techniques for an inadequate response.

7.0 REFERENCES/FURTHER READING

Awolola, M. D (2004). *Questionnaire Design in Agricultural Extension Research.* AESON Publication.

Contrill, H. (1944). *Gauging Public Opinion*. Cambridge: Princeton University Press.

Kahn, R. C. & Connell, C. F. (1957). *The Dynamics of Interviewing: Theory, Techniques and Cases*. New York: Wiley.

Peil, M. (1982). *Social Research Methods: An African Handbook*. London: Hodder and Stoughton.

UNIT 3 MEASUREMENT AND DATA COLLECTION METHODS

CONTENTS

1.0 INTRODUCTION

The design of the research project specifies both the data that are needed and how they are to be obtained. It means that once the research problems are formulated, it should become evident what kind of data will be required to study the problem, and what kind of analytical tools are appropriate to analyse the data.

2.0 OBJECTIVES

At the end of this unit, you should be able to:

- find out what kind of data will be required to study the research problems
- explain the different data collection devices, their strengths and weaknesses.

3.0 MAIN CONTENT 3.1 Data Elicitation Method

It is a way of getting information, so that you could gain information by observing what people do (observation), and what they do-think-or feel (self report). Some methods of data collection involve no direct contact with the object of research. They rely essentially upon archives (i.e. records) or artifacts as sources of data (archival). The data observation, self-report and archival information can then be recorded in many different ways. All these can be used to generate both qualitative or quantitative data records (and sometimes both at the same time). When we use observation we can choose to give a qualitative account of what happens or we can decide to report a quantified breakdown of what happens.

3.2 Nature of Data

Not all categories of data analysis are always as distinct as they might first appear. They can, depending on the particular research aims, be closely interlinked. For example, in exploratory research the data may be continuously analysed as they are collected. These analyses keep giving clues to the most fruitful area of further data collection and subsequent analysis. When a particular phenomenon is investigated according to a specific predetermined methodology, it might not even be possible to begin the analysis until all the relevant data have been collected.

When considering what data you might require, consider carefully the sources, the availability and the possible methods of collecting data. When considering analysis, think about the tools, techniques and resources required. The different research strategies have often distinctly different methods for data collection and analysis.

3.3 Quantitative Versus Qualitative Research

Most researchers use either qualitative or quantitative research. Qualitative relates to distinctions based on quality and quantitative relate to considerations of size.

The difference might be summarised by saying that quantitative research is structured, logical, measured and wide. Qualitative research is more intuitive, subjective, and deep. This implies that some subjects are best investigated using quantitative while in others, qualitative approaches will give better results. In some cases both methods can be used.

The main difference between qualitative and quantitative research is not of quality but of procedure. In qualitative research, findings are not arrived at by statistical methods or other procedures of quantification. In some studies data may be quantified, but the analysis itself is qualitative, such as with census data reports. It is quite common for researchers to collect their data through observation and interviews, the methods normally relate to qualitative research. Qualitative and quantitative methods are therefore not mutually exclusive. Qualitative research is mixture of the rational, explorative and intuitive, where the skills and experience of the researcher play an important role in the analysis of data. It is often focused on social process and not on social structures, which is frequently the focus in quantitative research. In quantitative research, variables need to be measurable (more about measurement) and the more precisely defined the variable, the easier it is for the data to be analysed and measured.

3.4 Collecting Data

The first step in data collection process is to look for secondary data. These are data that were developed for some purpose other than for helping to solve the problem at hand.

Economic and social statistics are collected incidentally or intentionally. A considerable body of data is available as a result of administrative acts e.g. statistics of crime, car accidents, numbers of wage-and salary earners incidental to payroll tax collections, and details of imports incidental to collection of custom duties. These statistics are not collected primarily for research purposes, although they may be very useful in the field of research. Since they are not collected primarily for research, they may not be in a form completely appropriate to research.

You will collect two kinds of data during your research. First, the historical data is collected from background reading, access to written records, audio disks, CDs, video and computer sources, letters, drawings and so on. Second, the first hand data that you generate during research process, interview information, questionnaire returns and so on.

In all studies, majority of researchers gather some primary data to answer their research question. Once the researcher has decided to collect information/data through primary sources, he/she has to decide what kind of data collection method to use. He/she could use observation, experiment, interview or survey. However, the choice of data collection will depend upon judgment on which type of data is needed for a particular research problem. One important aspect is to identify/know the unit of analysis. Here, we consider some details on how to review the literature to form the background to your study. By now, you should have gained considerable experience in the method of data collection and recording, but it will probably be useful at this point to go back and check on the main points again.

SELF-ASSESSMENT EXERCISE

i. Briefly explain qualitative research and quantitative research.

ii. List the steps in data collection.

3.5 Types of Data

To find information when one is doing a historical study (i.e. of any past events or ideas, even the very recent past), three major types of data can be collected. The first is the primary data found in the form of historic artifacts, such as building and ruins, commercial and domestic objects, human and animal remains, works of art, etc. The second is primary or secondary data in the form of literary sources, such as histories, commentaries, diaries, letters, etc. The third is records, which are contemporary, impersonal recording of events, situations and states, which may be descriptive or statistical in nature. These, again, may be classified as primary or secondary forms of data, depending on the types of information which you are extracting from them. Any of them can be quantitative or qualitative in nature.

3.5.1 Secondary Data

These are useful not only to find information to solve our research problem, but also to better understand and explain our research problem. In most research, we need to begin with a literature review (earlier studies on and around our topic of research). They include books, journal, articles, and online data sources such as web pages of firms, government, semi-governments organisations and catalogues. The first step is to locate these sources and then to evaluate the usefulness of the contents of each. Some research questions can be answered only through secondary data sources, where no further data collection is needed. The first and foremost advantage of using secondary data obviously is the enormous saving in time and money. The researcher needs to go to the library and locate and utilise the sources. This not only helps the researcher to better formulate and understand the research problem, but also broadens the base for which scientific conclusions can be drawn. In other words, the verification process is more rapid and the reliability of information and conclusions is greatly enhanced.

3.5.2 Primary Data

When secondary data are not available or are unable to help our research questions, we must ourselves collect the data that are relevant to our particular study and research problem. These data are called primary data. What we should look for, ask about and collect depend upon our research problem and research design. We have several choices as regards the means of collecting primary data. Normally this includes observation, experiments, surveys (questionnaires) and interviews. The main advantage of primary data is that they are collected for a particular project at hand. This means that they are more consistent with our research questions and research objectives.

3.6 Sources of Data/Information

One of the main problems faced by the researcher seeking historical and recorded data is that of locating and accessing them. Another fact is often that of authenticating these sources, and the question of interpretation.

Locating historical data is an enormous task. Activities can involve anything from unearthing city ruins in the desert to rummaging through dusty archives in an obscure library or downloading the latest government statistical data from the internet. It is difficult to limit a brief description of sources to those which might be relevant to student research, as the nature of the detailed subject of research determines the appropriate source and the possible range of subjects is enormous. However, here are some of the principal sources:

i. **Libraries and archives**: They are generally equipped with sophisticated catalogue systems which facilitate the tracking down of particular pieces of data or enable a trawl to be made to identify anything which may be relevant. International computer networks can make remote searching possible.
ii. **Museums and collections**: These often have efficient cataloguing systems which will help your search. However, problems may be encountered with searching and access in less organised and restricted and private collections. Larger museums often have their own research departments which can be of help.
iii. **Government department and commercial/professional bodies**: They often hold much statistical information both current and historic. iv. **The internet**: This is a rapidly expanding source of information of all types.
v. **The field**: Not all historical artifacts are contained in museums. Ancient cities, buildings, archaeological digs, etc. are available for study in situ. Here, various types of observation will be required to record the required data.

Two important factors that play a vital role are:

i. Authentication of historical data can be a complex process, and is usually carried out by experts. A wide range of techniques are used, for example, textural analysis, carbon dating, paper analysis, location checks, cross-referencing and many others. ii. Interpretation is an integral part of the analysis of the data. Although it can be argued that a current interpretation of the historical evidence is required before any real analysis can begin. A detailed historical analysis of an event will be worthless if the historical data have been incorrectly interpreted, for example, if the evidence was from a source whose bias was undeleted.

SELF-ASSESSMENT EXERCISE

i. List types of data.
ii. Mention two factors that play important role in sources of data.

4.0 CONCLUSION

To conduct a research study, two types of data can be gathered. They are primary and secondary data. Secondary data collection is the first step in data collection. Primary data, in contrast, is collected to solve the particular problem under investigation.

5.0 SUMMARY

In this unit, we have learnt that collection of data is a way of getting information after proper analysis. Also, we have learnt that the quality and nature of data collected will determine the information that will emanate from it. The issues of quantitative and qualitative were also discussed. In this, a particular research lends itself to the usage of both quantitative and qualitative.

6.0 TUTOR-MARKED ASSIGNMENT

1. What is data elicitation method?
2. Distinguish between quantitative and qualitative type of research.
3. Explain the difference between secondary and primary data.
4. What are the advantages of using secondary data?
5. Name two precautions that are required while collecting the data.

7.0 REFERENCES/FURTHER READING

Adler, P. A & Adler, P. (1994). *Observation Techniques.* Sage Publication.

Churchill, G. A. (1991). *Marketing Research: Methodological Foundations*. Dryden Press.

Richard, Y. C. & Matthew, E. N. (1995). *Continuous Improvement Tools*. London: Kogan Page Publications.

MODULE 3 STATISTICAL THEORY AND DIFFERENT STATISTICAL METHODS FOR HANDLING DATA

Unit 1 Sampling and Statistical Tools

UNIT 1 SAMPLING AND STATISTICAL TOOLS

CONTENTS

1.0 INTRODUCTION

The world is full of potential data. You will, however, only be interested in collecting data relevant to your study and specifically required in order to investigate your research problem. Even so, the amount of information you could collect on your specific subject is likely to be enormous, so a method must be used to limit the amount of data you must collect to achieve your aims. The main technique for reducing the drudgery of your data collection is to study a sample, i.e. a small section of the subject of your interest. There are several things one must consider in selecting a sample.

2.0 OBJECTIVES

At the end of this unit, you should be able to:

- list the various sampling techniques
- explain the different statistical tools and techniques.

3.0 MAIN CONTENT

3.1 Why Take Samples?

When the research problem is specified, and an appropriate research design and data collection instrument is developed, the next step in the research process is to select those elements from which the information will be collected. One possibility is to collect information from each member of the population. Another way is to collect information from a portion of the population by taking a sample of elements from the larger group. On this basis, infer something about the larger group. For quantitative studies in particular, sampling is extremely important. A well-known example is an opinion or exit poll, on the basis of a small fraction of all voters, to infer something of the voting intentions of all potential voters. There are at least two reasons for taking a sample instead of including all units or elements. The cost of including all units will often be prohibitive, and the time needed to do so will often be long. For example, if you want to study the crop specific problem faced by small and marginal farmers, you cannot study each and every farmer because you might not have the financial resources and time to visit each farmer to know about his crop specific problems. In this case, you will have to select perhaps a few hundred. The important issue is, however, that the farmers selected should be representative of the whole group.

Gathering information

Survey research depends heavily on the process of sampling and on asking questions, either through questionnaires, interviews or observations. You do need to acquire skills to do this properly. When conducting any kind of survey to collect information, or when choosing some particular cases to study in detail, the question inevitably arises: how representative is the information collected of the whole population?

3.2 Universe or Population

The two terms universe and population are used interchangeably. A population may be defined as the totality of a particular characteristic for any specified group of individuals or objects. Thus, all the science graduate students of a college, all the medical books in a library, all the workers in a factory, etc., are examples of population.

A population may be finite or infinite. A population is said to be finite if it consists of a limited number of elements so that it can be easily counted, e.g. the population of university students, the population of a village or city, etc. But an infinite population is one whose size is unlimited and therefore, its members cannot be counted. The population of fishes in a river, the number of microbes in the soil etc is examples of infinite population.

Sample

A sample may be defined as a selected number from a population to represent it. Generally, this selection is done according to some rule or plan. By studying the sample, some inferences may be made about the population.

Statistic and parameter

A statistic is a characteristic of a sample, but a parameter is a characteristic of a population. The word parameter is used to indicate various technical measures like mean, median, mode, standard deviation, correlation, etc. in the population or universe. Such statistical measures describe the characteristics of a population. For example, the population mean (g) is a parameter. Whereas, when we work out certain measures such as mean, median, mode or the ones from samples, then they are known as statistics as they are based on the characteristics of a given sample. The sample means (X) is a statistic.

Sampling frame

The element or object to be sampled or a large unit containing the objects is called sampling unit. But a list of other representation of elements in a population from which the sample is selected is called a sampling frame. The elements in a population must be represented in some way. Lists such as membership names, are often used. Other examples include organisation directories, telephone directories etc as a frame for concluding opinion survey in a city.

Sampling error

The basic difficulty which confronts us when we work with samples is that samples are virtually never identical to their parent populations. This difference between a sample and its population is referred to as sampling error. Furthermore, no two samples from the same population will be identical. They will be composed of different individuals, they will have different scores on a test or other measure and they will have different sample means. A sampling error is, therefore, defined as the difference between a parameter value and the sample estimate of a character which occurs when sampling is done from the population.

This is the underlying error involved in sampling, even if it is done scientifically.

Sample errors occur randomly and are equally likely to be in either direction (positive or negative). The magnitude of the sampling error depends upon the nature of the population. The more homogenous the population, the smaller is the sampling error. It is also inversely related to the size of the sample. This means that sampling error decreases when the sample size increases and vice-versa. The standard deviation of a sampling distribution of means is called standard error of mean.

Probability

Probability may be defined as the relative frequency of an event occurring, usually reported as a percentage or fraction. The probability of rolling 5 on dice is 1/6 because there are six possibilities on lines. The probability of a tail on a single toss of coin is ½ or 0.5. Probability of .05 means that an event is likely to occur 5 times out of every 100 times if it takes place in the same situation.

Method of selecting a random sample

For selection of a random sample, the Table of Random Numbers (Fisher and Yates, 1963) may be useful. The tables are reproduced in many other books on research methods and statistics. The table may be entered at any point and read in any direction - to the right, to the left, up the column, down the column, or diagonally using a different starting point and reading plan, a different but equally good set of random numbers may be obtained.

For selection of a sample up to two digits (1-99) the two-digit random number table may be consulted. If we intend to select a set of 10 subjects at random from a population numbered 1 to 98, by going through the first column from top downward, the following set of random numbers shall be obtained. The numbers which are repeated or exceed the size of population are deleted. If we continue, a second set of random numbers shall be obtained.

1st set:	03	97	16	12	55	84	63	33	57	16
2nd set:	26	23	52	37	70	56	31	68	64	27

Irrespective of the size of sample to be selected, for a population 1 to 999 and 1 to 9999, a three-digit and a four-digit random numbers respectively have to be consulted. For

arranging a set of numbers of objects randomly, the same method may be adopted. Nowadays, the random numbers can be generated through a computer.

SELF-ASSESSMENT EXERCISE

i. List reasons for using sample.
ii. Define the following:
- sample
- statistic
- sampling frame
- probability.

3.3 Types of Sampling Methods

Following Blalock (1960), most sampling method can be categorised into two groups:

- probability sampling methods
- non-probability sampling methods.

3.3.1 Probability Sampling Methods

In probability sampling methods, the size of the parent population or universe from which the sample is to be drawn must be known to the researcher. Besides, each element of individual in the population must have an equal chance of being included in subsequent sample. The positive point of probability sampling method is that the obtained samples are considered representative and therefore, the conclusions derived from such sample are worth generalisation and comparable to similar populations to which they belong. Probability sampling provides for calculating the standard error of the distribution. Probability sampling methods are of three types:

- simple random sampling stratified random sampling
- cluster sampling.

Simple random sampling: A simple random sample which is known as unrestricted random sample may be defined as a probability sampling method in which each element in the population has an equal and independent chance of being selected. A random selection of rural primary schools in the local government would have been such that every school in the area had equal chance of being selected. This sampling method is bias-free, thus the sample has a high probability of being representative of the population. The random nature of the sampling method is expected to act as control for all variables. For example, two groups randomly selected from the same population would be expected to have approximately the same average of physical, psychological, social and demographic characteristics.

One of the major disadvantages of simple random sampling is that it does not ensure that the elements which exist in small numbers in the population will be included in the given

sample. For example, in a population of 1000 university students only 10 students possess IQ more than 140 (i.e. genius) and the researcher is to draw only 50 students from 1000, students. If he wants to include few genius students in the sample, chances are very slim that such 'genius' students would be included.

Stratified random sampling: In stratified random sampling, the population is first divided into two or more homogenous subclasses or strata, which may be based upon a single criterion such as sex, yielding two strata - male or female or upon a combination of two or more criteria such as sex and level of education and so on. A simple random sample of the desired number may be taken from each population stratum using the table of random numbers. This stratification tends to increase the precision of the analysis because of the homogenous grouping (sub grouping results in reducing the variance within each subgroup while maximising the variance between groups). Stratified random sampling may be of two types.

i. **Proportionate stratified random sampling:** In this method, the researcher stratifies the population according to known characteristics of the population and subsequently, randomly draws the individuals in a similar proportion from each stratum of the population. For example, there are 1000 students in a university, comprising 600 BSc, 300 MSc and 100 PhD students. Now the researcher wants to draw a sample of 160 students from the three classes in similar proportion as they appear in the population. This is done in the following way:

	Class of Population	**Sample**	
	Number of Proportion in each class	**Number of students selected**	
BSc	600	0.60	160x.60 = 96
MSc	300	0.30	160x.30 = 48
PhD	100	010	160x.10 = 16
Total	**1,000**	**1.00**	**160**

Proportionate stratified random sampling increases the representativeness of the sample drawn from each stratum, because those elements that exist in a few numbers (PhD students in this example) are also included proportionately in the sample. In this method, the sampling error can also be minimised.

i. **Disproportionate stratified random sampling:** In disproportionate stratified random sampling method, the sample drawn from each stratum is not necessarily distributed according to their proportion in the population from which they are randomly selected. In other words, some of the strata of the population may be over represented and some under-represented. For example, out of 1000 university students, 700 students are male and 300 female. If the researcher

wants to draw a sample of 100 students from the set of 1000 and if he draws randomly both the males and females in equal number, say 50 each, it will constitute a disproportionate stratified random sample. The investigator puts equal weight to each stratum in selecting the sample and thereby over represents one stratum (female students) and under represents the other (male students). This method is comparatively less time consuming than proportionate stratified random sampling. But due to over-representation of some strata of the population, some bias in the sample may be introduced. Such type of sample may not be truly representative.

Cluster Sampling: This is a random sampling method in which the sampling units are not individual elements of the universe, but groups of elements or clusters. For example, a researcher wants to study the problems of rural school-going children and wants a sample of 10 percent children from the rural schools in the area. He may select at random 10 percent of the rural schools (clusters) in the area and then use as sample all the children attending those schools. This method has some practical advantage. It is easier to test all children in a few schools than to test a group of children scattered at random throughout, all the schools in the area. The limitation of this method is that it has large sampling error.

3.3.2 Non-Probability Sampling Methods

Non-probability sampling is that sampling procedure in which there is no way of assessing the probability of the elements of population being included in the sample. In this type of sampling, items for the sample are selected deliberately by the researcher instead of using the techniques of random sampling. Some of the important techniques of non-probability sampling methods are – quota sampling, purposive sampling, systematic sampling and double sampling.

Quota sampling: In this method, the elements of the sample are selected until the same proportion of selected characteristics which exist in the population is reached. For example, in a population of 10,000 individuals, 1000 people belong to upper class, 7000 and 2000 people respectively belong to middle and lower classes. If the investigator/researcher wants to select 1000 individuals, and finally selects 100 individuals from upper class, 700 from middle class and 200 from lower class according to his convenience (but not randomly), this constitutes quota sample. The main difference between quota sampling and proportionate stratified random sampling is that in the former, the final selection of individuals is not random, whereas in the latter, the final selection is random. This method is convenient, less costly and can include the individuals from different strata of population. But quota sampling being a non-random, potentially biased sampling method, can lead to large sampling error.

Purposive sampling: It is also non-random sampling method in which the sample is arbitrarily selected because characteristics which they possess are deemed important for research. In purposive sampling, the investigator has some belief that the sample being selected is typical of the population or is a very good representative of the population. This is also known as judgment sampling. For studying attitude of the people towards

the national issues, a sample of journalists, teachers and legislators may be chosen. This is an example of purposive sampling. This method ensures that those individuals will be included in the sample that is relevant to the research design. This is a popular method for student research project.

Systematic sampling: This is a non-random sampling method in which every n^{th} element is chosen from a list of numbered elements. Thus, every element does not have a chance of being drawn once the starting point is selected. The starting point is often chosen randomly and sometimes changed several times during the selection process to improve the chances of representativeness especially in ordered list. Selecting every 7th roll number in a class of 70 students, drawing every 5th name from a telephone directory constitutes examples of systematic sampling. This method is easier, faster and less expensive to carry out particularly with a large population. But it is potentially a biased sampling method. Bias and consequent misleading conclusions are particularly likely if lists are ordered on some large to small character or vice versa, or there is periodicity i.e. sampling a particular day over several weeks. If selecting every 10th name in a list once the starting point is selected every 10th name has a 100 percent probability of being selected, whereas the nine names in between have zero probability of being selected.

Double sampling: This is defined as drawing a small sample of individuals from a bigger sample of them. For example, the researcher wants to study the knowledge of newly married couples towards family planning through mailed questionnaires. For this purpose, he/she mails questionnaire to 1000 couples residing in different localities. He finds that only 50 percent that is 500 questionnaires are returned. From these 500 persons he draws a random sample of 100 and mails another set of questionnaire to get their in-depth knowledge about family planning. This method is known as double sampling. This method has the disadvantage of taking much time and labour of the researcher.

Points to note during sampling

Having selected a suitable sampling method, the remaining problem is to determine the sample size. There is no easy answer to this problem. If the population is very homogenous, and the study is not very detailed, then a small sample will give a fairly representative view of the whole. In other cases, you should consider the following:

1. The greater the accuracy required in the true representation of the population, the larger the sample must be.

1. The size of the sample also should be in direct relationship to the number of questions asked, the amount of details required in the analysis of the data, and the number of controls introduced. It is generally accepted that conclusions reached from the study of a large sample are more convincing than those from a small one. However, the preference for a large sample must be balanced against the practicalities of the research resources, i.e. cost, time and effort.

1. The amount of variability within the population (technically known as the standard deviation) is another important factor in determining a suitable sample size. Obviously, in order that every section of a diverse population is adequately represented, a larger sample will be required than if the population were more homogenous.

3.4 Statistical Tools

There are two meanings to the world 'statistics'. The first is the science of collecting and analysing numerical data, especially in, or for, large quantities, and usually inferring proportions in a whole from proportion in a representative sample. The second refers to any systematic collection or presentation of such facts; for example, population statistics are records of population numbers and make up. Statistical methods deal purely with quantitative data, or with qualitative data which are expressed in numerical terms.

As you well know by now, one of the primary purposes of scientific investigation is to discover relationships among phenomena in order to explain, predict and possibly control their occurrence. It is in the discovery and quantification of these relationships that statistical methods are valuable tool. We are talking here about correlation, rather than casual relationships.

Correlation techniques generally aim to answer three questions about two variables or two sets of data. Does a relationship exist between the two variables or sets of data? If so, what is the direction of the relationship, and what is the magnitude? A wide range of techniques can be used, depending on the nature of the variables being analysed, and they bear exotic names like Kruskal's gamma, Kendall's coefficient of concordance, Guttmann's lambda and Chi-square and KolmogorovSmirnov tests. However, there is no reason to be nervous about this for as Leedy suggested, a simple definition of statistics might be: a language that, through its own special symbols and grammar takes the intangible facts of life and translates them into comprehensible meaning. If statistical tests are to be used to analyse the data, there are usually minimum sample sizes specified from which any significant results can be obtained.

You do not even have to be a mathematician to use this special language, as user-friendly computer packages (such as SPSS) will do all the calculations for you. However, you must be able to understand the function and applicability of the various tests to your own sets of data. In this respect, researchers are advised to follow; the first rule of statistics, which is "always consult a qualified statistician". The second rule is, know enough about statistics to be able to view the advice critically.

There is no need to explain in detail the range of tests and their uses. It will help your understanding, though, is a description of the realm of statistics provided, and the various branches of the disciplined outlined. The diagram below portrays the main classes of statistics and their characteristics, and will serve as a guide to the explanations

which followed.

STATISTICS

Inferential
Statistic

Inference

Predictions

Hypothesis
testing

Estimations

Variation

Correlation

Central Tendency

Parametric

Statistics

Non
-
Parametric Statics

Randomness

Variance

Association

Correlation

Distribution

The major classes of statistics are parametric and non-parametric statistics. An understanding of the meaning of a parameter, which in this context refers to a function of the population, is essential in order to appreciate the difference between these two words. A parameter of a population is a constant feature which it shares with other populations: a common one is the 'bell' curve of the normal frequency distribution. Most populations display a large number of more or less 'average' cases with extreme cases tailing off at each end. For example, most people are of about average height, with those who are extremely tall or short being in a distinct minority. The distribution of people's heights shown on a graph would take the form of the normal or "Gaussian curve shown below. Although values vary from case to case, the generality of this type of curve amongst populations is so strong that statisticians take it as a constant – a basic parameter – on which the calculations of parametric statistics are based. For those cases, where this parameter is absent, non-parametric statistics may be applicable.

3.5 Descriptive Statistics

The two classes of parametric statistics are descriptive and inferential statistics. Descriptive statistics provide a method of quantifying the characteristics of the data, where their centre is, how broadly they speed and how one aspect of the data relates to another aspect of the same data. The centre of gravity of the data, their point of central tendency can be determined by finding the 'mode' or the 'median' and any one of several 'means'. These measures have their own characteristics and applications and should be chosen with regard to the data being analysed.

Frequency count

Mean Score

Mean and mode coincide with the mean

The measure of the dispersion (or spread) of the data, how flat or steep the Gaussian curve appears, is an indication of how many of the data closely resemble the mean. The flatter the curve, the greater is the amount of data that deviate from the mean i.e. the fewer that are close to the average. The horizontal length of the curve also gives an indication of the spread of values and the extent of the extremes represented in the data, while the occurrence of a non-symmetrical curve indicate skewness in the data values.

Apart from examining the qualities of a single set of data, the main purpose of statistical analysis is to identify and quantify relationships between variables. This is the type of research called correlation research. But remember, the mere discovery and measurement of correlation is not sufficient on its own to provide research answers. It is the interpretation of these discoveries which provides the valuable knowledge which will give answers to your research question.

Coefficient of correlation

The technical term for the measure of correlation is the coefficient of correlation. There are many types of these, the 'Pearson's' being the most common. It is possible to measure the correlation between more than two variables if you use the appropriate tests. However, one must be wary about assuming that, because a string statistical correlation between variables can be demonstrated, there is necessarily a casual bond between the variables. It may be purely chance or the influence of other factors that, say, leads to area as of high density development in cities having high crime rates. The researcher must carefully question the premises on which such a casual assertion is made, and review the facts to examine if such casualty is verifiable in other ways.

3.6 Inferential Statistics

Inferential statistics goes beyond describing the characteristics of data and the examination of correlations between variables. As the name implies, it is used to produce predictions through inference, based on the data analysed. What is not so clear from its name is that inferential statistics is also used to test statistically based hypotheses.

The predictive role is limited to estimating population parameters from simple statistics. This is not as abstruse as it might at first seem. Singly stated, this entails making predictions about the qualities of a total population on the basis of the qualities of a sample. This is an exercise which is commonly carried out in quality control production processes, where a sample of the production is tested in order to estimate the qualities of the total production. Three parameters (qualities) are commonly estimated: central tendency (proportion of products which are close to the norm, e.g., within permitted size tolerance; variability (e.g. range of sizes occurring), and probability (e.g. the proportion of acceptable products produced).

As with all predictions made from samples, the representative quality of the sample is crucial to accuracy.

Two types of estimate can be made of population parameters from sample statistics: points estimate and interval estimates. Point estimates attempt to pinpoint the population parameter through the sample statistic value (.e.g. the standard derivation and average size of a sample of manufactured components, giving an estimate of the standard deviation and average size of the whole production run of the same component). While this produces a precise estimate of the parameters of the population, the values are greatly dependent on the truly representative quality of the sample.

Interval estimates of parameters are the sample statistics to predict the band within which almost all (typically 95 percent or 99 percent) of the values will lie. The expected range of the statistical values of the population is established, and if any values fall with a significant frequency (i.e. more than five percent or one percent) outside this range, then this variability will be considered to be caused by something more than mere chance. For example, the range of sizes of a manufactured component is estimated from a sample to be normally within a certain tolerance, e.g. $\pm$ 1mm when the production machines are running correctly. If more than say five percent of the components are found to be beyond these sizes, and then the production is being changed (e.g. a fault in the production machine), obviously, the larger and more representative the sample from which the sample statistics were taken, the more likelihood that the prediction is correct.

3.7 Null Hypothesis

The statistically based hypothesis is commonly referred to as the null hypothesis. Inferential statistics are used to test this type of hypothesis. As a very simple example of the principles involved using the above case of the manufactured component, the engineer who designed the production machines could hypothesise that, according to his design, the components will be manufactured within the size tolerance $\pm$ 1mm. Assuming that the machines are properly built and working correctly, if samples taken of the components produced show that 95 percent of the component fall within the size tolerance (i.e. no significant different is observed between the predicted and the observed parameters: this is where the null comes from), then his hypothesis can be seen to be supported. If more samples are taken with the same result, then more support is received. If, however, the sample shows that more than five percent of the components exceed the size tolerance, then there is a cause for this beyond mere chance. The engineer's hypothesis will be rejected. For, his hypothesis was proved to be wrong either due to incorrect design or faulty calculation.

3.8 Non-Parametric Statistics

Not all data are parametric, i.e. samples and populations sometimes do not behave in the form of a Gaussian curve. Data measured by nominal and ordinal methods will not be organised in curve form. Nominal data tend to be in the dichotomous form of either/or (e.g. this is a cow or a sheep or neither), while ordinal data can be displayed in the form of a set of steps (e.g. the first, second and third positions on a winners' podium). Statistical test built around discovering the means, standard deviations, etc. of the

typical characteristics of a Gaussian curve are clearly inappropriate for analysing this type of data. Non-parametric statistical tests have been advised to recognise the particular characteristics of noncurve data and to take into account these singular characteristics of noncurve data as well as take into account these singular characteristics by specialised methods. In general, this type of test is less sensitive and powerful than parametric tests – it needs larger samples in order to generate the same level of significance.

Tests can be used to compare the qualities of two or more groups or samples (e.g. Wilcoxon matched pairs signed rank test, sign test, Kruskal Wallis test) or to analyse the rankings made by different judges (Kendall coefficient or concordance, Spearman rank order correlation), or to compare the data from observed and theoretical sources (chi-square test). Detailed information about which test to use for particular data sets can be obtained from specialised tests on statistics, and of course, your own expert statistical adviser. This is perhaps a good place to warn you that computer statistical packages (like STSS) will not distinguish between different types of parametric and non-parametric data. In order to avoid producing reams of impressive looking, though meaningless, analytical output, it is up to you to ensure that the tests are appropriate for the type of data you have.

This is a specific method of examining records of all kinds (e.g. radio and TV programmes, films, etc.), document or publications. A checklist is made to count how frequently certain ideas, works, phrases, images or scenes appear in order to be able to draw some conclusions from the frequency of their appearance (e.g. the perception of modern architecture in the media). Care must be taken with selection of the variables to be studied, and the method of their measurement. The data collected can be summarised and collected either on forms or by using statistical tests.

Qualitative analysis

By immersing him/herself in the data and then searching out patterns, surprising phenomena and inconsistencies, the researcher can generate new concepts and theory, or uncover further instances of those already in existences. The clues to new concepts and theory, though indistinct at first, will be strengthened by repetitions of incidents or works, irregularities (e.g. conflicting views offered by different groups of people), and other signs, such as particular emotions displayed when people see things.

SELF-ASSESSMENT EXERCISE

i. Identify the different sampling methods.
ii. Explain briefly:
a) Stratified random sampling
b) Cluster sampling
c) Purposive sampling.
iii. List the points to note during sampling.

4.0 CONCLUSION

A problem confronting the researcher determines which of the sampling methods available and the appropriate one to use. In this unit, the criteria the researcher can use when selecting the appropriate sampling method is discussed.

5.0 SUMMARY

This unit teaches you samplings of various natures. We have been able to look at population, sample, sampling frame, representative sample, sampling error, and non-sampling error.

6.0 TUTOR-MARKED ASSIGNMENT

1. Why do researchers use sampling procedures?
2. Briefly comment on simple random sampling and systematic sampling methods with examples.
3. Why should you not make generalisations about a group by the observation of a simple case? Are there instances when this is done? Give one example.
4. What do you mean by sampling error?
5. What is null hypothesis?

7.0 REFERENCES/FURTHER READING

Gouri, K. B. & Richard, A. J. (1979). *Statistical Concepts and Methods*. John Wiley & Sons.

Karmel, P. H. & Polasck, M. (1977). *Applied Statistics for Economists*. Pitman Publications.

Mahesh, C. & Mehrotra, P. N. (1980). *Practical Workbook in Elementary Statistics*. Macmillan.

Margaret, O. A. (2003). *Research Methodology with Statistics for Health and Social Sciences*. Ilorin: Nathadey Publication.

Ray & Sager, M. (1999). *Research Methods in Social Sciences and Extension Education*.

MODULE 4 PRESENTATION OF RESEARCH FINDINGS IN NARRATIVE, TABULAR AND GRAPHICAL FORMS

UNIT 1 DATA PRESENTATION

CONTENTS

1.0 INTRODUCTION

Data are for large samples where the parametric assumptions cannot be met. Data are either counted or ranked. Statistics is concerned with data collection, organisation, analyses and interpretation of data collected for decision making. Statistics is in a way making sense of variability. While descriptive statistics deals with describing a sample without making any generalisations, inferential statistics is about the confidence with which we can generalise from a sample to the entire population. In this unit, we are going to discuss how the data collected can be presented for possible decision making.

2.0 OBJECTIVES

At the end of this unit, you should be able to:

- handle data for decision making
- present research finding in narrative, tabular and graphical forms
- calculate measures of central tendency and dispersion.

3.0 MAIN CONTENT

3.1 Data Handling

When data have been collected in research, they have to be edited and coded in a numerical form ready to be summarised with tables, charts, diagram or group into frequencies before calculations are made. In this discussion, it is assumed that coding and editing have been done. The collected data are ready to be summarised. Data can be presented in four principal methods namely, tables, diagrams, charts and graph.

3.2 Tabular Presentation of Data

Tabulation is one way of presenting statistical data. A table is a detailed display of numerical information. Data are arranged in horizontal rows or vertical columns in a table.

Rules for tabulation

In constructing a table, the following rules are followed:

- table must be titled
- the title must be precise, concise and descriptive
- columns and rows must be divided by rulings table should not contain too much information
- table must be neat to convey its message.

Tables form the basis for reducing and simplifying the details of a mass of data into a form that is understandable. Tables also facilitate comparison of data.

A statistical table could be a simple one or a complex one. A simple table presents the number of measurements of a single set of items having the characteristics stated at the head of a column or row which forms the basis of the table. However, a complex table presents the numbers or measurement of more than a group of items set out in additional columns or rows and the table is often divided into sections. A simple table, (Table 1), shows scores obtained by 28 agricultural extension students in an agricultural statistics examination. While a complex table, (Table 2), presents array of examination scores obtained by 10 agricultural extension students in five agricultural extension courses.

Table 1: Students Scores in an Agricultural Statistics Examination

No. of Students	**Scores Obtained**
7	53
5	62
6	65
8	71
2	75
Total 28	

Note: Simple frequency distribution can also be called ungrouped frequency distribution

while grouped frequency distribution requires us to group values of the variables and therefore present results as grouped frequency distribution. A frequency distribution shares the frequency with which each value of the variable occurs in a distribution. In essence, we have simple frequency distribution and grouped frequency distribution.

Table 2: Students Scores in Five Agricultural Extension Courses

S/N	Program me Planning	Rural Sociology	Extension Communicati on	Extension Administrati on	Extensio n Research Methods
1.	60	56	62	61	58
2.	54	50	55	52	55
3.	64	63	62	60	61
4.	65	62	68	63	62
5.	58	55	56	54	55
6.	59	56	53	54	51
7.	68	64	66	63	65
8.	60	62	65	62	63
9.	64	63	64	61	68
10.	65	60	62	68	64

3.3 Diagrams and Charts

These are pictorial presentation of data. A chart does not present information as precisely as a table gives a quicker overall impression of findings. A glance at a chart might persuade or dissuade one to read the accompanying report. The most popular types or forms of graphs or chart are:

Histogram

A histogram is a graphical display of a frequency distribution. It looks more or less like a bar chart except that there are no spaces between the “bars” of the histogram. For example, if in a community, there are 38 poultry farmers (egg producers) and a survey was done to obtain their knowledge of some improved poultry production practices and the following scores were obtained:

84	82	75	70	71	65	66	68	59	54	55	57	58
49	50	49	51	44	46	47	48	48	44	39	40	41
41	41	34	34	29	29	24	25	19	20	22	15	12

The array of data could be presented as above. However, to obtain the histogram, first, we need to construct a frequency distribution as illustrated in Table 3.

Table 3: Frequency Distribution of Knowledge Scores of Poultry Farmers on Improved Poultry Practices

Score interval	Tallies	Frequency
84-88	I	1
79-83	I	1
74-78	I	1
69-73	II	2
64-68	III	3
59-63	I	1
54-58	IIII	4
49-53	IIII	4
44-48	IIII I	6
39-43	IIII	4
34-38	II	2
29-33	II	2
24-28	II	2
19-23	III	3
14-18	I	1
9-13	I	1

Sum 38

0
1
2
3
4
5
6
7
8
9

Frequency

9-13

14-18

19-23

24-28

29-33

34-38

39-43

44-48

49-53

54-58

59-63

64-68

69-73

74-78

79-83

84-88

Score interval

Fig. 1: Frequency Distribution

Frequency polygon

This is another method of representing a frequency distribution. In this, a curve is drawn by connecting the mid points of the columns. Then resultant frequency polygon usually shows the shape of a distribution. For example, from Table 3, a frequency polygon can be drawn for the knowledge scores of the 38 farmers.**Farmers**

0

2

4

6

8

10

9-13

14-18

19-23

24-28

29-33

34-38

39-43

44-48

49-53

54-58

59-63

64-68

69-73

74-78

79-83

84-88

Score Interval

Frequency

Fig. 2: Frequency Polygon

Frequency curve

A frequency curve shows the shape of a frequency distribution. It is necessary to always use small class intervals and large observations in order to have narrow and long columns. However, by joining the resultant histogram we will have a curve that describes how the variable is distributed in the population. Using scores on table 3, we can obtain a frequency curve as follows:

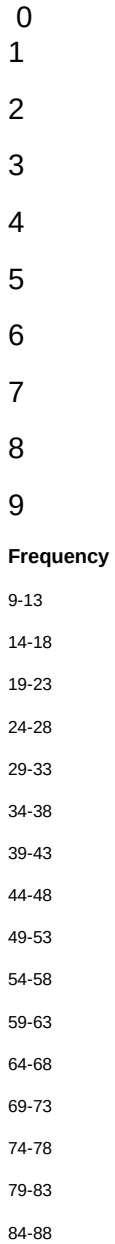

Fig. 3: Frequency Curve of Knowledge Score

Distribution shapes

A distribution is described as skewed if the cases tend to pile up at one end of a curve and with a long tail at the other. Skew is described in terms of the tails of the distribution. Generally, there are three main types of distribution shapes.

1. Positively skewed distribution

When there is piling up of cases on the left of the distribution and a long tail extending to the right, the distribution is said to be positively skewed. In this type of distribution, majority of people have low scores/values while few high scores or values. That is low values have the highest frequencies. This is the type of distribution obtained when difficult tests are given to students.

Fig. 4: Positively Skewed Distribution

2. Negatively skewed distribution

When there is a pile-up of so many cases on the right of the distribution with a long tail extending to the left of the distribution, the distribution is said to be negatively skewed. This kind of distribution is obtained when a teacher gives an easy test to a group of students and most of the students obtain high scores or values on the tests. In this type of distribution high values have the highest frequencies.

Fig. 5: Negatively Skewed Distribution

3. Normal distribution

When most students have similar scores with a few having low and high scores respectively in a test, a normal distribution is the result.

Normal distribution has a bell-shaped curve that is symmetrical about the arithmetic mean. The two ends of the curve do not touch the horizontal axis although they approach it as the values of the variables become extremely small or extremely large. The normal distribution has two parameters: they are the arithmetic mean and standard deviation.

These parameters determine the position and spread of distribution. There is a relationship between the normal curve and its parameters that enables the proportion of a population that lies between any two values of the variables in which we are interested to be found. Tables are available for these proportions.

We use Greek letter μ to denote the arithmetic mean of a population and the Greek letter σ to denote the standard deviation of a population. We use Greek letters to stand for population parameter and ordinary letter σ̄ for sample statistics for example, x.
In a normal distribution, almost all of the population lies between the values of the variable that are four standard deviations (4 σ) on either side of the arithmetic mean (μ) of the distribution. The range μ-4 σ to μ+4 σ spans 0.999 or 99.99 per cent of the population. There will be only a tiny proportion of the population (0.01 percent) outside the range.

As a normal curve is symmetrical, the mean, the median and the mode coincide at the centre of the distribution. Half of the population has a value of the variable above the mean and half lies below the mean.

Other proportions of the area under curve that can be found are:

0.6826 between (μ -1 σ) and (μ+1 σ)

0.9545 between (μ -2 σ) and (μ+2 σ)

0.9973 between (μ +3 σ) and (μ+3 σ)

Fig. 6: Normal Distribution

Cumulative frequency curve (Ogive)
Data can be pictorially represented by the cumulative frequency curve which is also called the Ogive. An Ogive is formed by obtaining the cumulative frequencies and converting them to the proportions of all the cases which is the relative cumulative frequency and then multiplies them by 100.

There are two types of Ogives, namely, “less than” Ogive and “More than” Ogive. To build an Ogive curve, we create additional columns for the score interval and frequency columns for the score interval and frequency columns. These additional columns are the cumulative frequency column (CF), cumulative proportion column (CP) and the cumulative percentage column (C %). This can be illustrated with a set of 50 farmers’ knowledge scores as shown in table 4.

Table 4: Frequency Distribution of Knowledge Score of 50 Farmers

Scores	Frequency	Cumulative Frequency (CCF)	Cumulative Proportion (CCP)	C%
30-39	6	6	6÷50=0.12	12
40-49	9	15	15÷50=0.3	30
50-59	9	24	24÷50=0.48	48

60-69	13	37	37÷50=0.74	74
70-79	8	45	45÷50=0.9	90
80-89	5	50	50÷50=10	100

In plotting and interpreting the "less than" Ogive, we use the upper class limits. The interpretation of Table 4 is that, for example, six farmers have knowledge score less than 39. Fifteen farmers have knowledge score less than 49 while 37 farmers have knowledge score less than 69 and so on.

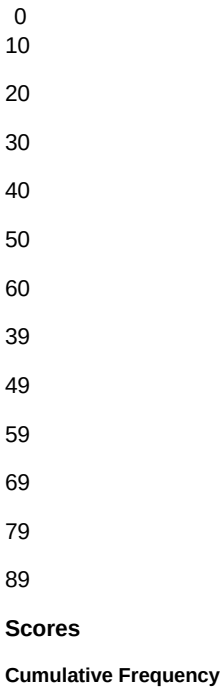

Fig.

7: Distribution of Knowledge Score of 50 Farmers

One advantage of either "less than" or "more than" curve is that we can estimate the number of farmers, who have certain particular level of knowledge.

In plotting "More than" Ogive we use the lower class limit. In Table 5 for example, 50 farmers have 30 percent or more knowledge; 44 farmers have 40 percent or more knowledge. Table 5 will be used to plot "More than" Ogive.

Table 5: **Frequency Distribution of Knowledge Scores of 50 Farmers**

Scores	**Frequency**	**Cumulative Frequency (CF)**	**Cumulative Proportion (CP)**	**C%**
30-39	6	50	50÷50 = 1	100

40-49	9	44	44÷50=0.86	86
50-59	9	35	35÷50=0.70	70
60-69	13	26	26÷50=0.52	52
70-79	8	13	13÷50=0.26	26
80-89	5	5	5÷50=0.1	10

Fig 8:A 'More Than' Ogive

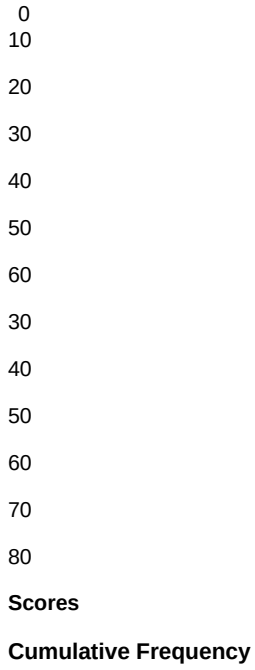

Fig. 8: A "More than" Ogive

Bar Chart

This is appropriate for displaying frequency distribution or other information related to qualitative or attribute data. This is because attribute data have distinct end points and cannot be merged as in the case of histograms. There are different types of bar charts namely – simple, multiple, component and percentage bar charts. Bar charts can be used for comparisons of two or more groups.

Tips for constructing bar

1. Thickness of the width must be regular throughout the construction.
2. Constant gap must be left between bars.
3. Start scale at zero point.
4. Bar charts can be presented in a horizontal form.
5. Description could be placed in bars if too rowdy.

Simple bar chart

A farmer incurred some expenditure on his farm operations; these are expressed as percentage of total expenditure. The data and figure of the expenditure are illustrated in Table 6 and Figure 9 respectively.

Table 6: A Farmer's Expenditure on his Farm Operations

Farm Operations	**Percentages (%)**
Clearing operations	10
Land preparations	15
Inputs	40
Labour	20
Miscellaneous	15
Total	**100**

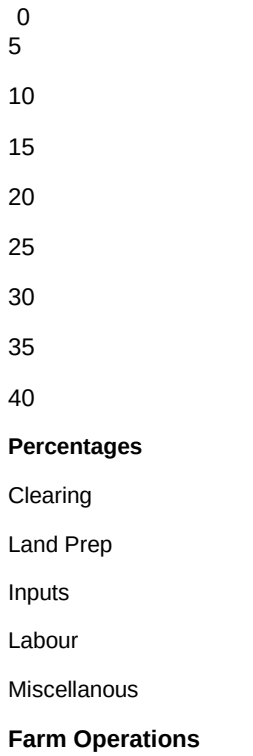

Fig. 9: A Farmer's Expenditure on his Farm Operations

Multiple bar charts

This is used to compare the changes in value of two or more variables. For instance, we might want to compare the earnings of male and female farmers within the same community over a period of time. The array of data is given in Table 7.

Table 7: Farmers Earnings in a Community

Sex	Jan	Feb	March	April	May	June
Male Female	500	550	400	450	500	600
	300	300	350	400	450	500

The array of data may be represented in Figure 10 using a multiple bar chart.

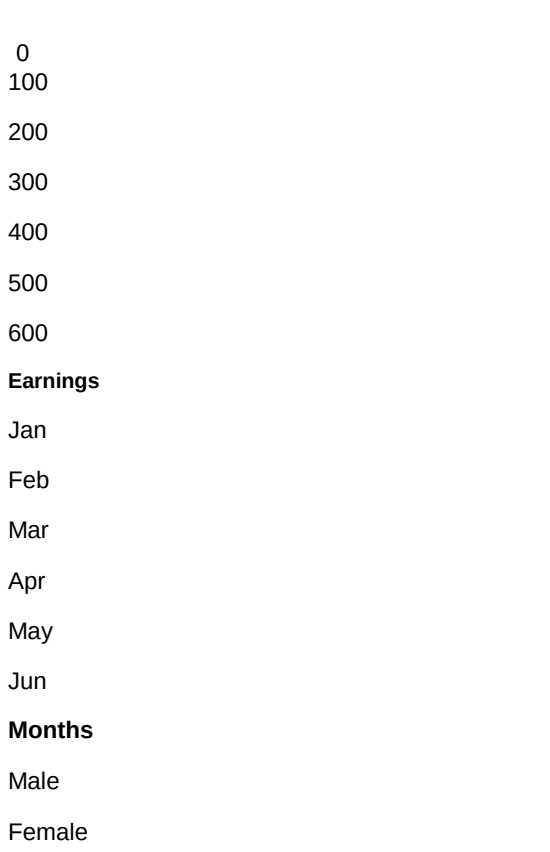

Fig. 10: Farmers Earnings in a Community

Component bar chart

This is used to show the total components of a variable over a period of time. Assuming the data in Table 8 illustrate the distribution of students in each department of a Faculty of Agriculture. Over a period of three years, a component bar chart as in Figure 11 could be constructed.

Table 8: Distribution of Students in a Faculty of Agriculture

Departments	1978	1979	1980
Extension ar Rural Development	20	25	22

Agricultural Economics	30	28	31
Agricultural Biology	10	10	10
Agronomy	13	12	12
Animal Science			
Total	**85**	**85**	**85**

Fig. 11: Distribution of Students in a Faculty of Agriculture

Percentage component bar charts

To obtain a percentage component, bar chart scores are converted to percentages before constructing bars. Each column in the graph constitutes 100 per cent of the time period of that variable. All columns in the graph are of the same height.

Pie chart

A pie chart is useful in representing proportional relationships of between magnitudes when these are more important than the magnitudes themselves. Pie chart is a circular diagram in which each class or attribute is represented by a

sector whose area is proportional to the percentage or relative frequency of the class or attribute. The relative frequency of a class is the ratio of the frequency of that class to the total number of occurrences or observation.

Steps in drawing a pie chart

i. Find the relative frequency or percentage value of the amount in each class. ii. Determine the angle of the sector corresponding to each class using the formula.

Angle of section: $\frac{fi}{\sum fi} \times \frac{360}{1}$

Where fi = Frequency or amount of class.

$\sum fi$ = sum of all frequencies or total amount in all classes.

The ratio $\frac{fi}{\sum fi}$ is called relative frequency

iii. Draw the pie charts.

For instance if there are 85 students in the example above and 20 students in Agric Extension Department, 10 students in Agronomy, 35 students in Agric. Economics, 15 students in Animal Science and 5 students in Agric. Biology, then we can construct a pie chart as in figure 12. This will be based on the following calculations:

Agric. Extension – 20 students = $\frac{20}{85} \times \frac{360}{1}$ = 84.7

Agronomy - 10 students = $\frac{10}{85} \times \frac{360}{1}$ = 42.35

Agric. Economics - 35 students = $\frac{35}{85} \times \frac{360}{1}$ = 148.24

Animal Science - 15 students = $\frac{15}{85} \times \frac{360}{1}$ = 63.53

Agric. Biology - 5 students = $\frac{5}{85} \times \frac{360}{1}$ = 21.18

20
35
5
10

15

Agric Ext

Agric Econs

Agric Biology

Agronomy

Animal Sc.

Fig. 12: A Pie Chart

3.4 Grouped Frequency Distribution

Sometimes due to the nature of the question of our research, we might wish to group values of variable and present results in grouped frequency distribution. If for example we have the knowledge score of improved practices of 50 farmers after a campaign and represented as follows:

45,	50,	70,	60,	30,	35,	50,	40,	80,	50,	35,	60,	40,
80,	70,	70,	55,	45,	80,	65,	40,	75,	50,	45,	70,	60,
60,	50,	40,	70,	60	65,	70,	35,	60,	80,	50,	65,	60,
50,	40,	80,	70,	60,	40,	30,	60,	50,	35,	50.		

For grouped frequency distribution, we can group scores in groups of 10. For instance,

30-39, 40-49, 50-59, 60-69, 70-79, 80-89 and so on.

Table 9: Frequency Distribution of Knowledge Scores of Farmers

Variable Scores	**Tally**	**Number of Farmers Frequency**
30-39	–IIII	6
40-49	I	9
50-59	IIII	9
60-69	IIII	13
70-79	IIII	8
80-89	IIII	5
	IIII IIII III	
	IIII III IIII	
Total	**50**	

In a group frequency distribution, class limits are the smallest and largest values that are included in a class. Hence, for group 40-49, 40 and 49 are the class limits. Class width is the difference between the lower class limit of any interval and the lower class limit of the next class interval. Class limits are also statistically defined as half a measurement unit above the largest observation value in the class and half a measurement unit below the smallest observable value in the class.

SELF-ASSESSMENT EXERCISE

i. List the different charts and diagrams for presenting research reports.
ii. What are the different steps to follow in order to draw the following charts and diagrams?
(a) Pie charts
(b) Histograms (c) Bar charts?

3.5 Measures of Central Tendency

Measures of central tendency or location are measures that help to locate where the data are concentrated, centered, clustered or packed. The most common measures of central tendency or location are mean, the mode and the medium.

The Mode

This is a category of a variable that contains more cases than can be found in other categories. In other words, the mode is the most frequent score. For instance, in a community, the following distribution represents the number of times maize plots were weeded.

Maize plot	**Frequency**
1	3
2	5
3	9
4	7
5	4
6	1
7	2

The most frequent value is 3 as it has a frequency of 9. Thus the mode is 3. Mode could be unimodal or bimodal. For grouped data, the midscore of the interval with the highest frequency is the mode or the interval with the highest frequency i.e. the modal interval. Mode is greatly influenced by skewed distributions.

Median

In a list when values of x can be arranged from highest to lowest score, the score that

falls in the middle is the median score. Median is the value of x such that there are as many score greater than the median as there are scores less than the median. The median is the midpoint of a distribution. That is the point where 50 percent of the score lie below and above. Actually, this is also a centile. A centile is that point below which a certain percentage of the cases lie. In this wise, the median is the fiftieth centile.

To calculate median remember to arrange score in a descending or ascending order before determining the median.

Apply either $\frac{N}{2} + 1$ for odd numbers or

Apply $\frac{N}{2}$ and $\frac{N}{2} + 1$ for even numbers

For instance, the median for odd numbered distribution of 55, 48, (39), 36, 25 is 39.

Median for even numbered distribution of 55, 48, 39, 36, 25, 10

The midpoint is the average of the two middle scores.

$$\frac{39 + 36}{2} = 37.5$$

Median for grouped data

The following ages of farmers were obtained in a farming community: 12, 14, 15, 16, 17, 19, 22, 24, 25, 24, 26, 27, 28, 28, 29, 30, 31, 32, 33, 33, 32, 36, 37, 37, 38, 39, 39, 38, 37, 42, 42, 43, 44, 46, 48, 48, 47, 51, 52, 52, 54, 58, 62, 62, 64, 65, 66, 68.

Find the median age of the farmers.

Procedure

1. Form the farmers' age into classes or groups.
2. Find the frequencies.
3. Find the cumulative frequency (or less than).
4. Apply the formula.

$$Mdn = L + \frac{(\frac{N}{2} - cfb)\, i}{fw}$$

where

Mdn = median
N = Number of cases in the distribution

L = Lower limit of the interval within which the median lies containing the median fw = Frequency of cases within the interval containing the median. i = interval size.

Table 10: Frequency Distribution of Farmers Ages in a Community

Interval	Tallies	Frequency	Cumulative Frequency (CF)
12-16	IIII	4	4
17-21	II	2	6
22-26	IIII	5	11
27-31	IIII I	6	17
32-36	IIII	5	22
37-41	IIII II	7	29
42-46	IIII	5	34
47-51	IIII	4	38
52-56	III	3	41
57-61	I	1	42
62-66	IIII	5	47
67-71	I	1	48

$$Mdu = L + \frac{\frac{N}{2} - cf6}{fw}\, i$$

$$= 32 + \frac{\frac{48}{2} - 17}{5}\, 5$$

$$= 32 + \frac{24-17}{5}\, 5$$

$$= 32 + \frac{7}{5}\, 5$$

$$= 32 + (1.4)\, 5$$

$$= 32 + 7$$

$$= 39$$

Mean

It is also called arithmetic mean

$$\bar{X} = \frac{\sum X}{N}$$

Weighted mean X = $\underline{Na\surd (Xa\surd) + Na\surd (Xa\surd)}$

Where

N_1 = Number of individuals in group 1

N_2 = Number of individuals in group 2

$X_{1 = Mean}$ of Group 1

X_2 = Mean of Group 2

For instance, in a sample of 100 farmers, 60 of the farmers were female while 40 were male and we determined their knowledge level to be 70 percent and 80 per cent male and female respectively, calculate the average knowledge of the farmer.

$$\text{Weighted mean } \bar{X} = \frac{Na\,(Xa) + Nb\,(Xb)}{Na + Nb}$$

$$\bar{X} = \frac{(60 \times 80) + (40 \times 70)}{40 + 60}$$

$$\bar{X} = \frac{4800 + 2800}{100}$$

$$= \frac{7600}{100} = 76$$

Comparison of mean, median, mode

Mean is more affected by skewed distribution, median tests for skewed distribution while mode is greatly influenced by skewed distribution.

Measures of dispersion

Measures of dispersion or measures of variability determine how far apart the observations are. It is possible for two different sets of data to have the same mean but different values of the same measure of dispersion. Measures of dispersion include range, qualities, mean deviation, variance and standard deviation.

Range

The range is the highest score minus lowest score plus one

$$HS - LS + 1$$

Range is affected by the number of scores. It is not true reflection of variability.

Quartiles

These refer to the point on a distribution below which a certain percentage of scores will be found. This is calculated exactly the same way as median.

There are actually 4 quartiles.

Q1 = (First quartile) 75 percent

This is the point below which 25% of the scores of the distribution will be found. This is also referred to as the lower quartile.

Q2 = (Second quartile) 50 percent

This is the point below which 50 percent of the scores will be found in the distribution.

Q3 = (third quartile) 2 percent

This is the point below which 75 percent scores of the distribution will be found. This is also called the upper quartile.

Q4 = (Fourth quartile)

This is the point below which 100 percent of the scores will be found.

The formulas for calculating Q3 (upper quartile) and Q1 (lower quartile) are:

$$Q3 = L + \frac{\left(\frac{3N}{4} - cfb\right)i}{fw}$$

While

$$Q1 = L \times \frac{\left(\frac{N}{4} - cfb\right)i}{fw}$$

The process of calculating them is the same as that for calculating median of a grouped data.

Mean deviation

The mean deviation is the average deviation of a distribution of scores from the mean.

$X_i = X_i - X$

$$MD = \frac{\sum(X1 - X)}{N} = \frac{\sum(Xi)}{N}$$

The mean deviation is obtained by subtracting each score from the mean, summing up all the differences and dividing by N. This arithmetic operation ignore all the signs (+ or -).

Variance

The variance is the squared deviation around the mean. There are two methods of calculating variance: sum of square method and raw score of method.

Sum of square method

In the sum of square method, variance is obtained by subtracting the mean from each score and squaring the difference. The differences are then summed up and it is referred to as sum of square (ss). When we divide the sum of squares by N, we obtain the variance. This sum of square method is illustrated on Table 11

Table 11: Sum of Square Method of Calculating Variance

X	**__ x(X-X)**	__ x^2 $(X-X)^2$
9	+3	+9
8	+2	+4
7	+1	+1
6	0	0
5	-1	+1
4	-2	+4
3	-3	+9
$\sum x>42$		$\sum x^2 = 28$

Variance σ

N 7

Raw score method of variance calculation

We calculate the sum of squares from raw scores as follows:

N

$\sum x$ = sum of raw scores.

$\sum x^2$ = sum of square of raw scores

) = correction term
N

$\sum x^{2=}$ also called the corrected sum of squares

For population

$$\frac{\sum x^2 - \frac{(\sum x)^2}{N}}{N}$$

For sample $S^2 = \frac{\sum x^2 - \frac{(\sum x)^2}{n}}{n-1}$

Table 12: Raw Score Method of Calculating Variance

X	X^2
9	81
8	64
7	49
6	35
5	25
4	16
3	9
$\sum x = 42$	280

$(\sum x)^2 = \underline{1764}$ N= 7

$$\frac{\sum x^2 - \frac{(\sum x)^2}{N}}{N}$$

$$\sigma^2 = \frac{280 - \frac{1764}{7}}{7}$$

$$= \frac{280 - 252}{}$$

$$= \frac{28}{7}$$

$\sigma^2 = 4$

For sample size, it will be

$$S^2 = \frac{\sum x^2 - \frac{(\sum x)^2}{n}}{n-1}$$

$$280 - \frac{252}{6}$$

$= 28/6$

$S^2 = 4.67$

Standard deviation

This is the square root of the variance. While the variance gives values in squared units of measurement, standard deviation gives values in the original units.

$$\text{or } \sigma_2 = -\sqrt{\sigma_2}$$

$$\frac{N}{n-1}$$

$$^2 \qquad _2 \text{ or } \sigma = \sqrt{\sigma}$$

Computational formula is

$$\frac{N}{N}$$

Or

For sample

$$\frac{n}{n-1}$$

From our earlier data used in calculating variance

$\sigma = \sqrt{4} = 2$

Or

$S = \sqrt{4.67} = 2.16$

Standard deviation allows us to describe the relative position of any observation in a distribution. Once we know the mean and standard deviation, we can say exactly the number of observations that are 1, 2, or 3 standard deviation above/below the mean.

Another important use of standard deviation is in the use of linear transformation of raw scores to standard "Z" scores. However, Z scores have 2 great limitations: having negative signs and decimals.

SELF-ASSESSMENT EXERCISE

i. Explain the following: central tendency, standard deviation, and variance.
ii. Express the calculation in formulas.

4.0 CONCLUSION

In data presentation, statistics is concerned with collecting organising, analysing and interpreting data collected for decision making. Statistics is in a way making sense of variability. While descriptive statistics deals with describing a sample without making any generalisation, inferential statistics is about the confidence with which we can generalise from a sample to the entire population.

5.0 SUMMARY

In this unit, we have learnt how the data collected can be presented for decision making. We were able to discuss among other things:

- rules for tabulation
- diagrams and charts
- skewed and normal distribution
- measures of central tendency measures of dispersion and
- grouped frequency distribution.

6.0 TUTOR-MARKED ASSIGNMENT

1. Describe the following:
a. Tabulation
b. Diagrams and charts
c. Normal distribution
d. Grouped frequency distribution.
2. Draw the pie charts to describe the following data:
a. Administration Department = 50 staff

b. Finance and Accounts Department = 35 staff
c. Rural Development and Gender Issues Department = 12 staff
d. Agricultural Development Management Department = 4.

7.0 REFERENCES/FURTHER READING

Andy, I. J. (1992). *Fundamental Statistics for Education and Behaviour Sciences.* Ibadan: Kraft Books Limited.

Best, J. W. (1981). *Research in Education*. New Jersey: Englewood Cliff Prentice Hall.

Kerlinger, N. F. (1973). *Foundations of Behaviour Research*. New York: Holf Rinehart and Winston Inc.

Sirkin, R. M. (1995). *Statistics for the Social Sciences*. London: International Educational and Professional Publishers Sage Publication.

UNIT 2 REPORT WRITING

CONTENTS

1.0 INTRODUCTION

One of the most difficult aspects in research is preparing a report, which is most often delayed for a variety of reasons. A research may have the most brilliant hypothesis, the most carefully designed and conducted study, the most striking findings but unless these are put together in a meaningful way and communicated to others, it will not have much relevance. For the fact that writing the report could be difficult, sometimes scares people.

Report writing is a process of analysing, understanding, appreciating and consolidating the findings and project a meaningful and coherent view of the phenomenon studied. The sole purpose of scientific writing is to inform, hence it should be accurate and truthful without exaggerations. The report should be factual rather than expressive of opinion, written in a passive rarely containing humour. Since the study is completed when the writing occurs, it is usually written in the past tense although the present tense may be used for those statements of continuing and general applicability. The purpose of the report is not communication with oneself but communication with the audience. The author makes an attempt to clarify his own thoughts in order to make it meaningful to the reader and for this purpose, he may have to clearly be aware of what the audience want or need to know about the study and how the information can be best presented.

Different procedures have been used to prepare research report depending on the type of audience. It may start with the statement of the problem, presentation of the research methodology, results, discussion, conclusions and implications.

There are some basic qualities of good scientific writing, which include accuracy and clarity. The first question that comes to the writer of the report is what information to convey and how the various points are related to one another. The requirement for this is to write a detailed outline of the proposed report and to go over it again to discover any omissions.

It would be desirable to have someone to go through the outline and make comments regarding the coverage, sequencing, the points of omission, etc. There are different outlines used for presentation of research report, a sample of which is presented later. This may also vary with the nature of the study and the particular discipline in which the report has been prepared.

As for the writing style, one should be guided by the objectives of the writing, the sole purpose of which is to convey information rather than to achieve a literary production.

2.0 OBJECTIVES

At the end of this unit, you should be able to:

- appraise the method of preparing a research report
- identify the target audience – sponsor of research, administrator or policy makers, academic community and to prepare the report accordingly

- provide a general framework for preparing a research report.

3.0 MAIN CONTENT 3.1 Target Audience

The form of a research report will depend upon the target audience. The research report can be prepared for the academic community, the sponsor of a research or the general public.

Reports prepared for academic community may take the form of a thesis or dissertation leading to some academic degree or it may take the form of a monograph incorporating in a detailed manner, the entire process of research or it may take the form of a research article dealing with different aspects of the study published independently in different journals. Here, the main objective is to indicate the scope, the methodological variations made in undertaking the study rather than the findings per se.

In the research reports prepared for the sponsors of a research, they become the target audience. The needs of the sponsor may be different as they may want to understand problem or solve a problem. The purpose of the research report is to highlight the various aspects of the problem studied based on the objectives in terms of fulfilling the requirements of the sponsor. It gives answers to the research questions proposed by them in an effort to solve or understand that problem.

As for the research reports published for the consumption of the general public, it may take the form of summary reports, articles or brochures. The main concern would be to know the salient features of the findings without bothering about technicalities. Based on the target audience, research reports can be classified into:

- comprehensive research report research articles
- summary report.

Each of these may take the form in which it meets the requirements of the target audience. However, it is assumed that attempting to write a multi-purpose report to serve all audience becomes difficult because what each one looks for in the report will be different. It is assumed that the most useful approach is to choose and write for a single target audience at a level appropriate to it.

The form, content and style of the research report should be chosen to suit the level of knowledge, experience and interest of the target audience as well as anticipate the likely uses for the research findings.

3.2 Suggestions for Preparing Reports

Comprehensive research report

The report should be an effort to narrate the total research process and experience. The

experience of the researcher should be documented in selective and organised way.

The objective of the research report is to communicate with the target audience and hence the level or knowledge, understanding of the audience should be kept in mind.

The major thrust of the report should be to communicate what actually happened throughout the research process and not simply what we hoped would happen at each stage.

Research experience and findings that may not seem to have any relation to the immediate objectives of the study should not be eliminated in a hurry. What may appear to be irrelevant information may turn out to be important to understand the difficult situation.

The report should not be merely selective – narration of our successes but may also contain the limitations along with the reasons for it.

It is more efficient to first prepare an adequate outline and then to follow it by drafting a report.

Since most readers may not go through the entire report, it is necessary to organise the chapters and sections effectively with meaningful titles so that those interested may read the concerned sections as required.

Research articles

It deals with different facets of a research problem in a more detailed fashion, emphasising the methodology adopted in the study. Also, the conclusions are to be related to the specific objectives.

Summary reports

These are generally meant for the general public. Since summary reports may receive wide distribution, they are frequently written in less technical language and may make greater use of better methods of presenting the data. It focuses on major findings, their elucidation in simple and clear language and categorical statements about the implication of the study and recommendations.

3.3 Reports for Administrators and Policy Makers

We may have to prepare reports for administrators and policy makers which may have implication for their activities.

If the research is sponsored by an organisation, they may ask the researcher to diagnose the situation, evaluate the action programme or advise them about alternative programmes to solve some particular problem.

However, in these reports, the technical details concerning design and execution of research are of subsidiary interest to the administrators or policy makers as they are primarily concerned about the diagnosis of the problem and recommendations. In case the research is aimed at evaluation of on-going action programme, the report should contain:

- an assessment or appraisal of the quality and quantity of activity or effort
- describe performance (results or efforts)
- evaluate adequacy of performance with regard to the total need
- report efficiency (in terms of cost, time and personnel)
- specify how and why the programme did not work or worked or what attributes of the programme made it more or less successful; which recipients are more or less affected, under what conditions the programme was more or less successful and the nature of the effects of the programme (unitary or multiple effects and cognitive, attitudinal and behavioural effects).

The report may be divided into two parts. In the first part, a brief introduction containing the background of the problem, the broad aspects of the programme, the terms of reference or objectives of the study may be given. It may be followed by a summary of the findings and suggestion and recommendations (sometimes this part is referred to as executive summary).

In part II, it may contain a detailed description of the programme, sources of data, procedures followed, statistical analysis of the data, discussion, conclusions and recommendations. This may be followed by selected references and appendices of technical matters. Language used for preparing the report for the administrators and policy makers should be carefully chosen and observations made or suggestions offered should have the findings as the base. The understanding of the practical problems of the implementing agencies should also be borne in mind while making the recommendations. By and large, the evaluation report should aid in understanding the problem areas, method of overcoming them, and contribute to the constructive development of the programme.

Finally, the broad sequence of contents of a comprehensive research report should be as follows:

i. **Title**: The title of the report should be clear, concise and indicative of its contents. In other words, it must state the problem adequately and in clear terms indicating the relationships of the variables studies.
ii. **Foreword:** Generally, a foreword of about a page indicating the content of the study, brief statement of the problem, objectives, agencies sponsoring the study etc to be written by a person who conducted the research or by the head of the research organisation or someone of importance associated with the type of research that was done.
iii. **Contents**: The purpose of a contents page is to enable the readers to identify the parts of the report and to see the relationship of the parts to one another and

indicate the title of each chapter and section, sub-section along with the page number.

iv. **Introduction**: It should contain information about the nature of the research problem, its importance in the larger context, the scope and the type of attempts made to answer the research question. It may also contain a brief narration of the way in which the material of the report is organised.

Review of previous studies

It is also necessary to provide a brief review of the existing information in relation to the research study. This may involve an analysis of the various findings related to the type of research in order to bring home the existing knowledge in the area and to indicate the knowledge gap leading to the formulation of the research hypothesis, for the present study. Sometimes, this can also form part of the introduction whenever information about previous findings is limited in scope. A separate chapter or section may also be given delineating the various aspects on which previous studies have been conducted and the conclusions arrived at, in order to put the research problem under study in the proper perspective.

Methodology

The methodology chapter may contain information about the objectives of the study in clear and precise terms or the hypothesis formulated, the method of formulating the hypothesis etc, followed by the brief description of the study area, the type of respondents covered (scripting), the nature of sample and the method of selecting the sample could be given in detail. It should contain information about the type of measuring tools used, the method of developing it, how the variables were identified, the decision about the nature of measuring instruments or the method followed to collect data, the type of secondary data collected along with their source and justification for using any particular method may be given in a more detailed form.

The method of collecting the data, the type of problems experienced and the limitations of the data may also be indicated in a detailed form to provide the context under which the data were conducted.

Following this, the method of treatment of the data, the procedure and statistical techniques adopted for treatment of the data, both qualitative and quantitative should be indicated. The need for using different statistical techniques and the type of inferences that are expected to be drawn should be emphasised in relation to the research question or hypothesis as the case may be.

Discussion and interpretation

While preparing the data, proper tabulations must be made giving information about the type of data contained in each table or the type of measures utilised, the way in which the table has to be read should be implied in the title and explanation of the figures contained may be given as footnotes.

The first table of the research is to draw inferences from the table to indicate the trends of the data and relate these trends to the research hypothesis or objectives as the case may be or to the findings of previous studies. The interpretation of the data can be based on one's general knowledge of the area, previous findings or the social context in which the study was done. In case the interpretation is repetitive, then it is necessary to combine a series of tables of results in order to derive the type of findings and then attempt an interpretation in the larger content.

One can use graph to predict the trends of the results as it would explicitly bring home the type of information more effectively than the table of content. The use of table, the statistical techniques relevant to the analysis of the data and for drawing inferences and relating to other larger findings should constitute a part of the discussion of the findings of the study.

However, three major aspects should be taken into account, viz.

1. A statement of the inferences drawn from the findings of a particular research may be expected to apply in similar situations
2. As a qualification of these inferences, the researcher should indicate the condition of his study that limit the extent of legitimate generalisation. It is also necessary to make an observation on a typical results obtained and related to the methodological contribution if any.
3. The discussion of the implications should also include relevant questions that are still unanswered or new questions that are raised.

SELF-ASSESSMENT EXERCISE

1. Why do you write reports?
2. List four suggestions for comparing comprehensive research reports.

4.0 CONCLUSION

The method of organising the research findings and experience in a coherent manner constitutes the research report. While writing the report, the researcher should keep in mind the target audience: academic community, administrators and policy makers or the general public. The approach followed to write the report depends on: for whom it is intended and for what purpose. The comprehensive research report is generally intended for academic purposes of obtaining a university degree. Care needs to be taken in the writing style, and language used as the purpose is to inform others based on one's research findings rather than show one's prowess in vocabulary.

5.0 SUMMARY

In this unit, we have learnt how to write the report of a research. We have been able to look at the different audience and their needs in terms of the structure and detail

expected in a research report. Suggestions were also offered for preparing comprehensive research report.

6.0 TUTOR-MARKED ASSIGNMENT

1. What is research report?
2. Who are the target audience?
3. What steps do you follow to prepare a research report for administrators and policy makers?
4. Recall the general format for the preparation of comprehensive research report.

7.0 REFERENCES/FURTHER READING

Helmstadder, G. C. (1970). *Research Concepts in Human Behaviour: Education, Psychology, Sociology*. New York: Appletorn Century Crofts.

Seltiz, C. (1965). *Research Methods in Social Relation*. London: Methuen & Co.

Shah, V. P. (1977). *Reporting Research*. New York: Agricultural Development Council.