NATIONAL OPEN UNIVERSITY OF NIGERIA
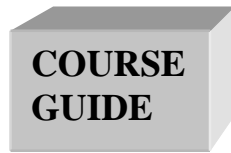
SCHOOL OF ARTS AND SOCIAL SCIENCES

COURSE CODE: POL 316

COURSE TITLE: POLITICAL EVALUATION

COURSE GUIDE

POL 316

POLITICAL EVALUATION

Course Writer: **Dr. Emmanuel Remi Aiyede**
Dept. of Political Science
University of Ibadan
Ibadan

Course Coordinator: **Abdul-Rahoof Adebayo Bello**
School of Arts & Social Sciences
National Open University of Nigeria

Course Editor: **Dr. Fatai Aremu**
Department of Political Science,
University of Ilorin

Programme Leader: **Prof. Remi Anifowose**
School of Arts & Social Sciences
National Open University of Nigeria

NATIONAL OPEN UNIVERSITY OF NIGERIA
Headquarters
14/16 Ahmadu Bello Way
Victoria Island
Headquarters
Lagos.

1.1   Introduction

Welcome to **POL 316: Political Evaluation. POL 316 is a three-credit unit course.** It introduces you to the basic concepts, perspectives and processes of political evaluation. It is designed to provide an introduction to the major concept, steps and processes in conducting political evaluation research.

1.2   What you will learn in this course

You will explore the concept, processes, and activities involved in the conduct of political evaluation and analysis. Important concepts in these processes discussed include: meaning and nature of Political Evaluation and Research; basic concepts in political research; measurement, variables, concepts etc; meaning of survey research; processes involved in survey research; techniques of data gathering; sample and sampling techniques; frequency distribution; calculation of mean, mode and median; variance and standard deviation and basic inferential statistics.

As a course in political research methods it deals with the various elements of political evaluation as well as the steps and instruments used in carrying out political evaluation or assessments. These usually involve understanding the key objects, subjects or issues that is being evaluated. It also involves breaking down of these concepts into measurable variables. This may be achieved through survey research which is used to determine how citizens evaluate political actors and institutions. You will examine the processes and activities involved in survey research such as sampling, data collection and analysis as well as the use of statistical techniques to analyze data and generate inferences.

This enables the student to understand political evaluation or assessment as a part of the larger process of doing political analysis or an analysis of political behaviour and institutions. The course will provide an overview of political research methods and techniques. In the world in which we live today, assessing governance and evaluating political systems are important because they enable us determine the performance of governments. Indeed, there are many organisations involved in doing political evaluation. Some of the notable ones with particular focus on Africa include Africa Peer Review Mechanism of the African Union, Governance Report of the United Nations Commission for Africa (UNECA), Mo Ibrahim Index of Governance and the Afro barometer surveys. At the global we have various data base on indices of political evaluation. Examples include the African Peer Review Mechanism (African Union), Governance Report (United Nations Economic Commission for Africa), Mo Ibrahim Index of Governance, Afro barometer Surveys, Failed states Index (Foreign Policy Magazine) and state fragility Index, Polity iv and freedom House Reports..

This course will be valuable for all who are interested in political development in their countries and particularly for those leaders or anyone planning a career in public or private organizations, nongovernmental agencies and community development organizations.

1.3   Aims

At the end of the course the student is expected to be conversant with the various concepts used in political evaluation. The student is also expected to have internalized the practical processes involved in political evaluation. The student is also expected to have acquired the capacity to conduct independent evaluation of political actors, institutions and regimes.

1.4   Objectives

The course seeks to enable the student to

1. Understand the meaning and nature of Political Evaluation, especially with basic concepts in political research; measurement, variables, and survey research application of basic and inferential statistics;

2. Demonstrate competency in carrying out political evaluation research

3. Become familiar with the techniques of data gathering, measurement and basic inferential statistics.

4. Recognize situations in which specific methods can be applied quickly and appropriately

5. Communicate the results of political evaluation research.

1.5   Working through the course

This course guide is written in modules (sections) and units. There are end of unit reviews and examination questions at the end of each unit or module. You are expected to answer these self-review questions unaided. Additional requirements for the course will include an appraisal paper and a final examination. The course guide tells you briefly what the course is all about, what you are expected to know in each unit, what course materials you need to use and how you can work your way through these materials.

1.6   Study Units

There are twenty-one study units in this course spread through four modules. These are as follows:

MODULE I: INTRODUCTION TO POLITICAL EVALUATION

Unit 1:　　　　Political Evaluation and its Uses

Unit 2:　　　　Political Evaluation Research

Unit 3:　　　　Political Evaluation Research Strategies

Unit 4:　　　　Types of Evaluation Research

Unit 5:　　　　Statistics and Political Evaluation

MODULE 2:　　MEASUREMENT

Unit 1:　　　　The Concept of Measurement

Unit 2:　　　　Levels of Measurement

Unit 3:　　　　Scales of Measurement

Unit 4:　　　　Reliability and Validity in Quantitative Research

Unit 5:　　　　Reliability and Validity in Qualitative Research

MODULE 3:　　 THE SURVEY PROCESS

Unit 1:　　　　The Survey

Unit 2:　　　　The Survey Instrument: Questionnaire

Unit 3:　　　　Data　and Data Collection Techniques

Unit 4:　　　　Population and Sample

Unit 5:　　　　Non-probability Sampling

Unit 6:　　　　Probability Sampling

MODULE 4: DESCRIPTIVE AND INFERENTIAL STATISTICS

Unit 1:　　　Variables and their Operationalization

Unit 2:　　　Classification of Variables

Unit 3:　　　Measures of Central Tendency: Mode, Median, and Mean

Unit 4:   Measures of Dispersion: Variance and Standard Deviation

Unit 5:   Inferential Statistics: Estimation and Hypothesis Testing

As noted earlier, each unit contains a number of self- assessment exercises (SAE). These self-assessment exercises are designed to test you on the materials you have just covered. They will help you to evaluate your progress as well as reinforce your understanding of the material. Together with tutor-marked assignments, these exercises will assist you in achieving the stated learning objectives of the individual units and of the course.

1.7   Set Textbooks

The following books are suggested for further reading:

Babbie, E. (2010) *The Practice of Social Research* (12th ed). Belmont: Wadsworth Pub. Co.

Baker, Therese 1999. *Doing Social Research*. New York: McGraw-Hill College.

Black, T (1999) *Doing Quantitative Research in Social Sciences: An integrated approach to research design, measurement and statistics*, London: Sage.

Bless, Claire; Craig Higson-Smith and Ashraf Kagee (2006). *Fundamentals of Social Research Methods: An African Perspective* (4th edition). Cape Town: Juta & Co. Ltd.

Bryman, Alan (2004) (2nd Edition) *Social Research Methods*, Oxford: Oxford University Press

Cosby, Paul C. 2006. *Methods in Behavioural Research*. New York: McDraw-Hill

Creswell, J (2003) *Research Design: Qualitative, Quantitative, and Mixed Methods Approaches*, (2nd ed.), Thousand Oaks, CA: Sage Publications.

Jones, E. Terrence 1971. *Conducting Political Research*. New York: Harper and Row Publishers

King, Gary; Robert Keohane, and Sidney Verba. 1994. *Designing Social Inquiry* Princeton: Princeton University Press.

McNabb, David E. 2010. *Research Methods for Political Science: Quantitative and Qualitative Methods*. 2nd edition. New York: M.E. Sharpe , Inc.

Marsh, David and Gerry Stoker (eds) 2010. *Theory and Methods in Political Science*. New York: Pelgrave Macmillan.

Meier, Kenneth J.; Jeffrey L. Brudney and John Bohte.2011. *Applied Statistics for Public and Nonprofit Administration* (6th edition). Belmont, CA: Thomson Higher Education.

Mitchell, Mark L. and Janina M. Jolley 2007. *Research Design Explained* (6th edition). Belmont, CA: Thomson Higher Education.

Patton, M. Q. 1990. *Qualitative evaluation and research methods*. (2nd ed.). Thousand Oaks, CA: Sage.

Taylor, Steven J. and Robert Bodgan 2006. *Introduction to Qualitative Research Methods: A Guide Book and Resource*. New York: John Riley and Sons Inc.

1.8   Assignment File

In your assignment file, you will find all the details of the work you must submit to your tutor for marking. The marks you will obtain for these assignments will count towards the final mark you obtain for this course. Further information on assignments will be found in the Assignment File itself, and later in this Course Guide in the section on assessment.

There are many assignments for this course, with each unit having at least one assignment. These assignments are basically meant to assist you to understand the course.

1.9   Assessment

There are two aspects to the assessment of this course. First, are the Tutor-Marked Assignments; second, is a written examination.

In tackling these assignments, you are expected to apply the information, knowledge and experience acquired during the course. The assignments must be submitted to your tutor for formal assessment in accordance with the deadlines stated in the Assignment File. The work you submit to your tutor for assessment will account for 40 per cent of your total course mark.

At the end of the course, you will need to sit for a final examination of three hours duration. This examination will account for the other 50 per cent of your total course mark.

1.10   Tutor-Marked Assignments (TMAs)

There are 21 tutor-marked assignments in this course. The best four (that is, the highest four of the 20 marks) will be counted. Each assignment counts for 20 marks but on the average when the four assignments are put together, each assignment will count 10% towards your total course mark. This implies that the total marks for the best four (4) assignments, will constitute 40% of your total course mark.

The Assignments for the units in this course are contained in the Assignment File. You will be able to complete your assignments from the information

and materials contained in your set books, readings and study units. However, it is always desirable at this level of your education to research more widely, and demonstrate that you have a very broad and in-dept knowledge of the subject matter.

When each assignment is completed, send it together with a TMA (tutor-marked assignment) form to your tutor. Ensure that each assignment reaches your tutor on or before the deadline given in the Assignment File. If, for any reason you cannot complete your work on time, contact your tutor before the assignment is due to discuss the possibility of an extension. Extensions will not be granted after the due date unless there are exceptional circumstances warranting such.

1.11   Final Examination and Grading

The final examination for to **POL 316: Political Evaluation** will be of three hours' duration and have a value of 60% of the total course grade. The examination will consist of questions, which reflect the practice exercises and tutor-marked assignments you have previously encountered. All areas of the course will be assessed.

Use the time between the completion of the last unit and sitting for the examination, to revise the entire course. You may find it useful to review your tutor-marked assignments and comment on them before the examination. The final examination covers information from all aspects of the course.

1.12 Course Marking Scheme

Table 1: *Course marking Scheme*

ASSESSMENT	MARKS
Assignments	Best four marks of the Assignments @10% each (on the average) = 40% of course marks
Final examination	60% of overall course marks
Total	100% of course marks

1.13 How to get the most from this Course

1. In distance learning, the study units replace the university lecture. This is one of the advantages of distance learning; you can read and work through specially designed study materials at your own pace, and at a time and place that suits you best. Think of it as reading the lecture instead of listening to the lecturer. In the same way a lecturer might give you some readings to do, the study units tell you when to read, and which are your text materials or set books. You are provided exercises to do at appropriate points, just as a lecturer might give you an in-class exercise.

2. Each of the study units follows a common format. The first item is an introduction to the subject matter of the unit, and how a particular unit is integrated with the other units and the course as a whole. Next to this is a set of learning objectives. These objectives let you know what you should be able to do by the time you have completed the unit. These learning objectives are meant to guide your study. The moment a unit is finished, you must go back and check whether you have achieved the objectives. If this is

made a habit, then you will significantly improve your chances of passing the course.

3.   The main body of the unit guides you through the required reading from other sources. This will usually be either from your set books or from a reading section.

4.   The following is a practical strategy for working through the course. If you run into any trouble, telephone your tutor or post the question on the web GS OLE's discussing board. Remember that your tutor's job is to help you. When you need assistance, do not hesitate to call and ask your tutor to provide it.

5.   Read this Course Guide thoroughly, it is your first assignment.

6.   Organise a Study Schedule- Design a 'Course Overview' to guide you through the Course. Note the time you are expected to spend on each unit and how the assignments relate to the units. Important information, e.g. details of your tutorials, and the date of the first day of the semester is available from the WebGS OLE. You need to gather all the information into one place, such as your diary or a wall calendar. Whatever method you choose to use, you should decide on and write in your own dates and schedule of work for each unit.

7.   Once you have created your own study schedule, do everything to stay faithful to it. The major reason that students fail is that they get behind with their course work. If you get into difficulties with your schedule, please, let your tutor know before it is too late for help.

8. Turn to Unit 1, and read the introduction and the objectives for the unit

9. Assemble the study materials. You will need your set books and the unit you are studying at any point in time.

10. Work through the unit. As you work through the unit, you will know what sources to consult for further information.

11. Keep an eye on the Website Up-to-date course information will be continuously posted there.

12. Well before the relevant due dates (about 4 weeks before due dates), access the Assignment File on the Website and download your next required assignment. Keep in mind that you will learn a lot by doing the assignment carefully. They have been designed to help you meet the objectives of the course and, therefore, will help you pass the examination. Submit all assignments not later than the due date.

13. Review the objectives for each study unit to confirm that you have achieved them. If you feel unsure about any of the objectives, review the study materials or consult your tutor. When you are confident that you have achieved a unit's objectives, you can start on the next unit. Proceed unit by unit through the course and try to pace your study so that you keep yourself on schedule.

14. When you have submitted an assignment to your tutor for marking, do not wait for its return before starting on the next unit. Keep to your schedule. When the Assignment is returned, pay particular attention to

your tutor's comments, both on the tutor-marked assignment form and also the written comments on the ordinary assignments.

15. After completing the last unit, review the course and prepare yourself for the final examination. Check that you have achieved the unit objectives (listed at the beginning of each unit) and the course objectives (listed in the Course Guide).

16. Tutors and Tutorials

There are 15 hours of tutorials provided in support of this course. However, tutorials are non - compulsory, demand driven and are meant to be problem-solving sessions. You will need to contact your tutor for more information on these tutorials, together with the name and phone number of your tutor.

Your tutor will mark and comment on your assignments, keep a close watch on your progress and on any difficulties you might encounter and provide assistance to you during the course. You must mail your tutor-marked assignments to your tutor well before the due 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 hesitate to contact your tutor by telephone, e-mail, or discussion board. The following might be circumstances in which you will find help necessary. Contact your tutor if -

- You do not understand any part of the study units or the assigned readings.
- You have difficulties with the exercises.

- You have a question or problem with an assignment, with your tutor's comments on an assignment or with the grading of an assignment.

- You should try your best to attend the tutorials if you have problems. This is the only chance to have face-to-face contact with your tutor and ask questions which are answered instantly. You can raise any problem encountered in the course of your study. To gain the maximum benefits from course tutorials, prepare a question list before attending them. You will learn quite a lot from participating in the discussions.

**MAIN
CONTENT**

POL 316: POLITICAL EVALUATION

MODULE I: INTRODUCTION TO POLITICAL EVALUATION

Unit 1: Political Evaluation and Its Uses

Unit 2: Political Evaluation Research

Unit 3: Political Evaluation Research Strategies

Unit4: Types of Political Evaluation Research

Unit 5: Statistics and Political Evaluation

UNIT 1:   POLITICAL EVALUATION AND ITS USES

Table of Content

1.0   Introduction

2.0   Objectives:

3.0   Main Content

      3.1   What is Political Evaluation?

      3.2   The Goals of Political Evaluation

4.0   Summary

5.0   Conclusion

6.0   Tutor Marked Assignments

7.0   References/Further Reading

1.0   INTRODUCTION

In this unit, we will define political evaluation and explore it as an important form of political analysis that is used to assess various aspects of governance and political life. We will also deal with the use purposes and uses of political evaluation.

2.0   OBJECTIVES

In this unit, you are expected to become

1. familiar with the meaning and various uses of political evaluation.

2. appreciate the importance of evaluating human behaviour as they relate to government, and

3. understand political valuation as a tool of decision making not only for government officials and politicians, but also for business as communities and individuals who are affected by government decisions.

3.0   MAIN CONTENT

The primary focus of this unit is political evaluation as an important form of political analysis. It will provide explanation of the activities that constitute political evaluation. It will also explain political evaluation as an aid to decision- making.

3.1   What is Political Evaluation?

Political evaluation is the analysis of social activity and behaviour relating to systems of political governance. The essence of political evaluation is to help strengthen governance. In order to strengthen governance, it is necessary to evaluate the status of governance as a basis for intervention. Such an evaluation has to provide objective, meaningful and comparative judgments on a nation-state's governance effectiveness and quality. This

assumes that governance is measurable. A series of studies have been provided that measures for various aspects of governance. These include measures of corruption, competitiveness, political risk, independence of the judiciary, human rights, political inclusiveness, bureaucratic effectiveness and general measures of human development.

There are many evaluation research reports that provide information on the status of governance across the world. Some of these reports include a ranking of countries according to their performances in various indicators of governance. Many government leaders, international organizations, investors and donors make use of these rankings in decision-making.

Political evaluation focuses on the various political goods that citizens expect their governments to provide for them. Political goods could be tangible or intangible. Tangible goods include schools, clinics, roads. Some of these goods are intangible. Intangible political goods include security, rule of law, essential freedoms such as the right to participate in politics and to compete for office, tolerance of dissent and difference and human rights. All citizens desire to be governed well. They expect the delivery of essential political goods that are tangible as well. These include sustainable economic opportunity, medical and health care, schools and educational instruction, roads, railways, communication networks, and an effective banking system presided over by a central bank. The provision of each of these goods can be measured by creating proxy indicators and sub-indicators. For instance, in a typical measure of the rule of law, the effectiveness and predictability of the judiciary, the number of judges per 1000 people (the more judges the less judicial delay), the number of political prisoners, the level of corruption, the extent of demonstrated respect for property rights, and the ability to enforce

contracts are used as proxies. This engagement with measuring and assessing performance of the political system falls within the realm of political evaluation. The goal is to determine the level of government accomplishment in terms of the provisions of the above public goods. Evaluation is very important in measuring performance. It is evaluation that enables us to know if a policy or government has achieved its stated objectives and intended effects.

Political Evaluation is the systematic acquisition and assessment of information to provide useful feedback about governance. Political evaluation like any other form of evaluation is a systematic endeavour. It involves the acquisition and assessment of information. All evaluation works involve collecting and sifting through data, making judgments about the validity of the information and of inferences we derive from it.

3.2   The Goals of Evaluation

The general goal of most evaluations is to provide "useful feedback" to a variety of audiences including sponsors, donors, client-groups, administrators, staff, and political actors, business entities, communities and other relevant constituencies. Most often, feedback is perceived as "useful" if it aids in decision-making. But the relationship between an evaluation and its impact is not a simple one -- studies that seem critical sometimes fail to influence short-term decisions, and studies that initially seem to have no influence can have a delayed impact when more friendly conditions arise. Despite this, there is broad consensus that the major goal of evaluation should be to influence decision-making or policy formulation through the provision of empirically-driven feedback.

Self-Assessment Exercises

1) Describe how political evaluation aid decision making.

2) State three goals of political evaluation.

4.0   SUMMARY

In this unit, we have provided a brief description of the meaning and context of political evaluation.   We have explained political evaluation as an important form of assessment of government performance.   Political evaluation is an assessment of government performance in providing both the tangible and intangible goods that governments are expected to provide for their citizens. The aim is to have a clear idea of the performance of government or to help improve policy implementation.

5.0   CONCLUSION

In conclusion, political evaluation is very important in any political process that wants to continue to improve.   Indeed, all government programmes should be subject to one form of evaluation or the other.   In real life, political evaluation is a continuous exercise that helps public accountability and performance of government.   It may also help the government in determining the preferences of citizens.   Indeed, it can be used to aid the responsiveness of government to the needs of the people.

6.0   TUTOR-MARKED ASSIGNMENT

1) Define political evaluation

2) State the role of evaluation in political life

3) List some situations when political evaluation may be necessary

7.0   REFERENCES/FURTHER READING

Jones, E. Terrence 1971. *Conducting Political Research*. New York: Harper and Row Publishers

King, Gary; Robert Keohane, and Sidney Verba. 1994. *Designing Social Inquiry* Princeton: Princeton University Press.

McNabb, David E. 2010. *Research Methods for Political Science: Quantitative and Qualitative Methods.* New York: M.E. Sharpe, Inc.

Marsh, David and Gerry Stoker (eds) 2010. *Theory and Methods in Political Science,* New York: Pelgrave Macmillan.

Meier, Kenneth J.; Jeffrey L. Brudney and John Bohte.2006. *Applied Statistics for Public and Nonprofit Administration* (6th edition). Belmont, CA: Thomson Higher Education.

Mitchell, Mark L. and Janina M. Jolley 2007. *Research Design Explained* (6th edition). Belmont, CA: Thomson Higher Education.

Patton, M. Q. 1990. *Qualitative evaluation and research methods.* (2nd ed.). Thousand Oaks, CA: Sage.

UNIT 2: POLITICAL EVALUATION RESEARCH

CONTENT

1.0   INTRODUCTION

In this unit, we focus on political evaluation as a scientific activity that involves constant study, investigation and inquiry into political reality. As a scientific activity, political evaluation seeks to explain and propound statements about the characteristics and behaviours of human beings and institutions. We will see that political evaluation research is an aspect of social science research that follows the essential attributes of the scientific process as much as possible.

2.0   OBJECTIVES

In this unit, you will

1. understand political evaluation as a scientific activity;

2. appreciate the fact that research is central element of political evaluation; and

3. be able to explain the nature of political evaluation research.

3.0 MAIN CONTENT

This unit focuses on political evaluation as a systematic activity that adopts elements of the scientific method. This usually involves some form

of research activity. It explains political evaluation as a scientific research activity aimed at achieving valid and reliable conclusions that are verifiable or generalizable.

3.1 Political Evaluation as a Scientific Activity

In the preceding unit, we established that political evaluation is a systematic assessment of governance and government activity. Our idea of a 'scientist' is that of a man found in the laboratory, fiddling with test tubes and performing complex experiments, or an engineer at a drawing board, designing a blue- print of a bridge, building or any other structure. A more sophisticated observer may conceive of a scientist as one who articulates complex theories, a thinker who quotes profusely, the brilliant ideas of other thinkers and make use of them to advance knowledge and improve the lot of humanity.

However, all ideas that we hold of a scientist share one thing: that scientists are engaged in the search for knowledge and perhaps the use of it (Asiaka, 1991). A scientist engages in constant study, investigation and inquiry into reality or different phenomena in order to explain and propound statements about the characteristics and behaviours of the phenomena. Political evaluation research is an aspect of social science research. Like social science research more broadly, it follows the essential attributes of the scientific process as much as possible. The essential attributes of the scientific process include; controlled experimentation, objectivity, sound theoretical foundation and construction of complex theories, predictability and generalisability of results (Fadeyi and Adelokun, 2004). In this regard political evaluation research must strive to:

• employ systematic, empirical methods that draw on observation or experiment;

• involve rigorous data analyses that are adequate to test the stated hypotheses and justify the general conclusions; and

• rely on measurements or observational methods that provide valid data across evaluators and observers, and across multiple measurements and observations.

3.2   What is Research?

The word 'Research' has two separate parts: the prefix-Re and the base word-search. The prefix "Re" means doing something again or repeating while the base word search is looking for something repeatedly. Research can be defined as a process of finding out the solution to a problematic situation. Finding out or searching for is a step towards solving problems. That is, the product of finding out is obtaining information, which will lead to solving problems or puzzles. Research then means a problem or puzzle - solving exercise.

The Oxford advanced learner's dictionary defines research as "investigation undertaken in order to discover new facts, or get additional information". It is a systematic attempt to develop a valid knowledge or understanding of a particular phenomenon as well as to provide answers or solutions to various questions or social problems in a more technical sense, it is a systematic process of investigation or inquiry carried out in accordance with laid down (scientific) procedures for the purpose of finding answers or solutions to a set of defined problems (Fadeyi and Adelokun: 2004).

According to Osuala (1990), research is the process of arriving at dependable solutions to problems through the planned and systematic collection, analysis and interpretation of data. Research is an important skill required of all political evaluators. Research, according to McNabb (2010), is "gathering, processing, and interpretation of data. It also means intelligently

and cogently communicating the results in a report that describes what was discovered from the research." This definition shows that research is an activity that is carried out after careful planning. There is the need for the researcher to spend time on planning in order to arrive at the expected end.

3.3   What is Political Evaluation Research?

In general, political evaluation research includes any research which has as its subject of investigation or inquiry, political phenomena, whether government or segments of government institutions, political organisations or candidates in a political contest. In a broad sense, political evaluation research is concerned with systematic gathering of data that can help us answer questions about various aspects of political life . It is carried out to enable us understand and explain various aspects or features of politics and governance. It seeks to provide answers to questions that relate to specific problems or to explore questions of theoretical interest in politics and governance.

Self -Assessment Exercises

1) Describe political evaluation as a scientific activity.

2) Political evaluation takes place within a political context. Discuss

4.0   SUMMARY

Political evaluation is a scientific activity. It is an aspect of social science and employs social science method. It involves research, finding out information in a systematic manner about political reality. It seeks to answer questions about politics and governance through scientific investigation.

5.0   CONCLUSION

We have examined political evaluation as a methodological area that is closely related to, but distinguishable from more traditional social science research. It utilizes many of the same methodologies used in traditional social science research, but because political evaluation takes place within a political and governance context, it requires group skills, management ability, political dexterity, sensitivity to multiple stakeholders and other skills that social research in general does not rely on as much.

6.0   TUTOR MARKED ASSINGMENT

1) Explain what you understand by research

2) Analyse the role research play in political evaluation

7.0   REFERENCES/FURTHER READING

Bless, Claire; Craig Higson-Smith and Ashraf Kagee 2006. *Fundamentals of Social Research Methods: An African Perspective* (4th edition). Cape Town: Juta & Co. Ltd.

Bryman, Alan (2004) (2nd Edition) *Social Research Methods*, Oxford: Oxford University Press

Cosby, Paul C. 2006. *Methods in Behavioural Research*. New York: McDraw-Hill

Creswell, J (2003) *Research Design: Qualitative, quantitative, and mixed methods approaches*, (2nd ed.), Thousand Oaks, CA: Sage Publications.

Jones, E. Terrence 1974. *Conducting Political Research*. New York: Harper and Row Publishers

King, Gary; Robert Keohane, and Sidney Verba. 1994. *Designing Social Inquiry* Princeton: Princeton University Press.

Marsh, David and Gerry Stoker (eds) 2010. *Theory and Methods in Political Science*. New York: Palgrave Macmillan

McNabb, David E. 2010. *Research Methods for Political Science: Quantitative and Qualitative Methods.* New York: M.E. Sharpe Inc.

Meier, Kenneth J.; Jeffrey L. Brudney and John Bohte 2011. *Applied Statistics for Public and Nonprofit Administration* (6th edition). Belmont, CA: Thomson Higher Education.

UNIT 3: POLITICAL EVALUATION RESEARCH STRATEGIES

CONTENT

1.0 INTRODUCTION

In this unit, we are concerned about the various strategies for carrying out political evaluation. We will look at the two main strategies of research, which also applies to political evaluation. These are qualitative evaluation and quantitative evaluation.

2.0   OBJECTIVES

In this unit, you will encounter the two main strategies of social research which are used in political evaluation. You will be able to identify the meaning and uses of qualitative and quantitative research strategies and their characteristics.

3.0 MAIN CONTENT

3.1   Research Strategies

Research strategies have to be clarified in the process of political evaluation. These strategies relate to whether the goal of the research can be presented as findings in numerical terms or limited to words.

3.2   Qualitative Strategies

Qualitative research is a strategy that seeks to study issues such as people's motives and subjective interpretation of their actions. It is appropriate for studies that do not lend themselves to quantifiable data. For instance, if we want to find out how powerful or powerless a people feel about the political system, we will need to ask questions that not only tell us the number of people who feel powerful or powerless. We will need to find out why said they are powerful or powerless. We will have to ask why they feel powerful and the responses may not be amenable to quantification. But we will be able to characterise the political system and conceptualise the relation between the citizens and the government.

Bryman (2004) describes qualitative research strategy as one that "emphasise words rather than quantification in the collection and analysis of data''. According to him, it is characterised by the following:

- An inductive approach to the relationship between theory and research in which the emphasis is placed on the generation of theory

- Rejects the practices and norms of the natural scientific model and of positivism in particular in preference for an emphasis on the ways in which individuals interpret their social world; and

- Embodies a view of social reality as a constantly changing object involving individual's construction (p.20).

3.3   Quantitative Strategies

Quantitative methods are about quantity. As Miller (1995:155) puts it, it seeks to answer the question: 'How many are there?' For Bryman (2004:19),

it emphasises quantification in the collection and analysis of data. It is characterised by the following:

- A deductive approach to the relationship between theory and research, in which the emphasis is placed on the testing of theories;

- Incorporates the practices and norms of the scientific model and of positivism in particular, and

- Embodies a view of social reality as an external, objective reality.

Quantitative studies may be experimental involving the manipulation of variables or they could be just observational. It also utilises a variety of data: aggregate data, household data and individual data. It usually involves enumeration and sampling and some form of measurement. It is also amenable to several forms of analyses - univariate, bivariate or multivariate models of analysis depending on the number of variables and their relationships.

Self-Assessment Exercises

Discuss the notion that quantitative strategies focus on numbers while qualitative strategies focus on words.

4.0   SUMMARY

Qualitative research strategy is one that emphases words rather than quantification in the collection and analysis of data. It views social reality as a constantly shifting emergent property of individuals' creation Quantitative strategies seeks to answer the question: 'How many are there?' It focuses on testing propositions. It may involve the use of experiments whereby variables are manipulated or it could be just observational. It uses enumeration and sampling and some form of measurement.

5.0 CONCLUSION

In this unit we have examined two important elements of evaluation research that relate to the methods and strategies of carrying out political evaluation. One of them focuses on quantification, the other on words and their meaning. Although these are two distinct strategies for carrying out research, both methods may be employed in a single research or evaluation.

6.0 TUTOR MARKED ASSIGNMENTS

1. Discuss a research strategy.
2. Distinguish between qualitative and quantitative research strategies.

7.0 REFERENCES/FURTHER READING

Bless, Claire; Craig Higson-Smith and Ashraf Kagee 2006. *Fundamentals of Social Research Methods: An African Perspective* (4th edition). Cape Town: Juta & Co. Ltd.

Bryman, Alan (2004) (2nd Edition) *Social Research Methods*, Oxford: Oxford University Press

Cosby, Paul C. 2006. *Methods in Behavioural Research*. New York: McDraw-Hill

Creswell, J (2003) *Research Design: Qualitative, quantitative, and mixed methods approaches*, (2nd ed.), Thousand Oaks, CA: Sage Publications.

Jones, E. Terrence 1974. *Conducting Political Research*. New York: Harper and Row Publishers

King, Gary; Robert Keohane, and Sidney Verba. 1994. *Designing Social Inquiry* Princeton: Princeton University Press.

Marsh, David and Gerry Stoker (eds) 2010. *Theory and Methods in Political Science*. New York: Palgrave Macmillan

Meier, Kenneth J.; Jeffrey L. Brudney and John Bohte 2010. *Applied Statistics for Public and Non profit Administration* (6th edition). Belmont, CA: Thomson Higher Education.

UNIT 4:   TYPES OF POLITICAL EVALUATION RESEARCH

CONTENT

INTRODUCTION

Evaluation research is a research done for a specific purpose: to assess some social or political activity or programme aimed at ameliorating a social or political problem (Baker 1999). To Ogundipe *et al* (2006), it is a type of research that involves the collection of data in order to make decisions about the value of a product, project or technique. Bryman (2004) defines it as a research concerned with the evaluation of such occurrences as social and organizational programmes or interventions. It seeks to provide objective assessments of past, present or proposed programs of actions. Suffice it to state here that most evaluation researches are sponsored by actors in the environment of the program, whether superordinate organizations, auditing organizations, peer organizations or donor organizations.

2.0   OBJECTIVES

In this unit, you are expected to

1.   identify two broad types of political evaluation research;

2.   explain formative evaluation and summative evaluation;

3. understand the various uses of the two broad types of evaluation research.

3.0   MAIN CONTENT

Evaluation research is conducted under different typologies. The two prominent types are formative and summative evaluation.

3.1:   Formative evaluation

According to Ogundipe *et al* (2006) formative evaluation refers to one in which data are collected about a programme, project or policy while it is still being developed. It is an evaluation which is focused on assessing program quality, implementation to provide feed-back and information for internal improvement without external consequences. Its main goal is to provide internal feedback to improve practice while the program is in progress rather than waiting until it is over. If it is discovered that a programme was not being implemented as intended and so did not have the desired results adjustment could therefore be made mid-way.

The point being made here is that formative evaluations strengthen or improve the program being evaluated by examining the delivery of the program or technology, the quality of its implementation and the assessment of the organizational context, personnel, procedures, inputs and so on. Formative evaluation includes several evaluation types such as needs assessment, evaluability assessment, structured conceptualization, implementation evaluation and process evaluation. These are explained below.

I.   *Needs assessment:* This is a study carried out to determine who needs the program, how great the need is and what might work to meet the need.

II. *Evaluability assessment:* This is carried out to determine whether an evaluation is feasible or not and how stakeholders can help to shape its usefulness.

III. *Structured conceptualization*: It helps stakeholders to define the program or project, the target population and the possible outcomes.

IV. *Implementation Evaluation*: This monitors the fidelity of the program or project delivery, whether it is faithfully delivered or not.

V. *Process evaluation*: This investigates the process of delivering the program or project including alternative delivery procedures.

According to Christine and Dalley (2007) formative evaluation seeks to answer the following questions: is the program or intervention in place? Is it reaching the people intended to assist or affect in some way? How is it changing over time during implementation? What are the challenges to implementing the program? How much does it cost to implement the program? Is it feasible? Do the benefits justify the costs? Where is the problem and how big is it? What is the definition and scope of the problem or issue?

Answers to these questions and other sundry issues can help to improve program services as well as understanding what appears to be working and making a difference.

3.2: Summative evaluation

Summative evaluation is a research designed to provide information on program impact to external agencies (Christine and Dalley 2007). It is done or carried out after the program or project has been fully developed or wound up. It aims to examine the effects or outcomes of a program or project on the target population. Findings here are usually reported through formal written reports usually coming together in a final report.

Summative evaluation enjoys the advantage of keeping stakeholders informed and satisfied.

Summative evaluation can be subdivided into different types; these are:

1. Outcome evaluations: This investigates whether the program caused effects on specifically defined target outcomes.

2. Impact evaluation: It is broader and assesses the overall or net effects—intended or un-intended— of the programme or project as a whole.

3. Cost-effectiveness and cost-benefit analysis: These address questions of efficiency by standardizing outcomes in terms of their monetary costs and values.

4. Secondary analysis: This re-examines existing data to address new questions or use methods not previously employed.

5. Meta-analysis: Meta-analysis integrates the outcome estimates from multiple studies to arrive at an overall judgment on an evaluation question.

Summative evaluation seeks to answer these questions: Did the program accomplish its goals? Is the program effective in addressing the problem as intended? And so on. Answers to these questions help the evaluator to assess the impact of a program or project on the target people or population.

Self -Assessment Exercise

1) Describe the broad classification of political evaluation research.

2) State the type of evaluation research that would be appropriate for someone who wants to determine whether a programme should continue or be scraped.

4.0: SUMMARY

Formative evaluation examines the delivery of a programme or technology, the quality of its implementation and the assessment of the organizational context, personnel, procedures, inputs and so on. Formative evaluation includes several evaluation types such as needs assessment, evaluability assessment, structured conceptualization, implementation evaluation and process evaluation. Summative evaluation examines the effects or outcomes of a program or project on the target population. It is done to keep stakeholders informed and satisfied.

5.0: CONCLUSION

As a type of research, political evaluation research collects data to measure the impact of a project, policy or programme. Though an important and indispensable part of the policy making process, evaluation regrettably is the most neglected or abandoned aspect of a project. All too often, government, firms and entrepreneurs busy themselves with making and implementing programmes without taking the pains to evaluate the impact of such programmes in the lives of the target population. For an effective evaluation of a project or programme to take place, both formative and summative evaluations should be used. Again, feasibility studies should be undertaken by stakeholders to know the basic needs of a target population before initiating policies or programmes to tackle them. Also, stakeholders should take evaluation very seriously as it is as important as initiation and formulation of policies and programmes.

6.0: TUTOR MARKED ASSIGNMENT

1) State the purposes of formative evaluation

2) Describe summative evaluation

3) Describe the essence of summative evaluation

7.0:   REFERENCES/FURTHER READING

Baker, T. 1999. *Doing Social Research* (3rd Edition) McGraw-Hill, McGraw Hill College, 300-332.

Bryman, A. 2004. *Social Research Methods* (2nd Edition) London, Oxford University Press.

Christine, A and Dalley, T. 2007. *A Guide for Education Personnel: Evaluating a Programme or Intervention*, Washington D.C: American Institute for Research.

Ogundipe, G.A., Eucas, E.O., & Sanni, A.I. 2006. Systematic Collection of Data, Eds A.I. Olayinka, V.O. Taiwo, A. Raji -Oyelade, I.P. Farai, eds *Methodology of Basic and Applied Research* (2nd Edition). Ibadan: University of Ibadan Press. Pp. 107-108.

UNIT 5: STATISTICS AND POLITICAL EVALUATION

CONTENT

1.0 Introduction
2.0 Objectives
3.0 Main Content
    3.1 What is statistics?
    3.2 Types of Statistics
    3.3 Importance of Statistics in Political Evaluation
4.0 Summary
5.0 Conclusion
6.0 Tutor Marked Assignment
7.0 References/further Reading

1.0 INTRODUCTION

Statistics is widely used in political evaluation. Statistics may be used to describe something or used to infer similar measurements in another larger group. This unit explains statistics as tools of political evaluation.

2.0 OBJECTIVES

In this unit, you will be able to do the following:

1. Understand the meaning and uses of statistics in political evaluation;

2. Identify different types of statistics and appreciate their application in political evaluation research, You will be able to use statistics for the purposes of making generalisation about a population from a sample, estimate or draw conclusion about a population or predict something about some future event or state of affairs.

3.0 MAIN CONTENT

The primary focus of this unit is to explain the meaning and types of statistics. It examines descriptive and inferential statistics and their uses in political evaluation.

3.1   What is Statistics?

As we have seen from the discussion of quantitative and qualitative research strategies, statistical data and techniques are employed in the process of political evaluation. What does an individual think of when he or she thinks of statistics? Statistics has two meaning depending on how it is being used at a given time. When statistics is used in singular form, it means the discipline or science, which deals with the collection, classification, analysis and use of numerical data. Brase and Corrine (2007:4) defines statistics as the "study of how to collect, organize, analyse, and interpret numerical information from data". But when used in a plural form, it refers to the numerical facts about data themselves e.g. GNP, divorce , population , salary , age , crime , voters turn out etc.

3.2   Types of Statistics

Statistics is equally divided into two major types namely: descriptive statistics and inferential statistics. Descriptive statistics are quantitative tools of analysis that enable social scientists to describe and summarise data. It involves organizing, picturing and summarising a large amount of information from samples or populations. Any average, for example, is a descriptive statistic. So, average daily rainfall, or average daily temperature are good examples of descriptive statistics. Examples of descriptive statistics   include measures of central tendency(mean, mode and median value of dataset) measures of variability in the dataset (such as standard deviation , range, and the interquartile range), measures of relative position in the set, such as percentiles and standard scores; and measures of correlation between two or more variables, such as correlation tests that are

41

used to show how strongly and in what direction two variables are related, if at all (See McNabb 2004:121).

The second type of statistics is inferential statistics. This involves using information from a sample to draw conclusions regarding the population (Brase and Brase 2007). For instance, we may want to know what the people of Lagos State think about the performance of Babatunde Fashola's administration. Because it will take forever to ask the over 9,000,000 citizens of Lagos state about their assessment of the performance of the administration, we could randomly select 200 citizens and ask them about the governor's performance. From this sample we will then infer the views of the citizens of Lagos State. Some basic types of inferential statistics commonly used in Political Science are:

1. The *t*-test for significant differences between means of dependent (uncorrelated) groups.

2. The *t*-test for significant differences between the means of paired or correlated groups.

3. Simple regression analysis for measuring the strength and the direction of of relationships between variables.

4. Analysis of variance (ANOVA) tests for difference on one variable for two or more groups.

5. Analysis of variance (ANOVA) tests for differences on two or more variables between two or more groups, and for any *interaction* that might result from the two variables.

6. Analysis of variance (ANOVA) as used in pre- and post-test experimental applications (See McNabb 2004:123).

3.3   Importance of Statistics in Political Evaluation

Numerical data and their analyses help us to make inference from a small collection of people or item about a larger collection of people or item. Statistics is therefore a central element of political evaluation. It is particularly useful when we deal with issues that have value or numerical measurement. Indeed, statistics helps us to understand complex situations and make decisions.

Self-Assessment Exercise

1) State the significance of statistics for evaluation research

4.0 SUMMARY

The word statistics is used to mean one or more specific measures or values describing something, a population or a sample of a population. They may be used to summarise a larger set of numbers, called a dataset or used to measure a smaller group, the sample, and then used for making assumptions about a larger group, the population.   Statistics is a central element of political evaluation, particularly useful when we deal with issues that have value or numerical measurement. Statistics helps us to understand complex situations and make decisions.

5.0 CONLUSION

In this unit, we have seen that statistics is very important for political evaluation research. They are used to describe and make inferences. There are computer programmes that are designed to aid our analysis of statistical datasets. Examples of such programmes include E-View and Statistical Package for the Social Sciences (SPSS).

6.0   TUTOR MARKED ASSIGNMENTS

1) Describe the role of statistics in doing political evaluation.

2) List the four basic types of descriptive statistics

3) Define inferential statistics in relation to when they are used

7.0 REFERENCES/FURTHER READING

Brase, Charles H. and Corrine P. 2007. *Understanding Basic Statistics* New York & Boston: Houghton Mifflin Company.

Bryman, Alan (2004) (2nd Edition) *Social Research Methods*, Oxford: Oxford University Press

Creswell, J (2003) *Research Design: Qualitative, Quantitative, and Mixed Methods Approaches*, (2nd ed.), Thousand Oaks, CA: Sage Publications.

Jones, E. Terrence 1974. *Conducting Political Research*. New York: Harper and Row Publishers

McNabb, David E. 2011. *Research Methods for Political Science: Quantitative and Qualitative Methods.* New York: M.E. Sharpe , Inc.

Marsh, David and Gerry Stoker (eds) 2010. *Theory and Methods in Political Science*. New York: Pelgrave Macmillan.

Meier, Kenneth J.; Jeffrey L. Brudney and John Bohte.2010. *Applied Statistics for Public and Nonprofit Administration* (6th edition). Belmont, CA: Thomson Higher Education.

Mitchell, Mark L. and Janina M. Jolley 2007. *Research Design Explained* (6th edition). Belmont, CA: Thomson Higher Education.

Patton, M. Q. 1990. *Qualitative Evaluation and Research Methods.* (2nd ed.). Thousand Oaks, CA: Sage.

Taylor, Steven J. and Robert Bodgan 2006. *Introduction to Qualitative Research Methods: A Guide Book and Resource.* New York: John Riley and Sons Inc.

MODULE 2: MEASUREMENT IN POLITICAL EVALUATION

Unit 1: THE CONCEPT OF MEASUREMENT

CONTENT

1.0   Introduction
2.0   Objectives
3.0:   Main Content
      3.1:   The Concept of Measurement
      3.2:   Why do we Measure?
4.0   Summary
5.0:   Conclusion
6.0   Tutor Marked Assignments
7.0   References/Further Reading

1.0   INTRODUCTION

In this unit we will be looking at the concept of measurement. We will also examine the concept of operational definition that enables us to measure a political reality in an indirect way, including the use of measurement in political evaluation research.

2.0: OBJECTIVES

In this unit, you should be able to do the following

1. explain the concept of measurement and importance in political evaluation research;

2. understand the idea of operational definition and indicators; and

3. be able to explain why it is necessary to measure in political evaluation.

3.0   MAIN CONTENT

The primary focus of this unit is measurement. This unit will explain the concept of measure and the procedure for doing so. These include the use of operational definition and indicators in measurement. The unit will also explain why it is necessary to measure.

3.1:   The Concept of Measurement

Measurement is the assignment of numbers to things or concepts in order to analyse them. According to Jones (1971:30) "when we measure something, we are assigning numerals to things according to their operational definition". Measurement is the assignment of numbers to things we are interested in analysing. It entails "a careful, deliberate observation of the real world for the purpose of describing objects and events in terms of the attributes composing a variable" (Babbie 2005:122).   In their opinion, Agbaje and Alarape (2005:22), see measurement as "the process of empirically observing, codifying and estimating the extent of presence of concepts related to the phenomena under investigation." The essence of measurement in any scientific research is to ensure accuracy and precision in gauging the consistency, validity and reliability of a particular phenomenon (Neuman: 2000). Nzeneri (1994) posits that 'to assign weight to variables, we must make use of measurement scales'.

Not all phenomena can be directly measured.   For instance, the effectiveness of the military, programme success, and independence of the judiciary cannot be measured directly. Concepts like these can only be measured indirectly through indicators specified by their operational definition. An operational definition is a nominal definition of a concept; it tells how a concept will be measured. The aim is to achieve a maximum clarity concerning the use of a concept in the context of a particular study. It

specifies the exact operation involved in measuring a variable. It goes beyond conceptualisation - the process through which we specify what we use a particular term to mean in a research work. An indicator is a set of observations that results from applying the operational definition. It is a reflection of the variable we wish to study. For instance, the guarantee of tenure of judicial officers may be used as an indicator of the independence of the judiciary. Sometimes a single indicator may not be sufficient to measure a concept. Hence, multiple or composite indicators may be used. This is often the case because a concept may have more than one dimension. Hence, the rule of law may be more adequately measured by multiple indicators like constitutional checks and balances status, leadership's respect for court order, police respect for human rights, citizen's confidence in law enforcement organs, monitoring of violations by police and prisons, civil society monitoring of violations by police and prisons, penalty for violations of human rights by police, and watch-dog organizations independence from executive.

3.2: Why Do We Measure?

There are several reasons why we engage in measurement in the process of research. Alan Bryman (2004:66) has identified three of the reasons: First, measurement allows us to delineate fine differences between people in terms of the characteristic in question. This is very useful, since, although we can often distinguish between people in terms of extreme categories, finer distinctions are much more difficult to recognize. The same point is made by Obasi (1999) when he stated the classificatory function of measurement enables researchers to differentiate between the objects being studied according to the properties they possess. This means that

individuals or objects being studied can be classified because of their possession of certain characteristics.

Second, measurement gives us a consistent device or yardstick for making such distinctions. A measurement device provides a consistent instrument for gauging differences. This consistency relates to two things: our ability to be consistent over time in its result and our ability to be consistent with other researchers - reliability.

Third, measurement provides the basis for more precise estimates of the degree of relationship between concepts. An example is correlation analysis. For example we can relate voter turnout in an election with expressed confidence of citizens in the electoral process.

Measurement helps researchers to provide accurate description of hypothesis about the phenomena. It makes it possible for data to be quantified thereby becoming amenable to statistical manipulation and treatment.

Lastly, measurement makes it possible for hypothesis and theories to be subjected to empirical verification much more easily.

Self –Assessment Exercise
1) Discuss the importance of measurement in political research.
2) Compile a list of indicators that can be used to evaluate political representation.

4.0: SUMMARY

In this unit we have examined the concept of measurement as the assignment of numbers to things we are interested in analysing. We cannot measure most political phenomena directly; we therefore measure indirectly

by using indicators specified in the operational definition. An indicator is a set of observations that results from applying the operational definition. It is a reflection of the variable we wish to study. The aim of measurement is to achieve a maximum clarity concerning the use of a concept in the context of a particular study.

5.0:   CONCLUSION

Measurement is a very important element of political evaluation. It is a process of assigning numbers to phenomenon according to specifications contained in the operational definition. An indicator may be used where the phenomenon cannot be directly measured. In the case of a phenomenon that has more than one dimension, multiple indicators may be used. Measurement enables us to achieve more precise description of variables and make more nuanced distinction objects of political evaluation.

6.0:   TUTOR MARKED ASSIGNMENTS

1.   Describe measurement in political evaluation.

2.   Identify a concept in political evaluation, such as professionalism or accountability and explain how this concept has been measured in the literature.

7.0: REFERENCES/FURTHER READING

Agbaje, A. and Alarape, A.I. (2005), *Introductory Lectures on Research Methodology.* University of Ibadan, Ibadan.

Asika, I. (1991). *Research Methodology in the Behavioural Sciences.* Lagos: Longman.

Babbie, E. (2007). *The Practice of Social Research,* (11th edition). USA: Wadsworth

Cengage Learning. Garret, H.B. (1962), *Elementary Statistics.* New York: David McKay Inc.

Bless, Claire; Craig Higson-Smith and Ashraf Kagee 2006. *Fundamentals of Social Research Methods: An African Perspective* (4th edition). Cape Town: Juta & Co. Ltd.

Brase, Charles H. and Corrine P. Brase 2007. *Understanding Basic Statistics* (4th edition), Boston & New York: Houghton Mifflin Company.

Bryman, Alan 2004. *Social Research Methods* (2nd Edition). Oxford: Oxford University Press

Jones, E. Terrence 1974. *Conducting Political Research.* New York: Harper and Row Publishers

Meier, Kenneth J.; Jeffrey L. Brudney and John Bohte 2006. *Applied Statistics for Public and Nonprofit Administration* (6th edition). Belmont, CA: Thomson Higher Education.

Nzeneri, I.S. (1994). *An Introduction to Educational Research Methods and Statistics.* Onitsha: Goodway Press Ltd.

Obasi, 1. (1999). *Research Methodology in Political Science. Enugu:* Academic Printing Press.

UNIT 2: LEVELS OF MEASUREMENT

CONTENT

1.0:   INTRODUCTION

In this unit we will identify the various levels of measurement. We will also look at the nature of measurement at the various levels, especially the extent to which measurement is precise and definitive. We will look at the various factors that determine which level of measurement can be attempted and what sort of hypothetical relationship can be formulated using the variable.

2.0:   OBJECTIVES

In this unit, you are expected to:

1. explore and grasp key concepts that describe the basics of measurement. These include origin, order, and distance;

2. become familiar with the various levels of measurement: nominal, ordinal, interval and ratio levels of measurement.

3.0   MAIN CONTENT

There are different levels of measurement in research. These levels of measurement have also been denoted as scales by some researchers, but these should not be confused with the conventional scales. The following are the various levels of measurement.

According to Hoover and Donovan (2004: 94-5) the following determines what level of measurement can be attempted in measuring a variable.

1. The properties or characteristics of the variable,

2. The measurement technique appropriate to the these properties; and

3. The levels of measurement that are possible in view of the variable's properties and available techniques.

3.1: Nominal Measurement

This is the lowest and most simple level of measurement. It involves the classification or categorization of subjects, observations, responses or variables into classes that have no direction to them or differentiation made. Agbaje and Alarape (2005) have identified the three Steven's typology of Nominal measurement which has the following characteristics:

- It makes no assumption about the values assigned to the data.

- Each assigned value is a distinct category and serves only as a label or name for each value.

- It makes no assumption of ordering or distances between categories.

The assignment or use of numbers here to represent the groups does not mean a score but is used to indicate the frequencies of the different groups or classes. For instance, in a study involving university students, we may be interested in gender - male or female. Again, we may also be interested in their religion - Christian or Muslim . The principle governing the classification, according to Obasi (1999), is simply what the researcher is interested in highlighting or for partitioning with respect to the composition of the subjects being studied. The nominal measurement is merely labelling. Variables whose attributes have only characteristics of exhaustiveness and mutual exclusiveness are subject to nominal measures. These include such

variables as gender, religious affiliation, political party affiliation, birthplace, college major, and hair colour. These attributes which comprise each of these variables are distinct from one another and exhaust the possibilities of difference among people. They have no additional structures.

3.2: Ordinal Measurement

This enables researchers to rank or order their subjects or responses of their subjects in such ways like 'greater than' and 'less than', etc. As a higher level measurement in relation to nominal measurement, the ordinal measurement enables researcher not only to differentiate groups but also to express them in 'greater than' or 'less than' relationships with numbers used to represent the groups. The numbers used show the relative positions of the differentiated groups. Examples of variables conforming to the ordinal scale are social class, ratings of universities, organizations etc.

Agbaje and Alarape (2005) enumerated some characteristics of the ordinal measurement to include:

• Each category used to measure the values of a variable has a unique place relative to other categories. It is either less than or more than others.

  • However, it conveys no information as to the extent of difference between or among the categories. In other words, there is no information or indication of the distance separating the categories.

• The only mathematical property of measurement at this level is that of ordering.

Obasi (1999) recognized one disadvantage of this measurement. Since the numbers assigned to subjects or responses only represent their position in a rank order, it is not possible to specify the magnitude of the differences between the numbers assigned to the elements.

3.3:   Interval Measurement

According to Black and Champion (1976), this level of measurement has all the properties of the nominal and ordinal levels of measurement in addition to showing equal spacing between the intervals. It not only tells us the order of things, it also tells us the interval or distance between them. The interval measurement takes care of the inadequacy of the ordinal scale. This is because any one unit difference in score represents the same amount of difference in the variable being quantified as any other unit difference in score (Obasi 1999). One limitation of interval measurement is that it does not have an absolute zero point to enable one make a ratio statement. Examples of interval measurements are IQ or temperature. If, for example, a person has an IQ of 150 and another has an IQ of 75, one can only say that there is a 75-point difference between their IQ levels. One cannot say that the one student is twice as intelligent as the other. This illustrates the fact that distances are assumed equal but, because there is no absolute zero level, relative comparisons cannot be made.

3.4:   Ratio Measurement

The ratio measurement has the properties of the highest level of measurement in terms of precision. Measures of extent (in inches), of weight (in pounds), of time (in seconds), are illustrations of ratio scales (Garret: 1962).

According to Black and Champion (1976), it is important to note that these different levels of measurement require 'particular set of statistical procedures and techniques of analysis that are permissible under certain scientific and mathematical rules'.

Self-assessment Exercises

1) State the determinants of the levels of measurement.

2) Describe the highest level of measurement in terms of precision.

4.0   SUMMARY

In this unit, we have seen that there are various levels of measurement. These are nominal, ordinal, interval and ratio measurements. Nominal level of measurement describes a variable that has attributes that are merely different. The ordinal level of measurement describes a variable that with attributes that enables us to rank-order along some dimensions. In addition to specifying that two or more people have different attributes we can rank also one more than the other. For instance that one is more religious or more conservative than the other. At the interval level of measurement, we describe variables whose attributes are rank - ordered and have equal distances between adjacent attributes. The Fahrenheit scale is at the level of interval measurement. This is because the distance between 14 and 15 is the same as that between 80 and 81. The actual distances are meaningful standard intervals. The ratio level of measurement describes a variable with attributes that have all the qualities of nominal, ordinal and interval measures and, in addition, are based on a "true zero" point. Age and income are examples of attributes that can subject to ratio level of measurement.

5.0 CONCLUSION

They express varying levels of specificity and precision in measurement. While nominal measure is the lowest level of measurement, ratio level is the highest and most precise level of measurement. This therefore calls for proper understanding of the nature of one's data so as to know which appropriate statistical technique of analysis could be applied.

6.0   TUTOR MARKED ASSIGNMENT

1) What do you understand by measurement level?

2) Categorise these measurements associated with the life of a politician in your constituency according to level: nominal, ordinal, interval or ratio.

(a) Length of time to complete a political debate session

(b) Major political parties.

(c) Candidate popularity scale: very popular, popular, not popular

(d) Age of candidate

(e) Votes won at the last party primary (based on hundred voting party delegates).

7.0 REFERENCES/FURTHER READING

Agbaje, A. and Alarape, A.I. (2005), *Introductory Lectures on Research Methodology.* University of Ibadan, Ibadan.

Asika, I. (1991). *Research Methodology in the Behavioural Sciences.* Lagos: Longman.

Babbie, E. (2007). *The Practice of Social Research.* USA: Wadsworth Carnegie Learning.

Garret, H.B. (1962), *Elementary Statistics.* New York: David McKay Inc.

Brase, Charles H. and Corrine P. Brase 2007. *Understanding Basic Statistics* (4th edition), Boston & New York: Houghton Mifflin Company.

Bryman, Alan 2004. *Social Research Methods* (2nd Edition). Oxford: Oxford University Press

Hoover, Kenneth and Todd Donovan 2004. *The Elements of Social Scientific Thinking.* Belmont, CA: Wadsworth.

Meier, Kenneth J.; Jeffrey L. Brudney and John Bohte 2006. *Applied Statistics for Public and Nonprofit Administration* (6th edition). Belmont, CA: Thomson Higher Education.

Nzeneri, I.S. (1994). *An Introduction to Educational Research Methods and Statistics.* Onitsha: Goodway Press Ltd.

Obasi, 1. (1999). *Research Methodology in Political Science. Enugu:* Academic Printing Press.

UNIT 3: SCALES OF MEASUREMENT

CONTENT

1.0: Introduction
2.0: Objectives
3.0: Main Content
        3.1   What is a Scale
        3.2   Types of Scale
        3.3   Likert Scale of Measurement
4.0: Summary
5.0: Conclusion
6.0: Tutor Marked Assignment
7.0: References/Further Reading

INTRODUCTION

In this unit, we will focus on the measurement scale. We will examine the meaning of scaling and the operations involved in the process. We will look at the various types of scale and their applications. We also look list some examples of scales frequently used in political evaluation and concentrate on the Likert scale.

OBJECTIVES

In this unit, you are expected to be able to

1. explain the idea of scaling;

2. be familiar with measurement scales, especially the Likert Scale.

3.0   MAIN CONTENT

In this unit, we define what a scale is as used in statistical operations. We examine the meaning of Scaling and the various types of scales. We also focus on the Likert Scale.

3.1:   What is a Scale?

Scales are devices constructed or employed by researchers to quantify the responses of a subject on a particular variable. Scales are tools used to show

a broader aspect of measurement. The scales can be used to obtain interval data concerning attitudes, judgments or perceptions about almost any subject or object. A scale is a type of composite measure composed of several items that have a logical or empirical structure among them. Scaling involves assigning scores to patterns of responses, recognising that some items reflect a relatively weak degree. Scales offer more assurance of ordinality by tapping intensity of structures among indicators.

Although there are various types of scales namely : Likert scale, Thurstone scale, Guttman scale, Bogardus social distance and Semantic differential scale, the most commonly used due principally to the easiness or convenience of its application, as well as the simplicity of interpreting its measures, is the Likert scale. Based on this, we shall discuss only the Likert scale here.

There are various types of scale corresponding to specific levels of measurement. The presence or absence of three properties or attributes in a variable determines the types of scale:

1. The existence of magnitude, which is the possibility of comparing different amounts or intensities so as to assess whether two values or levels of a variable are the same, or one is lesser or greater than the other.

2. The existence of equal intervals, which allows magnitude to be expressed by a certain number of units on a scale, all units on the scale being equal by definition.

3. The existence of an absolute zero, which is a value indicating that the measurement of a variable is meaningless in circumstances in which the variable is non-existent.

3.2: Types of Scales

Based on the three properties of a variable listed above, four measurement scales may be identified. These are nominal, ordinal, interval and ratio scales.

Nominal Scale: This is the most elementary form of scale. It does not measure in the strict sense of the word. What it does is to label or name a variable. A nominal scale is used to classify information into categories or groups. The essence is not to compare the categories or groups. This is because the classification does not make such comparison possible. They are quantitatively different. The variable gender may be classified into 'male' and 'female'. This does not in any way provide for a comparison of the two attributes, it merely classify the same according to the attribute.   The nominal scale is used for such attributes as marital status, or state of origin of respondents.

Ordinal Scale: This scale provides more information than the nominal scale.   It allows some measure of comparison and rank-order between different attributes of a variable.   In the ordinal scale, we can assess the magnitude of attributes to a point when can say that one attribute is more or less than another. For instance, while we classify at the nominal level that a respondent is happy or unhappy, in the ordinal scale we can say a respondent is very happy, happy, indifferent, unhappy and very unhappy. But this does not suggest that the very happy respondent is twice as happy as the happy respondent. This is the case because there is no unit of measurement to be used in comparing them. That is, only the property of magnitude is present in this scale. Similar forms of variables of only magnitude are the grade a student gets in an essay, and the tax bracket into which a person falls and so on.

Interval Scale: Interval scales are more precise than the ordinal scale in the sense that a comparison between the different occurrences of attributes can be done because of the existence of equal intervals or units of measurement. Interval scales have both order and distance but they do not have an original point unless one is assumed for them (Asika 1991: 56). Bless *et al* (2004) illustrate the point in this way: After the concept of employment has been operationalized, a person is classified as employed or unemployed on a nominal scale; employed fulltime, employed part-time or unemployed on an ordinal scale; employed for a certain number of hours per week on a scale possessing equal intervals (one hour). In the last case, a person employed for forty hours a week is employed twice as long as a person working only twenty hours a week. The unit underlying the scale is the hour. All the scales are based on a set of real numbers. However, the number does not have an absolute zero point. For instance, if the money owned by somebody is used as a unit of measure say N1,500.00 somebody who is debt may end up being classified as -N500.00 if the person is in debt to the tune of N2,000.00. This scale has magnitude and equal units of measurement. Although it's values of 0 and negative values are meaningful, it has no absolute zero point.

Ratio Scales: The variables that are subject to ratio scale have the three properties listed above: magnitude, equal intervals and an absolute zero. A very clear example of a variable with absolute zero is age.   A person can be 50 years old but no one is minus three years old. This means there is an absolute zero. Before birth is the absolute zero, that is non-existence. The unit or interval is the year. A comparison can be made between the age of the father and the child. If the child is 20 years and the father is sixty years, it means that the father is three times as old as the child.

Both the interval and ratio scale have ceiling and floor effects. Ceiling effects occurs when a scale does not permit sufficient high scores, resulting in all responses clustering at the top of the scale. A floor effect occurs when a scale does not permit sufficiently low scores and all responses cluster at the bottom of the scale. In the case of examinations, it may be difficult to distinguish between excellent students and average students.

3.3: The Likert Scale of Measurement

A Likert scale is a five-point scale in which the interval between each point on the scale is assumed to be equal, and it is used to register the extent of agreement or disagreement with a particular statement of an attitude, belief or judgment. It is derived after the developer's name, Rensi Likert. Furthermore, there are several ways of expressing the Likert scale measure-responses in questionnaire. The most widely used form is the Agree-Disagree response pattern. In this application, respondents are asked questions that require different or varying intensity of responses. The questions are usually put in a statement form such as "Goodluck Jonathan will be a good president if he wins the 2011 elections". The Likert format of responses will look like this:


Strongly Agree   Agree   Undecided Disagree   Strongly Disagree

In order to convert these qualitative responses into a form amenable to quantitative analysis, the use of figures is introduced.

The responses are assigned weight thus:

SA	A	Und	D	SD
5	4	3	2	1
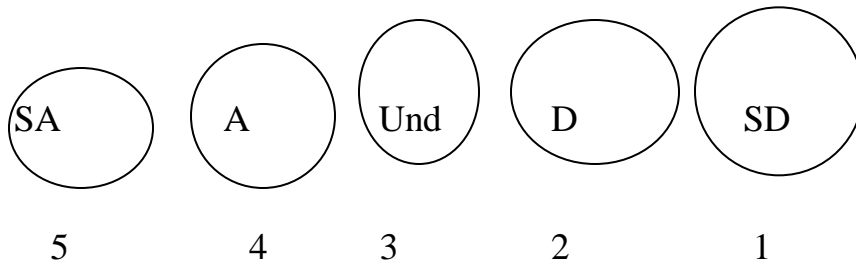
On the other hand, if the question is negatively phased, the weights will change direction. Example 'Goodluck Jonathan will not be a good president if he wins the 2011 elections". The weights will now be:

SA	A	Und	D	SD
1	2	3	4	5
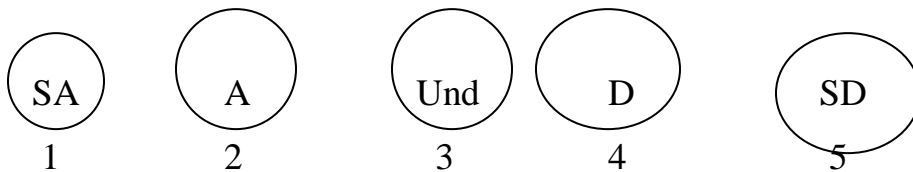
Likert scale makes it possible to transform feelings into an interval scale that can be subject to statistical analysis. Using Likert scale, we can compare the responses among individuals, and between groups using Chi-square analysis. Likert scale is also flexible and can be used to measure in minute detail, the degree of intensity of feeling or attitudes towards a phenomenon.

Self-Assessment Exercise

1) Describe scaling as a tool of political evaluation research
2) List at least four different scales that can be used for measurement.

4.0: SUMMARY

Scaling is a tool for measuring the amount of a property possessed by a class of objects or events. There are various types of scale corresponding to specific levels of measurement. The absence or presence of three properties or attributes in a variable determines the types of scale: magnitude, equal interval and absolute zero. Attitude scales, like the Likert scale, consists of a number of attitude statements with which the respondent is asked to agree or disagree along some continuum.

5.0: CONCLUSION

A political researcher needs to conduct a research by systematic, careful and deliberate observation to be able to understand the real world for describing objects and events in terms of attributes composing of variables. He further presents these data in a more understandable manner for other readers and researchers by adopting measurements at the appropriate levels and the appropriate scales of measurement.

TUTOR MARKED ASSIGNMENT
1. Explain scaling.
2. Discuss the properties of variables that determine the type of scale to use for measurement.
3. Assess the Likert scale and explain how it works.

4.4:   REFERENCES/FURTHER READING

Agbaje, A. And Alarape, A.I. (2005), *Introductory Lectures on Research Methodology.* University of Ibadan, Ibadan.

Asika, I. (1991). *Research Methodology in the Behavioural Sciences.* Lagos: Longman.

Babbie, E. (2007). *The Practice of Social Research.* Belmont, CA: Wadsworth Cengage Learning.

Babbie, E. (2005). *The Basics of Social Research,* Belmont, CA: Wadsworth Cengage Learning.

Baker, Therese 1999. *Doing Social Research*. New York: McGraw-Hill College.

Black, T (1999*) Doing Quantitative Research in Social Sciences: An Integrated Approach to Research Design, Measurement and Statistics*, London: Sage.

Bless, Claire; Craig Higson-Smith and Ashraf Kagee 2006. *Fundamentals of Social Research Methods: An African Perspective* (4th edition). Cape Town: Juta & Co. Ltd.

Hoover, Kenneth and Todd Donovan 2004.   *The Elements of Social Scientific Thinking.* Belmont, CA: Wadsworth.

Jones, E. Terrence 1971. *Conducting Political Research*. New York: Harper and Row Publishers

King, Gary; Robert Keohane, and Sidney Verba. 1994. *Designing Social Inquiry* Princeton: Princeton University Press.

UNIT 4: RELIABILITY AND VALIDITY IN QUANTITATIVE RESEARCH

CONTENT

1.0:   INTRODUCTION

Reliability and validity are important elements of measurement. In political evaluation, it is important to be sure that the measure of behaviour is stable and consistent. Thus, how to determine and ensure the stability and consistency of a measure is the focus of this unit. This is also concerned about validity. The focus of validity is to ensure that a measure actually measures what it is intended to measure. It examines the various ways of assessing the validity of a measure.

2.0:   OBJECTIVES

At the end of this unit, you are expected to be able to

1. evaluate the reliability of measurement instruments and techniques in quantitative research;

2. evaluate the validity measurement instruments and techniques; and

3. acquire the ability to select measurement strategies that maximise validity and reliability in quantitative research.

3.0. MAIN CONTENT

3.1:   What is Reliability?

Joppe (2000:1) defines reliability as "the extent to which results are consistent over time and an accurate representation of the total population under study is referred to as reliability and if the results of a study can be reproduced under a similar methodology, then the research instrument is considered to be reliable". Embodied in this citation is the idea of replicability or repeatability of results or observations. Bryman (2004:28) says reliability is concerned with the question of whether the results of a study are repeatable. It is concerned with how consistent a measure is. Kirk and Miller (1986:41-42) identify three types of reliability referred to in quantitative research, which relate to: (1) the degree to which a measurement, given repeatedly, remains the same (2) the stability of a measurement over time; and (3) the similarity of measurements within a given time period.

Charles (1995) adheres to the notion that consistency with which questionnaire [test] items are answered or individual's scores remain relatively the same can be determined through the test-retest method at two different times. This attribute of the instrument is actually referred to as stability. If we are dealing with a stable measure, then the results should be similar. A high degree of stability indicates a high degree of reliability, which means the results are repeatable. This common way of determining the stability of a measure is the test-retest method. It means that you administer a test or measure to a sample on one occasion and then re-administer the same test or measure on another occasion to the same sample.

Joppe (2000) detects a problem with the test-retest method, which can make the instrument, to a certain degree, unreliable. She explains that test-retest method may sensitize the respondent to the subject matter, and hence influence the responses given. We cannot be sure that there was no change in extraneous influences such as an attitude change that has occurred. This could lead to a difference in the responses provided. Similarly, Crocker and Algina (1986) note that when a respondent answer a set of test items, the score obtained represents only a limited sample of behaviour. As a result, the scores may change due to some characteristic of the respondent, which may lead to errors of measurement. These kinds of errors will reduce the accuracy and consistency of the instrument and the test scopes. Hence, it is the researchers' responsibility to assure high consistency and accuracy of the tests and scores. Thus, Crocker and Algina (1986:106) say, " test developers have a responsibility of demonstrating the reliability of scores from their tests".

The second element of reliability according to Bryman (2004: 71) is internal reliability. This relates to whether the respondents' scores on any one indicator tend to relate to their scores in other indicators.

Third element of reliability relates to inter-observer reliability. This relates to the consistency in the observations made by observers involved in recording observations. This is very important in studies that involve a great deal of subjective judgement.

Although the researcher may be able to prove the research instrument's repeatability and internal consistency, and, therefore reliability, the instrument itself may not be valid.

3.2:   Validity

Validity is a term used to describe a measure or indicator instrument that accurately measures the item or concept it is intended to measure.   The question that validity answers is what does an instrument really measure? What does the result of a research really mean? According to Babble (2005:148) validity refers to the extent to which an empirical measure adequately reflects the real meaning of the concept under consideration. Validity can pose serious challenge in political evaluation. This is because political concepts often carry more than one meaning.

For instance, power may be defined as a person's ability to make others do as he or she wants them to do. But an investigation of how power is exercised will lead to some problems. This is because power may include the use of force, the use of threat, persuasion and reward. But some people will argue that the use of force and threat shows a lack of power. Some political scientist will argue that persuasion or reward is not a form of power.   This illustrates how challenging it can be to operationalise a concept and achieve a valid measure. The description of validity and reliability discussed above relates largely to quantitative research. Indeed, Joppe (2000:1) provides the following explanation of what validity is in quantitative research:

Validity determines whether the research truly measures that which it was intended to measure or how truthful the research results are. In other words, does the research instrument allow you to hit "the bull's eye" of your research object? Researchers generally determine validity by asking a series of questions, and will often look for the answers in the research of others.

Three aspects of validity are important for consideration. These are face validity, criterion validity and construct validity.

3.2.1: Face validity

This means that an indicator seems to reasonably measure some variable. That is people with experience in the field of politics in terms of issues in political evaluation can be asked to tell, if on the face of it, a measure seems to reflect the concept in question.

3.2.2: Criterion Validity

This is the extent to which a measure relates to some external criterion. This can be achieved by relating a measure to another measure that is known to be valid. If the data collected using the instrument under consideration closely matches the data collected using the instrument assumed to be valid, then the researcher must conclude that the instrument is also valid.

3.2.3: Construct Validity

This refers to the accuracy of the operational definition of variables.   It is based on the logical relationships among variables.   According to Bless *et al* (2006:159), it is the extent to which scores on an instrument reflect the desired construct rather than some other construct.

Self-Assessment Exercise
1) Explain reliability.
2) Identify and discuss two forms of validity.

4.0:   SUMMARY

Reliability and validity in quantitative research represent two important issues in any scientific research. Reliability deals with the question whether the result is replicable. Validity has to do with whether the means of measurement (the measure) is accurate and whether they are actually measuring what they are intended to measure. A measure may enjoy face validity construct validity, and criterion validity.

5.0:   CONCLUSION

In this unit, we have looked at the concepts of reliability and validity as they relate to quantitative research. It is important to note that these concepts of reliability and validity are viewed differently by qualitative researchers, who strongly consider these concepts when strictly defined in quantitative terms as inadequate. In other words, these terms as defined in quantitative terms may not apply to the qualitative research paradigm. The question of replicability in the results does not concern them but precision, credibility, and transferability provide the lenses of evaluating the findings of qualitative research. In this context, the two research approaches or perspectives are essentially different paradigms (Winter, 2000, Hoepf, 1997, Glesne & Peshkin, 1992, Kuhn, 1970).

6.0:   TUTOR - MARKED ASSIGNMENT

1) Explain the difference between a measure and an indicator.

2) Analyse the main criteria for evaluating measurement validity.

7.0: REFERENCES/FURTHER READING

Babbie, E. (2007) *The Practice of Social Research* (10th ed). Belmont: Wadsworth Pub.Co.

Baker, Therese 1999. *Doing Social Research*. New York: McGraw-Hill College.

Black, T (1999) *Doing Quantitative Research in Social Sciences: An integrated approach to research design, measurement and statistics*, London: Sage.

Bless, Claire; Craig Higson-Smith and Ashraf Kagee 2006. *Fundamentals of Social Research Methods: An African Perspective* (4th edition). Cape Town: Juta & Co. Ltd.

Bryman, Alan (2004) (2nd Edition) *Social Research Methods*, Oxford: Oxford University Press

Cosby, Paul C. 2006. *Methods in Behavioural Research*. New York: McDraw-Hill

Creswell, J (2003) *Research Design: Qualitative, Quantitative, and Mixed Methods Approaches*, (2nd ed.), Thousand Oaks, CA: Sage Publications.

Hoover, Kenneth and Todd Donovan 2004. *The Elements of Social Scientific Thinking.* Belmont, CA: Wadsworth.

Jones, E. Terrence 1971. *Conducting Political Research*. New York: Harper and Row Publishers

King, Gary; Robert Keohane, and Sidney Verba. 1994. *Designing Social Inquiry* Princeton: Princeton University Press.

Kirk, J., & Miller, M. L. (1986). *Reliability and validity in qualitative research.* Beverly Hills: Suyvi Publications.

McNabb, David E. 2004. *Research Methods for Political Science: Quantitative and Qualitative Methods.* New York: M.E. Sharpe , Inc.

Meier, Kenneth J.; Jeffrey L. Brudney and John Bohte.2006. *Applied Statistics for Public and Nonprofit Administration* (6th edition). Belmont, CA: Thomson Higher Education.

Mitchell, Mark L. and Janina M. Jolley 2007. *Research Design Explained* (6th edition). Belmont, CA: Thomson Higher Education.

Patton, M. Q. 1990. *Qualitative evaluation and research methods.* (2nd ed.). Thousand Oaks, CA: Sage.

Taylor, Steven J. and Robert Bodgan 2006. *Introduction to Qualitative Research Methods: A Guide Book and Resource.* New York: John Riley and Sons Inc.

UNIT 5:   RELIABILITY AND VALIDITY IN QUALITATIVE RESEARCH

CONTENT

1.0:   INTRODUCTION

In the previous unit, we discussed the important concepts of reliability and validity in measurement. Our focus was on quantitative research. These concepts are not applicable only to quantitative evaluation. They relate to qualitative evaluation as well. However, they have different implications in qualitative research. In this unit, we will examine reliability and validity in qualitative political evaluation.

2.0:   OBJECTIVES

At the end of this unit, you are expected to be able to

1. understand how the concepts of reliability and validity are applied to qualitative research;

2. evaluate the relevance of this concepts to qualitative research;

3. acquire the ability to maximise validity and reliability in qualitative research.

3.0: MAIN CONTENT

3.1: Reliability in Qualitative Research

Although the term 'reliability' is a concept used for testing or evaluating quantitative research, the idea is most often used in all kinds of research. If we see the idea of testing as a way of information elicitation then the most important test of any qualitative study is its quality. A good qualitative study can help us "understand a situation that would otherwise be enigmatic or confusing" (Eisner 1991: 58). Patton (2001), states that validity and reliability are two factors, which any qualitative researcher should be concerned about while designing a study, analysing results and judging the quality of the study. This corresponds to the question that "how can an inquirer persuade his or her audiences that the research findings of an inquiry are worth paying attention to?" (Lincoln & Guba, 1985: 290).

To answer to the question, Healy and Perry (2000) assert that the quality of a study in each paradigm should be judged by its own paradigm's terms. For example while the terms reliability and validity are essential criteria for quality in quantitative paradigms, in qualitative paradigms the terms credibility, neutrality or conformability, consistency or dependability and applicability or transferability are to be the essential criteria for quality (Lincoln & Guba, 1985). To be more specific with the term of reliability in qualitative research, Lincoln and Guba (1985, : 300) use "dependability" in qualitative research which closely corresponds to the notion of "reliability" in quantitative research. They further emphasize, "inquiry audit" (p. 317) as one measure which might enhance the dependability of qualitative research. This can be used to examine both the process and the product of the research for consistency (Hoepfl, 1997). In the same vein, Clont (1992) and Seale (1999) endorse the concept of dependability with the concept of consistency or

reliability in qualitative research. The consistency of data will be achieved when the steps of the research are verified through examination of such items as raw data, data reduction products, and process notes (Campbell, 1996).

To ensure reliability in qualitative research, examination of trustworthiness is crucial. Seale (1999:200), while establishing good qualitative studies through reliability and validity in qualitative research, states that the "trustworthiness of a research report lies at the heart of issues conventionally discussed as validity and reliability" . When judging (testing) qualitative work, Strauss and Corbin (1990) suggest that the "usual canons of 'good science'...require redefinition in order to fit the realities of qualitative research" (p. 250).

In contrast, Stenbacka (2001) argues that since reliability issue concerns measurements then it has no relevance in qualitative research. She adds the issue of reliability is an irrelevant matter in the judgment of quality of qualitative research. Therefore, if it is used then the "consequence is rather that the study is no good" (p. 552).

To widen the spectrum of conceptualization of reliability and revealing the congruence of reliability and validity in qualitative research, Lincoln and Guba (1985) states that: "Since there can be no validity without reliability, a demonstration of the former [validity] is sufficient to establish the latter [reliability;]" (p. 316). Patton (2001), with regards to the researcher's ability and skill in any qualitative research, also states that reliability is a consequence of the validity in a study.

3.2: Validity in Qualitative Research

The concept of validity is described in a wide range of terms in qualitative studies. This concept is not a single, fixed or universal concept, but "rather

a contingent construct, inescapably grounded in the processes and intentions of particular research methodologies and projects" (Winter, 2000, p.l). Although some qualitative researchers have argued that the term validity is not applicable to qualitative research, but at the same time, they have realised the need for some kind of qualifying check or measure for their research. For example, Creswell & Miller (2000) suggest that the validity is affected by the researcher's perception of validity in the study and his/her choice of paradigm assumption. As a result, many researchers have developed their own concepts of validity and have often generated or adopted what they consider to be more appropriate terms, such as, quality, rigor and trustworthiness.

The discussion of quality in qualitative research emanated from the concerns about validity and reliability in quantitative tradition which involved substituting new term for words such as validity and reliability to reflect interpretive [qualitative] conceptions. Stenbacka (2001) argues that the concept of validity should be redefined for qualitative researches. The notion of reliability is one of the quality concepts in qualitative research, which is to be solved in order to claim a study as part of proper research.

In searching for the meaning of rigor in research, Davies and Dodd (2002) find that the term rigour in research appears in reference to the discussion about reliability and validity. Davies and Dodd (2002) argue that the application of the notion rigor in qualitative research should differ from those in quantitative research by "accepting that there is a quantitative bias in the concept of rigour, we now move on to develop our re-conception of rigor by exploring subjectivity, reflexivity, and the social interaction of interviewing" (p. 281).

Lincoln and Guba (1985) argue that sustaining the trustworthiness of a research report depends on the issues quantitatively discussed as validity and reliability. The idea of discovering truth through measures of reliability and validity is replaced by the idea of trustworthiness (Mishler, 2000), which is "defensible" (Johnson 1997, 282) and establishing confidence in the findings (Lincoln & cuba, 1985). If the issues of reliability, validity, trustworthiness, quality and rigor meant differentiating a 'good' from 'bad' research then testing and increasing the reliability, validity, trustworthiness, quality and rigor will be important to the research in any paradigm.

3.3:   Testing Validity and Reliability

So far, the concepts "of reliability and validity as they have been redefined for their usefulness in qualitative research have been presented. Now, the question, which remains to be answered, is 'How to test or maximize the validity and reliability of a qualitative study?' If the validity or trustworthiness can be maximized or tested then more "credible and defensible result" (Johnson, 1997; 283) may lead to generalizability which is one of the concepts suggested by Stenbacka (2001) as the structure for both doing and documenting high quality qualitative research. Therefore, the quality of a research is related to generalizability of the result and thereby to the testing and increasing the validity or trustworthiness of the research.

In contrast, Maxwell (1992) observes that the degree to which an account is believed to be generalizable is a factor that clearly distinguishes quantitative and qualitative research approaches. The ability to generalize findings to wider groups and circumstances is one of the most common tests of validity for quantitative research. In this sense, the validity in quantitative research is very specific to the test to which it is applied - while triangulation methods

are used in qualitative research. Patton (2001) states that generalizability is one of the criteria for quality case studies depending on the case selected and studied. Triangulation is typically a strategy (test) for improving the validity and reliability of research or evaluation of findings.

Mathison (1988:13) elaborates this by saying: ''Triangulation has arisen as an important methodological issue in naturalistic and qualitative approaches to evaluation [in order to] control bias and establishing valid propositions because traditional scientific techniques are incompatible with this alternate epistemology.'' Triangulation is defined to be "a validity procedure where researchers search for convergence among multiple and different sources of information to form themes or categories in a study" (Creswell & Miller 2000: 126). According to Patton (2001:247) "triangulation strengthens a study by combining methods.

This means using several kinds of methods or data, including using both quantitative and qualitative approaches". To acquire valid and reliable conclusions, multiple methods of searching or gathering data are in order. Engaging multiple methods, such as, observation, interviews and recordings will lead to more valid, reliable and diverse construction of realities. Triangulation may include multiple methods of data collection and data analysis, but does not suggest a fix method for all the researches. The methods chosen in triangulation to test the validity and reliability of a study depend on the criterion of the research.

Self-Assessment Exercise

1) Select a research article and discuss the reliability and validity of its argument.

4.0: SUMMARY

Reliability in qualitative research is about dependability and trustworthiness of research process and outcome. It relates to the rigour and thoroughness involved in the research. Validity in qualitative research is about the control of bias. A very important way to ensure reliability and validity in qualitative evaluation is to adopt multiple methods of searching or gathering data. Engaging multiple methods, such as, observation, interviews and recordings will lead to more valid, reliable and diverse construction of realities. This process is called triangulation.

5.0: CONCLUSION

Reliability and validity are conceptualized as trustworthiness, rigor and quality in qualitative paradigm. Validity and reliability of qualitative research is aimed to eliminate bias and increase the researcher's truthfulness of a proposition about some political phenomenon using triangulation.

6.0: TUTOR-MARKED ASSIGNMENT

1) Explain triangulation in qualitative research.

2) Discuss reliability and Validity as a challenge to qualitative research.

3) What are the advantages of the qualitative method?

7.0:   REFERENCES/FURTHER READING

Babour, R. S. (1998). Mixing Qualitative Methods: Quality Assurance or Qualitative Quagmire? *Qualitative Health Research, 8(3),* 352-361.

Bogaan, R. C. &,Biklen, S. K. (1998). *Qualitative Research in Education: An Introduction to Theory and Methods* (3rd ed.). NecuWm Heights, MA: Allyn & Bacon.

Campbell, T. (1996). Technology, Multimedia, and Qualitative research in education. *Journal of Research on Computing in Education, 30(9),* 122-133.

Charles, C. M. (1995). *Introduction to educational research (2 ed.).* San Diego, Longman.

Clont, J. G. (1992). *The concept of reliability as it pertains to data from qualitative studies.* Paper Presented at the annual meeting of the South West Educational Research Association. Houston, TX. 605 *The Qualitative Report* December 2003

Creswell, J. W. & Miller, D. L. (2000). Determining validity in qualitative inquiry. *Theory into Practice, 39(3),* 124-131.

Crocker, L., & Algina, J. (1986). *Introduction to classical and modern test theory.* Toronto: Holt, RineHart, and Winston, Inc.

Kirk, J., & Miller, M. L. (1986). *Reliability and validity in qualitative research.* Beverly Hills: Suyvi Publications.

Maxwell, J. A. (1992). Understanding and validity in qualitative research. *Harvard Educational Review, 62(3),* 279-300

Patton, M. Q. (2002). *Qualitative evaluation and research methods* (3rd ed.). Thousand Oaks, CA: Sage Publications, Inc.

Stenbacka, C. (2001). Qualitative research requires quality concepts of its own. *Management Decision, 39*(7), 551-555.

Wainer, H., & Braun, H. I. (1988). *Test validity.* Hilldale, NJ: Lawrence Earlbaum Associates. *Nahid Golafshani* 606

Winter, Glyn 2000. A Comparative Discussion of the Notion of 'Validity' in Qualitative and Quantitative Research The Qualitative Report, 4(3 & 4) Available: http://www.nova.edu/ssss/QR/QR4-3/winter.html

MODULE 3: THE SURVEY PROCESS

Unit 1:   The Survey

Unit 2:   The Survey Instrument: Questionnaire

Unit 3:   Data and Data Gathering Techniques

Unit 4:   Population and Sample
Unit 5:   Non-Probability Sampling
Unit 6:   Probability Sampling

UNIT 1:   THE SURVEY

CONTENT

1.0:   Introduction
2.0:   Objectives
3.0:   Main Content
   3.1   What is a survey?
   3.2:   Steps in the Survey Process
   3.3:   Advantages of the Survey
   3.4:   Limits of Survey Research
   3.5:   Error in Survey Research
4.0:   Summary
5.0:   Conclusion
6.0:   Tutor Marked Assignment
7.0:   References/Further Reading

1.0:   INTRODUCTION

This unit focuses on the survey method. The survey method is a very important method or research that is commonly used for political evaluation. This unit explains the reasons for conducting surveys. It explains the survey process and the limits and advantages of the use of survey for political evaluation.   The instrument used for conducting survey research: the questionnaire is also explored.

2.0: OBJECTIVES

In this unit, you are expected to:

1. understand the survey as important research method;

2. discuss reasons for conducting survey research, the survey process and the limits of survey research; and

3. explain the advantages and limits of the application of survey research.

3.0   MAIN CONTENT

3.1: What is a survey?

A survey is the systematic observation and recording of human behaviour in social systems for developing and testing social theories. Social survey covers a wide variety of enquiry. It might be the study of features (e.g. demographic characteristics, socio-economic characteristics) political and economic behaviour and study of attitude and opinion. Consequently, the activity of social scientists, market researchers and opinion pollsters and other similar enquirers are included in the social surveys.

Survey research focuses on populations or the universe. Data are collected from the population for intensive study and analysis. More often than not, the researcher finds that he cannot possibly study all the subjects or items in the population. Hence, the survey researcher selects a sample from or a subset of the population using some techniques of sampling which we shall discuss later.

Survey research is often referred to as sample survey because it involves large and small populations where samples are selected and studied in

order to discover relative incidence, distribution and interrelations of sociological and psychological variables. The sample selected for survey study should be an adequate representative of the totality of the individuals, objects, events and situations.

Babbie, (2007) describes survey as a method used in studying a segment of a population for making estimated assertions about the nature of the total population from which the sample has been selected.

As said earlier, survey research has to do with finding out the attitudes, ideas and disposition of a particular population. In survey, not everyone in a population can be studied. That is why the researcher must sample (select a portion of) from the population.

The increased use of survey has created changes in the way they are conducted and reported. More attention is now given to sample selection, questionnaire designs and error rates. This means surveys require careful planning and execution.

3.2: Steps in the Survey Process

The steps below are usually taken in a survey research.

1. Plan and design your survey: determine the issues(s) to be researched. State your objectives clearly. Plan a schedule for your survey. Define the population and estimate the sample size. Select your methods of data collection. Design and pre-test the questionnaire. Develop questions and devise answer alternatives for closed questions. Administer questionnaire/schedule to sample. Follow up non-respondents at least once. This aspect of the research will benefit from review literature/theory relating to topic/area.

2.	Data Collection and Organisation. To ensure that you have clean, unbiased and up-to-date data, you need to decide on the appropriate method to collect the data. Decide on mode of administration (face to face; telephone; postal; e-mail; web). Pay attention to the issues of non-response the entire questionnaire and item-non response in the questionnaire.

3.	Data Access: Transform completed questionnaires/ schedules into computer readable data (coding). That is read the data into analytic software for further processing. The software for data entry should be compatible with the analysis software to make this step easy. Enter data into statistical analysis program like Statistical Package for the Social Sciences (SPSS).

4.	Data Preparation and Management: At this stage, you extract useful information from your collected data. Try to understand the data. This involves setting up the code book, setting up multiple item indices and scales. The essence is to get the data ready for analyses.

5.	Data Analysis: Analysis involves extracting useful information to make informed decisions or conclusions. Summarise your data and secure accurate description of the variables of interest.

6.	Reporting: You report the results of your analysis.

7.	Submission or dissemination of report. Here the results of the survey are made available to the appropriate audiences.

## 3.3:	Advantages of Survey Research

1.	Survey instruments can be used to investigate problems in realistic settings (i.e. when you are engaged in experimental research you are not dealing with realistic setting rather you simulate or create

environment for such people). Newspaper reading, television viewing, radio listening and similar consumer behaviour pattern can be examined where they happened rather than in a laboratory or screening room under artificial condition.

2.   The cost of survey is reasonable considering the amount of information gathered.

3.   Large amount of data can be collected with relative ease from a variety of people. The survey technique allows the researcher to examine many variables-demographic, lifestyle information, attitudes motives, intention etc and to use multivariate statistics. Data which are helpful to survey research already exist in some cases. So one can go and collect the data where they exist or go and collect the data in research centres. It is important to clearly acknowledge the source of data whenever existing data are used.

3.4:   Limits of Survey Research

Survey research is not a perfect research methodology. When you are conducting an experiment in a laboratory, you eliminate variables capable of altering the result but in a survey research, you carry out the research in a natural environment.

1. The first (most) severe disadvantages is that independent variables cannot be manipulated as in laboratory. Without control of independent variables, the researcher cannot be certain whether the relationship between independent and dependent variables are causal or non causal(i.e. a survey may establish that A and B are related but it is impossible to determine that A solely causes B). In other words, in survey research,

causality is difficult to establish because many intervening extraneous variables are involved.

2. The wording (framing) and placement of items within a questionnaire can have a wrong effect on survey result (i.e the type of word you choose may create some bias). Again, great care must be taken to ensure that the data collected as well as the instrument are reliable and valid.

3. If the sample is largely unrepresentative, the result will have little relevance to other situations even though the sample size is fairly large. Surveys based on representative samples however produce reliable (useful) information. They are especially useful for collecting information on audiences and readership. However, they are not particularly suited for testing causal hypothesis.

3.5: Error in Survey Research

We can think of 'error' as a term that has been employed on a number of occasions, as being made up of four main factors.

- Sampling error. This kind of error arises because it is extremely unlikely that one will end up with a truly representative sample, even when probability sampling is employed.

- We can distinguish what might be thought of as sampling -related error. This is error that is subsumed under the category non-sampling error. But that arises from activities or events that are connected with the issue of generalizability or external validity of findings. Examples are an inaccurate sampling frame and non-response.

- There is also error that is connected with the implementation of the research process. We might call this data collection error. This type of error may result from poor question wording in self-completion questionnaires or structured interviews: poor interviewing techniques and flaws in the administration of research instruments.

- Finally, there is data processing error. This arises from faulty Management of data, in particular, errors in the coding of answers.

Types/ Forms of Error in Survey Research

| Sampling error | sampling -related error | Data collection error | Data processing error |
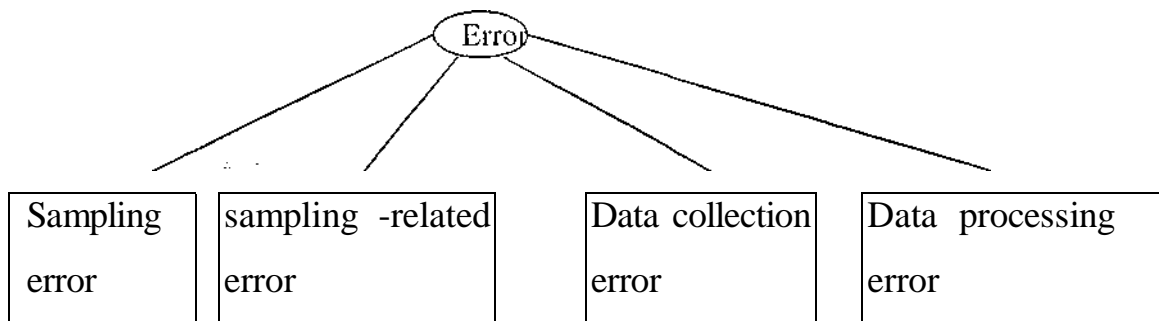
Self-Assessment Exercise

1) List the major steps in a survey.
2) What are the advantages of survey research?
3) State the limitations of survey research.

4.0: SUMMARY

Survey research involves large and small populations where samples are selected and studied in order to discover relative incidence, distribution and interrelations of socio-political variables. The sample selected for survey study should be an adequate representative of the totality of the individuals, objects, events and situations. The chief

instrument of the survey is the questionnaire. The questionnaire is used to generate responses to specific questions, including attitudes, opinion, motivations, knowledge or demographics. The administration of the questionnaire may be done by mail, personal interviews or telephone interviews.

5.0   CONCLUSION

The survey method is one of the most popular forms of data collection in political evaluation. It is particularly suitable when the data is to be derived from a large population. Surveys however suffer from some limitations. Thus, efforts must be made to minimise these errors in the way questions are structured and the survey instrument administered.

6.0   TUTOR - MARKED ASSIGNMENT

1.   Explain a survey.

2.   List the advantages and disadvantages of using the survey method.

3.   Describe the basic techniques of survey data collection.

7.0   REFERENCES/FURTHER READING

Alan, Bryman (2004). *Social research methods*. Second Edition: Oxford university Press.

Asika, Nnamdi (1991). Research methodology in the behavioural science: Longman Nigeria Plc.

Babbie, E. (2005) *The Practice of Social Research* (10th ed). Belmont: Wadsworth Pub.Co.
Baker, Therese 1999. *Doing Social Research*. New York: McGraw-Hill College.

UNIT 2:   THE SURVEY INSTRUMENT: QUESTIONNAIRE

CONTENT

1.0:   Introduction
2.0:   Objectives
3.0:   Main Content
    3.1   What is a questionnaire?
    3.2:   Elements of the questionnaire
    3.3:   Questionnaire construction procedure
4.0:   Summary
5.0:   Conclusion
6.0:   Tutor - Marked Assignment
7.0:   References/Further Reading

1.0   INTRODUCTION

The essence of instrument construction is the systematic collection of information and systematic analysis of such information. Research instruments vary as the research methods vary. For survey, we use questionnaires, interview guide (also called interview schedule) and focus group discussion guide.

1.0   OBJECTIVES

In this unit, you are expected to:

1. understand the questionnaire as a major instrument used in survey research;

2. list the various elements of the question, construct a survey questionnaire;

3. master the techniques of asking and organising the questions in your questionnaire; and

4. be familiar with issues associated with the administration of the questionnaire.

3.0   MAIN CONTENT

3.1   What is a Questionnaire?

The questionnaire is an organized assemblage of questions/items related to the problem of the study. The word 'questionnaire' refers to a collection or a form, containing a set of questions addressed to a statistically significant audience for which responses or information are elicited to a survey. Items come in 2 main forms:

- Structured items
- Unstructured

Structured items are those items for which possible responses are provided by the researcher. That is, the researcher gives a list of choices from which the respondent is expected pick whichever best represents his opinion. For example:

i.   SEX: a) MALE ( )

b) FEMALE ( )

Such items are easy to fill in and to analyze.

Unstructured items are those for which the respondent is given the free hand to react to the questions being asked in his own words. However, at the time of analysis, categories will have to be created for the responses.

The questionnaire helps to do the following

1. Validate the authenticity of a statement and findings for generalisation

2. Elicit information that are needed from a large audience or participants to facilitate statistical analysis of results from an opinion survey;

3. Test hypothesis to establish the continuous relevance of a theory;

4. Corroborate other findings;

5. Complement findings from other diagnosis and , or clinical findings;

6. Protect the identity or privacy of the respondents;

7. It may also be used where rigorous experimentation is not feasible; and

8. It is used for a small sample to save resources and time.

Elements of the Questionnaire

Demographic Items

These items deal with specific socio-economic status and personal data of the respondent. They cover issues like sex, age, educational qualification, income, and other demographic issues. Some researchers or respondents keep these questions/items last because they are personal.

Information Items

These questions test the respondents' knowledge of things. They relate to specific events or issues about which concrete ideas are expected. These always show the extent to which the respondent is informed about the subject matter under investigation.

Opinion and Attitudinal Items

These seek to elicit information about individuals' opinion, attitude and predispositions to certain phenomena.

Cheater Items

These are those items that are meant to reveal inconsistency or dishonesty on the part of a respondent.

3.3   The Questionnaire Construction procedure

In constructing a questionnaire, the researcher must follow a systematic procedure in order to achieve three broad objectives:

1. Successfully gather information that answers each study question

2. Motivate the respondents to answer all questions to the best of his or her ability.

3. Keep all potential error to a minimum.

1. Establish what information must be gathered.	4 select the question type and decide on question structure.	7. Pretest the questionnaire and revise if necessary.
2. Specify the appropriate data-gathering method.	5. Select the best working for each individual question.	8. administer the questionnaire and process the data.
3. Specify the procedure for analysing the gathered data.	6. Organise the questions into a logical order.	

You must bear the following in mind while constructing your questionnaire:

1. Every item on your questionnaire should be tied to the research problem/questions. If a question/item is not, delete it.

2. Demographic questions are NOT compulsory. If they are not connected with your questions, delete them.

3. In wording your questionnaire items, observe the principles of:

a. SIMPLICITY (items should be easy to understand).

b. COHERENCE (sensible, logical)

c. UNITY (sequential, internally united).

(4) Do you like classical music, jazz or Times *magazine?* Is it internally disunited?

4. Know that you respondents most likely know much less than you do. So, come down to their level.

5. Avoid bias in your framing of questions e.g. 'People have said that the removal of Evan Enwerem as Senate President is justifiable'. This is clearly biased. A better way to cast the question is this: What do you think about the removal of the former Senate

President?

6. Avoid double-barrelled questions. E.g. Do you like Fanta or Coke better?

7. A questionnaire should not be too long. Once you have dealt with the entire relevant questions, stop. The shorter a questionnaire is the better. Some research experts think a questionnaire should not exceed 30 items. If your questionnaire is too long, it will discourage respondents.

8. In the preamble to your questionnaire, do not spell out all you want to do. Do not say, "This questionnaire is out to measure the influence of alcohol on family unity." Instead, use the theme: This study deals with family-related issues.

Self-Assessment Exercise

1) Name the eight steps followed in questionnaire construction.
2) Describe the difference between close-ended and open-ended questions.

4.0   SUMMARY

The questionnaire is the most popular way of gathering primary data. It is most appropriate when research problems require a descriptive design. The questionnaire is a very flexible instrument and can be customised to meet the objectives of almost any type of research project. There are specific steps to be followed in constructing a questionnaire.   The questions in the questionnaire may be close-ended or open-ended.

5.0:   CONCLUSION

The survey is both a science and an art. In conducting a good survey, the researcher has to ensure that he put together a set of questions that get as precisely as possible the information he or she requires. The questions must be such that they are clearly

understood by the respondents. The researcher must also ensure that the questions constitute a logical whole and that they are pleasing enough for the respondent to be willing to spend time to complete it.

6.0: TUTOR-MARKED ASSIGNMENTS (TMAs)

1. Discuss factors a researcher should take into consideration when constructing a questionnaire.

2. Explain what you understand by survey error

3. List the steps in a survey process.

7.0: REFERENCES/FURTHER READING

Alan, Bryman (2004). *Social research methods*. Second Edition: Oxford university Press.

Babbie, E. (2005) *The Practice of Social Research* (10th ed). Belmont: Wadsworth Pub.Co.

Baker, Therese 1999. *Doing Social Research*. New York: McGraw-Hill College.

Asika, Nnamdi (1991) *Research Methodology in the Behavioural Science*. Ibadan: Longman Nigeria.

Neuman, W.L. (2001). *Social Research Methods: qualitative and quantitative approaches,* 4th edition. Massachusetts: Allyn and Bacon.

UNIT 3: DATA AND DATA GATHERING TECHNIQUES

CONTENT

1.0   INTRODUCTION

This unit deals with data types and data gathering methods. It explains the meaning of data as used in political evaluation .it also highlights and explains the various techniques of data collection.

2.0:   OBJECTIVES

In this unit, you are expected to be able to:

1. Explain the concept of data in political evaluation research;

2. List and discuss various techniques of data collection, document analysis and the field method of gathering data.

3.0   MAIN CONTENT

3.1:   Data and Types of Data Gathering

After the planning stage of a research, the next step for consideration is the data collection stage. There are various instruments and techniques for gathering data in social research. Data are the symbols, numbers and alphabetical characters used to describe one or more attributes such as age, sex, volume, growth rates, temperature, etc, of an entity or a social

phenomena. They are obtained by observing, counting, measuring, weighing, etc, which are the then recorded (Ogundipe *et al*., 2006:96).

It must also be added, as noted by Soyombo and Oyefara that in social research there are "quantitative and qualitative" data types (2007: 203). Quantitative data relates to figures and are numerical and mathematical/statistical in its procedures and uses such tools as chi-square, regression analysis, and standard deviation among other statistical tools. On the other hand, qualitative data as noted by Jegede entails use the of direct interaction, records, summaries, newspapers, popular literature, academic reviews and other sources (Jegede, 2006:114). It must be noted that while qualitative data may also involve numeric records , such are not subject to statistical analysis.

 Fagbohunge (2002:111) on his part identifies primary and secondary data. Primary data are raw data that a researcher goes out to directly collect himself while secondary data are those based on earlier works of other research findings as located in books, journals, annals, and other documented or recorded materials. In another vein, data could be cardinal of numerical, nominal or categorical, or even ordinal or continuous (Ogundipe et al, 2006:95). Cardinal data are also referred to as discrete data. They are "those which can only take certain values" as for instance the number of nodes on a cowpea plant which must be an integer such as 0, 1,2,3 and so on . Categorical data "are simply facts that can be sorted into classes and enumerated as colour and breed" while ordinal data "are those that have ordered relationship" (Ogundipe 2006:95-96). It must be however, added that qualitative data could be from primary or secondary sources (Jegede, 2006:114).

But regardless of its nature, data are the building blocks of information in social research. The terms data and information are often used interchangeably in everyday conversation as meaning the same things, but in practice, they have different meanings. Specifically, "data are the input raw materials from which information is produced" (Jegede 2006: 95). Therefore, information can be defined as data in textual, pictorial or vocal form that have been assembled, processed and interpreted in a meaningful way. Thus, the purpose of acquiring information is to increase knowledge, reduce uncertainty and facilitate decision-making in our everyday endeavours. Having said this, there are different techniques of data gathering. The next part of this unit discusses these techniques.

3.2:   Techniques of Data Gathering

There are different techniques of data gathering in social research. For instance, Soyombo and Oyefara (2007) identify two broad techniques. First, they identify qualitative techniques, which include census, sample survey (one-time/cross-sectional survey and longitudinal/multiple-time survey, experiments and simulation. Second, they identify qualitative techniques which includes observation (participant and non-participant observation), in-depth interview (IDI), key informant interview (KII), case study, focus group discussion (FDGs) and document studies (Soyinbo and Oyefara, 2007:203-207). On his part, Fagbohungbe identifies four techniques he considers "the most commonly used" to include questionnaire, interviews, observation and experimentation (Fagbohungbe, 2002:111-121). Yet Ogundipe *et al* identifies "three basic techniques" for data gathering: interviews (both face to face and via questionnaire), observation and examination of existing records (Ogundipe *et al*, 2006:98-99). However, in spite of these slight differences in the nomenclature of techniques identified

by the aforementioned writers, this paper interrogates these techniques within the context of experimentation, document analysis and field studies. These techniques will be discussed in sections as below.

3.2.1: Experimentation

Experimentation is a rigorous research design which attempts to gather or generate data in order to explain the cause of a social phenomenon in the real world (Isiugo-Abanihe 2002:57). In this method of data collection, a researcher sets up a controlled, quasi-artificial k laboratory research situation in which data is generated by observing the relationship between the two or more variables by deliberately manipulating one variable to see whether this produces a change in the phenomena under observation (Fagbohungbe, 2002:119). In experimentation, a real world situation or group is studied or measured, then new treatment or intervention is introduced, and the group is measured or studied the second time, noting the observed changes (Campbell and Stanley, 1963).

In experimentation, the manipulated variable is referred to as the independent variable because it is independently manipulated for the effects it has on the dependent variable. Experiments are useful here because they help to generate data that could assist to develop three crucial research evidences that are required to establish causation beyond reasonable doubt. First, it helps to establish evidence of concomitant variation between dependent and independent variables that suggest either that the two are associated or they are not associated. In other words, such evidence indicates the extent to which the variables concomitantly vary.

Second, it provides evidence of time-order, that such an association is temporarily continuous and that the presumed effect (dependent variable) did not occur before the presumed cause (independent variable). And thirdly,

it provide evidence of elimination of alternative explanations, to the effect those other factors that could be constructed as possible determining conditions of the dependent variable are eliminated from the research setting. The advantage of this technique is that it is the most effective in data collection as it offers maximum control over extraneous factors such as environmental factors that could contaminate the study (Fagbohungbe, 2002:119; Soyombo and Oyefara, 2007:204). But its limitation is located in the fact that its uses are still limited in the social sciences mainly in Psychology. In addition, not only is experiment suitable for small groups as opposed to large groups, but also the artificial environments are often believed to alter social behaviour (Soyombo and Oyefara, 2007:204).

3.2.2: Document Analysis

Unlike the experiment action technique, this is a technique by which a researcher extracts data from records and documents (such as print and electronic, audio and visual, published and broadcast). This section examines three types of document analysis as techniques of data gathering. These include: content analysis; analysis of existing statistics and comparative and historical research.

*(*a*)* Comparative and Historical Analysis

Social scientists use comparative and historical methods to discover patterns in the histories of different cultures. Although often regarded as a qualitative method, comparative and historical research can make use of quantitative techniques. The basic purpose of this technique is to enable the researcher to reconstruct the past systematically and objectively through the collection, evaluation, verification and synthesizing of record evidence in order to establish facts and reach defensible conclusions as required in relation to research questions, objectives and hypotheses.

Sources of comparative and historical and data are numerous. To begin with, historians may have already reported on a subject of research, and their analysis can give a researcher an initial grounding on the subject. In addition to personal sources, there are pubic records which are also revealing of family history and actions of a people. Beside, newspapers are especially rich in evidence on the educational, legal and recreational aspects of family life in the past as seen from a local point of view.

Magazines reflect more general patterns of family life, and as noted by Babbie, students often find them interesting to explore for data on perception and expectation of mainstream family values (Babbie, 2007:344). Also organizations generally document themselves, so that if one is studying the development of some organizations, one can examine its official documents, charters, policy statements, speeches by leaders, and so on. Similarly, official government documents provide the data needed for analysis. However, whatever resources the researcher uses, a note of caution is needed. One cannot trust the accuracy of existing historical data, official or unofficial, primary or secondary. Therefore, one needs to beware of biases in historical data sources.

(b)   Content Analysis

Content analysis is a social research method appropriate for studying human communications through social artefacts. Researchers can use it to study not only communication processes but also other aspects of social behaviour as well (Babbie, 2007:345). Content analysis, according to Agbaje and Alarape (2005:29) involves the objective, systematic, often quantitative use of manifest materials and documents to generate data. The method enables the researchers to distil from manifest, content elements of latent content,

influencing factors and intents of the material in question. No doubt, the researcher is often interested in the forces behind the content, but he codes contents only in terms of what he sees.

Content analysis is objective in that it prescribes that categories used to collect data must be defined so precisely that different researchers can analyse the same content using these definitions and arrive at the same results. It is systematic because it insists that the selection of content to be analysed must be based on an informal, pre-determined, unbiased plan. In effect, therefore, content analysis helps in:

- The study of attributes of content;
- Drawing conclusions about sources of content;
- Drawing conclusions about context, target and audience of content, and
- Drawing conclusions about intent of content (Agbaje and Alarape, 2005:30).

Agbaje and Alarape identifies two broad types of content analysis. These are; (i) Analysis of "What categories (ii) Analysis of "How" categories. Analysis of "What'' categories focuses on the substance. Second, analysis of "How" categories focus on the form of the content. However, common units of analysis in content analysis include elements of communication such as words, paragraphs, books, and so forth. Content analysis involves coding transforming raw data into categories based on some conceptual scheme. Coding may attend both manifest and latent content. Both quantitative and qualitative techniques are appropriate for interpreting content analysis data. The advantages of content analysis include economy, safety, and the ability to study processes occurring over a long time. Its disadvantages are that it is

limited to recorded communications and can raise issues of reliability and validity.

Bailey (1978: 276 -288) suggests two types of content analysis. These are the relatively unstructured and non-quantitative case-study approach, and the structured content-analysis approach that yields quantitative data. The goal of the latter content analysis is to "take a verbal, non-quantitative document and transform it into quantitative data". In such situations, the researcher may engage in counting how often certain words or phrases appear, identifying preferences, and patterns of voting and so on.

3.3:   Field Methods

Qualitative field research methods enable researchers to observe social life in its natural habitat: to go where the action is and watch. Thus, this type of research technique can produce a richer understanding of many social phenomena than can be achieved through other observational methods, if the researcher observes in a deliberate, well-planned and active way (Babbie, 2005). In field research observation, data processing and analysis are interwoven. Field methods are defined in terms of where much of the data collection associated with their application takes place - literally in the field (and not in libraries or laboratories). In essence, field methods involve the collection of data in the field. It involves the study of human institutions, characteristics or behaviour as they occur in their natural settings. (Agbaje and Alarape, 2005:32).

For a researcher that adopts field methods to collect data, the goal should not be to draw conclusion about cause-effect relationships, since that would be impossible to attain through collection of data in more or less natural settings. Rather, the goal is more of establishing co-relationship; that is to see the degree to which two variables co-related. Field methods allow for

more natural conditions of selection and exposure. There are four basic types:-

(a) Observation Method: direct observation, indirect observation, participant observation, non-participant observation, controlled observation, and uncontrolled observation.

(b) Interview Method: This could be in form loosely structured interview, highly structured interview, open interview, closed interview, face-to-face interview, telephone interviews, oral interviews, internet interviews, Focused Group Discussion (FGD), panel studies, and elite Interview (Agbaje and Alarape, 2005:33).

(c) Questionnaire (group questionnaire, privately filled face-to-face questionnaire, mail questionnaire, electronic questionnaire; and the combined method).

(d) The combined method is a combination of the preceding methods. However, according to Agbaje and Alarape (2005), highly structured interviews, the measuring instrument, called the ''interview schedule'', is filled by the researcher or his/her field assistant, whereas the questionnaire is filled by respondents (research subjects). This basic difference has implications for the construction of measuring instruments, since the questionnaire has to contain more instruction on how it is to be filled than the interview schedule.

Field observation technique of data collection differs from some other models of observation in that it is not just a data collecting activity. Frequently, perhaps typically, it is a theory generating activity as well. More typically, the researcher will attempt to make sense out of an ongoing process that cannot be predicted in advance making initial observations, developing tentative general conclusions that suggests particular types of

further observations, making these observations and thereby revising your conclusions, and so forth. In field method, observers can play any of several roles, including participating in what they want to observe. Hence, in most studies, the 'term' field research is used rather than the frequently used term 'participant 'because the field researcher need not always participate in what they are studying, though they usually will study it directly at the scene of the action (Babbie, 2007).

It must be stated that survey research, a popular social research method, is the administration of questionnaires to a sample of respondents selected from some population. Survey research is especially appropriate for making descriptive studies of large populations. It may be used for explanatory purposes as well. Questionnaires provide a method or technique of collecting data by; (i) asking people questions or (ii) asking them to agree or disagree with statements representing different point of view. Questions may be open-ended or closed-ended.

In an interview survey, the essential characteristic of interviewer is that he or she be neutral; their presence in the data collection process must have no effect on the responses given to questionnaire items. Interviewers must be carefully trained to be familiar with the questionnaire, to follow the question wording and question order exactly, and record responses exactly as they are given. Interviewers can use probes to elicit an elaboration on an incomplete or ambiguous response. Probes should be neutral. Survey research in general offers advantages in terms of economy, amount of data that can be collected, and the chance to sample a large population. The standardization of the data collected represents another special strength of survey research. However, survey research has several weaknesses: it is somewhat artificial, potentially superficial and relatively inflexible. Using surveys to gain full sense of

social process in their natural settings is difficult. In general, survey research is comparatively weak on validity and strong on reliability.

Self-Assessment Exercise

1) Describe the importance of data in political evaluation.
2) List the various data collection methods that you know.

4.0   SUMMARY

Data are the symbols, numbers and alphabetical characters used to describe one or more attributes such as age, sex, volume, growth rates, temperature, etc, of an entity or a social phenomenon. They are obtained by observing, counting, measuring, weighing, and so on and then recorded. Data could be obtained by observation method, experimentation, field methods such as Interview Method ad questionnaire administration.

5.0:   CONCLUSION

This unit has located data gathering techniques under three headings: experimentation, document analysis and field methods. It must however be concluded that not only are the aforementioned data gathering techniques different in terms of procedures but they also vary in terms of advantages and disadvantages. Yet all of them have a peculiar usefulness to political evaluation research.

As earlier stated, there are different techniques for data gathering in social research. These include census, sample survey (one time/cross-sectional survey and longitudinal/multiple-time survey), experiments, simulation, observation (participant and non-participant observation), in-depth interview (IDI), key informant interview (KII), case study, focus group discussion (FDGs), document studies, questionnaire, interviews. These techniques as for instance discussed by Soyombo and Oyefara (2007), Ogundipe *et al*

(2006) and Fagbohungbe (2002) could also be located under different broad categories.

6.0: TUTOR – MARKED ASSIGNMENT

1.   Discuss experiment as a data gathering techniques.

2.   List and discuss two field methods of data gathering.

3.   What is content analysis?

7.0 REFERENCES/FURTHER READING

Agbaje, A & Alarape, A. (2005). Introductory Lectures on Research Methodology. Faculty of the Social Sciences, University of Ibadan.

Babbie, E. (2007). *The Practice of Social Research*. Wadsworth: Cengage Learning.

Bailey, Kenneth 1978. *Methods of Social Research*. New York: Free Press.

Campbell, D.T. & Stanley, J.C. (1963). *Experimental and Quasi-Experimental Designs for Research*. Chicago: Rand McNdly.

Fagbohunbe, F.O. (2002) .*Research Methods for Tertiary Institutions* (2nd edn.) Lagos: KOTLEB Publishers.

Jegede, A.S. (2006). "Analysis of Qualitative Data." In A. I. Olayinka; V.O. Taiwo; A. Raji Oyelade; & I.P. Farai (eds) *Methodology of Basic and Applied Research. Ibadan:* The Postgraduate School University of Ibadan. Pp. 113-132.

Isiugo-Abanile, U.C. (2002). "Quantitative Research Techniques. "In U.C. Isiugo-abanihe; N. Austin; N. Isamah & J.O. Adesina (eds) *Currents and Perspectives in Sociology*. Lagos: Malthouse press Limited.

Ogundipe, G.A., Eucas, E.O., & Sanni, A.I. (2006). "*Systematic Collection of Data.*" In A.I. Olayinka; V.O. Taiwo; A. Raji-Oyelade; & I.P. Farai (eds) *Methodology of Basic and Applied Research. Ibadan: The Postgraduate School University of Ibadan*. Pp. 95-112.

Soyombo, O. & Oyefara, J.L. (2007). "*Research Methods in Social Sciences*. "In L. Olurode & K. Fagbohungbe (eds) Readings in Public and International Affairs. Laos: MPIA, pp. 199-220.

UNIT 4: POPULATION AND SAMPLE

CONTENT

1.0   INTRODUCTION

This unit deals with the sampling as important aspect of research, especially survey research. It begins with a discussion of why sampling is important, the key concerns in sampling and the characteristics of a good sample .The unit ends with a discussion of probability sampling techniques.

2.0: OBJECTIVES

In this unit, you are expected to:

1. understand what is meant by population and sample;

2. Know the key concerns in achieving a good sample and be able to itemise the various sampling techniques.

3.0 MAIN CONTENT

3.1:   Population

There is no way to design a research without defining its population and sample. Population is the total number of the subject. Sample is drawn from the population. The population in a study is the group of people or

object the researcher is studying. The term population could be people, schools, establishments, animals, specimens or even countries. The defined population must have at least one characteristic that differentiates it from other groups. However, if a researcher is carrying out a study on the attitude of Nigerian women to party politics, the researcher in this position will realize that it is impracticable to get every woman in Nigeria.

He will therefore need to decide to select a sample from a more narrowly defined (target) i.e. between 18 and 45 in order to save time and resources. In defining a research population, the researcher establishes a boundary of the conditions which stipulate who is to be included or excluded from the population. Seldom, if ever will researchers find themselves measuring entire populations. Rather, they are far more likely to draw a sample from the population and measure the elements in that sample. These results are then assumed to apply to the entire population. Researchers believe that similar results would be found if every element in the sample were measured.

Sample measurement is called statistics; population measurement is parameters. The process is known as inference, and the statistical tests that are used for this purpose are called inferential statistics. If all possible information needed to solve a problem could be collected, there would be no need to sample. Political science researchers seldom have this luxury; they are typically limited in time and money. Therefore, people making decisions based on research use data gathered from samples, and make their decisions based on probabilities that the sample data mirror what could be expected if it were possible to survey an entire population.

A population is made up of all conceivable elements, subjects or observations relating to a particular phenomenon of interest to the researcher (Asika, 1991:39). A population may be finite or infinite. Several factors has been given as reason d'être for sampling. Olayinka and Gbadegesin (2005:110) posits that we sample because sampling saves time, resources, energy and it preserves the items under study if they are fragile. Fadeyi and Adeokun (2004) and Asika (1991) agreed with the reasons above but added that sampling helps to estimate the population characteristics and in the circumstances afford a better supervision than with a complete coverage of the entire population. It helps to obtain quicker results than does a complete coverage of the population.

3.2: Sample

The concept of sample is derived from the ideas that it is not desirable even if it is possible to study the entire population. This is the major reason sample has to be taken from the population. This is a way of reducing the data to a manageable size or proportion. A well-drawn sample is a good representation of the entire population so it should not be taken that a sample is inferior to the population unless the sample drawn are too small. In order to select a sample to be studied the researcher has to follow the basic steps.

1. Identify the population.
2. Determine the required sample size.
3. Select the sample using a definite sampling technique.

3.3:   Sample Size

This varies. It is often difficult to determine the sample size, which is to be a good representation of the population. Generally it is accepted that the larger the sample size, the smaller the sampling error. If the sample is too small, the result may not be general sable.

There is no hard and fast rule on the percentage of the population that the sample represents, although some research recommend 10% while others recommend 20%. These are not based on sampling theory. To estimate the adequate size of the sample properly, a researcher needs to determine what level of accuracy is expected of the estimates, that is, how large a standard error is acceptable. The concept of standard error is central to sampling theory and to the determination of the size of a sample. Details are discussed in unit 18. It is worth emphasizing that, the acceptability of research finding should not be based on a single study but must be based on results of many studies using similar samples.

It must also be noted that larger samples could lead to erroneous conclusions if they are not properly selected. No matter how well a sample is selected, no sample will have identical composition as that of the population. Tuckman   (1978),   said   a   researcher   could   never   be certain of the representativeness of a sample unless the entire population is taken. However, this is always unlikely to be done.

3.4:   Characteristics of a Good Sample

Every researcher strives through sampling to have an accurate estimate of its population characteristics; if his estimate is 100% accurate then the sample is completely representative of the population from which it was drawn;

1) It is often said that the best sample is a complete census of the population itself because every element of the population is represented in the population itself. Representativeness then is the hall mark of a good sample. It makes sample characteristics valid estimates of the population characteristics.

2) A good sample must be quite representative of the population. Representativeness is directly related to:

   a) Precision, by which we ensure that random fluctuations or error variance or sampling error, is minimal.

   b) Absence of systematic variance or sampling bias, which is caused by some known or unknown influences that causes the scores to tend move to one side than the other (Asika, 1991).

3.5: Sampling Technique/Method

The skill in sampling determines to a considerable extent the degree to which accurate statement about a total population can confidently be made. Sample selection is an important step in any research work. Sampling method refers to the way the sample unit are selected from a parent population (Bryman, 2004).

Sampling techniques or design can be divided into two types of samples; probability and non-probability sampling designs. Authors have referred to these two types of samples with different concepts. Some have called them random and non-random samples, probability and purposive samples, strategic and non-strategic samples designs:

Probability sampling Methods	Non-Probability sampling methods
1. Random Sampling	1. Conventional/accidental
2. Systematic Sampling	2. Quota Sampling
3. Stratified sampling	3. Judgement sampling
4. Area/cluster sampling	

Source: Asika, (1991:42)

Probability sampling is one of the fundamental bases upon which all inferential statistics is built. A probability sample is one in which the sample units (people, parts, groups, homes, cities, tribes, companies, etc) are selected at random and all have equal chance of being selected. It is where every item in the population is given equal and independent chance of being included in the sample (Asika 1991).

In deciding whether to use probability or non-probability sample, a researcher should consider four points.

1.   Cost Vs. Value: The sample should produce the greatest value for the least investment. If the cost of a probability sample is too high in relation to the type and quality of information collected, a non-probability sample is a possible alternative.

2.   Time Constraint: In many cases, researchers collecting preliminary information or data operate within time constraints imposed by sponsoring agencies, management directives or publication deadline. Since probability sampling is often time - consuming a non-probability sample may provide a temporary list.

3. Purpose of the Study: Some researches are not designed for generalization to the population but rather to investigate variable relationships or to collect exploratory data for designing questionnaire or other measuring instruments. A non-probability sample is often appropriate in situation of this nature.

4. Amount of Error Allowed: In preliminary or pilot studies, where error control is not a prime concern a non-probability sample is usually adequate. Probability sampling, generally, incorporate some *type* of systematic selection procedure such as a table or round of number to ensure that each unit has an equal chance of being selected.

Self-Assessment Exercise

1. Define a sample

2. State the steps to be taken in selecting a sample.

4.0 SUMMARY

4.1 The entire set of relevant unit of analysis is the population. A sample is a subset of the population. To accurately estimate the parameter of a population from the sample the research must endeavour that the sample is representative of the population. This means that the researcher must be meticulous in way sample units are selected from a parent population. Available sampling techniques include probability and non-probability sampling methods.

5.0: CONCLUSION

Samples are used in place of a census of a population because it saves cost, time and can accurately be inferred to the population. Researchers must however decide on what type of sampling method is appropriate for the

research project at hand. While perfect match between sample and population is impossible, the researcher can achieve a high probability that the sample reflects the population.

6.0   TUTOR - MARKED ASSIGNMENT

1.   Discuss the main reason why researchers conduct samples

2.   Explain the difference between probability and non-probability sampling.

3.   Differentiate between the population and a sample.

7.0   REFERENCES/FURTHER READING

Asika, Nnamdi (1991) *Research Methodology in the Behavioural Sciences*, Lagos: Longman Nigeria Pic.

Bryman, Alan (2004) *Social Research Methods* (2nd Edition). Oxford: Oxford University Press.

Fadey, A.O. and Adedokun O.A. (2004) *Fundamentals of social research*. Lagos: Landmark global-links concepts.

Framkfort-Nachmias, Chava and David Nachmias1996. Fifth ed. *Research Methods in the Social Science*. New York: St Martin's Press.

Gbadegesin, Adeniyi (ed) (2005) *Statistics for the Social Sciences*, Ibadan: Ibadan University Press.

Mcnabb, David E. (2005) *Research Methods for political science: quantitative and qualitative methods*. New Delhi: Prentice hall of India limited.

UNIT 5:   NON-PROBABILITY SAMPLING

CONTENT

1.0:   Introduction
2.0:   Objectives
3.0 Main Content
      3.1:   Defining Non-Probability Sampling
      3.2:   Accidental or Convenience Sampling
      3.3:   Quota Sampling
      3.4:   Judgement Sampling
      3.5:   Panel Samples
      3.6:   Multi-stage sampling
      3.7:   Double Sampling
      3.8:   Haphazard Sampling
4.0:   Summary
5.9:   Conclusion
6.0   Tutor Marked Assignment
7.0   References/Further Reading

1.0   INTRODUCTION

This unit deals with non-probability sampling techniques/methods.   It examines such sampling techniques as panel sampling, judgement sampling, accidental or convenience sampling, quota sampling, double sampling and haphazard sampling. It emphasises that the non-probability sampling does not specify the probability of each sampling unit being included in the sample.

2.0   OBJECTIVES

In this unit you are expected to be able to:

1. explain the character of non-probability sampling methods; and

2. know the major non-probability sampling methods used in political evaluation research.

3.0   MAIN CONTENT

3.1:   Defining Non-Probability Sampling

This is a sampling method that does not guarantee randomness. In other words, the elements of the population do not have the privilege of having equal chance or known probability of being selected in the sampling process. Randomness may, however, occur by chance but it does not really matter whether randomness exists or not in non-probability sampling process because the population elements use not deliberately given equal chance of being selected. The following are types of non-probability sampling methods.

3.2: Accidental or Convenience Sampling

A researcher who is particularly interested in having a feeling or an idea of a phenomenon of interest may find convenience sampling very convenient. A correspondent of the Nigerian Television Authority (NTA) who, accompanied by this camera operator, interviews anybody he sees, regarding opinions about certain measures taken by the government, is definitely employing accidental or convenience sampling. Likewise, when a researcher stands before the entrance to a department store or super market and interviews any shopper that accidentally passes by, about his opinion on a particular brand or product, he is involved in accidental or convenience sampling. Accidental sampling is cheap and simple to use but often lead to an unreliable result because it lacks precision.

3.   Quota Sampling

Quota sampling attempts a fair representation of different classes that may exist in a given population. It is commonly used in survey having to do with public opinion and market research. In such surveys, the interviewers is required to ensure that the specified number of units in various classes like sex, age, income group geographical location etc are included in the sample.

The man advantage of quota sampling is that it reduces sampling costs where in-depth interviews are necessary, while the main disadvantage is the risk of bias as the interviews may consciously discriminate against certain types of people.

It is easy to confuse quota sampling with stratified sampling. The difference between them is the lack of well defined rule of selection in quota sampling. This is the major point against the wide spread use of quota sampling. In stratified sampling, apart from random selection from each stratum, the units to be included in the sample are known before sampling unlike what obtains in quota sampling.

3.4: Judgment or Purposive Sampling

In choosing some sample elements, the researcher may be guided by what he considers typical cases, which are most likely to provide him with requisite data or information. When a researcher chooses his sample under this condition, he is said to be involved in judgment sampling. The researcher selects sampling units subjectively in an attempt to obtain a sample that appears to be representative of the population. Thus, the chance that a sample will be selected depends on the subjective judgement of the researcher.

A research studying the use of birth control pills is not likely to get any good response from men, very old women, pre - high school age girls or good loyal members of extreme religious sects. In his own judgment, typical cases will be found among young females of between 16 and 35 years who are single and outgoing. These girls make up the population of study from which he can pick his sample subjects.

3.5: Panel Samples

A panel is defined as a permanent sample whose members are used repeatedly for successive interviewing. The method of selecting a panel is

either purposive or quota but certainly not any of the methods of probability sampling because there is a deliberate attempt here to screen individuals who are interested in cooperating in the panel. Panel sampling is also more of a procedure than a method of sampling.

3.7:   Double Sampling

A researcher, who has both time and money to his advantage, may like to duplicate sampling in order to achieve a high level of precision. He may resort to double sampling procedure. In doing this, he uses any of the sampling methods discussed earlier to choose a sample size larger than the size eh actually needs. From this extra-large sample, he may then choose the actual sample size by also applying any of the sampling methods discussed above. Double sampling is a modified version of multi-stage sampling procedure but its objective is a little different. Whereas multi-stage sampling aims at ensuring representativeness in a complex population, double sampling aims of a high level of precisions through sampling intensity.

3.8:   Haphazard Sampling

This is the case in which the selector thinks he is making a random selection. A good example is the sort of selection in public places often made by various press agents. For instance, press operators often stop people at random to interview them or ask them any question without following any prescribed rules like those that we have in other sampling methods.

Self-Assessment Exercise

1. Describe the main characteristics of non-probability sampling.
2. List three forms of non-probability sampling methods and discuss them.

4.0:   SUMMARY

This chapter has examined the main characteristics of various non-probability sampling methods.   These include panel sampling, judgement sampling,

accidental or convenience sampling, quota sampling, double sampling and haphazard sampling.

5.0   CONCLUSION

Sampling can provide an efficient and reliable way of obtaining information about large numbers of cases. Just how efficient and accurate depends on the type of sample used, the size of the sample and the method of collecting data from the sample. In the end, the decisions about sample will be a compromise between cost, accuracy and the nature of the research problem.

6.0   TUTOR - MARKED ASSIGNMENT

1.   Why are samples used to describe populations?

2.   Discuss the major types of non-probability sampling, with emphasis on their strengths and weaknesses.

7.0   REFERENCES/FURTHER READING

Asika, Nnamdi (1991) *Research Methodology in the Behavioural Sciences*, Lagos: Longman Nigeria Pic.

Bryman, Alan (2004) *Social Research Methods* (2nd Edition). Oxford: Oxford University Press.

Fadey, A.O. and Adedokun O.A. (2004) *Fundamentals of social research*. Lagos: Landmark global-links concepts.

Framkfort-Nachmias, Chava and David Nachmias1996. Fifth ed. *Research Methods in the Social Science*. New York: St Martin's Press.

Gbadegesin, Adeniyi (ed) (2005) *Statistics for the Social Sciences*, Ibadan: Ibadan University Press.

Mcnabb, David E. (2005) *Research Methods for Political Science: Quantitative and Qualitative methods.* New Delhi: Prentice hall of India limited.

UNIT 6:   PROBABILITY SAMPLING

CONTENT

1.0: INTRODUCTION

In this unit we will examine various techniques of probability sampling , We will look at the basis of choice among alternatives techniques. The techniques to be discussed include simple random sampling, systematic sampling, stratified sampling, cluster sampling and multi-stage sampling.

2.0: OBJECTIVES

In this unit you are expected to

1. understand and be able to use the simple random sampling method,

2. Explain and utilize systematic sampling;

3. Adopt the Stratified sampling, and

4. Understand the uses and basis of cluster sampling.

3.0   MAIN CONTENT

3.1: Simple Random Sampling

A Simple Random Sample is one in which each subject or unit with the population has an equal chance of being selected if the subject or unit is drawn from the population. The procedure is known as random sampling without replacement. Researchers usually use a table of random numbers to generate a simple random sample For example A researcher who wants to analyse 10 prime - time TV programmes out of a total population of 100 to determine how the medium portrays elderly people can take a random sample from the 100 programmes by numbering each show from 00-99 and then selecting 10 numbers from 2 table of random numbers. First, a starting point is an arbitrary decision. The researcher then selects the remaining 9 numbers by going up, down, left or right on the table or even randomly through the table. However, two rules are always applicable:

1.   Each unit of subject in a population must have an equal chance of being selected. 2. The selection procedure must be free from subjective intervention by the researcher.

2.   Random sampling method is the most fundamental method of probability sampling. Its principles, i.e. randomness, are applied in all probability sampling methods. Random sampling uses the principle of randomization, which is a procedure of giving every subject in a population equal chance of appearing in the selection. Each of the following methods does this:

(i) Writing all the names or numbers of the subjects or units on cards and shuffling the cards, and taking the top card each time the card are shuffled continuously until the required sample size is met. Someone who plays card would appreciate what the above method entails. One

123

notices that the cards are randomized in order to avoid two or more similar cards appearing in sequence before distributing the cards to the players.

(ii) The other method is the use of a table of random numbers. This is known to involve a very minimum of sampling bias due mainly to the method by which this table is compiled. A table of random numbers contains numbers chosen entirely without human intervention thus every number (1,2,3,4,5,6,7,8,9,0) has a chance of appearing in the selection.

3.2:   Systematic Sampling

This is similar in some ways to simple random sampling. This is random sampling with a system. From the sampling frame, a starting point is chosen at random, and thereafter at regular intervals. Here, every (n^{th}) subject or unit is selected from a population. Systematic sampling method involves the selection of the nth subject or item from serially listed population subjects or units: where n is any number usually determined by dividing the population by the required sample size. The population is given as N. Suppose in a population of 10,000 items or subjects; we want to select 1,000 items or subjects using systematic sampling procedure. The procedure is as follows;

Step 1: Number of items/subject serially up to 10,000

Step 2: Divine 10,000 by 1000 i.e. N/n = *10*.

Step 3: Randomly select a starting point, say number '10' on the population list.

Step 4: The select every 10th unit after the first. This list would include the following: 10th, 20th, 30th, 40th, 50th, 60th, 70th and so on.

For example, suppose you want to sample 8 houses from a street of 120 houses. 120/8=15, so every 15th house is chosen after a random

starting point between 1 and 15. If the random starting point is 11, then the houses selected are 11, 26, 41, 56, 71, 86, 101, and 116.

If there were 125 houses, 125/8=15.625, so should you take every 15th house or every 16th house? If you take every 16th house, 8*16=128 so there is a risk that the last house chosen does not exist. To overcome this, the random starting point should be between 1 and 10. On the other hand if you take every 15th house, 8*15=120 so the last five houses will never be selected. The random starting point should now be between 1 and 20 to ensure that every house has some chance of being selected.

In a random sample every member of the population has an equal chance of being chosen, which clearly not the case here is, but in practice a systematic sample is almost always acceptable as being random.

Systematic sampling often saves time, resources and efforts when compared to simple random sample. In fact since the procedure clearly resembles simple random sample, many researchers consider systematic sample equal to simple random procedure. The degree of accuracy of systematic sampling depends on the adequacy of the sampling frame. A sampling frame is a complete list of all the members of the population. Where there is an exhaustive list of the population, the research methods are correct.

Advantages of Systematic Sampling

1. Selection is easy.

2. Selection can be more accurate than in a simple random sample.

3. The procedure is generally in expensive.

Disadvantages of Systematic Sampling

1. A complete list of the population must be obtained.

2. Periodicity may bias the selection process.

Before deciding to use systematic sampling, one should consider the goals and purpose of the study as well as the availability of a comprehensive list of the population. If such a list is not available, systematic sampling is probably ill advised.

3.3:   Stratified Sampling

This is used when a researcher is interested in a particular characteristic segment or stratum of the population e.g. classification according to age, sex, education, level of income etc. Thus, instead of selecting a sample from the population at large the researcher identifies a significant variable, selects subject who have these traits and chooses a sub - sample from the group. The variable of interest might be age, sex, religion, education, income or political affiliation. However, the more variables are added to the stratification list the harder it becomes to identify subject meeting the criteria.

Advantages of Stratified Sampling

1. Representative (ness) of relevant variables is assured.

2. Comparison can be made to other population,

3. Selection is made from a homogeneous group.

4. Sampling error is reduced.

Disadvantages of Stratified Sampling

1. Acknowledgement of the population prior to selection is required.

2. The procedure can be costly and time - consuming.

3. It can be difficult to find a sample if incidence is low.

Stratified sampling ensures that a sample is drawn from homogeneous sub- set of the population i.e, from a population with similar characteristics. Homogeneity helps the researcher to reduce sampling error.

3.4.: Area Sampling or Cluster Sampling

The usual sampling procedure is to select one unit of subject at a time but that requires that the researcher will have a complete list of the population. In some cases there is no way of having the complete list. One way to avoid this problem is to select the sample in groups or categories. This procedure is known as cluster sampling For example, analysing newspaper readership, Television. Viewership or radio listenership in Lagos state will be time - consuming and complicated. If individual subjects were to be randomly selected with cluster sampling, one can divide the state into clusters-senatorial districts, local government councils, wards and select respondents from the clusters. Cluster sampling may however create 2 errors. In addition to the error involved in determining the initial clusters, errors may arise in selecting from the clusters.

This sampling method is used when the researcher recognizes that some populations are distributed in clusters or pockets of settlement and he or she wants to use the clusters as a basis for selection. If the sub-population represented by each cluster is known, this can be used as a basis for proportional selection of samples such that the number of subjects selected from each cluster will represent its share of the entire population.

This method of sampling is used mainly in geographically distributed population. A reader who knows the city of Lagos very well will notice that some Nigerian ethnic groups live in clusters. If a researcher is interested in ethnic representativeness in his study population, he will use cluster or area sampling. If he chooses to use cluster sampling, then he will find that the Igbo are clustered in Ajegunle, Yoruba in Mushin and Hausa in Idi-Araba. He can then use this clustering as a basis for his sample selection. The steps involved in cluster sampling are as follows:

Step 1:   Identify the population to be sampled, e.g. all household unit in Lagos.

Step II: identify the salient characteristics that you think would enhance representativeness, e.g. ethnic groups within the population.

Step III: Locate the areas where units or subjects with the characteristics cluster and know their respective sizes (population subsets).

Step IV: Use random selection procedure to select your sample units selected from each cluster in proportion to the cluster's share of the total population. There are several valid techniques of selecting a sample from a population. The sampling may be different but the basic steps in sampling are mainly the same. They are:

1. Identification of the population.
2. Determination of the required sample size.
3. Selection of the sample.

3.4.1: Problem with the Use of Cluster

It can create false homogeneity. It is biased because it is possible to group people of similar backgrounds into one. For instance, a ward area may comprise mostly residence of a low socio economic status, who are unrepresentative of the remainder of the state if selected for analysis. Such a group may confound the research result. To control such error, it is best to use small areas of cluster both to decrease the number of elements in each cluster and to maximize the number of clusters selected.

3.4.2: Advantages

1. Only a part of the population needs be enumerated.
2. Costs are reduced if clusters are well defined.
3. Estimate of cluster parameter is made and compared to the Population

3.4.3: Disadvantages

1. Sampling errors are likely to occur.
2. Cluster may not be representative of the sample.

Each subject or unit must be assigned to a specific cluster. Note that in many nationwide studies, researchers use a form of cluster sampling called multi-stage sampling in which case households or persons are selected, not groups.

3.5: Stratified Sampling

Stratified sampling procedure is an applied random sampling method. In this method, the population is just grouped into some definite characteristics. These groups are called 'strata'. From these strata, the sample is chosen by applying random selection technique on each stratum.

In choosing a particular number of items from the various strata, the researcher must be proportional to the stratum's share of the total population.

Suppose the researcher is interested in finding the effect of Structural Adjustment Programme (SAP) on the worker in Lagos, he might stratify workers by income as follows:

Stratum I: Low income (below ₦ 5,000)

Stratum II: Middle Income ₦ 5000 - ₦ 10000

Stratum III: Upper Income ₦ 10000 - ₦ 15000

Stratum IV: Super Upper Income over ₦ 15000

He may then proceed to randomly select his samples from each level (stratum) and the number of employees he selects from a particular stratum must be proportional to the stratum's share of the total population.

This sampling method is superior to both simple random sampling and systematic sampling procedure because it uses extra methods or representativeness by first identifying some characteristics that are being researched and then using these characteristics as a basis for further random sampling of the entire population.

3.6   Multistage Sampling

The researcher who may be interested in precision and thoroughness may discover that none of the methods of probability or non-probability can d give him the exact sample size he requires. He may find that the distribution of the population is so complex that he needs more than one

sampling technique to select his sample. He definitely has to resort to sampling in stages. This is the principle behind multistage sampling. The multistage sampling involves three processes which are listed in the figure below:

Stage 1	• Random selection of Census Tracts identified by socio-econmic characteristics
Stage 2	• Random Selection of Residential Blocks in Selected Census Tracts
Stage 3	• Random Selection of Individual Households in Each Block
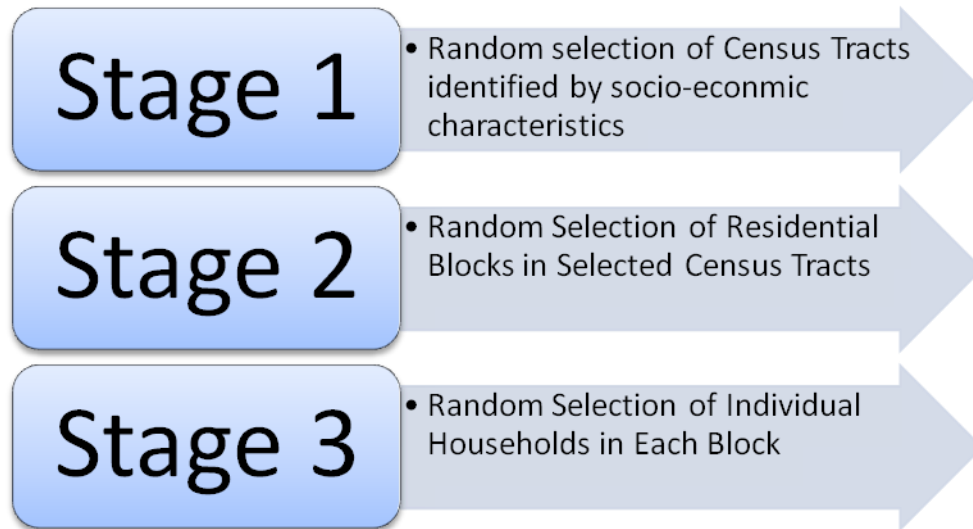
Source: Adapted from David McNabb (2005: 129).

4.0   SUMMARY

We have discussed several techniques of probability sampling.   These include simple random sampling, systematic sampling, stratified sampling, cluster sampling and multistage sampling. The simple random sampling method selects sampling units by using a table of random digits or assigning a unique number to each sampling unit. Systematic sampling determines the sampling interval (N/n), select the firs sample unit randomly, and select the remaining units according to the interval. Stratified sampling determines the strata and from each stratum selects a random sample proportionate to the stratum in the population. Cluster sampling determines the number of levels

131

of clusters and from each level or cluster selects a sample randomly. A multi-stage sample is one in which sampling is done sequentially across two or more hierarchical levels, such as first at the county level, second at the census track level, third at the block level, fourth at the household level, and ultimately at the within-household level.

5.0   CONCLUSION

Probability sample is one in which the sample units are selected at random and all have equal chance at being selected. The choice between a probability and non-probability sample is often based on the cost –versus – value principle. The researcher would select a sampling technique that provides the greatest value over cost margin.

6.0   TUTOR-MARKED ASSIGMENT

1.   Explain why a researcher use multistage sampling over the single-stage sampling

2.   State the difference between stratified and non-stratified samples

3.   Analyse the problems associated with cluster sampling.

7.0   REFERENCES/FURTHER READING

Gbadegesin, Adeniyi (ed) (2005) *Statistics for the Social Sciences*, Ibadan: Ibadan University Press.

Bryman, Alan (2004) *Social Research Methods* (2nd Edition). Oxford: Oxford University Press.

Mcnabb, David E. (2005) *Research Methods for political science: quantitative and qualitative methods.* New Delhi: Prentice hall of India provide limited.

Fadey, A.O. and Adedokun O.A. (2004) *Fundamentals of social research*. Lagos: Landmark global-links concepts.

Framkfort-Nachmias, Chava and David Nachmias1996. Fifth ed. *Research Methods in the Social Science*. New York: St Martin's Press.

Asika, Nnamdi (1991) *Research Methodology in the Behavioural Sciences*, Lagos: Longman Nigeria Pic.

MODULE IV: DESCRIPTIVE AND INFERENTIAL STATISTICS

UNIT 1:   VARIABLES AND THEIR OPERATIONALIZATION

CONTENT
1.0: Introduction
2.0: Objectives
3.0: Main Content
      3.1: Defining a Variable
      3.2: Importance of Variables
     3.3: Variable Operationalization
4.0: Summary
5.0: Conclusion
6.0: Tutor - Marked Assignment
7.0: References/Further Reading

1.0: INTRODUCTION

In this unit, we will focus on the idea of variable and its importance in research. We will examine the operationalization of a variable in political evaluation, which is how a variable is defined for the purpose of measurement.

2.0: OBJECTIVES

In this unit, you are expected to be able to

      1 define a variable;

      2   appreciate the importance of variables in political evaluation research; and

3 understand the meaning of the idea of variable operationalization.

3.0: MAIN CONTENT

3.1: Defining a Variable

A variable is something that can change, such as 'gender' and are typically the focus of a study. Asika (1991:6) defines variable as "a construct or concept to which numerical values can be assigned". Neuman (2000) notes that variables take two or more values. Bryman (2004:29) summarizes variable thus:

> A variable is simply an attribute on which cases vary. Cases can obviously be people, but they can also include things such as households, cities, organizations, schools, and nations. If an attribute does not vary, it is a constant... Constants are rarely of interests to social researchers.

Asika (1991) argues that numerical values cannot be assigned to most concepts because they just do not vary. Such invariable concepts can be referred to as constants or parameters. In scientific research, variables refer to factors or conditions that can change during the course of an experiment. For example, in political evaluation a political scientist may study factors that affect change in electoral behaviour. For experimental purposes, communication through the mass media may be used to manipulate public opinion to change the voting pattern. In a natural science experiment, certain variables may be manipulated to see how different conditions affect the temperature at which water boils. The size of the burner and pot used, amount of water, temperature at which the water is heated and any other item may be manipulated. These items are all variables. Scientists attempt to change only one of these variables at a time so that there is no confusion

about what caused a change. Variables have attributes, which are sub-values of a variable, such as 'male' and 'female'.

Causal research examines the world in terms of variables (those things that reveal variation within a population). In computer science and mathematics, a variable is a symbol denoting a quantity or symbolic representation. In mathematics, a variable often represents an unknown quantity; in computer science, it represents a place where a quantity can be stored. Variables are often contrasted with constants, which are known and unchanging.

In other scientific fields such as biology, chemistry and physics, the word variable is used to refer to a measureable factor, characteristic or attribute of an individual or a system. In a scientific experiment, so called "independent variables" are factors that can be altered by the scientist. For example, temperature is a common environmental factor that can be controlled in laboratory experiments. "dependent variables" or "response variables" are those that are measured and collected as data.

Variables can be used in open sentences. For instance, in the formula: $x + 1 = 5$, x is a variable which represents an "unknown" number. In mathematics, variables are usually represented by letters of the Roman alphabet, but are also represented by letters of other alphabets; as well as various other symbols. In computer programming, variables are usually represented by either single letters or alphanumeric strings.

Variables may have the following characteristics:
- Period: When it starts and stops.
- Pattern: Daily, weekly, ad-hoc, etc.
- Detail: Overview through to 'in depth'.

- Latency: Time between measuring dependent and independent variable (some things take time to take effect).

3.2   The Importance of Variable

Research scientists manipulate variables in order to test their hypotheses and learn more about how the world works. Unlike in algebra, where the word "variable" refers to an unknown quantity that the mathematician is trying to identify, in research variables are any factor or condition that are changeable during the course of an experiment such as temperature, time or composition. Scientists attempt to change only one variable at a time so that the reasons for the results of an experiment are clear (Silberstein, 2010). Variables are useful in quantitative study because they allow instructions to be specified in a general way. If one were forced to use actual values, then the instructions would only apply in a more narrow and specific set of situations. For example: a mathematical definition for finding the square of any number can be represented as $X^2 = x * x$. Now, all we need to do to find the square of a number is to replace *x* with any number we want.

Square (X) = $x * x = y$ OR $X^2 = x*x$, let x=1,2or3

Square (1) = 1 * 1 = 1

Square (2) = 2 * 2 = 4

Square (3) = 3 * 3 = 9

In the above example, the variable x is a "placeholder" for any number. One important thing we are assuming is that the value of each occurrence of x is the same - that x does not get a new value between the first x and the second x. A very useful example is the effect of exercise on pulse and breathing rates.

3.3 Variable Operationalization

The operationalization section of a proposal or research tells the reader precisely what the writer means when he refers to the variables in his evaluation. It clears up any confusion others may have about the meaning and measurement of the variables in the study. When the researcher has chosen the variables for his research questions, he must also *operationalize* those variables. Operationalization answers two key questions:

    1) How is each variable defined?

    2) How is each variable measured?

Defining each variable is often a technical exercise based on specific jargon and uses of words that are characteristic to the area of research. One place to begin is to consider how other researchers defined your variables in their research. The researcher may adjust or adapt other definitions to meet his own needs, or borrow the definitions that other researchers have used.

The research must also indicate how the variable is to be measured in the research. Some variables, like *age* or *height,* are pretty straightforward, but the researcher must still state the method of measurement: "age will be measured in years," or "height will be measured in inches." For more complex variables, like *political persuasion* or *racism,* you must indicate how you will quantify these concepts so they can be used as numbers for statistical analysis. Political persuasion might be a nominal variable determined by something as simple as a single question on a survey that asks:

    How would you describe yourself (check one):

        ___ Liberal   ____Conservative

Or, the researcher may use a method which produces a ratio variable by asking several questions that would be combined into a score so that the researcher, could determine quantitatively *how much* of a characteristic (such as liberalism) that each subject possesses. However it is done, the researcher must describe his method for doing so in the *operationalization* section (describe it in a way that other researchers would make the same decisions about each subject if they followed your method). Operationalization therefore means putting a concept into a form that permits some kind of measurement.

It is a process through which we turn concepts into usable variables. According to Meier (2006) *et al* an operational definition is a statement that tells us how a concept will be measured by the analyst. An indicator is a variable or set of observations that results from applying the operational definitions. Operational definitions are often not stated explicitly but implied from the research report or briefing. In some cases multiple indicators are used to measure a single concept. This happened when the concept being measured has more than one dimension.

Examples of operational definitions provided by Meier *et al* (2006) include the following:

1. Educational attainment for Head Start participants is defined by the achievement scores on the IOWA Tests of Basic Skills.

2. A convict is a considered a recidivist if, within 1 year of release from jail, the convict is rearrested and found guilty.

3. An active volunteer in the Environmental Justice Association is defined as a person who donates her or his time to the association at least 5 hours per week on average.

One very interesting study of social capital, done by Robert Putnam (1993), operationalized social capital as social group membership.

Self-Assessment Exercise

1) Describe a variable.

2) Explain the concept of operationalization.

3) Identify important concepts in political evaluation such as bureaucracy, accountability, responsiveness and explain how they have been measured in the literature.

4.0 SUMMARY

A variable is a construct or concept to which numerical values can be assigned. It may take two or more values. A variable is simply an attribute on which cases vary. It is used to refer to a measurable factor, characteristic, or attribute of an individual or a system. Research scientists manipulate variables in order to test their hypotheses and learn more about how the world works. They are able to do these by turning concepts into usable variables through operationalization. An operational definition is a statement that tells us how a concept will be measured by the analyst

5.0   CONCLUSION

As we have noted earlier, measurement is the assignment of numbers to some phenomenon. A variable is a construct or concept to which numerical values can be assigned. An operational definition tells how concept will be used. It enables us to transform a concept into a useable variable. Operationalisation is very crucial for the measurement of variables and whatever they represent.

6.0   TUTOR-MARKED ASSIGNMENT

1. Define a variable

2. Explain the operational definition of a concept

3. Suggest new indicators that might improve upon the measurement of the concept.

7.0   REFERENCES/FURTHER READING

Asika, J. Nnamdi (1991) *Research Methodology in the Behavioural Science.* Lagos: Longman

Bryman, Alan (2004) *Social Science Research Method* (second edition). Oxford: Oxford University Press

Jacobson, Gary C. (1978) "The Effects of Campaign Spending in Congressional Elections". *American Political Science Review* 72:469-91.

Kelejian, Harry H. (1973) "Two-Stage Least Squares and Econometric Systems Linear in Parameters but Nonlinear in the Endogenous Variable". *Journal of the American Statistical Association* 66:373-374.

Green, Donald P., and Jonathan S. Krasno (1988) "Salvation for the Spendthrift Incumbent". *American Journal of Political Science* 32:844-907.

Gerber, Alan (1998) "Estimating the Effect of Campaign Spending on Senate Election Outcomes Using Instrumental Variables". *American Political Science Review* 92:401-12.

Meier, Kenneth J.; Jeffrey L. Brudney and John Bohte (2006). *Applied Statistics for Public and Nonprofit Administration* (6th edition). Belmont, CA: Thomson Higher Education.

Mitchell, Mark L. and Janina M. Jolley (2007). *Research Design Explained* (6th edition). Belmont, CA: Thomson Higher Education.

Neuman, W. L (2000) *Social Research Methods: Quantitative and Qualitative Approaches* (4th edition). Boston: Allyn and Bacon.

Patton, M. Q. (1990). *Qualitative evaluation and research methods.* (2nd ed.). Thousand Oaks, CA: Sage.

Putnam, Robert D. (1993) *Making Democracy Work: Civic Traditions in Modern Italy* (Princeton: Princeton University Press.

Taylor, Steven J. and Robert Bodgan (2006). *Introduction to Qualitative Research Methods: A Guide Book and Resource.* New York: John Riley and Sons Inc.

UNIT 2: CLASSIFICATION OF VARIABLES

CONTENT

INTRODUCTION

This unit deals with the various classifications of variables. It focuses on the basic ways variables are classified such as dependent and independent variables. It will also focus on classification according to levels of measurement, which is quantitative, numeric variable and qualitative, non-numeric variables. It also examines other forms of classification of variables.

OBJECTIVES

In this unit, you are expected to be able to:

1. provide a general classification of variables;

2. identify which class a variable falls into; and

3. understand the basis of the various classifications of variables.

3.0   MAIN CONTENT

3.1   General Classification of Variables

There are some ambiguities and debates about how to classify variables. While the meaning of the concept is clear in scientific communities, there are various classifications and representations of variables. Various fields classify and apply them differently. While attempt is made in this unit to provide varied classifications, a researcher should be careful enough to understand the classification in his own research community and employ them accordingly in order to reduce ambiguity and enhance reader's understanding of his work.

However, it is pertinent to distinguish between two broad types of variables under which other classifications are made: quantitative (or numeric) and qualitative (non-numeric). Each is broken down into two sub-types: qualitative data can be ordinal or nominal, and quantitative data can be discrete (often, integer) or continuous. Bryman (2004:29) states that "it is common to distinguish between different types of variables. The most basic distinction is between independent and dependent variables. The former are deemed to have causal influence on the latter". Silberstein (2010) observes that in any science experiment, there are three types of variables: independent, dependent and controlled variables.

3.1.1 Independent Variables

The independent variable is the variable that the scientist manipulates. An independent variable is typically the cause, while a dependent variable is the effect. The independent variable is that variable assumed to be the causal variable. In experimental research, the investigator manipulates the

independent variable. The effect (the dependent variable) is dependent on the causal variable (Asika, 1990, Bryman 2004). If unemployment is thought to cause crime rates to increase, unemployment is the independent variable (it can vary between high and low) and crime rates the dependent variable. Something which is an independent variable at one time can be a dependent variable at another. In natural science for example, if scientists are studying how putting salt in cold water affects how long it takes to boil water, the presence of salt is the independent variable. An independent variable is one that is manipulated by the researcher. It is like the knob on a dial that the researcher turns. In graphical representation, the independent variable is put on the X-axis.

A graphical presentation of the X and Y Axis
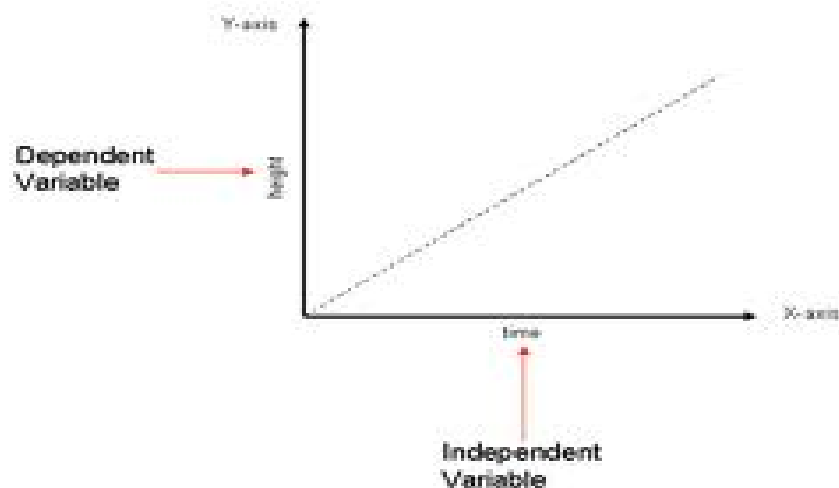
3.1.2. Dependent Variables

The dependent variable is a variable that changes as a result of the independent variable. If candidate A wins an election few days after opinion poll indicates that candidate B was leading with a very high margin, because the police

declared candidate B wanted a day before the election, then the change in the voting pattern is the dependent variable. Dependent variable is the outcome of the independent variable being changed or manipulated. In graphical representation, the dependent variable is put on the Y-axis. The Holy Grail for researchers is to be able to determine the relationship between the independent and dependent variables, such that if the independent variable is changed, then the researcher will be able to accurately predict how the dependent variable will change.

3.1.3. Controlled Variable

Controlled variables are variables that the scientist does not want to change. They however help the scientist in determining the effects of the stimulus applied to the experimental group. For instance, if a political scientist wants to determine the effects of appeal to ethnicity in a campaign rally on voting behaviour , he could set up an experimental group that is then exposed to a campaign rally with a dose of appeal to ethnic identification. He would however administer a questionnaire regarding their voting preferences before and after the campaign rally. He would however also set up another group, the controlled group that is not exposed to the campaign rally. The questionnaire will also be administered to the controlled group as well in a similar manner. The comparison of the responses of the control and experimental group at the end of the experiment points to the effect of the experimental stimulus of exposure to a campaign rally with a heavy appeal to ethnic identity.

3.1.4. Extraneous variables /or intervening / or Lurking variables

These are additional variables, which could provide alternative explanations or cast doubt on conclusions.

145

3.1.5 Instrumental Variable

Whenever an explanatory variable is endogenous, and thus, 'X' does not equal zero in theory, this means that one of the independent variables is not fixed, and that it is potentially correlated with the errors. This can happen in a number of different ways. In political science, the classic example of instrumental variable use can be thought of as necessitated by an omitted variable. Suppose we want to estimate the effects of campaign contributions received by a candidate upon that candidate's vote share.

A fundamental challenge for scholars estimating the impact of money on votes is that both variables may be influenced by perceptions about the threat posed by a challenger, a factor that is notoriously difficult to measure (leaving us with an omitted variable). Even when one accounts for past performance of the challenger's party and candidate quality, a link between the unexplained variance in naira and votes remains. This shared error most likely reflects the unmeasured perceptions of the challenger's chances. Researchers have attempted to deal with this simultaneity by using two-stage least squares estimations (Jacobson, 1978; Green and Krasno, 1988; Gerber, 1998). These procedures first predict candidate finances from factors that are (in theory) not directly related to election outcomes, and then use these systematic figures — purged of their candidate-specific information - to explain vote totals.

3.1.6. Proxy Variable

In social science research, certain concepts such as belief, joy, peace, justice etc. cannot be measured. This is because they are value based and highly normative in nature. Such concepts are often philosophically contested. Researchers have tended to study such concepts using other closely related

concepts that could serve as pointer and indicate the presence of such concepts that are being examined. These closely related concepts that could be used to replace and indicate others are referred to as proxy variables.

3.2: Classification of Variables by level of Measurement

We can distinguish between two types of variables according to the level of measurement:

1. Continuous or Quantitative Variables.
2. Discrete or Qualitative Variables.

A quantitative variable is one in which the variables differ in magnitude, *e.g.* income, age, GNP, etc. A qualitative variable is one in which the variables differ in kind rather than in magnitude, *e.g.* marital status, gender, nationality, etc.

3.2.1 Continuous or Quantitative Variables

Continuous variables can be classified into three categories:

• *Interval - scale Variables:*

Interval scale data has order and equal intervals. Interval scale variables are measured on a linear scale, and can take on positive or negative values. Interval scale variables are measured on a linear scale, and can take on positive or negative values. It is assumed that the intervals keep the same importance throughout the scale. They allow us not only to rank order the items that are measured but also to quantify and compare the magnitudes of differences between them. We can say that the temperature of 40°C is higher than 30°C, and an increase from 20°C to 40°C is twice as much as the increase from 30°C to 40°C. Counts are interval scale measurements, such as counts of publications or citations, years of education, etc.

- *Continuous Ordinal Variables*

They occur when the measurements are continuous, but one is not certain whether they are on a linear scale, the only trustworthy information being the rank order of the observations. For example, if a scale is transformed by an exponential, logarithmic or any other nonlinear monotonic transformation, it loses its interval - scale property. Here, it would be expedient to replace the observations by their ranks.

- *Ratio - scale Variables*

These are continuous positive measurements on a nonlinear scale. A typical example is the growth of bacterial population (say, with a growth function). In this model, equal time intervals multiply the population by the same ratio, (hence, the name ratio - scale). Ratio data are also interval data, but they are not measured on a linear scale.

With interval data, one can perform logical operations, add, and subtract, but one cannot multiply or divide. For instance, if a liquid is at 40 degrees and we add 10 degrees, it will be 50 degrees. However, a liquid at 40 degrees does not have twice the temperature of a liquid at 20 degrees because 0 degrees does not represent "no temperature" — to multiply or divide in this way we would have to use the Kelvin temperature scale, with a true zero point. (U degrees Kelvin — -273.15 degrees Celsius), in social sciences, the issue of "true zero" rarely arises, but one should be aware of the statistical issues involved.

There are three different ways to handle the ratio-scaled variables.

- Simply as interval scale variables. However this procedure should be avoided as it can distort the results.
- As continuous ordinal scale.

- By transforming the data (for example, logarithmic transformation) and then treating the results as interval scale variables.

3.2.2: Qualitative or Discrete Variables

Discrete variables are also called categorical variables. A discrete variable, *X,* can take on a finite number of numerical values, categories or codes. Discrete variables can be classified into the following categories:

1. Nominal variables
2. Ordinal variables
3. Dummy variables from quantitative variables
4. Preference variables
5. Multiple response variables

1. *Nominal* Variables

Nominal variables allow for only qualitative classification. That is, they can be measured only in terms of whether the individual items belong to certain distinct categories, but we cannot quantify or even rank order the categories: Nominal data has no order, and the assignment of numbers to categories is purely arbitrary. Because of lack of order or equal intervals, one cannot perform arithmetic (+, -, /, *) or logical operations (>, <, =) on the nominal data. Typical examples of such variables are:

Gender:	1. Male 2 Female
Marital Status:	1. Unmarried 2 Married 3. Divorcee 4. Widower

2. *Ordinal Variables*

A discrete ordinal variable is a nominal variable, but its different states are ordered in a meaningful sequence. Ordinal data has order, but the intervals between scale points may be uneven. Because of lack of equal distances, arithmetic operations are impossible, but logical operations can be performed on the ordinal data. A typical example of an ordinal variable is the socio-economic status of families. We know 'upper middle' is higher than 'middle' but we cannot say 'how much higher'. Ordinal variables are quite useful for subjective assessment of 'quality importance or relevance'. Ordinal scale data are very frequently used in social and behavioural research. Almost all opinion surveys today request answers on three-, five-, or seven-point scales. Such data are not appropriate for analysis by classical techniques, because the numbers are comparable only in terms of relative magnitude, not actual magnitude.

Consider for example a questionnaire item on the time involvement of scientists in the 'perception and identification of research problems'. The respondents were asked to indicate their involvement by selecting one of the following codes:

1	=	Very low or nil
2	=	Low
3	=	Medium
4	=	Great
5	=	Very great

Here, the variable 'Time Involvement' is an ordinal variable with 5 states. Ordinal variables often cause confusion in data analysis. Some statisticians treat them as nominal variables. Other statisticians treat them as interval scale variables, assuming that the underlying scale is

continuous, but because of the lack of a sophisticated instrument, they could not be measured on an interval scale.

3. Dummy Variables from Quantitative Variables

A quantitative variable can be transformed into a categorical variable, called a dummy variable by receding the values. Consider the following example: the quantitative variable *Age* can be classified into five intervals. The values of the associated categorical variable, called dummy variables, are 1, 2,3,4,5:

[Up to 25]	1
[25, 40]	2
[40, 50]	3
[50, 60]	4
[Above 60]	5

4. Preference Variables

Preference variables are specific discrete variables, whose values are either in a decreasing or an increasing order. For example, in a survey, a respondent may be asked to indicate the importance of the following factors in voting decision by using the code for the most important factor and for the least important factor:

1. Party manifesto
2. Party candidate
3. Ethnic group of candidate
4. Status of the party relative to other parties
5. Religious affiliation of the candidate

Note that preference data are also ordinal. The interval distance from the first preference to the second preference is not the same as, for example, from the sixth to the seventh preference.

5. Multiple Response Variables

Multiple response variables are those, which can assume more than one value. A typical example is a survey questionnaire about the use of computers in research. The respondents were asked to indicate the purpose(s) for which they use computers in their research work. The respondents could score more than one category.

1. Statistical analysis
2. Lab automation/ process control
3. Data base management, storage and retrieval
4. Modelling and simulation
5. Scientific and engineering calculations
6. Computer aided design (CAD)
7. Communication and networking
8. Graphics

Self-Assessment Exercises

1) Explain the basis on which variables are classified.
2) Distinguish between dependent and independent variable.

4.0 SUMMARY

There are various classifications of variables. Descriptive variables are those that will be reported on, without relating them to anything in particular. Categorical variables result from a selection from categories, such as 'agree' and 'disagree'. Nominal and ordinal variables are categorical. Numeric variables give a number, such as age. Discrete variables are numeric variables that come from a limited set of numbers. They may result from

answering questions such as 'how many', 'how often', etc. Continuous variables are numeric variables that can take any value, such as weight. Variables are very important aspect of research.

5.0   CONCLUSION

A researcher's success or failure in his research endeavours is highly dependent on his ability to understand and make use of appropriate variable in time and places vis-a-vis his research fields and object. Variables are often specified according to their type and intended use. In many studies more than one variable is recorded per case or individual. It is often the purpose of the study to determine if and/or how one or more variables affect another. A fair-test investigation is one in which we change the independent variable and observe the effect that it has on the dependent variable, while, as far as possible, controlling all other relevant variables at constant values. Whether a variable is the independent variable, the dependent variable or a fixed variable in an investigation depends both on what is being investigated and on the approach being adopted. There are numerous ways various disciplines classify variables and their usage.

6.0.   TUTOR- MARKED ASSIGNMENT

1.   List at least three different variables according to their levels of measurement.

2.   Identify the independent and dependent variables in the following descriptions of experiments.

(a) Students watched a political debate either alone or with others and then rated how interesting they found the debate to be.

(b) Some citizens were told that a politician had university education, and other citizens were told that the politician had not finished high school; they then rated the politicians' leadership potential.

7.0 REFERENCES/FURTHER READING

Asika, J. Nnamdi (1991) *Research Methodology in the Behavioural Science.* Lagos: Longman

Bryman, Alan (2004) *Social Science Research Method* (second edition). Oxford: Oxford University Press

Jacobson, Gary C. (1978) "The Effects of Campaign Spending in Congressional Elections". *American Political Science Review* 72:469-91.

Kelejian, Harry H. (1973) "Two-Stage Least Squares and Econometric Systems Linear in Parameters but Nonlinear in the Endogenous Variable". *Journal of the American Statistical Association* 66:373-374.

Green, Donald P., and Jonathan S. Krasno (1988) "Salvation for the Spendthrift Incumbent". *American Journal of Political Science* 32:844-907.

Gerber, Alan (1998) "Estimating the Effect of Campaign Spending on Senate Election Outcomes Using Instrumental Variables". *American Political Science Review* 92:401-12.

Neuman, W. L (2000) *Social Research Methods: Quantitative and Qualitative Approaches* (4th edition). Boston: Allyn and Bacon

Online Dictionary of Social Science (2010) "Independent Variable" (http://bitbucket.icaap.org/dict.pl7ternFlNDEPENDENT%20VARIABLE [retrieved on 26th October, 2010]

UNIT 3: DESCRIPTIVE STATISTICS AND MEASURES OF CENTRAL TENDENCY

CONTENT

1.0: Introduction

2.0: Objectives

3.0: Main Content

    3.1: Descriptive Statistics

    3.2 Frequency Distribution

    3.3: Mean

    3.4: Median

    3.5: Mode

4.0 Summary

5.0: Conclusion

6.0: Tutor-Marked Assignment

7.0: References/Further Reading

1.0 INTRODUCTION

This unit deals with the measures of central tendency: the mean, median and mode. It focuses on how to calculate these measures from raw data and interpret them.

2.0: OBJECTIVES

In this unit, you are expected to achieve the following:

1. understand and discuss issues relating to measures of central tendency in statistics;

2. compute mean, median and mode from raw data; and

3. be able to interpret what mean, mode and media tell.

3.0: MAIN CONTENT

3.1 Descriptive Statistics

Descriptive statistics is a term used to refer to numbers that summarise a group of data. This data may be about the number of votes won by a political party in several elections or the number of children immunised monthly in Ibadan metropolis or the number of students admitted into a University annually in five years. See table 13.1 below:

The Performance of Political Parties in Several Elections

Election	PDP	ACN	ANPP	Total No. Of votes (million)
Presidential elections	45	20	25	90
National Assembly Election	40	12	20	72
Governorship	45	15	10	70
House of Assembly	30	25	20	75

*Please note that these are hypothetical data.

The data are presented in their raw form. They do not tell us anything significant about the relative performance of each political party. Nor do they tell us how close the parties are in terms of each election. They need to be organised and interpreted using a variety of statistical techniques in order to tell us more about the performance of the parties in the elections.

A measure of central tendency is a number or score or data value that represents the average in a group data. There are three measures of central tendency. These are the mean, median and the mode. However, before these

156

could be estimated as a particular dataset, there is usually a need to order the data frequency.

3.2: Frequency Distribution

Frequency distribution is the most basic re-ordering of raw data to facilitate understanding. A frequency table pairs data values or ranges of data value with their frequency of occurrence. When summarizing large masses of raw data, it is often useful to classify the data. The classification of data is usually based on a particular trait, characteristic or variable. A class is one of the group categories of the variable. The class frequency is the number of observations or occurrences of the variable within the class. A tabular arrangement of data by classes together with the corresponding class frequencies is called a frequency distribution. The tabular summary of data for a single variable is a frequency distribution.

Therefore, a frequency distribution shows the number of data values in each of several overlapping classes. It can also be summarized in a graphical structure. Frequency distribution is, in most cases, a tabular summary of a set of a data showing the frequency (for number) of items in each of several overlapping classes (Anderson, et al., 1981:97). It represents data in a relatively compact form, give a good overall picture, and contain adequate information for many purposes.

Frequency distribution of height of 100 male students at Dome University

Height (in)	Number of student
60 – 62	5
63 – 65	18
66 – 68	42

69 – 71	27
72 – 74	8
	100

Data organized and summarized as in the above frequency distribution are often called grouped data. Although the grouping process generally destroys much of the original details of the data, an important advantage is gained in the clear overall picture that is obtained and in the vital relationship that are made evident.

The class interval is distance between the upper limit of one class and the upper limit of the next higher class while the class mid point is the point half way between the upper and the lower class boundaries. The class boundaries are the lowest and highest values that fall within the class. These are the end number 60 and 62. They are also called class limits. The smaller number (60) is the lower class, and the larger number (62) is the upper class limit. A symbol definition of a class, such as 60-62 is called a class interval. The term class interval and class are often used interchangeable, although the class interval is actually a symbol for the class. If heights are recorded to the nearest inch, the class interval 60 - 62 theoretically includes all measurement from 59.5 to 62.5 inches. These numbers indicated briefly by the exact numbers 59.5 and 62.5, called class boundaries or true class limits; the smaller number (59.5) is the lower class boundary, and the larger number (62.5) is the upper class boundary. In other words, subtract 0.5 from the lower class and the larger number (62.5) is the upper class boundary. In other words, subtract 0.5 from the lower class and add 0.5 to the upper class.

     e.g. $60 - 05$ $=$ 59.5 $=$ Lower class boundary

$$62 + 0.5 \quad = \quad 62.5 \quad = \quad \text{Upper class boundary}$$

The size or width of a class interval is the difference between the lower and the upper class boundaries and is also referred to as the class width, class size or class length. If four class intervals of a frequency distribution have equal width, these common widths are denoted by (C). In such case, C is equal to the difference between two successive lower class limits or two successive upper class limits.

$$C = 65.5 - 62.5 = 3$$

General rules for forming frequency distribution

1. Review the data to find the lowest and highest values, the largest and smallest numbers in the raw data and thus find the range (the difference between the largest and smallest numbers).

2. Make a list of the values from the lowest to the highest and then mark as follows. Divide the range into a convenient number of class intervals having the same size. If this is not feasible, use class interval of different side or open class interval. The number of class interval is usually taken between 5 and 20, depending on the data. Class intervals are also chosen so that the class marks (mid - point) coincide with the actual observed data. This tends to lessen the so-called grouping error involved in further mathematical analysis. However, the class boundaries should not coincide with the actual observed data.

3. Determine the number of observation following each class interval: that is, find the class frequency. This is best done by using a tally, or score sheet.

4. Avoid classes so narrow that some intervals have zero observation

5.   Make all the class intervals equal unless the top or bottom class is open ended.

6.   Use open ended intervals only when closed intervals would result in class frequencies of zero. This usually happens when some values are extremely high or extremely low.

7.   Try to construct the intervals so that the mid - points are whole numbers.

3.3:   Mean

This is a measure of central tendency which is central to probability statistics. It is commonly referred to as "average" It is the arithmetic average of a set of numbers. It is useful both in indicating the characteristic value of a distribution and the simple index of a variable. There are three types of mean (1) Arithmetic mean (2) Geometric mean and (3) Harmonic mean. Our focus here is on arithmetic mean for ungrouped and grouped data.

Arithmetic mean is the sum of series of figures divided by the number of times the figure appeared. That is the summation of the values of the data divided by the number of the value making up the total. The arithmetic mean is the commonest known and the most widely used among the three averages. Symbolically, the arithmetic mean is represented by :

$$\bar{x} = \frac{\sum x}{n}$$

$\bar{x}$ =   Arithmetic mean

$\sum$ =   Sigma in Greek letter which means "Sum of"

x   =   number of figures or element

Example: The age of ten students in 100 levels in the Department of Political Science is given below: find the mean age.

14, 15, 16, 16, 17, 17, 22, 22, 22, 22

Solution: $\underline{14 + 15 + 16 + 17 + 17 + 22 + 22 + 22 + 22}$

$$10$$

$$\bar{x} \quad = \quad \underline{183} \quad = \quad 18.3$$
$$10$$

Mean for Frequency Data

a.   **Ungrouped data**

If element or observations repeat themselves we apply frequencies. Frequencies are the number of times the value occurs. The formula of arithmetic mean for ungrouped data is given as:

$$\bar{x} = \sum \underline{fx}$$

$\bar{x}$ =   Arithmetic mean

$\sum$ =   Sum of notation

$\underline{f}$ =   corresponding frequency

x   =   Observed values

Example: Representing the ages of ten students in 100 level Political Science Department.

x (age)	f
14	1
15	1
16	2
17	2
22	4

Solution

x	f	fx
14	1	14
15	1	15
16	2	32
17	2	34
22	4	88
	10	183

$$\bar{x} = \frac{\sum fx}{\sum f} = \frac{183}{10}$$

$$\bar{x} = 18.3$$

3.3.1: Grouped Data

The ungrouped frequency distribution becomes inappropriate where a long list of scores also has repeated scores. The table becomes long, unwieldy and meaningless. To solve this problem, scores can be grouped into separate classes with frequencies of occurrence of scores in each class matched against them. This is a very useful method for compressing very large data into desirable number of classes. The following should be noted in grouped frequency distribution:

      i.   Class width: It is commonly denoted by 1 to represent the number of scores.

      ii.   Class: A class is a range of scores defined by a lower limit and an upper limit.

iii.   Number of groups: The number of classes in the distribution

iv.   Mid-point value: The average of the lower and upper limits of a given class.

Upper class + lower class

2

Thus, in calculating the means for grouped data we apply the logic for ungrouped data. The mean for grouped data refers to the sum of all the values divided by the number of values. Whenever grouped data are used for calculation, it is assumed that all values are spread evenly throughout the interval. Thus, the mean of the first class, or any class, is equal to the midpoint of the class. In the example below, he mid-point for the first class is 54.

Example:

Given the weight of 34 students in the Department of Political Science, University of Ibadan.

Weight kgs	f
52 – 56	3
56 - 60	6
60 - 64	10
64 - 68	4
68 - 72	8
72 - 76	2
76 - 80	1

To calculate the mean for this data we seek for mid - value for each of the Interval Mid value = upper value + lower value

2

Weight	Mid value	Frequency
52 – 56	54	3
56 – 60	58	6
60 – 64	64	10
64 – 68	66	4
68 – 72	70	8
72 – 76	74	2
76 – 80	78	1

To calculate the mean

Weight	Mid point X	F	Fx
52-56	56	3	168
56-60	58	6	348
60-64	62	10	620
64-68	66	4	264
68-72	70	8	560
72-76	74	2	148
76-80	76	1	76
		34	2184

$$\overline{X} = \frac{\Sigma x}{N}$$

$\overline{x} = \qquad \dfrac{\Sigma fx}{\Sigma f} \qquad \dfrac{2184}{34}$

$\overline{x} \quad = \quad 64.23$

Advantages

1. It is easy to calculate.

2. The principle of arithmetic mean is easy to understand.

3. Its calculation is clear and precise.

4. It provides a good measure of comparison.

Disadvantages

1. For qualitatively classified or nominal data, the mean is meaningless

2. Computational complication arises when there are unbounded classes

3. There are situations when it is used as a summary measure that it is not particularly meaningful

4. It cannot be obtained graphically.

3.4: Median

According to Meier *et al* (2006: 77) the median is "the middle observation in a set of numbers when the observation is ranked in order of magnitude". It is obtained by rank-ordering the values of the observations by magnitude and then choosing the value that is the middle of the rank order. When the

number of n of observation is odd, and the observations are arranged in ascending order, the median is the observation that is simply the middle value.

Example: Find the median of the following marks; 44, 40, 79, 42, 51, 59, 71, 44, 45, 51, 59, 65, 71, 79.

The median value is = 51

For even number of observation, the median is taken as the mean of the two middle value i.e. if n is even the median

Example: The number of attendances in twelve lectures in POS 702 is as shown: Find the median 40, 32, 30, 24, 40, 38, 35, 40, 28, 32 and 37.

Solution: We arranged in ascending order

24, 28, 30, 32, 35, 37, 38, 40, 40, 40

Median $= \dfrac{35 + 37}{2}$ = 36

3.4.1: Median for Grouped Data

As we have noted earlier, on arranging the data in ascending or descending order median is the middle - most observation. If the number of observations are odd then the median is $(n+1 / 2)^{th}$ observation where 'n' is the number of observations.

If the number of observations are even then median is the average of $(n / 2)^{th}$ and $(n / 2 +1)^{th}$ observation.

For group data, the median is obtained as the $n^{th}/2$ observation whether n is even or odd. In making the calculation, the first thing is to identify the class n which the mean falls. Suppose it falls in a class which begin at B_i and ends at B_{ii} that is the median lie between B_i and B_{ii}. if the cumulative frequency preceding the class containing the median is cf_p , the frequency for the interval B_i -B_{ii} is F_m , and the width of the interval containing the median is i then the median is:

$$Median = B_l + \left(\frac{\frac{n}{2} - cf_p}{f_m} \right) i$$

B_l =   lower bound of class containing the median

n   =   sample size

cf_p =   cumulative frequency of **all** classes preceding the class containing the median

f_m =   frequency of the class containing the median

i   =   width of the interval containing the median

B_{ii} =   upper bound of the class containing the median

Let us consider the table below

Weights (Kgs)	Frequency
52-56	3
56-60	6
60-64	10
64-68	4
68-72	8
72-76	2
76-80	1

From the table, the middle value is the 17th value since n = 34 which is in the interval of 60-64.

From here, B$_i$ = 60, B$_{ii}$ = 64, F$_m$= 10 and cf$_p$ = 3 + 6 = 9

Using the equation

$$\text{Median} = 60 + (\frac{(17-9)}{10})4$$

$$\text{Median} = 60 + (\frac{8}{10})4$$

$$\text{Median} = 60 + (\frac{32}{10})$$

$$= 60 + 3.2$$
$$= 63.2 \text{ or } 63$$

3.5: Mode

The mean and the median are not appropriate tools for distributions that are relatively asymmetrical. The mode is the value that occurs most frequently in a distribution. Mode offers two main advantages. One, it requires no calculation, only counting and two, it can be determined even for qualitative or nominal data.

Example:

The 20 meetings of the Faculty of the Social Sciences Board of Examiners were attended by 25, 25, 28, 23, 25, 24, 24, 21, 23, 28, 26, 24, 32, 25, 27, 24, 23, 24 and 22 of its members. Find the mode:

From the result, 24 occur five times and is the modal attendance.

Mode from Grouped data

Weight kg	Frequency
52-56	3
56-60	6
60-64	10
64-68	4
68-72	8
72-76	2
76-80	1

From the table above we have the modal class to be an interval of 60-64, to obtain the mode we use the formula

$$\text{Mode} = B_i + (\frac{F_m - f_p}{2F_m - f_p - f_s})i$$

B_l = lower bound of the interval containing the medal class
n = sample size
f_p = frequency preceding the modal class
f_s = frequency of succeeding the modal class
f_m = frequency of the class containing the modal
i = width of the interval containing the modal
B_{ii} = Upper bound of the interval of the modal class

i = width or size of class intervals

$$\text{Mode} = 60 + (\frac{10 - 6}{2(10) - 6 - 4})4$$

Mode = 60 + (4)/(20-10) x 4

Mode = 60 + 4/10(4)

Mode = 60 + 0.4 x 4

Mode = 60 + 1.6 =   61.6 Or 62

Modal class = 62

Self-Assessment Exercises

1. State the uses of the measures of central tendency.

2. Explain mean, median and mode and illustrate with calculations from any set of scores.

8.0   SUMMARY

Measures of central tendency are used to summarise a body of data by indicating the middle (or central) points in the distribution. These are the mean, the median and the mode. The arithmetic mean is the average of all data. The median is the data value that is greater than 50% of all the data points. The mode is the data value that occurs most often in the data points. There are calculations for all three measures of central tendency for both grouped and ungrouped data.

5.0:   CONCLUSION

The measures of central tendency, the mean, median and mode, are used to describe an entire sample or population. There is a distinct meaning and calculation for each measure of central tendency.

6.0:   TUTOR MARKED ASSIGNMENT

1. You are the research assistant to the administrator of a small bureau in the federal government. Your boss has received some criticism that the bureau does not respond promptly to requests of the National Assembly. The only information you have is the day the agency received the request and the day the agency mailed the response. From those figures, you have calculated the number of days the agency took to respond as shown in the table below.

9	1	6	10	8	12	9	14	15	7
19	8	21	10	50	37	9	4	28	44
9	18	8	39	7	1	4	15	7	28
47	9	6	7	24	10	41	7	9	29
6	4	12	7	9	15	39	24	9	2
20	31	18	9	33	8	6	3	7	16
20	26	9	9	16	5	3	12	36	11
8	6	28	35	8	10	11	20	3	10
16	8	12	4	6	9	10	10	9	16
4	14	11	8	5	8	11	9	7	6
11	9	7	8	10	9	11			

From the data contained in the table, do the following:

    (a) Prepare the frequency distribution.

    (b) Prepare a cumulative frequency

    2. The Nigerian army is allowed only five test firings of the Tomahawk Missiles. The following figures represent the number of feet the missiles missed the target by: 25, 145, 34, 62, and 50. Calculate the mean and median. Which should the army report?

    3. The Department of Welfare wants to know the average outside income for all welfare recipients in the state. Calculate both the mean and the median from the data in the accompanying table.

Income	Number of families
0-300	25
300-600	163
600-900	354
900-1200	278
1,200-1,500	421
1,500-1,800	603
1,800-2,100	211
2,100-2,400	84
2,400-2,700	32
2,700-3,000	5

(Note: these assignments are taken or adapted from Meier *et al* 2006)

4. What is a frequency distribution?

5. What is a measure of central tendency?

6. Distinguish between the mean, median and mode.

7.0: REFERENCES/FURTHER READING

Anderson, D.R. Sweeney D.J. (1981) *Introduction to statistics: An Application Approach*, Minnesota, West Publishing Co. International Inc.

Brase, C.H. and C. P. Brase 2007. *Understanding Basic Statistics*, Boston and New York: Houghton Miffling Company.

Bryman, Alan (2004) *Social Research Methods*, (2nd Edition) Oxford: Oxford University Press

Cosby, Paul C. 2006. *Methods in Behavioural Research*. New York: McDraw-Hill

Creswell, J (2003) *Research Design: Qualitative, quantitative, and mixed methods approaches*, (2nd ed.), Thousand Oaks, CA: Sage Publications.

Jones, E. Terrence (1971). *Conducting Political Research,* New York: Harper and Row Publishers

Meier, Kenneth Meier, Kenneth J.; Jeffrey L. Brudney and John Bohte 2006. *Applied Statistics for Public and Nonprofit Administration* (6th edition). Belmont, CA: Thomson Higher Education.

Spielgel M and Stephens L. (1999): Theory and problems of statistics, London, McGraw-Hill Companies.

UNIT 4: MEASURES OF DISPERSION

CONTENT

1.0   INTRODUCTION

This unit deals with the measures of dispersion. It explains the processes of determining how much a body of data do or do not cluster about the mean. It explains how to calculate the variance and standard deviation of a statistical data.

2.0:   OBJECTIVES

In this unit, you are expected to be able to:

1. find the range, variance and standard deviation of a set of data; and

2. acquire the skills to calculate the variance and standard deviation of a statistical data.

3.0   MAIN CONTENT

3.1: Measures of Dispersion

As we have noted in the previous unit, in order to use frequency distribution we need to know where the centre of the distribution is located, the *central tendency*, i.e. the **averages**, notably the *arithmetic mean, median* and *mode* amongst others. We also need to know the **dispersion** of the data, that is,

how spread out from the mean is the values (e.g. are they all closely clustered around the mean or are they well scattered). The three most usual measures of dispersion are:

- *range:* the distance between the smallest value and the largest value.
- *variance:* the most sophisticated and useful measure, leading to:
- *standard deviation:* which is the square root of the variance.

3.2: Variance and Standard Deviation

Variance and Standard Deviation are two standard statistical tools, which are related, They fall under Measures of Dispersion or variation. Variation or dispersion here is about standard and this is the average, especially the mean. The measure of dispersion seeks to tell how much the data do or do not cluster around the mean. Among dispersion measures are the range, the average deviation, inter-quartile deviation and the standard deviation. The Range is the difference between the highest value and the lowest value in the distribution of data.

The standard deviation is the commonest measure of dispersion. It is the square root of the average squared deviation of data from the mean. As a univariate analytical tool, it describes one variable. Examples of univariate analysis are frequency distribution, averages and measures of dispersion that express, on the basis of average deviation from the mean, the degree of variation within a variable. It is the analysis of a single variable, for purposes of description, describing a case in terms of a single variable that is the distribution of attributes that it comprises (Babbie, 2005:409). For instance if war were measured, it would be important to know how many of the case were conventional war and how many were non-conventional war.

When a single figure is used to sum up a set of data, some of the information contained in the original data is lost. The lost information is the spread of the data and this spread in the figures is the dispersion (Adamu and Johnson, 2007:63). The Variance tells us on average how much a given data point (of a distribution of attributes) differs from the mean of all data points, thereby useful for comparison across samples that contain different numbers of observation, though squaring of data changes the unit of measurement.

The average error or the variance is statistically represented by the symbol below:

$$s^2 = \frac{\sum\left(x - \bar{x}\right)^2}{n-1}$$

This average error is remedied by converting back into the original units of measurement by taking the square root of the variance; this is the standard deviation measure.

The standard deviation is statically represented below:

$$s = \sqrt{\frac{\sum_{i=1}^{n} (x_i - \bar{x})^2}{n-1}}$$

Standard deviation, like the variance, is a measure of variability of data. The standard deviation is an index of this variability (Babbie, 2007:414). It requires an interval of ratio level of measurement. It is based on the mean and gives an "average distance" between all scores and the mean.

(a)   High standard deviation = spread out values

(b)   Low standard deviation = highly clustered

A higher standard deviation as in (a) means that the data are more dispersed; a lower standard deviation (b) means that they are more bunched together. Therefore, if the mean represents the data accurately, then most of the scores will cluster close to the mean and the resulting standard deviation is low relative to the mean, but where the mean is a less accurate representation of the data, the scores cluster more widely around the mean and the standard deviation is large or higher.

3.2:   Steps in Calculating the Variance and Standard Deviation

(1)   List the values of variable (x)

(2)   Compute the mean of the variable (x)

(3)   Compute deviation from the mean for each value: from each value or observation subtract the mean $(x - \bar{x})$

(4)   Square the deviation from the mean $(x - \bar{x})^2$

(5)   Dividing the sum of deviations (squared) by the number of cases gives us the variance S^2

$$s^2 = \frac{\sum \left(x - \bar{x}\right)^2}{n-1}$$

(6)   The square root of the variance provides us the standard deviation of the sample.

$$s = \sqrt{\frac{\sum_{i=1}^{n}(x_i - \bar{x})^2}{n-1}}$$

Example:

Find the variance and standard deviation of the distribution 3,4,5,5,5,6,7,8,8,9

Raw score x_1	Mean $\bar{x}$	Deviation $(x_1 - \bar{x})$	Square Dev. $(x_1 - \bar{x})^2$
3	6	-3	9
4	6	-2	4
5	6	-1	1
5	6	-1	1
5	6	-1	1
6	6	0	0
7	6	1	1
6	6	2	4
8	6	2	4
9	6	3	9
		0	34

$\bar{x} = 6$

$\sum (x - \bar{x})^2 = 34$

(I) Variance $= S^2$ $=$ $34 / (10 - 1)$

$=$ $34 / 9$

$=$ 3.8

(ii) Standard deviation $= S$ $=$ $\sqrt{3.8}$

$=$ 1.94

The diagram above (Fig 1.3), shows the number of battles each of ten army battalions fought in Sierra Leone with the mean value as 6. There are vertical lines connecting each observed value to the mean representing the differences between what the mean or mode predicts and actual collected data. The deviations are these differences.

Given that the 10th Army battalion fought 9 battles therefore the difference is $x - x = 9 - 6 = 3$, as this positive number shows the underestimation of our model of battles fought by the 10th battalion, predicting 6 battles when 9 was actually observed. The accuracy of the estimation can be estimated by addition of the deviations as in example 1.1 to give us an estimate of the total error. Doing this shows that the total deviations add up to zero. Erroneously we way conclude that because the total of all deviations is zero implies a watertight model of predicting the number of battles, this is so because the model measures from the mean. Half of the scores are greater than the mean and half are less, half of deviations are positive and half-negative cancelling each other out on addition. To square each of the deviations helps to alleviate this problem as in example 1.1, adding these squared deviations gives us the sum of squared deviation or sum of squared errors which is a good measure of the accuracy of the mean (Davey, 2007:679).

It should be noted that the standard deviation is a helpful tool when we need to compare scores around a standard, the mean. Perhaps that is why it is known as the "root mean squared deviation". In employing variance and standard deviation, analysis provides us a clearer picture than we would get from averages like mean, mode and median.

The standard deviation is of limited use by itself, used primarily to compare. It can tell a researcher the similarities and differences of a set of data.

Numerical data, explained here under variance and standard deviation provides us vital information for proper decision making, but it is said that it is the action taken as a result of the decision that is important because this leads to consequences. This decision-making is very important in the governmental process because they are made on the basis of available information and numerical information is used to assess consequences, This aspect is statistical which flows from numerical information.

Self-Assessment Exercises

1) State the uses of the measures of dispersion.
2) Outline the steps in calculating the variance and standard deviation.

4.0   SUMMARY

The standard deviation of a statistical population data set or probability distribution is the square root of its variance. Apart from the use of standard deviation to express the variability of a population, it is also used to measure confidence in statistical conclusions.

5.0:   CONCLUSION

Standard deviation is a commonly used measurement of variability or dispersion used in statistics. It shows how much variation or "dispersion" there is from the average (mean, or expected value). A low standard deviation indicates that the data points tend to be very close to the mean, whereas high standard deviation indicates that the data are spread out over a large range of values.

6.0:   TUTOR MARKED ASSIGNMENT

1. Analyse a measure of variability

2. Distinguish between the standard deviation and the range.

3. From the scores below:

65	65	40	65	50	80	48	59	79	85	72	70
	69	70	45	65	76	77					

(a)   Find the mode.

(b)   Compute the median.

(c)   Rank the scores and find the range.

(d)   Present an f table with interval size of 10 beginning with 40.

(e)   Find the grouped mean.

(f)   Calculate the S, and state the social information it conveys.

7.0: REFERENCES/FURTHER READING

Abadegesin, Adeniyi, Resaq Olopoenia and Afeikhena Jerome; (2005) *Statistics for the Social Sciences* (Ibadan University Press).

Babbie, E. (2005) *The Practice of Social Research* (10th ed). Belmont: Wadsworth Pub.Co.

Baker, Therese 1999. *Doing Social Research*. New York: McGraw-Hill College.

Black, T (1999*), Doing Quantitative Research in Social Sciences: An integrated approach to research design, measurement and statistics*, London: Sage.

Bless, Claire; Craig Higson-Smith and Ashraf Kagee 2006. *Fundamentals of Social Research Methods: An African Perspective* (4th edition). Cape Town: Juta & Co. Ltd.

Bryman, Alan (2004) (2nd Edition) *Social Research Methods*, Oxford: Oxford University Press

Meier, Kenneth J.; Jeffrey L. Brudney and John Bohte.2006. *Applied Statistics for Public and Nonprofit Administration* (6th edition). Belmont, CA: Thomson Higher Education.

Mitchell, Mark L. and Janina M. Jolley(2007). *Research Design Explained* (6th edition). Belmont, CA: Thomson Higher Education.

UNIT 5: INFERENTIAL STATISTICS: ESTIMATION AND HYPOTHESIS TESTING

CONTENT

1.0: Introduction

2.0: Objectives

3.0: Main Content

    3.1: What is a statistical inference?

    3.2: Types of Inferential Statistics

    3.3: Application of Inferential Statistics

4.0: Summary

5.0: Conclusion:

6.0: Tutor - Marked Assignment

7.0: References/Further Reading

1.0 INTRODUCTION

This unit deals with statistical inference. It examines the different types of inference: estimation and hypothesis testing. It however focuses on estimation issues, such as standard error of the mean, the principle of confidence limits, and hypothesis testing.

2.0: OBJECTIVES

In this unit, you are expected to:

1. understand what is meant by statistical inference;

2. be able to calculate and use the standard error of the mean;

3. understand the principles of confidence limits;

4. be able to explain how researchers use inferential statistics to evaluate sample data; and

5.   be able distinguish between the null hypothesis and the research hypothesis.

3.0   MAIN CONTENT

3.1:   What is a Statistical Inference?

A statistical inference is the process by which conclusions are drawn about some measure or attribute of a population based upon analysis of sample data. Inferential statistics involves the use of mathematical methods that employ probability theory for deducing (inferring) the properties of a population from the analysis of the properties of a set of data (sample) drawn from it. Inferential statistics is concerned also with the precision and reliability of the inferences it helps to draw. Thus, inferential statistics are quantitative tools of analysis in the social sciences that enable us to analyse data, test hypothesis, and draw conclusions especially about populations from studies of samples. Good examples of these are correlation co-efficient analysis, multiple correlations and multiple regressions, factor analysis (f-test), and others. When these statistical tools are applied to social data, it is called social statistics.

The heart of statistics is inferential statistics. Inferential statistics are used when we want to draw conclusions about a population from the sample drawn from that population. For example, when we want to determine if a candidate is more popular than another over a period of time, or if there are differences in how two political parties perform in presidential elections. Inferential statistics are often complex and may have several different interpretations.

However, the goal of inferential statistics is to discuss some property or general pattern about a large group by studying a smaller group of people in the hope that the results will generalize to the larger group. For example, we may ask residents of Abuja city their opinion about senatorial elections in Abuja. We would probably poll few thousands individuals in Abuja city in an attempt to find out how the city as a whole views the issue of senatorial election in Abuja. This leads us to apply basic inferential statistics.

Inferential statistics is normally used to compare two or more groups and attempts are normally made to figure out if the two groups are different from one another. For example, a drug company has developed an anaesthetic pill that can increase recovery time from common cold. How do we find out if the pill works or not?

Using inferential statistics, what we can do is to get two groups of people from the same population. For example, we select two groups of people from a small part of Abuja who had just caught cold, and administer the pill to one group (group N), and give the other group a placebo (group K). We could then measure how many days each group took to recover. When we measure something in a population, it is called a parameter. When we measure something in a sample, it is called a statistic. To be sure, if I got the average age of parents in single-family homes, the measure would be called a parameter. If I measure the average age of a sample of those individuals, it would be called a statistics. Thus, a population is to a parameter as a sample is to a statistics. This distinction between samples and population is important because with inferential statistics, we want to draw inferences about population from samples. Thus this unit is mainly concerned with the rules or logic of how a relatively small sample from a large population could be tested, and the results of those tests can be

inferred to be true for everyone in the population. A good example is that, we may wish to test whether *Extra Panadol* is better than *Septrin* at relieving pains. To do this, we cannot give these drugs to everyone in the population; it is not practical since the general population is too large. Instead, we would give it to a couple of hundreds of people and see, which one works better with them. With inferential statistics, we can infer that what was true for a few hundred people is also true for a very large population of hundreds of thousands of people.

In our test case of a drug company that administered pill to two different groups, they could now measure how many days each group took to recover. What is to be done at this juncture is to calculate the mean of each group. Let us assume that the mean recovery time for the group with the drug was 4.4days, and the mean recovery time for the group with the placebo was 4.8 days. The question then is, is the difference due to random chance, or does taking the pill actually helps one to recover from the cold faster? The means of the two groups alone does not help us determine the answer to this question. We certainly need additional information. This information is obviously from the sample size.

3.2:   Types of Inferential Statistics

There are basically two types of statistical inference. These are estimation and hypothesis testing. Estimation is concerned with the estimation of population characteristics, such as the mean and the standard deviation, from the population characteristics.   Hypothesis testing is the process of setting up a theory or hypothesis about some characteristics of the population and then sampling to see if the hypothesis is supported or not.

3.2.1 Statistical Estimates

As we have noted earlier, a population is the total set of items we are concerned about. Any set of people or objects with something in common can be referred to as population. Anything could be a population. We could have a population of university students. We might be interested in the population of the elderly. Other example include among others, single parent families, people with depression, or bum victims. For anything we might be interested in studying, we could define a population. In most cases, we would like to test something about a population. A population is the entire group of people you would like to know something about.

For example, we might work to test whether a new drug might be effective for a specific group. It is impossible most of the time to give everyone a new treatment to determine if it worked or not. Instead, we commonly give it to a group of people from the population to see if it is effective. This subset of the population is called a sample. A sample is a subset of the population. You may wish to sample different types of cream in a Shoprite store. A sample of a population should be just a similar version of the entire population. It is extremely important to understand how the sample being studied was drawn from the entire population.

The sample should be as representative of the population as possible (see module 3, unit 4 on sampling procedure). There are however, several valid ways of creating a sample form a population, but inferential statistics works best when the sample is drawn at random from the population. A random sample is a sample in which every member of the population has an equal chance of being included. If a sample is not a random sample, then the rules of statistical inference under consideration do not necessarily hold. Given

large enough samples, drawn at random; ensures a fair and representative sample of a population.

A statistic is a measure that is used to summarise a sample. The mean, the standard deviation, and the median of a sample are all statistics. A measure that is used to summarise a population is called a parameter. The population characteristics are known as parameters while the sample characteristics as known as sample statistics. To improve clarity, statisticians use different symbols for the mean and the standard deviation for parameters and statistics.

The mean is calculated the same way regardless of whether the data are taken from a sample or population. The standard deviation is however calculated differently.

Recall the formula for standard deviation of a population

$$\sigma = \sqrt{\frac{\sum_{i=1}^{n}(x_i - \mu)^2}{N}}$$

Recall that the formula for standard deviation of a sample is

$$s = \sqrt{\frac{\sum_{i=1}^{n}(x_i - \bar{x})^2}{n-1}}$$

The difference is that the sum of the squared deviation from the mean is divided by n-1 rather than N.

Properties of good estimators

There are four properties of a good estimator, these are

1. An estimator must be unbiased. This implies that the mean of the sample means **x** of all possible random samples of size **n**, drawn from a population size **N**, equals the population parameter **μ**. Thus, the

mean of the distribution of sample means would equal the population mean.

2. An estimator must be consistence. This means that as the sample size increases, the precision of the estimate of the population parameter also increases.

3. An estimator must be efficient. This means that in the repeated sampling, its variance is smaller.

4. An estimator must be sufficient. This means that it uses all information in the sample in estimating the required population parameter.

It must, however, be noted that all the population values are usually not available so we have to sample in order to estimate the population parameters. Hence, it is necessary to distinguish between the symbols used for ample statistics and population parameters as shown in the table below.

Symbols for Parameters and Statistics

Measure	Population parameter	Sample statistics
Mean	μ	x
Standard deviation	Σ	S
Number of cases	N	N

Estimating a population mean

The best estimate of a population mean σ is the mean of the sample x. The use of the sample mean $\ddot{x}$ to make inferences about the population mean is a common procedure in statistics. For instance, assuming we have a table

189

showing the number of arrest in 2011 made by 10 police officers at the Sango Police Station in Ibadan.

Table: Number of Arrests by Police Officers in Ibadan (2011)

Police officer	Number of Arrests (2011)
1	14
2	16
3	10
4	18
5	8
6	15
7	17
8	20
9	19
10	13

The mean number of arrest by police officers is 15.0. But in a situation where the population parameter is not known and therefore cannot be calculated or too large and therefore unwieldy to calculate, a mean of a sample can be used to estimate the population mean. If we take a random sample of the five police officers from Sango police station, using a random number table, and calculate the mean for those five as stated below:

Officers in sample	Arrests	Mean
1,3,2,8,4	14, 10, 16, 20, 18	15.6

It is immediately obvious that the sample mean is an estimate of the population mean. This is not exact but close. The discrepancy is the result of sampling error. This is because the sample is not a perfect representative

of the population. However, if we took numerous samples of five, the average sample mean would approach the population mean.

The standard deviation for mean estimates is called the standard error. To find the standard error of the mean, we calculate a mean for each sample and then calculate the standard deviation for the mean estimates.

3.2.2 Hypothesis Testing

A hypothesis is a statement about the world that may be tested to determine whether it is true or false. It is some testable belief or opinion. Hypothesis testing is the process whereby such beliefs or statement are tested by statistical means. Examples of testable hypothesis include the following:

1. Following the mounting of road blocks by the police the number of high way robberies dropped.

2. After implementing the civil service reform, the productivity of service has increased.

Hypothesis expressed in the negative is called the null hypothesis. The research hypothesis is expressed in the positive. In order to test a null hypothesis, you must first state the research hypothesis. A research hypothesis is always paired with a null hypothesis. The hypothesis to be tested is the null hypothesis designated H_O.

If the research hypotheses states that an outcome has been realised, the null hypothesis states that an outcome has not been realised. If the research hypothesis states that a change has occurred, the null hypothesis states that change has not occurred.

Results of Hypothesis Testing

There are only four possible results when we test hypothesis, H_O.

1. We accept a true hypothesis – a correct decision

2. We reject a false hypothesis – a correct decision

3. We reject a true hypothesis – an incorrect decision. This is known as type I error
4. We accept a false hypothesis – an incorrect decision. This is known as type II error.

Significance Levels

When a sample is taken to test a hypothesis there is no guarantee that the information from the sample data completely supports the hypothesis. This may be due either to the fact that the original hypothesis is wrong or the sample is slightly unrepresentative. Generally, all samples are to a greater or lesser extent unrepresentative. It is therefore important to test which of the two is the case. The test will show whether any difference can be attributed to ordinary random factors or not. If the difference is probably not due to chance factors, the difference is said to be statistically significant. Since we cannot say for 100% certainty that a difference is significant because we are dealing with samples and random factors, various levels of significance are chosen, most commonly 10%, 5% or 1%.

Estimate of the Sample Size

Statistics is normally used to compare two or more groups and attempts are normally made to figure out if the two groups are different from one another. But if our sample size only consisted of two people, that is, one from the drug group and one from the placebo group as in the drug case we alluded to earlier, there would be so few participants that we would not have much confidence that there is a difference between the two groups. That is to say, there is a high probability that chance explains our results as many explanations account for this. For example, one person

might be younger, and thus have a better immune system. However, if our sample consisted of 1,000 people in each group, then the results become much more robust while it might be easy to say that one person is younger than another, it is hard to say that, 1,000 random people are younger than another 1,000 random people.

If the sample is drawn at random from the population, then these 'random' variations in participants should be approximately equal in the two groups, given that the two groups are large. This is why inferential statistics works best when there are lots of people involved. Even if we have a large enough sample size, we still need more inferential statistics to reach conclusion. What we need is some measure of variability.

Variability

It takes about 5-6days to recover completely from cold. But, it is pertinent to ask if everyone around takes 5-6days, or do some people recover in 1 day, and others recover in 10 days? Understanding the spread of the data will tell us how effective the pill is. If everyone in the placebo group takes exactly 4.8 days to recover, then it is clear that the pill has a positive effect. But if people have a whole variability in their length of recovery (and they probably do), then the picture becomes a little fuzzy. It is only when the mean sample size variability has been calculated that a proper conclusion can be made. In the case under review, if the sample size is large, and the variability is small, then we would receive a small value (probability-value). Small p-value is good and this term is prominent enough to warrant further discussion.

P-Values

In classic inferential statistics, two hypotheses are made before the commencement of study. These are the **null hypothesis,** and the **alternative hypothesis.**

The null hypothesis states that the two groups we are studying are the same, while the alternative hypothesis states that the two groups we are studying are different. The goal of classic inferential statistics is to prove the null hypothesis wrong. The logic says that if the two groups are not the same, then they must be different. A low p-value indicates a low probability that null-hypothesis is correct, thus providing for the alternative hypothesis.

What a p-value actually means is that the 'p' value one obtains from a test like this tells you precisely the following:

It is the probability that you would obtain these or more extreme results assuming that the null hypothesis is true. For example, if we obtained a p -value of 0.01 or 1% for our drug experiment, it would then mean that the probability of obtaining a difference between these two groups, that is, this larger is 1% assuming that the two group are in fact not different.

When interpreting p-values, it is important to understand that it does not tell you the probability that the null hypothesis is wrong.

Self-Assessment Exercises

1) List the properties of a good estimator
2) What does confidence level mean?

4.0   SUMMARY

A statistical inference involves drawing inference or conclusions about the population from the samples. Estimation is concerned with estimating population parameters from sample statistics. The central limits theorem states that the means of samples tend to be normally distributed almost

regardless of the shape of the original population. The standard deviation of the distribution of means is called the standard error of the mean. Hypothesis or significant testing is testing a belief or statement by statistical methods. Based on the null hypothesis H_O, a type I error is rejecting a true hypothesis and type II error is accepting a false hypothesis. Significance levels are complementary concepts to confidence limits. 1%, 5% and 10%, are the most usual levels.

5.0   CONCLUSION

It is good to conclude our discussion on the basic inferential statistics by summarizing the usefulness of inferential statistics in social sciences and indeed in political evaluation research as follows:

(i)   It empowers us to read and understand professional literature;

(ii)   It helps us to carry out more scientific and better quality researches, especially when writing term papers, articles or projects;

(iii)   Inferential statistics equip us with the knowledge and skill of using statistical techniques needed in pursuit of higher education and training;

(iv)   It is essential for carrying out professional training and assignments;

(v)   It encourages the cultivation of the habit of disciplined thinking and accurate, precise and logical reasoning;

(vi)   It enables a researcher to summarize and describe his data and results more meaningfully, interestingly and exactly using tools such as means, standard deviation, frequency tables and polygons etc.

(vii)   It enables researchers make predictions, forecasts and projections;

(viii) It helps us estimate or determine relationship between variables using chi-square or correlation analysis and to explore causal

195

relationship between variables using regression, multiple regression, step-wise regression etc;

(ix) It brings research methods, data analysis, and research findings to near to scientific methods and spirit;

(x) Finally governments need high quality advice and information on (social) issues for policy formation and execution. Facts, figures, and analytical outputs, and well collected data and information (from the public at times) are needed for such advice and information. Social statistics, including basic inferential statistics enables researchers to acquire the relevant facts, figures, data and information that will be daily useful to move the society forward.

6.0: TUTOR-MARKED ASSIGNMENT

1. Explain statistical inference
2. Discuss the uses of statistical inference in political evaluation
3. Distinguish between the null hypothesis and the research hypothesis.
4. List when a researcher decides to reject the null hypothesis

7.0:   REFERENCES/FURTHER READING

Abadegesin, Adeniyi, Resaq Olopoenia and Afeikhena Jerome; (2005) *Statistics for the Social Sciences* (Ibadan University Press).

Babbie, E. (2005) *The Practice of Social Research* (10th ed). Belmont: Wadsworth Pub.Co.

Bryman, Alan (2004) (2nd Edition) *Social Research Methods*, Oxford: Oxford University Press

Cosby, Paul C. 2006. *Methods in Behavioural Research*. New York: McDraw-Hill

Jones, E. Terrence (1971). *Conducting Political Research,* New York: Harper and Row Publishers

Meier, Kenneth J.; Jeffrey L. Brudney and John Bohte.2006. *Applied Statistics for Public and Nonprofit Administration* (6th edition). Belmont, CA: Thomson Higher Education.

Mitchell, Mark L. and Janina M. Jolley 2007. *Research Design Explained* (6th edition). Belmont, CA: Thomson Higher Education.

Glossary of Terms

Sample: A sample is the segment of the population that is selected for investigation. It is a subset of the population whose method of selection may be based on a probability or a non-probability approach.

Sampling: Sampling is the process of selecting a sub-set of observations from among many possible observations for the purpose of drawing conclusion about that larger set of possible observation

Element: An element is that unit about which information is collected and which provides the basis of analysis e.g. people, families, social clubs etc.

Universe: A universe is the theoretical and hypothetical aggregation all elements, as defined for a given study. If the individual Nigerian is the element for a survey, then Nigeria would be the universe.

Survey Population: A survey population is that aggregation of element from which the actually selected.

Sample Unit: This is the element or set of elements considered for selection in some stages of sampling.

Sampling Frame: The listing of all units in the population from which the sample will be selected.

Representative Sample: A sample that reflects the population accurately so that it is microcosm of the population.

Sampling Error: The difference between a sample and the population from which it is selected, even though a probability samples has been selected.

Non Sampling Error: The difference between the population and the sample that arise either from deficiencies in the sampling approach or from such problems as poor question wording, poor intervening etc

Non-Response: A source non-sampling error that is particularly likely to happen when individuals are being sampled.

Variable: A variable is a set of mutually exclusive attributes; sex, age, employment, status, height etc. The element of a given population may be described in terms of their individual attributes on a given variable.

Parameter: This is the summary description of a given variable in a population. The mean income of all families in a city and the age distribution of the city's population are examples of parameter.