STATISTICAL ANALYSIS IN EDUCATION

Editors:
Martins Fabunmi
Babatunde Adeniyi Adeyemi
Oluwabunmi Blessing Adelaja

STATISTICAL ANALYSIS IN EDUCATION

Editors:
Martins Fabunmi
Babatunde Adeniyi Adeyemi
Oluwabunmi Blessing Adelaja

ISBN: 978-978-60153-4-7

R&DH-UK

Published by the Education Dialogue Association (EDUDIA) and the Research and Dialogue Hub UK

Review Process

Fifteen articles were submitted for publication as chapters in this book. The articles were subjected to a thorough process of double-blind peer review. The professionals in EDUDIA's team of Reviewers were from universities in different countries. They were strictly guided by the EDUDIA's Review Criteria. They were also requested to look at the manuscripts with the view to assisting authors produce quality articles that best meet learners' needs.

Following the review process, the editorial committee considered the reviewers' comments; and all the articles were suitable for publication. The authors of the suitably qualified articles were given the reviewers' reports and asked to use the suggestions to strengthen their papers. After receiving the corrected manuscripts, the editorial committee finally accepted 12 of the 15 qualified articles for inclusion in this Book: *Statistical Analysis in Education*. That means that the acceptance rate was about 80.0%.

Preface

This book, titled *Statistical Analysis in Education*, is derived from responses to a call for contribution of chapters from erudite scholars sometimes in January 2022. The following three accomplished scholars edited the book, the first two are full professors and the third is a doctoral degree student: Martins Fabunmi, Modibbo Adama University, Yola, Nigeria; Babatunde Adeniyi Adeyemi, Institute of Education, Obafemi Awolowo University, Ile-Ife, Nigeria; Oluwabunmi Blessing Adelaja, Department of Information Resource Management, Babcock University, Ilisan, Ogun State, Nigeria. Virtually all the authors are very senior academics with rich experiences in writing scholarly articles. Some of them are even professors, while others are budding erudite lecturers.

This book comprises twelve chapters with the following titles: Statistical Analysis in Education; Statistical Variable; Measurement Scales; Testing Hypotheses; Parameter Estimation; Organizing Data; Measures of Central Tendencies; Measures of Variability; Measures of Association; Comparing Variables; Computer Applications for Statistics in Education; and Data Interpretation in Statistics.

EDUDIA addresses the shortage of e-books of African backgrounds. The association has started with management in education, with this book dealing with the *Statistical Analysis in Education* aspect of the discipline. The book contains articles on exigent issues in *Statistical Analysis in Education*. EDUDIA continues to solicit scholarly articles from authors in different disciplines from all over the world through conferencing. I, therefore, implore all teacher educators to be part of this great mass movement and have their articles published in globally visible publishing outlets.

In view of these developments, I recommend this book as a must-read for everybody, irrespective of discipline. The book deals with *Statistical Analysis in Education* in a way that makes it a useful learning material for learners, an instructional and dependable guide for teacher educators, researchers, administrators, and everyone, no matter the professional calling. It is useful for individuals, researchers, family heads, religious leaders, community leaders, heads of organisations, and members of different groups. Finally, I am grateful to the co-editors for making the publication of this book, *Statistical Analysis in Education*, possible.

Professor Martins Fabunmi (FGEN),
Department of Educational Foundations, Modibbo Adama University, Yola, Nigeria.

CONTENTS

Chapter 1

Statistical Analysis in Education

Oluwabunmi Blessing Adelaja
Department of Information Resource Management
Babcock University, Ilisan, Ogun State, Nigeria

Martins Fabunmi
Research and Dialogue Hub-UK
Barnsley, United Kingdom

Introduction

Statistical analysis in education involves the use of statistical methods to compute, illustrate and display educational data (Degroot & Scherish, 2012). Statistics, which is the base word for 'statistical' is used to describe a function that returns its actual value (Spektrum, 2024). Statistics is a discipline that deals with data, facts and figures with which meaningful information can be inferred. Statistical data may be a numerical value, in the form of quantitative data, or a label, as with qualitative data (e.g., categorical variables such as gender or occupation). There are two major types of statistical analysis in educational research. These are descriptive and inferential statistics. Descriptive statistics summarises data characteristics, while inferential statistics enables researchers to draw conclusions about a larger population based on a sample. Attention is first devoted to these two types and subsequently, we will discuss other types of statistics.

In a descriptive study, measures of central tendency and measures of dispersion may be used. Examples of measures of central tendency, otherwise referred to as measures of centre or central location, include: mean, median, and mode. It is a summary measure that attempts to describe a whole set of data with a single value that represents the middle or centre of its distribution. While the measures of dispersion are: range, interquartile range, standard deviation and variance. They explain the deviation scattering of data, that is, the disparity of data from one another. Thus, giving a precise view of their distribution. Descriptive statistics include: Measures of central tendency, (such as mean, median, mode), Measures of variability (such as standard deviation, range, interquartile range) and Frequency distributions: which entail tabulating how often different values occur. Data visualisation, which involves using tables, graphs and charts, is used to present complex data in a clear and understandable way to communicate research findings effectively.

Inferential statistics entails methods like hypothesis testing, correlation analysis, and regression analysis. Due to its probabilistic nature, statistics has been described as 'the science of uncertainty and the technology of extracting information from data (Hand, 2010). Inferential statistics interprets data from a population sample with a view to inducing statements and making predictions about a population (Britannica, 2024; Tanton, 2005; Hand, 2010). In educational research, common inferential tests include testing for: differences between two or among more sets of scores; relationship between two or among more sets of scores; association between two or among more sets of scores; and whether certain variables can determine or predict the occurrence of another variable. Inferential statistics are often used when testing hypotheses in educational research. This entails testing a claim about a population parameter using statistical tests like t-tests, ANOVA, chi-square tests. Inferential statistical analysis includes the following.

i. **Correlation analysis**
This involves testing the relationship between two variables or among more than two variables. For example, relationships between student study habits and test scores.

ii. **Regression analysis**
It entails predicting the value of one variable based on the values of other variables. That is, how changes in one variable might affect another, such as how a new teaching method could impact student achievement.

iii. **Analysis of Variance (ANOVA)**
It is used to compare means across multiple groups to identify significant differences between them. Such as comparing student performance across different classrooms with varying teaching methods.

iv. **Exploratory data analysis**
This is an initial analysis to identify patterns and trends in data. Exploratory data analysis could be used to examine and visualise data to understand the characteristics, identify patterns, spot anomalies, and test hypotheses. This enables the researcher to summarise the data and uncover insights before applying more advanced data analysis techniques.

v. **Factor analysis**
It involves identifying underlying factors that contribute to a set of observed variables. Factor analysis is a statistical method that enables researchers to identify patterns in data by reducing many variables into a smaller set of factors. This is a useful method for simplifying complex data and understanding the underlying structure of the data set.

vi. **Time series analysis**
This entails analysing data collected over time. In educational research methods, time series analysis is a statistical technique that is used to examine and interpret data that are collected over time. Thus, allowing researchers to identify patterns, trends, and relationships within a sequence of observations taken at regular intervals, often to forecast future values based on historical data patterns. It is a statistical method for studying how variables change over time by analyzing data points collected chronologically.

vii. Causal analysis
It is an attempt to establish cause-and-effect relationships between variables. A causal analysis involves looking for relationships between causes and effects. This entails looking closely at each cause and identifying the relationships between them.

Other types of statistics are:

Parametric Statistics

Parametric statistics refers to statistical methods that rely on assumptions about the underlying population distribution, typically assuming the data follow a specific pattern like a normal distribution and use parameters like mean and standard deviation to analyze the data; essentially, it means making inferences about a population based on the characteristics of a known distribution. It is a branch of statistics that relies on models based on a fixed set of finite parameters (Oxford University Press, 1478). Unlike nonparametric statistics, parametric tests assume the data are taken from a specific population distribution, often a normal distribution. The parametric tests rely on population parameters like mean and standard deviation to calculate test statistics. Examples of parametric tests include: student's t-test, ANOVA (Analysis of Variance), Pearson correlation coefficient (Mann, 1995; Purdue, 2025; Fabunmi, 2023).

Nonparametric statistics

Nonparametric statistics refers to a set of statistical methods that are used to analyze data without making assumptions about the underlying population distribution (Oxford University Press, 1478, 2008). This implies that these methods can be used even when the exact shape or form of the data distribution is not known. They often rely on the ranking or order of data points rather than their numerical values. This is different from the parametric statistics which require specific distribution assumptions about the data. Examples of nonparametric tests include: The Wilcoxon test, Spearman's rank correlation coefficient, Mann-Whitney U test, Kruskal-Wallis test, and Chi-square test (Mann, 1995; Fabunmi, 2023; Purdue, 2025).

Validity and reliability of instruments

Reliability and validity are very important while constructing a survey instrument (Mann, 1995; Fabunmi, 2023; Purdue, 2025). The term reliability refers to the extent which an instrument yields the same results over multiple trials, that is, persistently obtains the same result in multiple investigations. While validity refers to how precisely the instrument measures what it was designed to measure. When assessing the validity and reliability of a research instrument, common methods are: test-retest reliability, parallel forms reliability, split-half reliability, and internal consistency. While the statistics often used include correlation coefficients (for reliability like test-retest or internal consistency), factor loadings (for construct validity), and comparisons to external criteria (for criterion-related validity), with higher values generally indicating greater validity and reliability. The specific statistical method to be used depends on the type of instrument, its scale of measurement and the chosen validation method. Scholars often wrongly use reliability statistics like Cronbach Alpha for testing the internal consistency of instruments without considering the scale of the measurement. Where the scale of instrument is ordinal, the appropriate statistic is Ordinal Alpha. However, if it's continuous data, the appropriate statistic is Cronbach Alpha (Fabunmi, 2023).

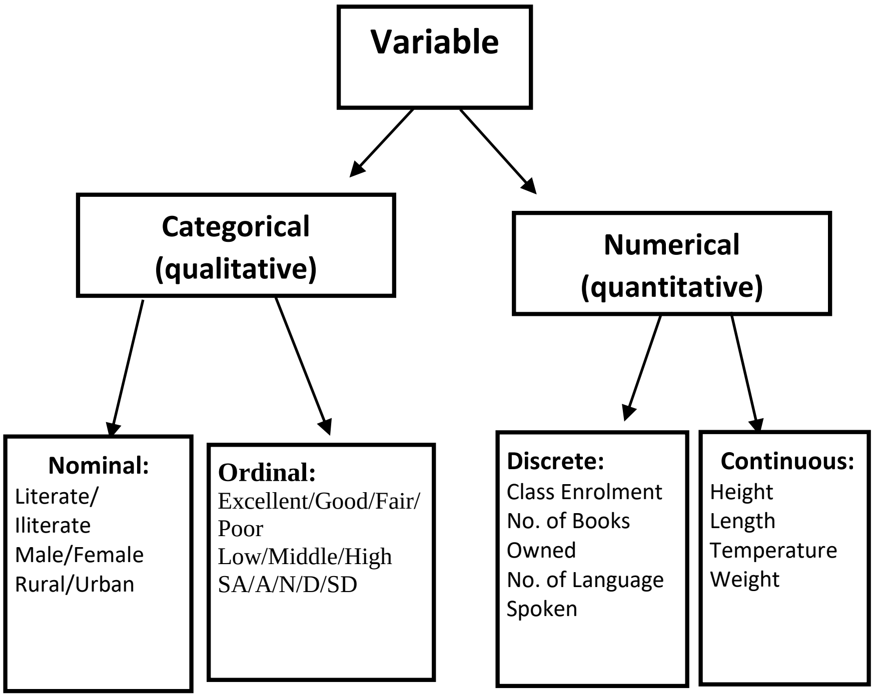
Types of Data

The type of data determines the appropriate statistical analyses. There are several types of data. These include (Fabunmi, 2023):

1) Nominal data refer to a type of qualitative data that groups variables into categories, such as male – female, rural – urban classifications.
2) Ordinal data are a type of qualitative data that group variables into ordered categories, which include a natural order or rank based on a hierarchical scale like the Body Mass Index (BMI), which is categorised as underweight, normal weight, overweight, and obese.
3) Discrete data are those that only take certain values, often in the form of whole numbers or integers. It should be noted that a set of discrete data can be counted and has a finite number of values. Examples of discrete data include the number of students in a class, test scores, number of books in the library, etc.
4) Continuous data are those that can take any value within a given range. They are quantitative data that describe data points that are not separated by distinct intervals. Such include: height, length, temperature and weight.

Statistics and scale of measurement

In educational research, the utmost consideration should be given to using appropriate statistics for different scales of measurement. Nominal data only allows for the use of mode as a central tendency measure. Ordinal data can use median and mode. Interval and ratio data can use all three measures of central tendency: mean, median, and mode, and also most measures of variability: range, variance, and standard deviation, due to their inherent properties of order and equal intervals. Researchers should always consider the level of measurement when selecting statistical tests (Fabunmi, 2023). Figure 1 illustrates the appropriate statistics for different scales of measurement.



Key to Ordinal Scale Abbreviations:
SA = Strongly Agree
A = Agree   N = Neutral
D = Disagree   SD = Strongly Disagree

Figure 1: Scale of Measurement

Data interpretation in educational research

Data interpretation is the process of identifying the inherent patterns and trends to draw inferences or conclusions from a completed data analysis. It is an important aspect of educational research. Data interpretation helps researchers understand what their results mean and how practical such results are. It is essential to do the following while interpreting data (Fabunmi, 2023):

i. Utilise the level of significance for making decisions on the statistical results obtained.
ii. Use models for illustration: Models categorize data into groups based on its characteristics, which can help identify events and scores
iii. Identify patterns in the results that were generated: Identify trends, clusters, and anomalies in the data
iv. Illustrate with charts: Charts can help identify trends and outliers
v. Compare data based on predetermined criteria or categories.
vi. Compare data with similar data from other places, times, or scales.

Summary

Statistics are used in educational research to analyse and interpret data, thus allowing researchers to test hypotheses, identify patterns, and draw meaningful conclusions from large datasets. They also provide a solid foundation for supporting or rejecting research claims thereby making it a crucial tool for drawing reliable inferences from studies across various fields. Statistics are important because they enable researchers to collect, analyze, and interpret data upon which meaningful conclusions about populations could be drawn. All these will enable informed decision-making by the management and education authorities.

Revision Questions

1) Write short notes on:
    a) Statistics
    b) Measures of central tendency
    c) Measures of variability
    d) Frequency distribution
2) Explain any six inferential statistical methods.
3) Differentiate parametric statistics from non-parametric statistics.
4) Discuss the concepts of validation and reliability of instruments.
5) Explain any four scales of measurement and the types of data that are appropriate for data analysis while using such scales.
6) Discuss the essential issues in data interpretation.

References

Britannica (2024). Statistics, Definition, Types, & Importance, Retrieved from: "Statistics | Definition, Types, & Importance | Britannica". www.britannica.com.

DeGroot, M. H., & Schervish, M. J. (2012). Probability and statistics. Pearson Education.

Fabunmi, M. (2023). Guided Thesis Writing. University of Cape Coast, Ghana: IEPA.

Hand, David (2010). "Statistics: An Overview". In Lovric, Miodrag (Ed.). International Encyclopedia of Statistical Science. Springer. pp. 1504–1509.

Hoad, T. F., ed. (1996). The Concise Oxford Dictionary of English Etymology. Oxford University Press. p. 460.

Mann, Prem S. (1995). Introductory Statistics. Retrieved from: Introductory Statistics (2nd ed.). Wiley.

Oxford University Press (1478). Statistics. Retrieved from: "statistics". Oxford English Dictionary

Oxford University Press (2008). Statistics. Retrieved from: "Statistics". Oxford Reference. Oxford University Press.

Purdue University (2025). Basic inferential Statistics. Retrieved from: "Basic Inferential Statistics - Purdue OWL - Purdue University".

Sheynin, Oscar (2010). "Statistics, History of". In Lovric, Miodrag (Ed.). International Encyclopedia of Statistical Science. Springer. pp. 1493–1504.

Spektrum (2024). Statistik. Retrieved from: "Statistik". www.spektrum.de (in German). Retrieved 2024-12-30.

Tanton, James (2005). "Statistics". Encyclopedia of Mathematics. pp. 478–484.

Chapter 2

STATISTICAL VARIABLE

Shikha Trivedi (PhD)
Department of Early Childhood, Botswana Open University
E-Mail: strivedi@staff.bou.ac.bw

1.0 Introduction

A statistical variable is a characteristic that can be measured and can assume different values among a population. It may also be called a data item since it is essential for understanding and describing data. The concept of 'variable' is underlined by the fact that the value may vary between data units in a population. In statistics, variables are fundamental concepts that provide the information necessary to analyse relationships, trends, patterns, and predictions. A variable is not only what we can measure, but we can also manipulate and control it, particularly in experimental studies. Variables are crucial in study design selection, statistical method application, and accurate result interpretation.

Some examples of variables are:

- Gender is a variable with three values: male, female or transgender.
- Family type is a variable with values such as nuclear family, extended family, blended family, single-parent family, and child-headed family.
- Marital status is a variable with values such as, single, married, divorced, or widowed.
- A person's attitude toward women's leadership is a variable, ranging from strongly agree to strongly disagree.

Statistical variables can be broadly categorised according to types and forms, each serving distinct purposes in data analysis and interpretation. Variables can be classified based on type (categorical vs. numerical) or nature (qualitative vs. quantitative). Categorisation according to forms is based on their role in a study (independent, dependent, control etc.), focusing on how they are used and analysed in research.

2 Types of Statistical Variables

Variables may be classified into two main categories: categorical and numerical. Each category is then classified into two subcategories: nominal or ordinal for categorical variables and discrete or continuous for numeric variables.

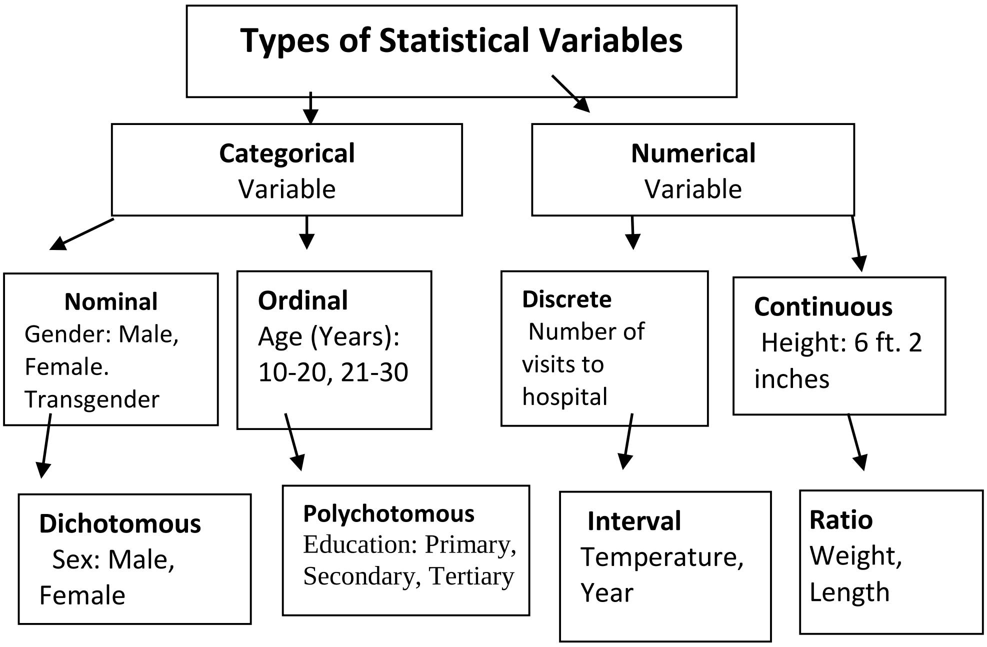


Figure 1: Types of Statistical Variables

1) **Categorical Variables** (qualitative variables)

Derived from its name, categorical variables represent categories or groups. It is a collection of information that has been classified. This data is called categorical because it consists of distinct groups. The attributes of a categorical variable vary in kind rather than degree, level, or quantity (Nwankwo and Emunemu, 2014).

Examples of categorical variables include gender (male/female), marital status (single/married/divorced), or eye colour (Black, Green, Brown). Here, all attributes of a category or subset are recognised as the same and assigned the same values. In statistics, categorical variables have a fixed number of possible values. Another way of explaining the categorical variable is when an institution acquires biographic data on its employees based on variables such as gender, level of education, income, and so on; the resulting information is categorical. These variables are typically assigned names or labels. The categorical data can be best represented using bar graphs and pie charts.

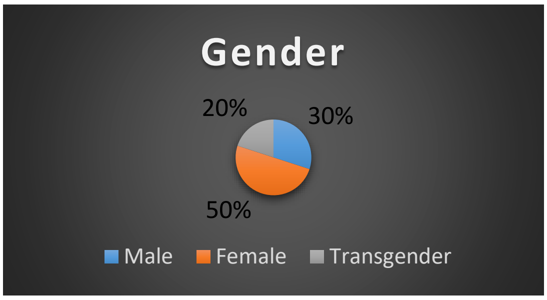


Figure 2: Pie Chart of showing gender

Categorical variables can be further categorised as either **nominal** or **ordinal**.

i) Nominal Variables

In Latin "Nomen" means "Name". Nominal variables have groups or values but do not have an essential order, inherent hierarchy, or structure. Examples include blood type, ethnic group, gender, etc. It can further be divided into **dichotomous** and **polychotomous** variables.

a) **Dichotomous variables** have only two categories or levels and may have similar features. They do not have intrinsic orders as nominal variables. An example could be the academic and non-academic cadre of staff in an institution of learning.

b) **Polychotomous variables** have more than two categories of categories. For example, marital status could have categories like single, married, and widowed.

ii) Ordinal Variables

Ordinal variables are statistical data in which the values follow an order or scale**.** One of the most distinguishing characteristics of ordinal data is that differences between data values cannot be determined**.** The most common way to present ordinal data is a bar chart. Examples of data that can be represented in ordinal variables are socioeconomic status (low income, middle income, high income), education level (primary, high school, bachelor's degree, master's degree).

2) **Numerical Variables** (quantitative variables): as the name implies, numerical variables represent measurable quantities. They can further be divided into discrete and continuous variables. Discrete Variables take on a finite numerical values, such as the number of children in a family (1, 2, 3, etc.). Continuous variables can take on any value within a range and often represent measurements. Examples include height, weight, temperature, and income. Example: 3 ft., 5 inches, 4 ft., etc. For instance, in a study that seeks to explore the relationship between teachers' income and job satisfaction at ABS High School.

   A researcher can conclude on factors influencing the workplace by analysing how income levels affect job satisfaction ratings across academic and non-academic groups. It should be recalled that income is a numerical variable, and Job satisfaction is a categorical variable with options like satisfied, Neutral and unsatisfied.

Continuous variables can be further divided into interval or ratio variables.

(**a) Interval variables are** variables measured at intervals on a scale. An interval variable has an equal distance between each value. It has an order and can be measured along a continuum in a scale. Subsets have a numerical value but an absolute zero point. The difference between the two values has a meaning. For example, the difference between 30°F and 40°F is the same as the difference between 50°F and 60°F, demonstrating equal intervals. However, because temperature in Fahrenheit does not have a true zero, it is classified as an interval variable.

(b) Ratio variables have the same properties as an interval variable, but it clearly defines zero. When the variable equals 0.0, it means none of that variable exists. An example is time durations, such as meeting length measured in minutes or hours. Zero duration means no time has passed. Another example is weight since it is a ratio variable. The weight of 10 grams is twice as heavy as a weight of 5 grams. We cannot apply the same theory for interval variable. For example, a temperature of 10 degrees C should not be considered twice as hot as 5 degrees C. A conflict would be created if it were because 10 degrees C is 50 degrees F and 5 degrees C is 41 degrees F. Clearly, 50 degrees is not twice 41 degrees. Weight is another example, 0 lbs. is a meaningful absence of weight.

Each type of variable is analysed using different statistical techniques. Measures such as frequencies, proportions, and chi-square tests are commonly used for qualitative variables. Quantitative variables are analysed using measures of central tendency (mean, median, mode), measures of dispersion (variance, standard deviation), and correlation coefficients.

3 Forms of Variables

Variables can be categorised based on their role in the study (e.g., dependent and independent variables) and their relationship to other variables (e.g., confounding or control variables). The five variables are independent, dependent, control, confounding, and composite.

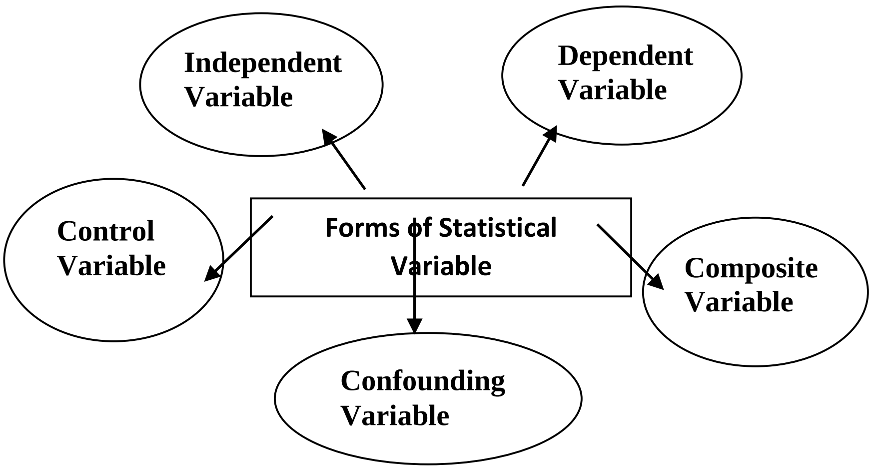


Figure 3: Forms of Variables

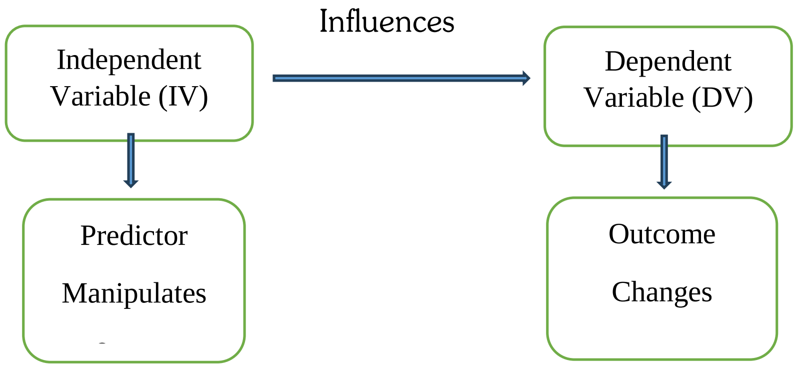
i) Independent Variables

Variables that are changed/manipulated so that their impact on the dependent variables can be monitored. They are sometimes also called predictor variables.

ii) Dependent Variables

These variables rely on something else (the independent variables) to occur/change before they can have a result (outcome variable).

Dependent variables are usually the variables the researcher is interested in.



For example, in a research that seeks to investigate the effect of family background on the academic performance of young children, the family background would be the independent variable in this study, and academic performance would be the dependent variable because it is measured and expected to change based on the family background factors (independent variable).

3. Control /Extraneous Variables:

Control variables are held constant to prevent them from influencing the relationship between the independent and dependent variables. For instance, in a study on the effect of using technology (independent variable) in a classroom on student performance (dependent variable), factors like students' prior knowledge about technology, socioeconomic status, or teacher experience with technology might be controlled to isolate the use of technology impact.

4. Confounding Variables:

Extra variables that you did not account for and can affect the whole experiment fall under confounding variables, which means extraneous factors can influence the outcome of an experiment or study, potentially leading to incorrect conclusions about the relationship between the studied variables. These variables can obscure the actual effect of the independent variable on the dependent variable, making it difficult to determine if an observed association is due to the variable of interest or some other factor.

For example, in a study examining the effect of exercise on weight loss, diet could be a confounding variable because it is related to exercise habits. Another example could be the effect of smoking (independent variable) on lung cancer (dependent variable). The confounding variable could be age (older individuals are more likely to have smoked more and thus likely to develop lung cancer). To address the confounding variable, you can design the study to minimise its impact or use statistical methods, such as regression analysis, to adjust for its effects.

5. Composite Variables:
A composite variable is a single variable created by combining two or more individual variables, often to summarise or reduce the complexity of the data. Composite variables can be created by adding the individual variables together or applying weights to each variable based on their relative importance or contribution to the composite. For example, BMI can be a composite variable requiring a person's height and weight.

4. Critical Traits of Variables in Research

A variable trait is a specific characteristic, quality, or tendency that someone or something has. Understanding these traits helps researchers design robust studies, choose appropriate measurement tools, and apply correct statistical analyses, leading to more accurate and meaningful results. The traits of variables depend upon their role in the research. Some of the key traits of variables are: nature of variables (quantitative or qualitative), role in research (independent variable, dependent variable, control variable, and confounding variable), level of measurement (nominal, ordinal, interval, and ratio), measurement scale (discrete and continuous), direction of relationship (positive, negative, and no relationship), distribution (normal distribution and skewed distribution). This knowledge is a catalyst for producing more accurate and meaningful results in your research.

5. Importance of Statistical Variables

1) **Analysis**: A researcher analyses and interprets data using variables. Different types of variables require different statistical methods for analysis.

2) **Comparison:** Comparing variables can help one understand similarities and differences between groups or populations.

3) **Prediction:** Variables are used in predictive modelling to forecast outcomes or trends based on historical data.

4) **Decision Making:** Variables provide the basis for making informed decisions in various fields, including business, healthcare, education, etc.

Summary

The following summarizes key aspects of statistical variables:

- Statistical variables are characteristics or attributes that can take on different values.
- Variables can be qualitative or categorical (nominal and ordinal) or quantitative or numerical (discrete, continuous).
- Categorical variables represent groups and can either be in a specific (ordinal) or no inherent order (nominal).
- Extraneous variables are not studied purposely but can create an error in the results.
- The interval variable has no true zero, and the ratio has a true zero on scales.

- The traits of variables depend upon their role in the research.

Revision Questions

1. What are some examples of nominal variables in a demographic survey?
2. Compare and contrast qualitative (categorical) and quantitative (numerical) variables.
3. Give examples of continuous variables and explain why they are considered continuous.
4. Describe any three traits of a variable.
5. State the difference between the variable type and the variable's form.

References

Nwankwo, J. I. and Emunemu, B. 0. (2014). Handbook on Research in Education and the Social Sciences. Ibadan: Giraffe Books.

https://www.classgist.com/blogs/206/how-to-identify-independent-and-dependent-research-variables

https://statistics.laerd.com/statistical-guides/types-of-variable.php

https://www.quora.com/What-is-the-difference-between-the-type-of-variable-and-the-classification-of-variable-in-statistics

https://www.researchprospect.com/variable-and-its-types-in-statistical-analysis/

https://testbook.com/maths/categorical-data

https://atlasti.com/research-hub/types-of-variables-in-research

Oyebanji, O.J.A., (2017). Research Variables Types, uses, and Definitions of terms. https://www.researchgate.net/publication/342897909

https://www.graphpad.com/support/faq/what-is-the-difference-between-ordinal-interval-and-ratio-variables-why-should-i-care/

https://web.pdx.edu/~newsomj/pa551/lecture1.htm

Chapter 3

MEASUREMENT SCALES

Memory Queensoap, *PhD*
Dogitimiye Memory, *PhD*

Introduction

Statistics and data interpretation are fundamental constructs in education. Statistics deals with the description and summarization of data which leads to the making of reasonable interpretations as well as drawing reliable, valid conclusions about such data. Decision making is key to educational statistics. According to Queensoap and Memory (2014), statistics is the science of collection, presentation, analysis and interpretation of numerical data which are used in making decisions about a population of study. This definition of statistics implies that data should be collected which primarily requires a tool or method. There are various methods of data collection such as questionnaires, interviews, observations, documentaries/records, etc. However, each method is determined by the type of data to be collected, which requires an appropriate measurement scale. This chapter discusses measurement scales, their types, properties, and applications in educational research.

Measurement of Scale

Quantification has been defined as a numerical method of describing observations of materials or characteristics. Adediwura et al. (n.d.) opined that when a defined portion of the material or characteristics or traite is used as a standard for measuring any sample, a valid and precise method of data description is provided.

Measurement is a fundamental step in the conduct of research. Adediwura et al. (n.d) defined measurement as the process through which observations are translated into numbers. It is the assigment of numerals to objects or events according to certain rules, Agbnifoh and Yomere (1999) posited that the numbers so assigned can form the basis for mathematical manipulations which then provide the basis for making inferences and generalisation about the variables under investigation. The definition indicates that numbers assigned should follow a rule which must be applied by the researcher. This also indicates that the rule of asignment is peculiar to the study and the researcher's discretion. It means therefore that the rule applied in one research should not be made a general rule for another. Nevertheless, whenever a rule of assignment of number is established, that rule must be used throughout the research work. Measurement is seen as the process of classifying or assigning numbers to observation with respect to some criterion or property. It is a fundamental aspect in educational research, especially quantitative research.

Measurement scale, also known as scaling, refers to the measurement of a variable in such a way that it can be expressed on a continuum (Queensoap, 2006). The scale can be expressed as a 5, 4, or 3-point continuum. A measurement scale is an established set of rules for assigning numbers or scores to variables, to indicate the quality or quantity of such variables (Adekunle et al., 2016). In educational research, there are four basic ways of assigning numbers to variables which include nominal scale, ordinal scale, interval scale and ratio scale.

Nominal Scale

This is the simplest measurement scale. It refers to classification of objects into categories. According to Adekunle et al. (2016), nominal scale is the simplest of all scales and it has only the property of identity. It is used to differentiate one variable from another. It implies that numbers may be assigned to the classification, but the numbers may not have numerical value. They will stand as codes or symbols. For example, the Nigerian postal codes for 10 states include:

Table 1: Nigeria Postal Code List for 10 States

S/n	Name of States	Postal Codes
1	Abia	440001
2	Adamawa	640001
3	Akwa Ibom	520001
4	Anambra	420001
5	Bauchi	740001
6	Bayelsa	561001
7	Benue	970001
8	Borno	600001
9	Cross River	540001
10	Delta	320001

Source: Nigeria Postal Code List for 36 States (Isoa, 2020).

Table 1 indicates different states assigned numbers as postal codes, however these postal codes have no mathematical value. It denotes that Bayelsa that has 561001 does not necessarily depict that it is more populated or bigger in size than Delta that has 320001. The codes are for identification and classification purposes in order to ease postal services.Other examples of nominal scales of measurement include blood groups (A, B, AB, or O), marital status (single, married, divorce/separated or widow), classification of persons into male and female, into rural and urban categories, into different ethnicities, etc. This scale, however, can only classify to show differences in variable but has no ability to show the direction of difference. In other words, it does not indicate which variable is greater or less than the other.

In educational research, nominal scale is mostly used to measure the demographic information from respondents. This is because they are used to name, identify and classify. According to Sidhu (2007), the only obtainable mathematical operations in data obtained with nominal scale are counting and statistical techniques such as mode, percentage, charts and non-parametric tests. All these arithmetical operations has no weight attached. Furthermore, Sidhu (2007) asserted that in nominal scale, a given class is sub-divided into a set of mutually exclusive sub-classes. This implies that members of the sub-group are equal in respect to quality or quantity of the variable being measured. It also possess the characteristics of flexibility where symbols or codes designating different groups may be interchanged without altering the essential information provided by the scale. For example, in a class divided into male and female with codes such as 001 and 002 respectively may be interchanged as male 002 while female may take 001, yet the information of classifying the calss into gender categories will not be affected. This is obvious in a football team where one player may take a number for one season and the next season will bear a different number on the jersey. Nevetheless, Kpolovie (2011) argued that nominal scale is not really a scale at all since it does not scale objects or psychological constructs along any dimension or continuum. The scale can simply place individual or object into qualitative recognition and distinction.

Ordinal Scale

This is also called rank order. Onunkwo (2003) states that an ordinal scale is next to the nominal scale in complexity. This is because nominal scale can only show that things are different while the ordinal can show the direction of differences. It is a scale that has an order of magnitude such as greater than and equal to. Examples of ordinal scales include grading in examinations (A, B, C, D, E, and F), comparative adjectives (tall, taller, tallest), and academic staff rankings (Instructor, Graduate Assistant, Assistant Lecturer, Lecturer II, Lecturer I, Senior

Lecturer, Reader, and Professor). However, this scale is limited by its inability to identify the difference between rankings. It means therefore that this scale has the properties to classify and indicate the direction of difference between comparing variables but does not have the ability to show equal intervals.

Sidhu (2007) pointed out that the fundamental difference between a nominal scale and an ordinal scale is that the nominal scale incorporates the relation of equivalence only while ordinal scale incorporates the relation of equivalent as well as of 'greater than'. For example,the use of 1st , 2nd , 3rd , 4th, 5th, 6th, 7th, and so on in reporting students performances in a class indicates that the student who took 1st position is topping the class while the 2nd follows until the last however these rankings do not show how much less or more is the 1st position and 2nd position as well as 3rd position and 4th position. The ranks or successive intervals, that is, distance between classes, on the scales are not equal. This is the reason why data obtained through an ordinal scale cannot be subjected to arithemetic permutations but like nominal scale can use percentage, chi-square, etc. as statistical techniques in educational research.

The prominent characteristics of ordinal scales can include the ability to order object or variables form most to least with respect to trait, there is no indication of "how much" in an absolute sense any of the objects possesses the trait, and there is no indication of how far apart of the objects are with the attribute (Dibu-Ojerinde, 2012). Ordinal scale assigns numerals to subjects in such a coherent manner that values correspond to relative position of the subjects in terms of measured attribute or trait.

Ordinal scale is useful when it is possible to express a characteristics quantitatively but impossible to define units of measurement which are equal at all points on the scale. Most of the measurements in education can be done at the level of ordinal scales, unless assumptions about the uniformity of unit is to be made (Sidhu, 2007).

In ordinal measurement, the empirical procedure used for ranking elements or objects must satisfy the criterion of transivity postulate. This postulate states that the relationship must be such that such that ;if object **a** is greater than object **b**, and **b** is greater than object **c**, then object **a** is greater than object **c**(Adediwura et al., n.d). This can be expressed as if (a > b) and (b > c), then (a > c).

Interval Scale

This scale is superior to nominal and ordinal scale since it possesses all the properties of the both scales as well as incorporating equal interval and to show an order of magnitude among comparing variables. Kpolovie (2011) stated that interval scale applies equal units of measurement to correctly depict quantity or magnitude of the variable under investigation. It relates that a score of 20 is halfway between 10 and 30, in other words, the difference between any two points, say 40 and 60. This scale satisfies the property of equal intervals but lacks the characteristics of absolute zero (0). This simply means that zero (0) is not empty. For instance, a student who scores 0 in a test does not mean that student had zero knowledge. Again, in the grading system of students performance such as 0-39 (F), 40-44 (E), 45-49 (D), 50-59 (C), 60-69 (B) and 70-100 (A), implies that a student who scored zero and another student who scored 39 are in the same interval, this means that zero can be interpreted as 39. Another example is a temperature of 35°F, which is equivalent to 0°C in measuring temperature while 273 kelvin is equal to 0 degree centegrade. That is zero is arbitrary.

Data generated in this scale can allow fractions and decimal points hence can be described as continuous data for this purpose and arithmetical computation like addition, subtraction, division and multiplication can be applied to data obtained at this level of measurement. In educational research, data collected at interval scale can be subjected to statistical tools such as mean, standard deviation, variance and parametric statistics. In a nutshell, interval scale is a scale that has the ability to classify variables, indicate differences between variables, possess order of magnitude and is limited by lack of absolute Zero.

It is also possible to transform an ordinal scale into an interval scale as in the case of a continuum scale like likert scale with response format as (5) Strongly agree, (4) Agree, (3) Undecided, (2) Disagree, (1) Strongly disagree; or (4) Always, (3) Sometimes, (2) Rarely, (1) Never. Since these response levels are regarde as continuum where it is possible to choose any point then we can regard it as interval scale.

Ratio Scale

Ratio scale, the most complex level of scaling, has a defined or absolute zero point which can be interpreted as nothingness. (Kpolovie, 2011). In this scale, all characteristics of interval scale are found, and in addition has the true zero-point. It denotes that any two scale points, unlike the interval scale, are independent of the unit of measurement. Each number that is assigned can be observed as a distance measured from zero. Examples of ratio scale include blood pressure, a baby's weight at birth, height, length, etc., if a person's has weight of 100 kg it can be said with a degree of accuracy that they weigh twice as much as 50 kg. The concept of zero is a definable concept.

Similarly, Sidhu (2007) stated that an individual who has an IQ of 120 is not twice as intelligent as one who has an IQ of 60. The concept of zero intelligence is a meaningless concept just as is the concept of zero achievement as measured by a particular test. Furthermore, Sidhu (2007) identified three advantages of ratio scale which are:

- It has a true zero point, hence can present a variable as meaningless or nothingness.

- It is the most precise type of measurement among all measurement scales.
- Arithmetical operations are permissible on the numerical values assigned to the variable as well as on the intervals between numerials.

Ratio scale is mainly used in physical measurement due to its advantage over the other scales.

Applications of Measurement of Scale in Educational Research

Educational research is a scientific research that involves empiricism; it is empirical in nature. It involves the collection of data that can be used to draw conclusions. Such conclusions are not based on feelings of the author or what the author might think of rather it depend on concrete evidence derived from the data obtained by careful observation of the events being investigated. For this reason, data collection requires appropriate methods, including questionnaires, interviews, and observations. Nevertheless, the nature of data required determines the measurement scale needed. For instance, if the study needs nominal data, it means that a nominal scale must be used to measure the data. Hence measurement scales are vital in educational research.

Another application of measurement scale in educational research is the influence on the choice of statistical tools to be used in analyzing the data obtained for the study or investigation. For example, nominal data analysis requires descriptive statistics such as percentages, charts, modes, etc., while inferential statistics used are non-parametric such as chi-square, binomial test, etc. For ordinal scale which is used for ranking and stating preference requires the median as a typical descriptive statistics while non-parametric statistics such as Mann Whitney U, Friedman two-way ANOVA, Spearman rho correlation, etc. For higher measurement scales, such as the interval and ratio scales, descriptive statistics include mean, standard deviation, and variance. Inferential statistical techniques for these scales include parametric tests such as the t-test, ANOVA, and Pearson's product-moment correlation.

Measurement scales in educational research play a vital role in translating higher scale of measurement to lower scale measurement and vice versa. For example, a scale to identify age of a class 15 years ( ) 16 years ( ), 17 years ( ), etc. represent a ratio scale while a scale such as under 15 ( ); 15-18 ( ), over 18 ( ) can be described as ordinal scale. Similarly, we can measure level of agreement on Nigerian Protest eitheragree or disagree. Those agree or disagree with the protest can be classified as nominal scale but considering it as strongly agree, agree, undicided, disagree and strongly disagree turns the scale to be ordinal.

Summary

Table 2 shows the utility of the levels of measurement of scale in educational research.

S/No	Scale	Key Characteristics	Statistical Methods
1	Nominal	Classify variables	Mode, frequency distribution, simple percentage, and non-parametric tools
2	Ordinal	Classify and show direction of difference but no order of magnitude	Median, percentile, rank-correlation

3	Interval	Classify, show direction of difference, has order of magnitude but lack absolute zero	Mean, standard deviation, variance, parametric tools
4	Ratio	Classify, show direction of difference, has order of magnitude and absolute zero.	Geometric mean, coefficient of variation, parametric tools.

Source: Adapted from Sidhu (2007)

Revision Questions

Question 1
(a) Define measurement scales
(b) State the four scales of measurement in the order of complexity

Question 2
Academic titles (Instructor, Graduate Assistant, Lecturer II, Lecturer I, Senior Lecturer, Reader, and Professor), follow which type of measurement scale? Justify your answer.

Question 3
Describe nominal, ordinal, interval and ratio scales of measurement with two examples in each case.

Question 4
Compare and contrast the four scales of measurement.

Question 5
Propose a researchable topic and select a suitable scale for the collection of data.

Refrences

Adediwura, Y., Olude, A. O., Makinde, S. A., Ismail, Y., & Ojerinde, A. A. (n.d). Statistical methods in education. In T. K. Adeyanju, H. Tuku, & M. Oyedeji, *Postgraduate diploma in education* (pp. 41-55). National Teachers' Institute.

Adekunle, Y. A., Adio, A. K., Efuwape, B. T., Maitanmi, S. O., Toyinbo, R. O., Ebiesuwa, S., . . . Ayanlowo, E. (2016). *Basic university statistics (2nd ed.).* Simplex Publishing .

Agbnifoh, B. A., & Yomere, G. O. (1999). *Research methodology in the management and social sciences* . Uniben Press.

Dibu-Ojerinde, O. O. (2012). Basic concepts: Testing, measurement, assessment and evaluation. In E. R. Afolabi, & O. (. Dibu-Ojerinde, *Educational test measurement* (pp. 1-20). Obafemi Awolowo University Press.

Isoa, S. (2020, December 18). *Nigeria postal code list for 36 states*. Retrieved from simblogshare.com: https://www.simblogshare.com

Kpolovie, P. J. (2011). *Statistical techniques for advanced research.* Springfield Publishers.

Onunkwo, G. I. (2003). *Fundamentals of evaluation.* Cope.

Queensoap, M. (2006). *Construction and standardization of corruption sensitivity test (Unpublished master's thesis).* University of Port Harcourt, Rivers State, Nigeria.

Queensoap, M., & Memory, D. (2014). *Basic biostatistics for public health and allied medical science students.* Pearl Publishers.

Sidhu, K. S. (2007). *New approaches to measurement and evaluation.* Ssterling Publishers.

Chapter 4

TESTING HYPOTHESES

Shikha Trivedi (PhD)
Department of Early Childhood, Botswana Open University
E-Mail: strivedi@staff.bou.ac.bw

Introduction

Testing a hypothesis is a fundamental step in scientific research. It involves evaluating a proposed explanation or theory about a phenomenon supported by empirical evidence. Before discussing the testing of a hypothesis, let us first look at the concept of hypothesis.

"A hypothesis is a statement that specifies an expected relationship between variables. It is a prediction that the researcher makes about what the researcher expects to find." (Creswell & Creswell, 2018) In Kumar's words, "A hypothesis is a tentative statement that proposes a possible explanation to some phenomenon or event. A useful hypothesis is a testable statement which may include a prediction." (Kumar, 2019). A hypothesis is a tentative answer to the research question or an expected relationship between specific and testable variables. It is based on previous knowledge of traceable facts. The hypothesis is always presented in declarative sentence form and should be directional e.g. 'Teachers' strategies are related to the academic performance of primary school children'.

Types of Hypotheses

Hypotheses are classified into various types based on their nature, purpose, and formulation. The two primary types of hypotheses are the null hypothesis and the alternative hypothesis. While this broad classification might be correct, it would be better to distinguish the most common forms mentioned in the context entirely. Below are some of the main types of hypotheses commonly used in research:

1. Null Hypothesis (H_0): A null hypothesis is a statement asserting no significant difference between variables in a particular situation. It serves as the baseline assumption that researchers aim to test against. It is denoted by H_0.

Example 1

Suppose you wish to test whether two samples from a population are similar. In that case, the null hypothesis can be stated as

H_0: There is no significant difference between the two samples.

Example 2

In a research that seeks to test whether parents' attendance in Parent Teacher Association meetings influences children's academic performance. The null hypothesis will be stated as:

H_0: Parental involvement in the Parent Teacher Association does not affect the children's academic performance.

2. Alternative Hypothesis (H_1 or H_A): The alternative hypothesis is the statement opposite the null hypothesis. It predicts the study outcome and proposes a significant effect or difference, which the researcher aims to support based on prior literature and studies suggesting a potential outcome.

The alternative hypothesis can either be directional or non-directional .

a) Directional hypothesis: A hypothesis that directs the statement such as high, low, positive, negative etc. It is also known as a one-tailed hypothesis. For example:

Example 1

H_1 or H_A: Drug A patients recover faster than those who take Drug B.

In this case, the researcher predicts that Drug A will assist patients in recovering faster than Drug B.

Example 2

H_1 or H_A Boys score higher in Mathematics than Girls in Secondary school.

This means that researchers predict that boys are better at mathematics than girls in secondary schools.

3. Non-directional hypothesis: A non-directional hypothesis only claims an effect on the dependent variable. It does not specify the direction of the predicted relationship between variables, which means it does not clarify whether the result would be positive or negative. It is also known as a two-tailed hypothesis. The researcher is open to the possibility of any outcome.

Example 1

H_1 or H_A: There is a relationship between attending preschool and academic performance.

In this case, the researcher predicts the relationship between the two variables but does not specify whether it is positive or negative.

Example 2

H_1 or H_A: There is a relationship between gender and technology usage.

In this case, the researcher predicts that there is some form of relationship between age and technology usage. However, it does not specify whether males are more likely to use technology or techno-phobic or whether females are more likely to use it.

Using a non-directional hypothesis, researchers can explore the potential relationship between gender and technology usage without assuming a specific pattern. A researcher uses a non-directional hypothesis because it allows for flexibility, and preconceived notions do not constrain the research. This can lead to more objective and unbiased findings. Also, using a non-directional hypothesis may be more appropriate when limited research or evidence supports conflicting results.

3. Complex Hypothesis: A complex hypothesis states a relationship between multiple independent and/or dependent variables.

Example 1
"People with high-sugar diets and less physical activity levels are more likely to develop diabetes."

Example 2
"A higher level of physical activity and a balanced diet together lead to a healthy heart."
You will notice two independent variables and one dependent variable in both examples of complex hypotheses.

4. Causal Hypothesis: A causal hypothesis is a predictive statement that indicates a potential cause-and-effect relationship between two or more variables. It suggests that one variable (the independent or cause variable) directly affects or causes a change in another variable (the dependent or effect variable).

Example1
"Consuming high-sugar diets causes an increase in blood glucose levels."
You will notice that the independent variable (high-sugar diets) is affecting the dependent variable (blood glucose levels).

As hypotheses are being stated, the following factors should be put into consideration**:**

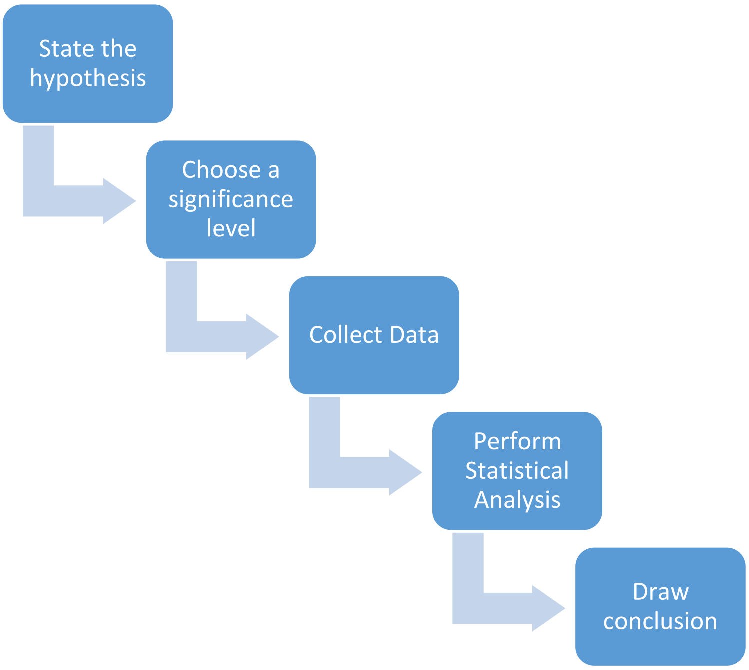
1. The focus of the hypothesis can be tested.
2. The stated hypothesis has both independent and dependent variables.
3. The variables can be manipulated.
4. Can the stated hypothesis be tested without violating ethical standards?

Hypothesis Testing

Hypothesis testing is a statistical procedure in which a choice is made between a null hypothesis and an alternative hypothesis based on information in a sample. Hypothesis testing is used to assess the plausibility of a hypothesis by using sample data. The end result of a hypotheses testing procedure is a choice of one of the following two possible conclusions:

1. Reject H_0 (and therefore accept H_a), or
2. Fail to reject H_0 (and therefore fail to accept H_a)

Here are the steps for testing the hypothesis.



1. **State the hypothesis**
2. **Choose a Significance Level (α):** The significance levels commonly used are 0.05, 0.01, or 0.10. The significance level represents the probability of making a Type I error (rejecting the null hypothesis when it is true).
3. **Collect Data:** Gather the required data to test the hypothesis. You can do this through surveys, experiments, etc.
4. **Perform Statistical Analysis:** Based on the data and research design, use statistical tests (e.g., t-tests, chi-square tests, ANOVA, etc.) to analyse the data and determine whether sufficient evidence exists to assume or not the stated hypothesis.
5. **Draw Conclusions:** Based on the statistical analysis, decide whether to assume or fail to assume the hypothesis. Failure to accept the null hypothesis supports

the alternative hypothesis, suggesting that the observed effect or difference is statistically significant.

In summary, testing a hypothesis is a structured process that allows researchers to make informed conclusions based on empirical evidence, thereby advancing knowledge in various fields.

Importance of Hypothesis Testing

1. Hypothesis testing provides a clear direction for research, focusing on specific questions.
2. It helps assess the validity of a theory or assumption, ensuring that data support conclusions.
3. Hypothesis testing provides a framework for decision-making based on data rather than personal opinions.

Limitations of Hypothesis Testing

1. Hypothesis testing relies exclusively on data and doesn't comprehensively understand the subject being studied.
2. The accuracy of the results depends on the quality of the available data and the statistical methods used. If data is inaccurate and the stated hypothesis formulation is not as per the study design and data, it may lead to incorrect conclusions or failed tests.

Concluding on a Hypothesis Test

Concluding the hypothesis involves interpreting the statistical analysis results and determining whether there is sufficient evidence to support or reject the hypothesis. In conclusion, summarising your findings clearly and accurately is essential for communicating the results of your hypothesis test and their implications for the broader research question.

For concluding the test, it would help if you looked at the p-value to the significance Level (α). The p-value is the probability of observing the test results under the null hypothesis. It helps determine the significance of your results.

- **If p-value ≤ α:** Reject the null hypothesis (H_0). This indicates that the observed data is unlikely under the null hypothesis, and there is evidence to support the alternative hypothesis (H_1). For example, if you were testing whether a new drug is effective, rejecting the null hypothesis would suggest that the drug has a significant effect.
- **If p-value > α:** Fail to reject the null hypothesis (H_0). This means insufficient evidence supports the alternative hypothesis(H_1). This does not prove that the null hypothesis is true, only that the data did not provide strong enough evidence to reject it.

Reporting the results of a hypothesis test

Include the following when reporting the results.

- **Test Type:** Specify the type of test used (e.g., t-test, chi-square test).
- **p-value:** Provide the p-value to indicate the level of statistical significance.
- **Decision:** State whether you rejected or failed to reject the null hypothesis.
- **Interpretation:** Provide an interpretation of what the results mean in the context of the research question.

Example Conclusion

Suppose you surveyed whether parental involvement in children's school activities significantly influences their academic performance. Your hypotheses were:

- **Null Hypothesis (H_0):** Parental involvement in children's school activities does not influence their academic performance.

- **Alternative Hypothesis (H_1):**Parental involvement in children's school activities significantly influences their academic performance.

After performing the statistical test, you obtained a p-value of 0.03 with a significance level (α) 0.05. You reject the null hypothesis since the p-value (0.03) is less than the significance level (0.05).

Conclusion Statement

Based on the statistical analysis, we reject the null hypothesis at the 0.05 significance level. There is sufficient evidence to conclude that Parental involvement in children's school activities significantly influences their academic performance. Therefore, we support the alternative hypothesis that Parental involvement in children's school activities significantly influences their academic performance.

Summary

- All types of hypotheses provide a framework for researchers to structure their inquiries and make predictions that can be tested and validated through research.
- Each hypothesis type serves a specific purpose depending on the nature of the study and the questions being addressed.
- The alternate hypothesis can either be directional or non-directional.
- The main difference between these two types of hypotheses is the level of specificity. Directional hypotheses specify the direction of the relationship between variables, whereas non-directional hypotheses leave the prediction open-ended.

- The five steps of hypothesis testing include stating the hypotheses, formulating an analysis plan, collecting data, analysing the sample data, and confirming the result.
- To conclude the result of hypothesis testing, it is essential to have sufficient evidence to reject or accept the hypothesis.

Questions

1. What is the difference between a one-tailed and a two-tailed test?
2. Explain a p-value?
3. How do you decide whether to reject or fail to reject the null hypothesis?
4. Describe the steps of testing the hypothesis.
5. What should you conclude if you fail to reject the null hypothesis?

References

Creswell, J. W., & Creswell, J. D. (2018). Research Design: Qualitative, Quantitative, and Mixed Methods Approaches (5th ed.). SAGE Publications.

Fay, M.P. & Brittain, E.H. (2022). Statistical Hypothesis Testing in Context: Reproducibility, Inference, and Science.Cambridge University Press.

https://2012books.lardbucket.org/books/beginning-statistics/s02 acknowledgements.html

https://resources.nu.edu/statsresources/hypothesis

https://workplacehero.co.uk/blog/understanding-the-fundamentals-of-a-non-directional-hypothesis#google_vignette

https://www.amazon.com/Everything-Psychology-Book-2Nd/dp/1440506914

https://www.examples.com/english/causal-hypothesis.html

Ning-Zhong Shi and Jian Tao (2008). Statistical Hypothesis Testing: Theory and Methods. https://doi.org/10.1142/6846

Chapter 5

PARAMETER ESTIMATION

Memory Queensoap, *PhD*

Introduction

Educational research is a systematic approach to finding a solution to an educational problem. It is systematic and scientific in making either a qualitative or quantitative inference about the population of the study. This is dependent on the data collected. Once the data are collected, either qualitative or quantitative, they are subjected to content and statistical analysis to draw valid inferences about the population using the sample information. Statistical inference is the process of generalizing from the sample to the population (Rangaswamy, 2007). Statistical inference can take, generally, in two forms namely estimation and hypothesis testing. In both processes, the sampling distribution is used.

This chapter considers parameter estimation which involves estimating population parameters from sample values. Estimation simply denotes the process of obtaining an estimate of the unknown value of a parameter by using a statistical tool. According to Adekunle, et al., (2016), the process of estimation entails calculating from sample observation some statistics that are offered as an approximation of the corresponding parameter of the population from which the sample is drawn. A population parameter is a measure that describes some characteristics of the population, for example, mean (μ), standard deviation (σ), variance (σ^2), etc. A distinction is commonly made between two types of estimates, point estimates and interval estimates.

Definition of Basic Terms in Parameter Estimation

There are some basic terms used in the computation of parameter estimation which include sample, population, estimate, confidence limits, and confidence interval. For purpose of clarity, we can define them as they are applied in parameter estimation.

A sample is a subset of measurements of interest selected from a population using a statistical method. It means therefore that sample is part of the population that is selected through a definite procedure from the defined population. Sampling is the process in which the sample is being selected from the population. In parameter estimation, sample mean (X, bar X) is the average mathematical value of a sample's values. It is a statistical indicator used to analyse various variables over time (Ahmed, 2024). Furthermore, the mean value of a sample of data calculated from within a large population of data is called sample mean. It is a good statistical tool to assess the population mean if sample size is large and the researchers randomly take fragments from the population, remarked Ahmed (2024).

Population is set of measurements of interest to a researcher. Population can be either infinite or finite. An Infinite population has indefinite number of elements which become impossible to determine or count while the finite population has a definite number of measurements of interest that are countable. A population parameter is a number that describes characteristics about an entire group or population. It is denoted by θ (read 'theta') and the corresponding statistic is denoted by θ^ (theta cap). Meanwhile the **population mean** is an average of group statistics denoted by μ. Population mean differs from sampling mean in the following ways:

1. Sample mean is the average of all the values of sample collected from a large population to estimate the population, while population mean is the mean of complete values contained in a population.

2. Sample mean is easy to calculate whereas population mean is not.
3. Population mean is more accurate than sample mean.

Estimate simply refers to the numerical value of an estimator calculated from the sample drawn from a population.

Confidence limits are the upper and lower bounds on a confidence interval for the estimation of a population parameter. The limits are determined in accordance with the conventional level of confidence that these limits will include the population parameter being estimated, for example, 95 percent confidence that the population mean falls within limits 20.05 and 25.34 (Prabhakara, 2006).

Confidence Interval refers to the interval within which a population parameter lies with a stated degree of confidence.

Point Estimation: An estimator of a population parameter is a sample statistic used to estimate the parameter. An estimate of the parameter is a particular numerical value of the estimator obtained by sampling. When a single value is used as an estimate, the estimate is called a point estimate of the population parameter (Aczel & Sounderpandian, 2006). A point estimate of a parameter θ, denoted by θˆ, is a single number that can be considered as a possible value for θ. Since it is computed from the sample $X = (X_1, . . . , Xn)$, it is a function of X, that is, $θˆ = θˆ(X)$. This process of computing a point estimator is known as Point Estimation.

Properties of Estimators

It is quite clear that sample statistics are used to estimate the population parameters, there is need to select the best estimator. According to Rangaswamy (2007), the best estimator is one that possesses the properties of unbiasedness, consistency, efficiency, and sufficiency.

Unbiasedness

The relationship between the true value of the parameter and the mean of all possible values of an estimator is described as unbiasedness. A method of estimation provides an unbiased estimate when the mean of many sample values, obtained by repeated sampling, approaches the population value in the limit as the number increases. This simply denotes that a statistic is unbiased when it displays no systematic tendency to be either greater than or less than the parameter. Symbolically it means that an estimator $\theta\hat{}$ (theta cap) of a parameter θ is said to be unbiased if, $E(\theta\hat{}) = \theta$ or $E(\theta\hat{}) - \theta = 0$. The bias, if any, is given by $E(\theta\hat{}) = 0$. If this value is negative, it is said to be negatively biased but if it is positive the estimator is flagged to be positively biased.

In another perspective, an estimate is unbiased when its expected value is equal to the parameter it purports to estimate. The expected value of a variable is obtained by multiplying each value of the variable by its associated probability and summing the products. For example, we can consider test scores for 100 students and their corresponding probabilities as indicated in Table 2.

Table 2: Frequency Distribution Table Showing Students' Test Scores with corresponding Probabilities

Test Scores (X)	*Frequency (F)*	*Probabilities (P)*
6	25	.25
7	40	.40
8	20	.20
9	15	.15
Total	100	1.00

Table 2 showed the expected value X, that is, E (X) as 7.25 which is the same with the mean of X in the distribution. Here, E (X) is the sum of $X_1(P_1)....X_n (P_n)$ which is obtained by6 x .25 + 7 x .40 + 8 x .20 + 9 x .15 = 7.25. So, the sample mean x is an unbiased estimator of the population mean μ, that is, E(x) = μ. The sample variance

$$S^2 = \Sigma (x - x)^2 / n - 1$$

is an unbiased estimator of the population variance σ^2, that is, $(s^2) = \sigma^2$. The sample variance given by

$$s^2 = \Sigma (x - x)^2 / n$$

is a biased estimator of σ^2. For this reason, s^2 is always used as an estimator of σ^2. This simply means that a statistic is a biased estimate when its expected value does not tend toward the population value but departs in a systematic fashion from it.

Consistency

Consistency is another property a best estimator should possess. This property refers to a point where the estimate approaches the population parameter more closely as the sample size increases. This relates that the property of consistency is concerned with the variability of the estimator. Simply put, an estimator may be unbiased or consistent or both, still it may not be the best estimator. An estimator will be best, if, in addition to unbiasedness and consistency, it is an efficient estimator.

Efficiency

The sampling variance of an estimator relates to the efficiency of the estimate. This denotes that an estimator with a smaller variance will deviate less from the true value of the parameter than the estimator with a larger variance. It implies that the smaller the standard error, the greater the precision of an estimator. In a normal distribution, the mean and median are estimates of the same population parameter, μ.

Sufficiency

Some scholars have observed that sufficiency as a method of estimation is somewhat intuitive. It is stated that a method of estimation is sufficient if it is more efficient than any other possible method of estimation, that is, if its sampling variance is less. It implies that all the information in the sample is made in the method of estimation. Rangaswamy (2007) presented that there is a problem in identifying the best estimator in point estimation however highlighted that there are some criteria used in identifying best estimators. These include method of moments, method of least squares, method of minimum χ^2 and method of maximum likelihood. It has been observed that method of maximum likelihood is most important.

Interval Estimation

Point estimation is not an absolute method of estimation because it can estimate a single value and such single value may not be exactly the estimate of the parameter. The difference between these two values is small with large samples but the difference may be observable with small samples. Point estimation does not take the size of the sample into consideration which serves as a limitation to its statistical value. To fill this gap, interval estimation is preferred to point estimation.

Interval estimation involves the determination of an interval within which the population value must lie with a specified degree of confidence. An interval estimator, otherwise called confidence interval is a formula that tells us how to use sample data to calculate an interval that estimates a population parameter (McClave & Sincich, 2000). In another vein, a confidence interval can mean a range of values, w_a to w_b, believed to include an unknown population parameter θ. Associated with the interval is a measure of the confidence we have that the interval does indeed contain the parameter of interest (Aczel & Sounderpandian, 2006). For example, in a research report that presents 95% confidence, intervals are values between 95% confident that the population parameter θ being estimated will lie. Therefore, carrying out a significant test and reject null hypothesis that a mean is really the same as an expected mean, it becomes necessary to estimate a confidence interval for the true value.

Confidence Interval for Means

We have discussed that the mean or difference between means is a sample statistic for point estimation. In this section we need to establish the interval likely to include the population parameter and the degree of confidence associated with the interval. If we are to draw a random sample of size *n* from a population whose parameter θ is to be estimated. From the sample, the statistics θ^ (theta cap) and its standard error SE (θ^) can be calculated. This can help us know the sampling distribution of θ^.

For instance, if the sampling distribution of theta cap (θ^) is a standard normal distribution, we can calculate the numbers θ^ - Z.SE (θ^) and θ^ + Z.SE (θ^). These numbers are called lower limit, L and upper limit, U of an interval. This interval covers the parameter with a specified probability 1 – α. The interval, L to U is known as a 100(1-α) percent confidence interval for the parameter θ. Meanwhile, the specified probability, (1 – α) is called level of confidence or confidence coefficient.

For a large sample size, a 100 (1 – α) percent confidence interval for θ is given by θ ± Z.SE (θ^). In such cases Z denotes the critical value of Z for α percent level of significance. For example, in a case of 95% confidence level, the critical value of Z is 1.96 and the critical value of Z for 99% confidence level is going to be 2.58.

Things to consider on confidence interval estimation

- We can check the true mean difference in confidence interval estimation irrespective of proving null hypothesis.

- We usually construct an interval that contains sample estimate in confidence interval estimation.
- The width of the confidence interval may be adjusted through sample size and confidence level
- If the confidence interval includes zero as a true difference the critical ratio will not be significant (null hypothesis accepted)
- If the confidence interval does not include zero as a true difference the critical ratio will be significant (null hypothesis rejected)
- The confidence interval estimation reflects the sampling variation (Prabhakara, 2006).

Study Problems

Question 1: In a comparative study of lecture method and team-teaching method of teachings on academic performances 200 level university students, the following data were available: n = 50; mean (x) was 52 and 45 for team-teaching group and lecture method group respectively; standard deviation was 2.5 and 1.7 for team – teaching and lecture method groups respectively. i) estimate standard error (SE) of means of groups ii) find out 95% confidence intervals on true means.

Solution

i) SE_x of lecture method group

$= SD/\sqrt{n} = 1.7/\sqrt{50} = 1.7/7.1 = 0.24$

SE_x of team-teaching method group

$= SD/\sqrt{n} = 2.5/\sqrt{50} = 2.5/7.1 = 0.35$

ii) 95% confidence interval for lecture method group:

For $\alpha = 0.05$, the critical value of Z is 1.96. This implies that the 95% confidence intervals for group of lecture method are given as:

$\mu = x \pm Z\,SE(x)$

$= 45 \pm 1.96\ (0.24)$

$= 45 \pm 0.47$

Therefore 95% confidence interval for population mean μ = 44.53 to 45.47

For team-teaching group:

$\mu = x \pm Z\ SE(x)$

$= 52 \pm 1.96\ (0.35)$

$= 52 \pm 0.69$

Therefore, 95% confidence interval for population mean μ = 51.31 to 52.69

Similarly, for α = 0.05 and 49 *d.f.* (degree of freedom), the critical value of t is 2.010. Hence, we can use the formula:

$\mu = x \pm t\ SE(x)$

This implies for lecture method as,

$45 \pm 2.01\ (0.24)$

$= 45 \pm 0.48$

$= 44.52$ to 45.48

While for team-teaching,

$= 52 \pm 2.01\ (0.35)$

$= 52 \pm 0.70$

$= 51.3$ to 52.7

Question 2: Identify the relationship between standard deviation and sample size

Solution

Sample size is the subset of measurements interest chosen from the population of the study through statistical method. It is a fraction of the study population. Meanwhile, standard deviation is the root-mean-squared deviation which is equal to the square root of the variance and values expressed in same scale as observational values. Standard deviation is an inherent property of a population. Hence, sample size does not influence standard deviation. If sample size is big, there will be more deviation around the mean and the numerator of the variance increases as well as the denominator. It is rather the standard error decreases as sample size increases because SE = SD/√n

Summary

This chapter considered parameter estimation. Parameter estimation was discussed in two perspectives namely point estimation and interval estimation. A point estimate of a parameter θ, denoted by θˆ, is a single number that can be considered as a possible value for θ. The properties of point estimate were discussed as unbiasedness, consistency, efficiency and sufficiency. Meanwhile the interval estimation involves the determination of an interval within which the population value must lie with a specified degree of the confidence. The interval estimate fills the gap which point estimate could not do. A few examples were presented.

Revision Questions

1. A random sample of 200 observations has a mean of 40 and a standard deviation of 12. Estimate the 95 and 99 percent confidence limit for the mean
2. Estimate the 95 and 99 confidence limits for p = .75 where N = 169
3. Distinguish point estimation and interval estimation and express it mathematically.
4. How does the standard error of mean is affected by tripling sample size?
5. Discuss the four properties of estimates.

References

Aczel, A. D., & Sounderpandian, J. (2006). *Complete business statistics (6th ed.).* McGraw-Hill.

Adekunle, Y. A., Adio, A. K., Efuwape, B. T., Maitanmi, S. O., Toyinbo, R. O., Ebiesuwa, S., Ayanlowo, E. (2016). *Basic university statistics (2nd ed.).* Simplex Publishing .

Ahmed, K. (2024, January 8). *Sample mean*. Retrieved from WallStreetMojo: http://www.wallstreetmojo.com

McClave, J. T., & Sincich, T. (2000). *Statistics (8th ed.)*. Prentice-Hall, Inc.

Prabhakara, G. N. (2006). *Biostatistics*. Jaypee Brothers Medical Publishers.

Rangaswamy, R. (2007). *A text book of agricultural statistics.* New Age International (P) .

CHAPTER 6

ORGANIZING DATA

Professor Babatunde Adeniyi ADEYEMI
Institute of Education, Faculty of Education
Obafemi Awolowo University, Ile-Ife, Nigeria
adeyemibabs2003@yahoo.com

Introduction

Data is a set of facts, ideas, and figures that provide a partial picture of reality. Whether data are being collected with a certain purpose or collected data are being utilized, questions regarding what information the data are conveying, how the data can be used, and what must be done to include more useful information must constantly be kept in mind. Since most data are available to researchers in a raw format, they must be summarized, organized, and analyzed to usefully derive information from them. Furthermore, each data set needs to be presented in a certain way depending on what it is used for. This is to ensure that the usability of the data is well-defined, structured, and articulated (Saunders, Lewis, & Thornhill, 2019). In further exploration of the findings by Saunders *et al* (2019), the main goal of any data in research is to answer the given questions. It can also be focused on testing hypotheses. This is why the process of coding and interpreting data should be in tune with the main focus of research. For instance, descriptive data are used to explain situations while inferential data shows the relationship between variables.

In line with the above-discussed, planning how the data will be presented is essential before appropriately processing raw data (In, Junyong & Lee, Sangseok, 2017). Planning helps in the coordination and categorization of data. A good researcher is concerned about the coding and decoding of data. When this is done properly, it is easy to understand data settings and their implications. Hence, data organization involves categorizing and classifying data to make it more usable and readable. Data organization is a process of organizing raw data, by classifying them into different categories. It is the way of arranging raw data in an understandable order. Organizing data includes classification, frequency distribution table, picture representation, graphical representation, etc. For example, arranging the marks obtained by students in different subjects is data organization. It is difficult to work on or do any analyses of raw data. There is a need to organize the data presenting it in a proper way (Byju's, 2024). As shown in Figure 1 (below), some of the methods of collecting data that will be discussed as the study unfolds are in-person interviews, phone interviews, focus groups, online surveys, panel surveys, and mail-in-polls.

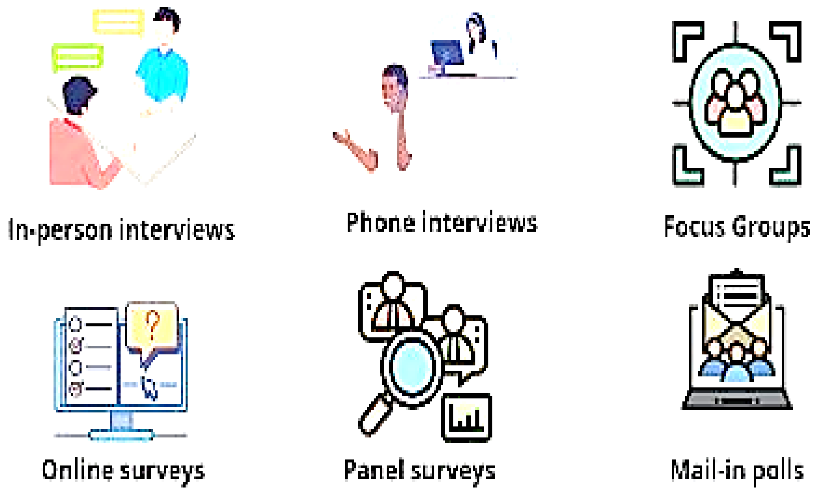


Figure 1: Methods of Data Collection (Adopted from Vedantu, 2024)

According to Vedantu (2024), the organization of data in statistics is a tool or a process that organizes the collected factual materials that are considered necessary in the scientific community to validate research findings. Research data are collected, observed, or created for analysis to produce original research results. Data are intended to showcase facts that will become meaningless if not properly preserved and interpreted. The collection of data and its analysis are so important in research activities to discover answers to research questions and make inferences on hypotheses. In some cases, it even predicts future outcomes. As this chapter unfolds, the focus is on data classification, frequency distribution, and percentiles. In the same vein, empirical studies and practical illustrations will be engaged to drive home points in the study.

Data Classification

Classification of data brings order to raw data. Large datasets can be classified based on the need or purpose. More so, this is concerned with how data are coded and decoded for interpretation. Data are also classified based on their nature, type, and usability. More so, data can be numerical and non-numerical. In other words, data can be qualitative or quantitative. Data can also be a combination of qualitative and quantitative information- this is called mixed data. As this section unfolds, the different types of data, based on which they are organized are listed below:

- **Chronological Data**: Chronological data are grouped or classified according to the time, such as days, weeks, months, and years. In the same vein, chronological data follow chronological order. By implication, listing is allowed in the process of sorting and clarifying the data. The data can also be descending or decreasing in nature. In other words, the data can increase or decrease. For example, the growth of population with time in years. Also, the number of students in a school in different years can be presented based on a period. An example of data that can decrease is the continuous recording of the weight of sportsmen for a couple of days. This shows that chronological data can increase or decrease, depending on how the type of data used by the researcher. Another example is the set of scores that have been gotten by students in a given term. These scores can be recorded based on the performance of learners. The students with the best scores can be recorded and the ones with poor performances are at the later part. This can include scores such as 77, 72, 70, 65, 50, 47, 34, and 22. This is a chronological arrangement, from "top to bottom".

- **Spatial Data**: This is any type of data that directly or indirectly references a specific geographical area or location. Sometimes called geospatial data or geographic information, spatial data can also numerically represent a physical object in a geographic coordinate system. Spatial data are classified based on geographical locations or areas such as cities, states, countries, etc. For instance, spatial data can be classified by cities such as Osogbo, Akure, Lagos, Ibadan, and Abeokuta. These are the states that make up Southwestern Nigeria. Other states can also be mentioned which will make up the Eastern or Northern regions of the country.

  Another example of spatial data is the names of states in a country. Some of the names of states in Nigeria include Osun, Oyo, Enugu, Abia, and Borno, among others. So, depending on the study that is carried out, the names of states, towns, or cities are mentioned and listed based on the data to be analyzed. The objectives that will be tested also contribute to this development. In the same vein, when spatial data are listed and used in a study, the size of the geographical settings can also be mentioned. For instance, the area size of Osun State is 14,875 (square kilometers). The estimated population of the people of Osun State is 4.7 million. These are all indications of parameters by which spatial data can be measured.

- **Qualitative Data**: Qualitative data are categorized under different attributes like nationality, gender, religion, marital status, etc. Such data cannot be measured but can be classified based on their presence and absence of qualitative characteristics. Qualitative data are information that cannot be counted, measured, or easily expressed using numbers. It is collected from text, audio, and images and shared through data visualization tools, such as word clouds, timelines, graph databases, concept maps, and infographics. Qualitative data analysis tries to answer questions about what actions people take and what motivates them to take those actions. One of the main advantages of engaging this type of data in research is the fact that it gives room for full expression by the participants. They will be able to say all that they know about the subject that is being discussed. They can express their views without any form of distraction (Woiceshyn & Daellenbach, 2018). However, the participants involved in this kind of data collection process are very few. This is to ensure that extensive data are collected from them. On the other hand, the process of data collection takes time.

  The researcher and participants must be able to make time available to attend to the process of collecting data. In their study, Woiceshyn and Daellenbach (2018) stressed that when a researcher uses this process of data collection, they must be attentive enough to ask questions in tandem with the study objectives. Some of the research instruments that are used in collecting qualitative data are interviews, open-ended questionnaires, focus group discussions, and narrative experiences among others. For instance, interviews can be in different forms such as in-person interviews and phone interviews. These are based on the process by which the interviews are conducted, that is, online or offline. In the same process, focus groups are engaged in research so that people can talk about their experiences in a discourse collectively. Some forms of contributions and augments are made among the participants towards answering the research questions that have been asked. It must be noted that two types of analyses are always engaged in this type of data collection method which are thematic and content analyses. Thematic analysis has to do with the process of analyzing based on the themes raised. The processes involved in thematic and content analysis are cording, sorting, decoding, interpretation and discussion.

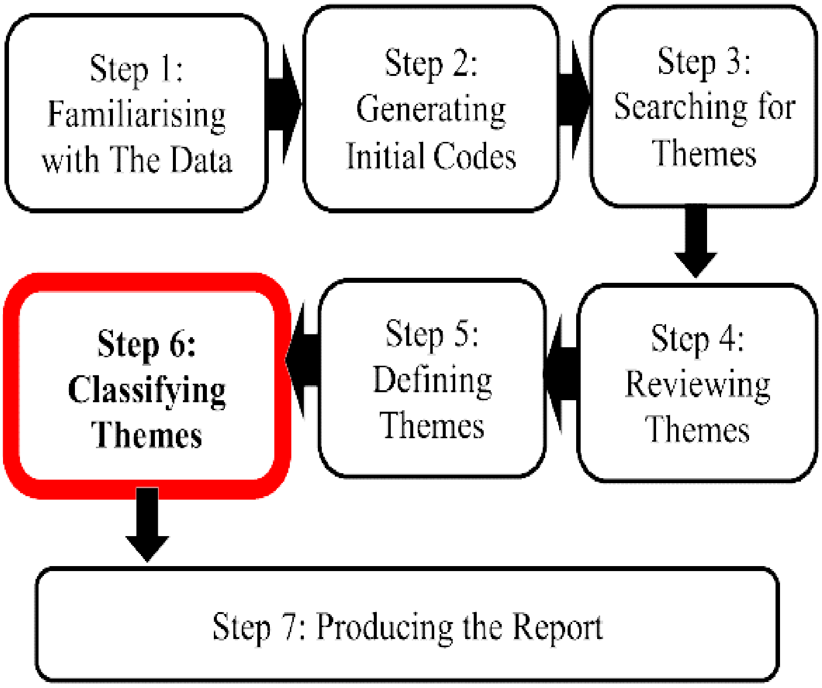


Figure 2: Steps in Thematic analysis (Adopted from Woiceshyn and Daellenbach, 2018)

As shown in Figure 2 (above), the themes engaged in thematic analysis are to ensure that the process of drawing themes from data is seamless. In the last stage, the main aim is to produce reports.

- **Quantitative Data**: Quantitative data can be understood as something that can be counted and measured. It is a fundamental concept that offers insight into the number of required variables. Quantitative data is the type of data when the above attributes (in the case of qualitative classification) are further categorized into number-based data such as height, age, marks of students, salary, etc.

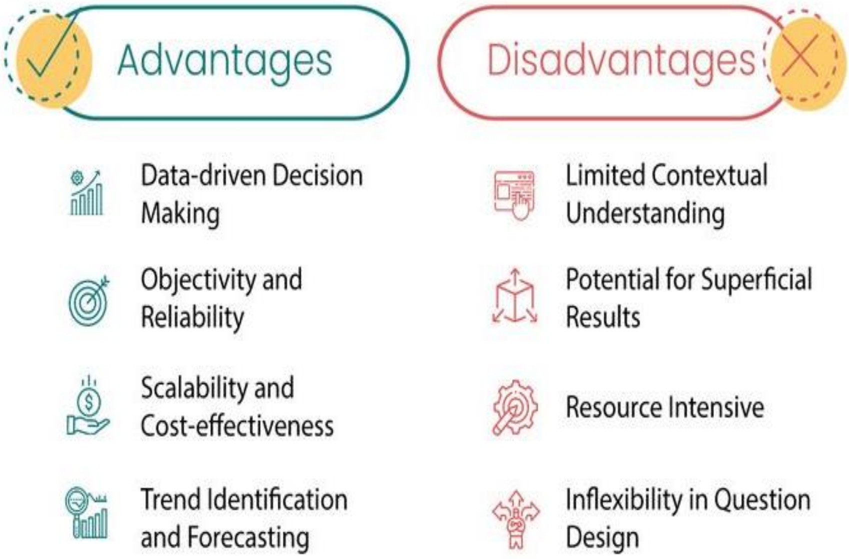


Figure 3: Pros and Cons of Quantitative Data (Adopted from Woiceshyn and Daellenbach, 2018)

Quantitative data has a lot of advantages over qualitative data (Ahmad et al., 2019). Some of these are based on the size of the data to be collected, the ease of analysis, and the ability to duplicate results. When using quantitative data, more than 10,000 respondents can be carried along in data collection. Recently, through the aid of Google Forms, it is more viable to collect data and analyze data in the same process. Using Google Forms, data can be collected from participants across different parts of the world. The main area of concern is ethical consideration. All the participants must be properly informed before they are engaged in data collection processes. This therefore raises the need for ethical concerns in qualitative and quantitative data as discussed below.

Ethical concerns are the moral considerations and dimensions of every research. Ethical consideration involves the process of making sure that there is a proper and thorough collection of data without causing any harm to its subject. The ethical concerns that will be taken care of in this research are informed consent, confidentiality, voluntary consent, and beneficence (BERA, 2018). In research work, the main focus of informed consent is to respect the fundamental human rights of people (Arifin, 2018). That is, participants should not be forced into taking part in the research. The teachers will be allowed to make decisions on their own, whether to take part in the research or not. Confidentiality has to do with the privacy of the participants and their information. To encourage participation in this research, the personal information of the participants should be kept private and used for the research alone (Hall, 2020). In tune with voluntary consent, no form of inducement or payment will be done to enable people to take part in the study as their participation will be done willingly.

After every aspect of ethical issues has been sorted, the collection of quantitative data is done through the aid of closed-ended questionnaires, online surveys, panel surveys, mail-in-polls, and observation checklists among others. With this data are collected and analyzed using descriptive and inferential analyses. For instance, descriptive analysis is done using frequency, percentage, mean scores, and standard deviation.

Since discussions have been done on the qualitative and quantitative data, a comparison is done between the two in Table 1 (below). After this, the place of mixed data in research is also discussed.

Table 1: Comparison of Qualitative and Quantitative Research Methodologies

Qualitative Research	Quantitative Research
It is keen on the collection of non-numerical data.	Numerical data are collected for analysis.
The research philosophy that is largely engaged is interpretivism. This gives room for the analysis of multiple realities and content.	Positivism and pragmatism are principally engaged in this research. The data can be measured and results can be easily duplicated.
Theory building is done through the aid of observations and interpretation of data.	To build and develop theories, measurement of figures and data is done.
Findings are based on a very small population. The sample can only cover a few numbers of people.	It allows for the collection of large data cutting across many people.
Tools for data collection include interviews, observations, and questionaries (open-ended).	Instruments for data collection include surveys, standardized tests, and questionaries (close-ended).
Open-ended tools are used for data collection.	Close-ended tools are used for data collection.
The process of data analysis is mostly through thematic/content analytic tools.	The process of data analysis is largely through thematic/content analytic tools.

(Adapted from Woiceshyn and Daellenbach, 2018; Saunders, Lewis, and Thornhill, 2019; Ahmad *et al.*, 2019)

- **Mixed Data**: As the name implies, this is the process of using both qualitative and quantitative data in a study. In other words, it is using both numerical and non-numerical data in research. This can be the combination of interviews and closed-ended questionnaires. When this is done, some parts of the data are to be answered using descriptive analysis and other, thematic analytical tools. Though the process of collecting the data might take time and involve a lot of money, the results are always very robust. Every aspect of the main research question is also answered. The main consideration of this kind of data is based on the type and nature of the research questions that have been raised. More so, the research design and study population determine if mixed data will be used or not. For instance, when a study adopts mixed methods, mixed data can also be used.

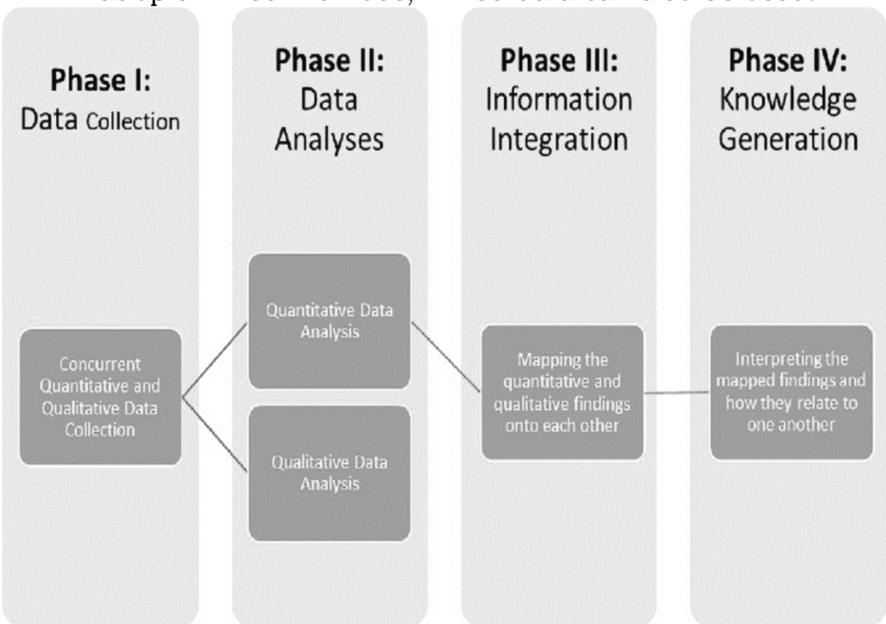


Figure 4: Stages in Mixed Methods

The phases in the mixed data are presented in Figure 4. Here, the collection of quantitative and qualitative data is done concurrently. After this is done, data analysis is carried out. In the third first, mapping the quantitative and qualitative findings onto each other. In the last phase, the interpretation of mapped findings is done.

Frequency Distribution

Frequency distribution is the graphical or tabular representation of data. It shows the number of observations based on the intervals that have been given. Frequency is therefore a value that represents patterns of variables in a given data set. One of the main advantages of using frequency in a study is to explain the number of items that are represented in a given data. For instance, in a garage, the numbers of red, yellow, and blue cars can be counted. Red cars, for example, can be 20, yellow cars can be 12, and blue cars can be 10. The percentages of cars can also be calculated and reported. To calculate a percentage, the formula that is engaged is number/total*100. With this, the frequency that has been retrieved is calculated using percentages. In most cases, the frequency distribution is represented with numbers, scales, and scores. For instance, the scores of learners can be categorized into performance levels. The percentage of learners that had a particular grade can be given to know where they fall. An example is in Table 2 and Table 3 presented below.

Table 2: Ungrouped Frequency Distribution

S/N	Number of Students	C.A Score
1	5	20
2	10	15
3	15	10
4	22	25
5	8	36

Table 3: Grouped Frequency Distribution

S/N	Frequency	Number of Students
1	0-10	01
2	11-20	12
3	21-30	22
4	31-40	05
5	41-50	02

In tune with the Table above, it has been shown that ranking can be in different areas. This is based on the type of data to be scored as well as the intention for the data. Ranking can also include remarks such as low, very low, high, and very high. There are different types of frequency distribution which include cumulative frequency distribution, relative frequency distribution, and common representations of frequency distributions. These are discussed as follows.

- **Cumulative Frequency Distribution**: This is when the addition or subtraction of frequencies is done. It is also carried out based on some particular class intervals. In cumulative frequency distribution, the calculation of frequencies is done for every category. For instance, the age, height, and grades of learners can be done differently. More so, the representation of data can be ascending or descending. When data is ascending, it is increasing. When data is descending, it is decreasing. Examples of the two types of data are presented below.

Table 3: Cumulative Frequency Distribution

S/N	Ascending Data	Descending Data
1	35	89
2	42	77
3	45	65
4	56	56
5	65	45
6	77	42
7	89	35

- **Relative Frequency Distribution**: This is a statistical application that is used every day. It is the distribution of data in different categories and in the way they are. There is no need for any format in the ways by which the data are distributed or represented. Unlike the ascending or descending format of data representation, relative frequency distribution takes any format. Some of the different relative frequencies that we have decimal points, percentages, or fractions.

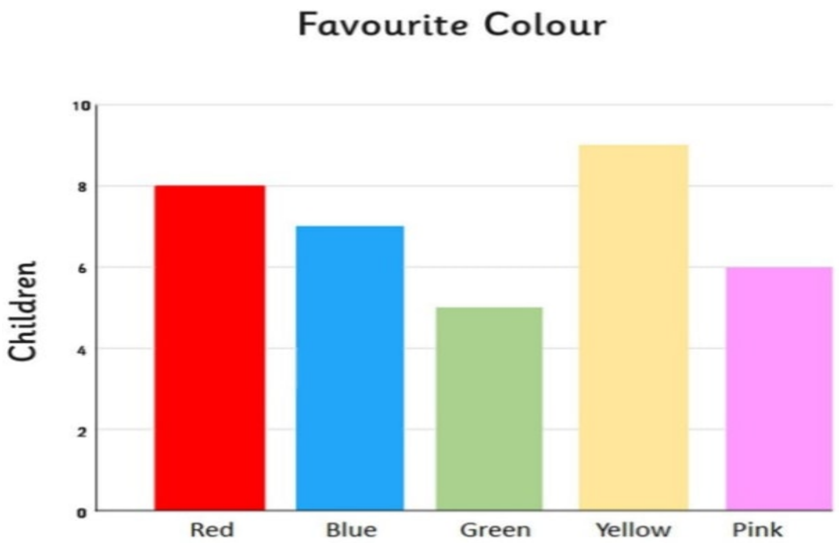


Figure 5: An example of a bar chart showing the favorite colors of children

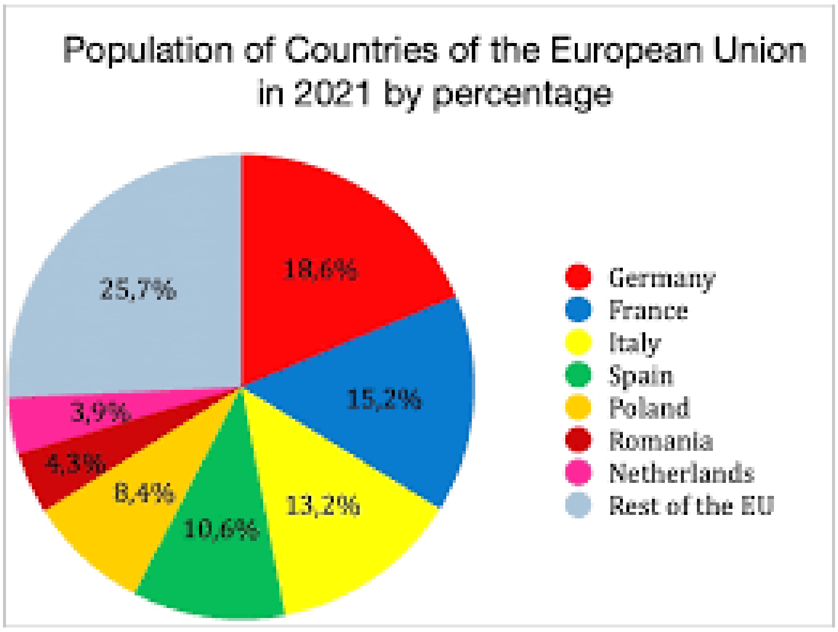


Figure 6: An example of a pie chart showing the population of countries of the European Union by 2020 by percentage

- **Common Representations of Frequency Distributions**: Here, the representation of data is done the way they are. The use of bar charts (see Figure 5) and pie charts (see Figure 6) are used in this area. An advantage of this is to ensure that data distributions can be observed at a glance.

Percentile

This is the representation of percentages in different centiles or scores. The centiles can also be broken down into frequencies before the percentages are derived. It is also known as the k-th percentile which helps in creating the values of range. By doing this, a threshold of acceptance is ensured. On this note, a percentile range is the difference that exists between two percentiles. An example is 10-30, 40-60, or 10-90. An additional example is in Figure 6.

	A	B	C
1	**Products**	**Sales**	
2	Product 1	$ 48,947	
3	Product 2	$ 42,781	**85th percentile**
4	Product 3	$ 42,306	
5	Product 4	$ 41,471	**66th - 85th percentile**
6	Product 5	$ 40,556	
7	Product 6	$ 38,602	
8	Product 7	$ 34,735	
9	Product 8	$ 27,916	**33rd - 66th percentile**
10	Product 9	$ 22,309	
11	Product 10	$ 18,576	
12	Product 11	$ 18,043	**20th - 33rd percentile**
13	Product 12	$ 14,430	
14	Product 13	$ 13,804	
15	Product 14	$ 13,631	**20th percentile**
16	Product 15	$ 13,469	
17			

Figure 7: An example of a percentile

A percentile is a statistical measure used to determine the percentage of scores or values in a dataset which are below a specified threshold. Its calculation involves dividing the number of data points falling beneath that score by the total number of datapoints, then multiplying it with 100. For instance, being at the 65th percentile indicates that approximately 65% of data inside your set fall short on this value limit

Another way by which percentile is used is to determine the number of learners who score above, for instance, 90^{th} percentile. To calculate the percentile is P=n/N * 100%. Here, P represents the percentile. The "n" in lowercase is the "number of data points". The "N" in upper case is the number of data points available in the dataset.

Various ways percentiles can be used

1. Percentiles are useful in characterizing the distribution of scores, revealing whether it is skewed or normal. A dataset featuring a prolonged tail of high scores could have a significantly elevated 90th percentile compared to its median, which implies an asymmetric distribution.
2. Percentiles provide a reliable way to assess scores across multiple datasets, regardless of differing means or standard deviations. For instance, if two students possess unequal GPAs, percentile comparison can be employed to ascertain which student has more grades above an established threshold.
3. Percentiles can serve as a means for establishing acceptance or rejection boundaries in diverse scenarios, including assessing job applicants and controlling product quality. To illustrate, an enterprise might implement an 80th percentile benchmark when evaluating candidates to fill vacancies. This implies that solely aspirants with scores surpassing the 80th percentile will qualify for consideration of appointment to the position at hand.
4. Other statistical measures can be derived using percentiles, including the interquartile range (IQR). This is determined by finding the difference between Q3 (the 75th percentile) and Q1 (25th percentile). The IQR serves as an alternative measure of dispersion that gives less weight to outliers compared with standard deviation.

$$IQR = Q3 - Q1$$

Stages of Percentiles

1. The first step is to gather and organize the data either in ascending or descending sequence. This is essential for percentile computation that relies on the ranking of data points rather than their real magnitude.
2. Data analysis is performed to ascertain the study's objectives in the second phase. This entails computing

summary statistics, visually representing information or detecting patterns and trends within the data.

3. The third step is data interpretation, which entails comprehending the outcomes of the analysis. This could include contrasting percentiles among various groups, scrutinizing score distributions, or discovering any outliers that might be present.
4. The last step involves communicating and reporting the discovered information to relevant parties. This can be done using various methods such as creating visual aids, preparing a written document or presenting findings during an audience presentation.

Types of Percentiles

1. Quartiles: Quartiles divide a dataset into four equal parts, similar to percentiles. The 25th percentile (Q1) and the 75th percentile (Q3) are respectively known as the first quartile and third quartile, dividing the set equally in quarters. Additionally, Q2 or the median represents the middle value at 50%.
2. Deciles: Deciles divide a dataset into ten equal parts and represent percentiles.

   Example
   D1- 10th percentile
   D9 - 90th percentile
3. Percentiles based on raw scores: Calculating percentiles using raw scores instead of ranks is useful in cases where the data lacks normal distribution or when its distribution cannot be determined.

Advantages of Percentiles

1. Percentiles provide a standardized way to compare values within a distribution, making it easier to understand relative standings.
2. Percentiles are useful for establishing reference ranges and thresholds in various fields such as education, healthcare, and economic analysis.

3. Percentiles exhibit resilience against outliers as they rely on ordering, rather than absolute values.

Disadvantages of Percentiles

1. Percentiles are susceptible to skewed distributions and even a minor shift in the distribution can lead to significant fluctuations in percentile values.
2. Percentiles are inappropriate for continuous data as ranks, the foundation of percentiles, have an ill-defined application to continuous data.
3. Percentiles may not be apt for small samples because the ranks might lack proper definition, thus yielding erroneous computations.

Summary

In summary, this section examined how data are organized and used in education. Different areas such as data classification, frequency distribution, and percentile have been discussed. It has been found that descriptive data are used to explain situations while inferential data shows relationships between variables. In the same vein, data organization involves categorizing and classifying data to make it more usable and readable. The representation of data can be ascending or descending. Further findings have shown that data can also be a combination of qualitative and quantitative information- this is called mixed data. When a study adopts mixed methods, mixed data can also be used. It was also discussed that percentile is the representation of percentages in different centiles or scores. It is also known as the k-th percentile which helps in creating the values of range. In conclusion, all these forms of data representation aid the viability of data collection and interpretation.

Revision Questions

1. What do you understand by data classification?

2. Mention and explain the types of data classification that we have.

3. What are the advantages of classifying data?

4. With copious examples, explain what you understand by chronological data.

5. What are the differences and similarities between qualitative and quantitative data?

6. Mixed data aid extensive findings. Discuss this with different examples.

7. What is percentile and what are the stages involved?

8. What are the advantages and disadvantages of using percentiles?

References

Ahmad, S., Wasim, S., Irfan, S., Gogoi, S., Srivastava, A., & Farheen, Z. (2019). Qualitative v/s Quantitative Research. 6, 2828–2832. https://doi.org/10.18410/jebmh/2019/587.

Byju's, (2004). Need of Organization of Data. Available online: https://byjus.com/maths/data-organization.

In, Junyong & Lee, Sangseok. (2017). Statistical data presentation. *Korean Journal of Anesthesiology,* 70(3), 1 -10.

Saunders, M. N. K., Lewis, P., & Thornhill, A. (2019). Research methods for business students (Eighth Edition). New York: Pearson.

Woiceshyn, J., & Daellenbach, U. (2018). Evaluating Inductive versus Deductive Research in Management Studies: Implications for Authors, Editors, and Reviewers. Qualitative Research in Organizations and Management: An International Journal, 13(1). https://doi.org/10.1108/QROM-06-2017-1538.

Vedantu (2024). Data Organization. Available online: https://www.vedantu.com/maths/data-organization.

Chapter 7

MEASURES OF CENTRAL TENDENCY

Opeyemi Seyi OLASUNKANMI (Ph.D)
Department of Educational Management,
University of Ibadan, Ibadan, Nigeria.
Email: oskomus@gmail.com
Tel: +2347037710690

Omolola Olukemi OGUNNIRAN (Ph.D)
Directorate of General Education Studies,
Adeleke University, Ede, Osun State.
Email: fitela907@gmail.com
Tel: +23408107708315

Introduction

Measures of central tendency are statistical measures that describe the centre of a given distribution of values. It aims to provide a single value that represents the entire dataset or distribution. A measure of central tendency is seen as a number used to represent the value in a collection of numbers. There are three primary measures of central tendency: the mean, the median, and the mode. Measures of central tendency will be explained under the grouped and ungrouped data.

Distribution of Ungrouped Data

Ungrouped data is a type of data that has not been organized or grouped into specific categories. It is also known as raw data or unorganized data. This section will explain the measures of central tendency (mean, median, and mode) under the distribution of ungrouped data.

Mean Distribution of Ungrouped Data

Mean is an average of a given distribution or arithmetic mean/average of a distribution. To calculate mean of a given distribution under the ungrouped data, two methods can be adopted, these are: Direct Method; and Assumed Mean Method.

Direct Method

Arithmetic mean can be calculated through direct method. This is done by adding all the observations in a distribution and then divides it by the number of observations.

i.e. Mean $= \frac{Sum\ of\ all\ data\ in\ the\ observation}{Nnumber\ of\ observation} = \frac{\Sigma X}{N}$

Where: Σ = Summation

N = Number of Observation

X = Observation under investigation

Example 1: Find the mean of the following marks obtained by students in a Mathematics test

35, 22, 36, 15, 19, 8, 15, 27, 18, 29, 31, 26, 14, 23 and 33

Solution: Using direct method,

Mean is derived by: $\frac{\Sigma X}{N}$

$$\frac{35 + 22 + 36 + 15 + 19 + 8 + 15 + 27 + 18 + 29 + 31 + 26 + 14 + 23 + 33}{15}$$

$= \frac{351}{15}$

$= 23.4$

Therefore, mean of the distribution is 23.4

Example 2: Marks obtained by 10 students in a Mathematics test is given in a table below, using direct method, calculate the mean of mark of the students

Number of Students	1	2	3	4	5	6	7	8	9	10
Marks	5	4	8	5	2	9	10	6	4	7

Solution: Mean of the marks of students can be obtained using $\frac{\Sigma X}{N}$

$$\frac{5+4+8+5+2+9+10+6+4+7}{10}$$

$$\text{Mean} = \frac{60}{10}$$
$$\text{Mean mark} = 6$$

Assumed Mean Method

The second method that can be used to calculate an arithmetic mean is called Assumed Mean method. This is done by assuming a particular figure among the distribution as mean value, then, deviations from the assumed mean from each of the observation is taken and added together, then, divide the total by the number of observations. Symbolically, assumed mean can be calculated using the formula:

$$\bar{x} = A + \frac{\Sigma d}{N}$$

Where: A = assumed mean

Σ = Summation

d = X minus Assumed Mean

X = Observation under investigation

N = Number of Observation

Example 3: The following are the weights of 10 patients in kg as measured in a hospital, 45, 56, 59, 39, 60, 72, 65, 48, 50 and 58. Use the assumed mean method, find the mean of the weights of the patients.

Solution:

Use $A + \frac{\Sigma d}{N}$

Pick any figure in the distribution as an assumed mean and substitute, (assumed mean = 59)

S/N	Weights (X)	d = X – A (Assumed Mean =59)
1	45	-14
2	56	-3
3	59	0
4	39	-20
5	60	1
6	72	13
7	65	6
8	48	-11
9	50	-9
10	58	-1
N=10	ΣX = 552	Σd = -38

Using assumed mean method, Mean = $A + \frac{\Sigma d}{N}$

$$= 59 + \frac{-38}{10}$$

$$= 59 + (-3.8)$$

$$= 55.2$$

Therefore, using assumed mean method, Mean = 55.2.

Note: This can be verified using direct method.

Advantages of Mean

The following are some advantages of mean.

1. Mean is very easy to calculate and understand.
2. It provides a central value representing the dataset.
3. It is useful for comparing datasets or trends.
4. It can be used for forecasting and prediction.
5. Facilitates data visualization (e.g. charts and graphs).

Disadvantages of Mean

1. Sensitivity to outliers i.e. extreme values can skew the mean
2. Small changes in data can significantly impact the mean

3. Assumes normality: Mean is optimal for normal distributions, but may not represent skewed data
4. Mean only provides central tendency, neglecting variability

Median Distribution of Ungrouped Data

The median is the middle number when the data is arranged in order of magnitude (either ascending or descending order). It can be seen as a positional value of the variable which divides the distribution into two equal parts, one part comprises all values greater than or equal to the median value and the other comprises all values less than or equal to it.

If a distribution is given, the data are arranged in order of magnitude and the middle number is selected as the median. In case of even distribution, the mean of the two middle observations becomes the median.

Example 4: Find the median of the following observations: 12, 24, 14, 16, 26, 17, 10

Solution:

12, 24, 14, 16, 26, 17, 10

Rearrange the data in descending order: 26, 24, 17, ***16***, 14, 12, 10

The middle observation in the given data is 16, therefore, the median of the observation is 16.

Example 5: The following scores were obtained by 10 students in Mathematics examination. Calculate the median score. 45, 53, 60, 39, 72, 55, 61, 49, 50, 69

Solution:

45, 53, 60, 39, 72, 55, 61, 49, 50, 69

Rearrange the data in ascending order

39, 45, 49, 50, ***53, 55***, 60, 61, 69, 72

It will be observed that the distribution is even, i.e. 10 observations, then, the median will be the mean of the two middle numbers,

i.e. $\frac{53+55}{2}$

$= \frac{108}{2}$

$= 54$

Therefore, the median is 54.

Advantages of Median

The following are the advantages of median.

1. Median is less sensitive to outliers and skewed distributions.
2. It remains stable despite unusual data points.
3. Median often represents typical values more accurately.
4. It doesn't assume normality or specific distribution.
5. Median is simple to understand communicate and interpret.

Disadvantages of Median

Here are the disadvantages of using median.

1. The median can be affected by the sampling method used, which can lead to biased results.
2. The median is most useful for symmetric distributions, in skewed distributions, it may not accurately represent the data.
3. The median ignores extreme values or outliers, which can be important in understanding the data.
4. Medians don't follow algebraic properties, making it challenging to perform mathematical operations.
5. Medians are less useful in inferential statistics, where means and standard deviations are often required.

Mode Distribution of Ungrouped Data

The mode is the value that appears most frequently in a dataset or distribution. It is a measure of central tendency, used to describe the typical value in a dataset.

Mode can be (a) Unimodal – one mode in a distribution; (b) Bimodal – two modes in a dataset; (c) Multimodal – multiple modes in a dataset and (d) No mode - no value appears more than once in a distribution.

Consider the distribution: 2, 4, 7, 8, 8, 9, 10. The mode in this distribution is 8 because it appears most frequently.

Example 6: Find the mode in the below distribution: 12, 1, 0, 7, 8, 18, 20, 12, 5, 8, 12, 6, 12.

Solution

Since the mode is the most occurrence frequency in a distribution, from the distribution, the mode is 12, because it appears 4 times.

Example 7: Given the table below table, find the mode in the distribution

Variable	11	12	13	14	15
Frequency	3	6	4	8	5

Solution:

Here, from the distribution, the most occurrence frequency is 8, and the value of mode as against 8 is 14. The implication of this is that 14 appears 8 times, which makes it a mode of distribution.

Advantages of Mode

The following are the advantages of mode:

1. Easy to understand and calculate
2. Useful for categorical data
3. Robust against outliers
4. Visual representation through histograms or bar charts

Disadvantages of Mode

The following are the disadvantages of mode:

1. Not suitable for continuous data with unique values
2. Limited algebraic properties
3. May not represent the entire dataset

DISTRIBUTION OF GROUPED DATA

The distribution of grouped data refers to the organization and presentation of data that has been categorized into groups or intervals, often with a specific range or width. The grouping helps to simplify complex data, reveal patterns and trends of data and also enhances visualization. A grouped data distribution shows the number of observations in each group, displays the accumulated frequencies and exhibits proportions or percentages of a given distribution. By grouping data effectively, one can uncover valuable insights, identify trends, and make informed decisions.

Mean Distribution of Grouped Data

The mean distribution of grouped data is the process of finding the average of a set of data that are grouped together in different categories. To determine the mean of grouped data, a frequency table is required to set across the frequencies of the data which makes it simple to calculate. Two methods will be considered in calculating mean of a grouped data, direct method and assumed mean method.

Direct Method of Calculating Mean Distribution of Grouped Data

To calculate mean using direct method, the below formula is used:

$$\text{Mean} = \frac{\Sigma FX}{\Sigma F}$$

Where:

ΣFX = sum of the product of variables and frequencies

ΣF = sum of frequencies

Example 8: Compute an arithmetic mean of the following information, using the direct method.

X	5	3	9	5	2	4	5
F	20	10	8	11	9	7	15

Solution

Using a direct method, the arithmetic mean is given as:

Mean $= \frac{\Sigma FX}{\Sigma F}$

X	**F**	**FX**
5	20	100
3	10	30
9	8	72
5	11	55
2	9	18
4	7	28
5	15	75
	ΣF =**80**	Σ FX = **378**

Mean $= \frac{\Sigma FX}{\Sigma F}$

Mean $= \frac{378}{80}$

Mean = 4.73

Example 9: Given the below information, calculate the mean.

Class interval	1 – 10	11 – 20	21 – 30	31 – 40	41 – 50
Frequency	17	12	6	18	13

Solution:

Mean is given by $\frac{\Sigma FX}{\Sigma F}$

Class Interval	Frequency	X (Mid Class)	FX
1 – 10	17	5.5	93.5
11 – 20	12	15.5	186
21 – 30	6	25.5	153
31 – 40	18	35.5	639
41 – 50	13	45.5	591.5
	ΣF =**66**		ΣFX = **1663**

Mean = $\frac{\Sigma FX}{\Sigma F}$

Mean = $\frac{1663}{66}$

Mean = 25.2

Assumed Mean Method of Calculating Mean Distribution of Grouped Data

The formula to calculate mean distribution of grouped data using assumed mean method is given below:

Mean = A + $\frac{\Sigma Fd}{\Sigma F}$

Example 10: Use the information in example 9, using assumed mean method, calculate the mean of the distribution

Class interval	1 – 10	11 – 20	21 – 30	31 – 40	41 – 50
Frequency	17	12	6	18	13

Solution:

Mean under Assumed mean method is given by A + $\frac{\Sigma Fd}{\Sigma F}$

Using 35.5 as assumed mean, the table is transformed to

Class Interval	Frequency	X (Mid Class)	d = (X – AM)	Fd
1 – 10	17	5.5	-30	-510
11 – 20	12	15.5	-20	-240
21 – 30	6	25.5	-10	-60
31 – 40	18	35.5	0	0
41 – 50	13	45.5	10	130
	ΣF =66			**ΣFd = -680**

A + $\frac{\Sigma Fd}{\Sigma F}$

35.5 + $\frac{-680}{66}$

35.5 + (–10.303)

35.5 – 10.303

Mean = **25.2**

Note: To calculate mid-class which is X, lower class interval is added to upper class interval, then divided by 2. e.g. $\frac{1+10}{2}$

= $\frac{11}{2}$ = 5.5

Compare the result of examples 9 and 10, the result is the same using direct method and assumed mean method.

Median Distribution of Grouped Data

Median of ungrouped data refers to the middle number in a given distribution after arranging in order of magnitude. In grouped data, it is difficult to find the median for the given observation by checking the cumulative frequencies, because the middle value of the given data will be in some class interval. It is necessary to find the value inside the class interval that divides the whole distribution into two halves. In this case, the median class has to be found. To find the median class, it is important to find the cumulative frequencies of all the classes, and divide it by 2 (n/2), Then, locate the class whose cumulative frequency is greater than but nearest to n/2, the class is called the median class.

To calculate the median of a grouped data, the below formula is used

$$L + \frac{(\frac{n}{2} - c.f)}{f} * c$$

Where: L = Lower limit of the median class
n = Number of observation
f = Frequency of the median class
c.f = Cumulative frequency before the median class
c = Class size or the magnitude of the class interval

Note: The first thing to do when calculating median of a grouped data is to identify the median class, this is done through $\frac{n}{2}$, then other things follow

Example 11: Find the median weight of the following data which represents the weights of some selected factory workers in kg.

Weights	**Frequency**
60 – 64	4
65 – 69	5
70 – 74	13
75 – 79	20
80 – 84	14
85 – 89	8

Solution:

Weights	Frequency	Cumulative Frequency
60 – 64	6	6
65 – 69	5	11
70 – 74	13	24
75 – 79	20	44
80 – 84	14	58
85 – 89	12	70

Using L + $\frac{(\frac{n}{2} - c.f)}{f} * c$

Begin with $\frac{n}{2}$

$\frac{70}{2}$ = 35 – Therefore, the observation lies between the class interval of 75 – 79, this is called the median class.

L + $\frac{(\frac{n}{2} - c.f)}{f} * c$

L = 75

N = 70

c.f = 24

f = 20

c = 5

Then, all these can be substitute in the formula

L + $\frac{(\frac{n}{2} - c.f)}{f} * c$

75 + $\frac{(\frac{70}{2} - 24)}{20} * 5$

75 + $\frac{(35 - 24)}{20} * 5$

75 + $\frac{(11)}{20} * 5$

75 + $\frac{(55)}{20}$

75 + 2.75

Median = 77.75

Mode Distribution of Grouped Data

The mode of grouped data is the interval or group with the highest frequency. Since grouped data does not provide individual values, the mode is an interval rather than a specific value. Modal class is the class with largest frequency. Mode can be calculated by using the formula:

$\text{Mo} = \text{L} + \frac{\text{D1}}{\text{D1}+\text{D2}} * c$

Where:

L = Lower class limit of the modal class

D_1= Difference between the frequency of the modal class and frequency before the modal class

D_2 = Difference between the frequency of the modal class and frequency after the modal class

C = Class size or the magnitude of the class interval

Example 12: Given the information below as the monthly income of some selected workers in a company, calculate the modal monthly income of the workers.

Class Interval	Frequency
1 – 10	12
11 – 20	15
21 – 30	20
31 – 40	17
41 – 50	25
51 – 60	13
61 – 70	9

Solution:

Using $\text{Mo} = \text{L} + \frac{\text{D1}}{\text{D1}+\text{D2}} * c$

To find the modal class, the modal class is the class that has the highest frequency, this is class interval 41 – 50, this makes it the modal class.

L = 41

D_1 = 25 – 17 = 8

D_2 = 25 – 13= 12

C = 10

Then, substitute in $L + \frac{D1}{D1 + D2} * c$

$$41 + \frac{8}{8+12} * 10$$

$$41 + \frac{8}{20} * 10$$

$$41 + \frac{80}{20}$$

$$41 + 4$$

$$45$$

Then, mode is 45

Example 13: The following marks were scored by 50 students in Mathematics examination.

33, 70, 36, 51, 82, 60, 61, 53, 44, 92, 64, 56, 55, 34, 53, 62, 41, 58, 95, 71, 51, 70, 26, 76, 78, 34, 75, 40, 82, 68. 64, 53, 49, 79, 16, 62, 29, 73, 68, 15, 48, 55, 64, 75, 43, 61, 50, 75, 22, 18

a. Tally the scores in a frequency distribution table using class size of 5

b. Calculate the:

i) mean, using direct method and assumed mean method

ii) median, and

iii) mode of the distribution

Solution

Class Interval	Tally	Frequency	Cumulative Frequency	X – Mid Point	FX	d = X– Assumed Mean	fd
15 – 19	///	3	3	17	51	-65	-195
20 – 24	/	1	4	22	22	-60	-60
25 – 29	//	2	6	27	54	-55	-110
30 – 34	///	3	9	32	96	-50	-150
35 – 39	/	1	10	37	37	-45	-45
40 – 44	////	4	14	42	168	-40	-160
45 – 49	//	2	16	47	94	-35	-70
50 – 54	~~////~~ //	7	23	52	364	-30	-210
55 – 59	///	3	26	57	171	-25	-75

60 – 64	~~////~~ ///	8	34	62	496	-20	-160
65 – 69	//	2	36	67	134	-15	-30
70 – 74	////	4	40	72	288	-10	-40
75 – 79	~~////~~ /	6	46	77	462	-5	-30
80 – 84	/	1	47	82	82	0	0
85 – 89	/	1	48	87	87	5	5
90 – 94	/	1	49	92	92	10	10
95 – 99	/	1	50	97	97	15	15
TOTAL		**50**	**N = 50**	**969**	**2795**	**-425**	**-1305**

Note: To calculate mid-point which is X, lower class interval is added to upper class interval, then divided by 2. e.g. $\frac{15+19}{2}$

$= \frac{34}{2} = 17.$

bi. To calculate the mean using direct method, use

$\text{Mean} = \frac{\Sigma FX}{\Sigma F}$

$= \frac{2795}{50} = 55.9$

Mean = 55.9 – With Direct Method

Using Assumed Mean method

$\text{Mean} = A + \frac{\Sigma Fd}{\Sigma F}$

Assumed mean as calculated on the table is 82

$\text{Mean} = 82 + \frac{-1305}{50}$

$= 82 + (-26.1)$

$= 82 - 26.1$

Mean = 55.9 – With Assumed Mean Method

bii. Median

To calculate median, use

$\text{Median} = L + \frac{(\frac{n}{2} - c.f)}{f} * c$

To determine the median class, find $\frac{n}{2}$

$\frac{n}{2} = \frac{50}{2} = 25$

Median class falls in the class interval of 55 – 59, then, substitute the formula

$$L + \frac{(\frac{n}{2} - c.f)}{f} * c$$

$$55 + \frac{(\frac{50}{2} - 23)}{3} * 5$$

$$55 + \frac{(25 - 23)}{3} * 5$$

$$55 + \frac{(2)}{3} * 5$$

$$55 + \frac{(10)}{3}$$

$$55 + 3.3$$

Median = 58.3

biii. Mode

To calculate mode, use

$$Mo = L + \frac{D1}{D1 + D2} * c$$

Determine the modal class by finding the class interval with highest frequency

The highest frequency is 8 with class interval of 60 – 64

Then,

$L = 60$

$D_1 = 8 - 3 = 5$

$D_2 = 8 - 2 = 6$

$C = 5$

$$Mo = L + \frac{D1}{D1 + D2} * c$$

$$60 + \frac{5}{5 + 6} * 5$$

$$60 + \frac{5}{11} * 5$$

$$60 + \frac{25}{11}$$

$$60 + 2.3$$

$$62.3$$

Then, mode is 62.3

Summary

Measures of central tendency provide a single value that represents the typical value of a dataset. Measures of central tendencies help to summarize and describe the central position of the data. The three main measures of central tendencies are mean, median, and mode. Mean is the average value of a distribution or dataset, it is obtained by adding all observations divided by the number of observations. Mean can be calculated through direct method and assumed mean method. The mean of ungrouped data can be obtained through:

$\frac{\Sigma X}{N}$ – Direct Method

A + $\frac{\Sigma d}{N}$ – Assumed Mean Method

Mean of the grouped data can be obtained by:

$\frac{\Sigma FX}{\Sigma F}$ – Direct Method

A + $\frac{\Sigma Fd}{\Sigma F}$ – Assumed Mean Method

Median is the middle numbers when the data is arranged in order of the magnitude, in case of even distribution, the mean of the two middle observations becomes the median. Median of the grouped data can be calculated using L + $\frac{(\frac{n}{2} - c.f)}{f} * c$.

Mode is the value that appears most frequently in a dataset or distribution. Mode of the grouped data can be calculated by Mo = L + $\frac{\text{D1}}{\text{D1+ D2}} * c$.

Exercises

1. Comprehensively define mean, median and mode. State three advantages and three disadvantages for each of them.
2. The frequency table below represents the distribution of students' scores in an examination. Compute the mean (use direct and assumed mean methods), median and mode for the distribution

Value	**Frequency**
12	38
13	46
14	58
15	22
16	65
17	73

3. The table below gives information about the production of a product in a company, calculate the mean median and mode values of the distribution

Production	55–57	58–60	61–63	64–66	67–69	70–72	73-75	76–78	79–81
No of Product	13	18	24	30	45	21	18	15	20

4. The following marks were scored by 50 students in Mathematics examination.

45, 55, 63, 52, 23, 16, 28, 61, 19, 24, 25, 45, 49, 43, 21, 11, 71, 45, 65, 32,
66, 61, 22, 23, 25, 45, 55, 62, 85, 91, 33, 50, 40, 46, 59, 39, 34, 60, 52, 51,
43, 28, 16, 12, 17, 25, 63, 54, 28, 29, 43, 45,33, 70, 36, 51, 82, 60, 61, 53,
44, 92, 64, 56, 55, 34, 53, 62, 58, 95, 71, 51, 70, 26, 76, 78, 34, 75, 40, 82,
68. 64, 53, 49, 16, 62, 29, 73, 68, 15, 48, 55, 64, 75, 43, 61, 50, 75, 22, 18

a. Tally the scores in a frequency distribution table using class size of 10
b. Compute the cumulative frequency of the table
c. Calculate the mean, median and mode of the distribution

References

"MLB Standings – 2012." Available online at http://espn.go.com/mlb/standings/_/year/2012.

Hand, D.J., F. Daly, A.D. Lunn, K.J. McConway, and E. Ostrowski. A Handbook of Small Datasets. London: Chapman & Hall, 1994, pg. 118.

Introduction to Statistics http://www.uop.edu.pk/ocontents/chapter%203.pdf

Mackowiak, P. A., Wasserman, S. S., and Levine, M. M. (1992), "A Critical Appraisal of 98.6 Degrees F, the Upper Limit of the Normal Body Temperature, and Other Legacies of Carl Reinhold August Wunderlich," Journal of the American Medical Association, 268, 1578-1580.

www.https://openstax.org/details/books/introductory-statistics-2e/

www.https://www.math.fsu.edu/~wooland/hm2ed/Part3Module2/Part3Module2.pdf

Introduction to Statistics Online Edition retrieved from https://onlinestatbook.com/Online_Statistics_Education.pdf

Chapter 8

MEASURES OF VARIABILITY

Veronica Folasade T. Babajide
Faculty of Education, University of Lagos, Lagos, Nigeria.

Introduction

Measure of variability is a measure of the extent to which a score differs from another score in an array of scores. It also refers to the measure of the extent of spread or cluster of scores or data. Measure of variability is a fundamental concept in research and statistics and has practical applications in man's day-to-day activities. This chapter discusses the measure of variability under: the range, variance and standard deviation.

Range

The range is the difference between the highest score and the lowest score in an array of scores, and it is obtained by inspection.

Example 1**:** What is the range of the following numbers? 1, 2, 3, 4, 5, 6, 7, 8 and 9?

Solution: The lower number is 1 and the highest number is 9
The range =highest – lowest   9-1= 8. The range of the numbers is 8.

Example 2**:** The following scores are obtained from 10 students in a science course. What is the range of their scores? 65, 90, 70, 92, 80, 76, 95, 89, 84 and 91

Solution:
The highest score is 91 and the lowest score is 65.
Range = highest -lowest.
91-65 =26.
The range of the scores is 26.

***Example 3*: ** Calculate the range of the weights of the following 30 students in a higher degree class: 74, 66, 76, 76, 64, 58, 68, 45, 66, 77, 54, 70, 72, 75, 61, 57, 54, 68, 81, 61, 55, 50, 52,50 ,52, 58, 57, 45, 66 & 70

***Solution*:**
Highest weight =77; lowest weight=45
Range = 77-45= 32.
The range of the weight is 32

The range only takes care of the two extreme numbers, scores, or data and does not consider or provide information on the number, scores, or data in between or within the dataset. The range is very useful for a quick check of spread. In education, the range can be used to quickly check the spread of performance of students in a test or examination to identify areas of improvement and strength. In the sector of health, the range is used to study the variabilities of the effectiveness of treatment to identify patient recovery times. In the sector of business, the range of sales figures over a given period can be used to identify market stability and performance.

Variance

Variance is another statistical measure of dispersion or spread of data points in a dataset or array of scores. It is the sum of the square deviation from the mean divided by the total number of scores which is the total frequency. Variance is a measure of the extent to which values in a dataset differ from the mean (average) of the dataset. The formula to calculate variance is:

$\boldsymbol{\sigma^2} = \frac{\sum(x-\mu)2}{N}$

Where σ^2 is population variance; x is individual score or data; μ is population mean; N is the number of scores or data in a population; Σ is summation (that is, sum of)

Sample Variance **(s^2)** $=\Sigma.\frac{(x-\bar{x})2}{N-1}$

Where $\mathbf{s^2}$ is sample variance; Σ is summation; x is individual score or data; and N is the number of scores in the data.

Procedure for computing the variance:

1. Compute the frequency distribution table of the data
2. Compute the mean.
3. Determine the mean deviation of each score $x - \bar{x}$
4. Find the square of each deviation from the mean $(x - \bar{x})^2$
5. Find the product of each square of the deviation from the mean and the corresponding frequency of each score /data. f $(x - \bar{x})^2$
6. Divide the product by the total number of students who have the scores or data.
7. Variance (s^2) = $\frac{f\ (x-\bar{x})2}{N}$

Example 4**:** Compute the variance of the following scores of students obtained in a class test. 3,7, 2,10,3.9, 9, 9, 10, 8, 8, 8, 2,4, 7, 7,7,6, 6, 6, 6, 6, 6, 6,4, 4, 4,4, 5,5,5,5,5,10,3

Step 1. Compute the frequency distribution table. (Let it be Table 31.1

Table 31.1: Table of frequency distribution of scores obtained by students in a class test.

Score x	Frequency (f)	Fx	$(x - \bar{x})$	$(x - \bar{x})^2$	$f(x - \bar{x})^2$
10	3	30	4.03	16.24	48.72
9	3	27	3.03	9.18	27.54
8	3	24	2.03	4.12	12.36
7	4	28	1.03	1.06	4.24
6	7	42	0.03	0.00	0.00
5	5	25	-0.97	0.95	4.75
4	5	20	-1.97	3.88	19.4
3	3	9	-2.97	8.82	26.46
2	2	4	-3.97	15.76	31.52
TOTAL	Σf=35	Σfx=209			$\Sigma f(x - \bar{x})^2 = 174.99$

Step 2: Calculate the mean. $= \frac{\Sigma fx}{N} = \frac{209}{35} = 5.97$

Step 3: Compute the deviation of each score from the mean.

Step 4: Find the square of each deviation from the mean

Step 5: Compute the product of the frequency and the square of each deviation from the mean

Step 6: Find the Sum of the product of frequency and the square of the deviation of each score from the mean

Step 7: Apply the formula and calculate the Variance

1. Variance $(s^2) = \frac{f(x - \bar{x})^2}{N} = \frac{174.99}{35}$

   Variance= 4.99

Example 5:

The following are the scores of students in EDU 802 in University of Lagos:

73, 78, 74, 68, 52, 50, 63, 68, 56, 65, 70 81, 74, 60, 60, 78, 63, 70, 68, 60, 72, 62, 72, 52, 61, 61,65, 68, 40, 52, 68, 60, 72, 50, 80, 52, 66, 60, 52, 56, 45, 60, 63, 63, 74, 60, 75, 60, 53, 54,

50, 70,40, 64, 70, 71, 60, 50, 65, 58, 50, 50, 52, 74, 60, 72, 63, 60, 53, 60, 40, 47, 51, 40, 50, 70, 60, 75,54, 53, 57, 66, 51, 55, 70, 52, 45, 56, 63, 60, 60, 52, 41, 45, 45, 75, 67, 57, 61, 68, 43, 57, 54, 63,57, 58, 66, 62, 74, 61, 64, 40, 51, 68, 64, 65,60,41, 70,55, 71,65,71,64,,70,51,59,64,80,55,62,66,52.
Using a class interval of 5 prepare a frequency distribution table and compute the Variance of the scores of the students in the distribution.

Step 1: Prepare a frequency distribution table.

Table 31.2: Frequency distribution table of the scores of students in EDU 802

Class interval	Frequency (f)	Mid point (x)	Fx	$x - \bar{x}$	$(x - \bar{x})2$	f $(x - \bar{x})2$
40-44	8	42	336	-19.02	361.76	2894.08
45-49	5	47	235	-14.02	196.56	982.80
50-54	26	52	1352	-9.02	81.36	2115.36
55-59	13	57	741	-4.02	16.16	210.09
60-64	35	62	2170	0.98	0.96	33.6
65-69	17	67	1139	5.98	35.76	607.92
70-74	21	72	1512	10.98	120.56	2531.76
75-79	5	77	385	15.98	255.36	1276.80
80-84	3	82	246	20.98	440.16	1320.48
	133		Σ8116			Σf $(x - \bar{x})2$=11972.67

Step 2 Calculate the Mean $=\frac{\sum fx}{N} = \frac{8116}{133}$ =61.02
= 61.02

Step 3: Compute the deviation of each score from the mean.
Step 4: Find the square of each deviation from the mean
Step 5: Compute the product of the frequency and the square of each deviation from the mean

Step 6: Find the Sum of the product of frequency and the square of the deviation of each score from the mean
Step 7: Apply the formula and calculate the variance

$$\text{Variance } (s^2) = \Sigma \frac{f\,(x - \bar{x})2}{N} = \frac{11972.89}{133} = 90.02$$

31.30 Standard Deviation

Standard deviation is defined as the square root of the variance. It provides a measure of the average distance from the mean in the same unit as the data in a distribution.
The formula for calculating the standard deviation of an array of scores or data in a distribution is

$$\text{Standard Deviation} = \surd\, \Sigma \frac{f\,(x - \bar{x})2}{N}$$

Where $\Sigma \frac{f\,(x - \bar{x})2}{N}$ is the formular for calculating the variance as explained in 4 above. All other definitions of all the symbols remain the same.

The value of the standard deviation in the example above is Standard Deviation √ 90.02= 9.49

Standard deviation means how much an individual score or data deviates on average from the mean of the score or data set. A high standard deviation means the scores are far from the mean; spread out over a wide range of values while a low standard deviation means that the scores are closer to the mean. Standard deviation is widely used to measure variability or dispersion of scores of data in a distribution. The knowledge of standard deviation can be applied in all areas of life. In education, it is used to evaluate the teaching effectiveness of a teacher in any subject or course. If the standard deviation of students' scores is calculated and it is very low, it means the teacher has done well. If it is high, then the teacher has not done well, he or she may have to re-teach the students in the subject or course. Generally, the standard deviation must be relatively low to the mean.

***Example 6*:** Find the standard deviation of the scores in example 31.4

Solution Follow the steps provided to calculate the variance. From the variance calculated as 4.99, the standard deviation is the square root of the variance.
Standard Deviation=$\sqrt{4.99}$ = 2.23

***Example 7*:** Calculate the standard Deviation of the students in EDU 802

***Solution*:** Follow the steps provided above to calculate the variance in example four.
The calculated value of variance in example 31.22 =90.02
Standard Deviation =$\sqrt{90.02}$= 9.49

***Example 8*:**. Calculate the value of the standard deviation of the scores obtained by 9 students in a mathematics course.
95. 90. 85, 80, 75,70,65, 60 & 55

Solution
Step 1. Calculate the mean of the scores

X	$(x - \bar{x}$	$(x - \bar{x})^2$
95	20	400
90	15	225
80	5	25
75	0	0
70	-5	25
65	-10	100
60	-15	225
55	-20	400
Σx =675		$\Sigma(x - \bar{x})^2$= 1375

Mean $\bar{x}$ = $\frac{675}{9}$ =75

Step 2: Find the deviation of each score from the mean

Step 3: Find the square of each deviation from the mean

Step 4: Find the sum of the square of each deviation from the mean

Step 5:. Divide the sum of the square of deviation from the mean by the total number of scores= variance

Variance $=\frac{\sum(x-\bar{x})2}{N}==\frac{1375}{9}= 152.7$

Standard deviation $=\sqrt{152.7}= 12.37$

Summary

This chapter has discussed the range, as a measure of variability but focused on the extreme scores only. It does not account for the other scores; it is used as a quick check. Variance and standard deviation as measures of variability consider other scores in the distribution. They provide an adequate measure of variability by accounting for all scores in the distribution. The variance is the average sum of the square of the deviation from the mean, and the standard deviation is the square root of the value of the variance. They are used majorly in education as a statistical tool for the analysis and the interpretation of data collected, for the assessment of teaching effectiveness, to improve teachers' methods of instruction and students' learning outcomes.

Exercises

1. What do you understand by variability?
2. Mention and discuss any two measures of variability.
3. A classroom teacher conducted an examination to assess students' level of students' learning outcomes. The value of the standard deviation of the scores is greater than the mean of the distribution, what will you say about the quality of the teaching and the way forward?

4. Compute the variance and the standard deviation of the following scores: 27. 27, 25, 24, 20, 18, 16, 16, 14, 12, 10 & 7.
5., Calculate the range, variance, and standard deviation.
90. 86. 85, 80, 77, 75,71,65,60 & 58.

References

Best, J. W & Kahn, J. V. (1993). Measure of Spread or Dispersion. Research in Education, 7th Edition. Printed in the United States of America pp. 283-287

Joe, A. I. (1991) Measures of Variability. *Fundamental Statistics for education and the behavioral Sciences. pp* 50-69

Chapter 9

MEASURES OF ASSOCIATION

Babatunde Kasim OLADELE
Department of Science and Technology Education,
Faculty of Education, University of Johannesburg,
Auckland Park, 2006, Gauteng, South Africa,

9.1 Introduction

Measures of association are statistical tools that quantify the strength and direction of the relationship between two or more variables. It is a quantitative tool used to assess the strength and direction of the relationship between two or more variables [Britannica. (n.d.)]. These measures go beyond simply observing a connection and provide a numerical value to express how much one variable changes to another. These measures are crucial in statistics because they allow us to:

i. **Understand relationships:** They go beyond simply observing that two variables seem linked. They tell us how strong that connection is.
ii. **Predict trends:** By quantifying the association, we can make predictions about how one variable might change based on the values of another.
iii. **Compare findings:** They allow researchers to compare the strength of relationships across different studies or datasets.

9.2 Role in Quantifying Relationships Between Variables

Measures of association play a critical role in quantifying relationships between variables by providing a **numerical value** that captures both the **strength** and **direction** of that connection. Imagine you're investigating the relationship between study hours and exam scores. You might notice a general trend where students who study more tend to score higher. But simply observing this trend isn't enough. Measures of association come in to precisely quantify this connection.

Strength: They tell you how much of a change in one variable (study hours) is associated with a change in the other (exam scores). A strong association might indicate that even small increases in study time led to significant score improvements. Conversely, a weak association suggests little impact of study time on scores.

Direction: These measures not only show the strength but also the direction of the relationship. In our example, a positive association indicates scores increasing with study hours. A negative association would be very unusual, but it could suggest other factors like test anxiety impacting scores negatively despite studying more.

9.3 Distinction Between Association and Causation

Measures of association highlight a relationship between variables, but they do not necessarily prove that one **causes** the other. The explanation is as follows:

Correlation does not equal causation: Just because two things seem linked (association), it does not mean one directly influences the other (causation). Imagine finding an association between ice cream sales and shark attacks. While the numbers might correlate, it is highly unlikely ice cream causes shark attacks! There might be a hidden variable, like hot summer weather, influencing both.

Reverse causation: The relationship could even be reversed. Maybe people who go swimming more (higher chance of shark attacks) also buy more ice cream on hot days.

Confounding variables: Other unseen factors might be influencing both variables, creating a false appearance of association.

9.4 Types of Measures of Association

The types of measures of association are classified into two. We have measures classified under continuous variables and categorical variables. For the continuous variables, Pearson Correlation Coefficient and Spearman Rho correlation are used while for the categorical variables Chi-Square Test and Phi Coefficient (Φ) and Cramer's V. The measures are illustrated as follows:

For continuous variables we have

9.4.1 Pearson's Correlation Coefficient (r) is a statistical measure that quantifies the strength and direction of a linear relationship between two continuous variables. It ranges from -1 to 1, where: 1: Perfect positive correlation, 0: No correlation (variables are not linearly related), -1: Perfect, negative

correlation
Thus, this can be calculated mathematically as given:
Step 1. Prepare the data
There is a need for a dataset with values for both variables, X and Y.
Organize your data in a table with columns for X, Y, XY (product of X and Y), X^2 (X squared), and Y^2 (Y squared).

Step 2. Calculate intermediate values

Mean (average): Calculate the mean of X by adding all X values and dividing by the number of data points (n).
Do the same for Y to find the mean of Y.

Deviations from the mean:

For each data point, subtract the mean of X from the corresponding X value. This gives you $X - \bar{X}$ (X minus X-bar).
Do the same for Y, calculating $Y - \bar{Y}$ (Y minus Y-bar).

Product terms and squares:

Multiply the corresponding deviations for each data point $(X - \bar{X}) * (Y - \bar{Y})$ and put the product in the XY column.
Square each deviation $(X - \bar{X})^2$ and put it in the X^2 column. Do the same for Y, squaring deviations and placing them in the Y^2 column.

Step 3. Apply the formula

The formula for Pearson's correlation coefficient (r) is:
$r = \Sigma(XY) / \text{sqrt}(\Sigma(X^2) * \Sigma(Y^2))$
Σ (sigma) represents the sum of all the values in that column.
XY, X^2, and Y^2 refer to the respective columns in your table.

Step 4. Interpretation

The value of r tells you the strength and direction of the linear relationship.
Positive r indicates a positive correlation (as X increases, Y tends to increase).
Negative r indicates a negative correlation (as X increases, Y tends to decrease).
A value closer to 1 (positive or negative) indicates a stronger correlation.
A value closer to 0 indicates a weaker or no linear correlation.

Direction of the Correlation
Positive ($r > 0$): As one variable increases, the other tends to increase.
Negative ($r < 0$): As one variable increases, the other tends to decrease.

Examples:
r=0.8r=0.8 suggests a strong positive linear relationship.
r=−0.6r=−0.6 suggests a moderate negative linear relationship.
r=0r=0 indicates no linear relationship.

Limitations and Considerations
Assumes Linearity: Pearson's correlation coefficient assumes a linear relationship; it might not capture nonlinear associations.

Sensitive to Outliers: Outliers can strongly influence the correlation, leading to potential misinterpretation.

Does Not Imply Causation: Correlation does not imply causation; a strong correlation does not prove one variable causes the other.

Significance Testing: Conducting a hypothesis test on the correlation coefficient helps determine if the observed correlation is statistically significant or could have occurred by chance.

Pearson's correlation coefficient (r) measures the linear relationship between two variables, X and Y. It ranges from -1 (perfect negative correlation) to +1 (perfect positive correlation), with 0 indicating no linear relationship.

Illustration 1: calculate the correlation coefficient (Pearson's r) to measure the association between these two variables.

X	**Y**
3	70
5	85
2	60
4	75
6	90

Calculate the following values:
(\sum{X} = 20), (\sum{Y} = 380)
(\sum{XY} = 1400), (\sum{X^2} = 54), (\sum{Y^2} = 18100)
(n = 5)
Substitute these values into the formula for Pearson's correlation coefficient:
r=5(1400)−(20)(380)[5(54)−(20)2][5(18100)−(380)2]r=[5(54)−(20)2][5(18100)−(380)2]5(1400)−(20)(380)
First, let's calculate the values in the numerator and denominator:
Numerator:
5(1400)−(20)(380)=7000−7600=−6005(1400)−(20)(380)=7000−7600=−600
Denominator:
[5(54)−(20)2][5(18100)−(380)2][5(54)−(20)2][5(18100)−(380)2]
Now, let's calculate each part:
5(54)−(20)2=270−400=−1305(54)−(20)2=270−400=−130
5(18100)−(380)2=90500−144400=−539005(18100)−(380)2=90500−144400=−53900

So the denominator becomes:

(−130)(−53900)=7017000(−130)(−53900)=7017000

Now, substitute these values back into the original expression:

r=−6007017000*r*=7017000−600

Now, simplify the fraction:

r=−6007017000=−111695*r*=7017000−600=11695−1

So, the value of *r* is −11169511695−1.

r =0.00. this shows no relationship between x and y

9.4.2 Spearman's Rank Correlation Coefficient (rho)

Spearman's Rank Correlation Coefficient (rho), also known as Spearman's rho, is a non-parametric statistic used to measure the **monotonic relationship** between two ranked variables. It assesses how well the relationship between two variables can be described by a **monotonic function**, which means as one variable increases (or decreases), the other tends to increase (or decrease) as well, not necessarily in a perfectly linear way.

Key Points about Spearman's rho

Data Type: Applicable to ordinal or continuous data ranked from lowest to highest.

Value Range: -1 to +1. +1 indicates a perfect positive monotonic relationship (as one variable ranks higher, the other tends to rank higher as well). -1 indicates a perfect negative monotonic relationship (as one variable ranks higher, the other tends to rank lower). 0 indicates no monotonic relationship.

Interpretation: The closer rho is to ±1, the stronger the monotonic relationship. A value near 0 suggests little to no association.

Advantages

Doesn't require the data to be normally distributed (unlike Pearson's correlation).

Useful for identifying non-linear relationships.

Disadvantages:

Less powerful than Pearson's correlation for detecting linear relationships.

Ignores the magnitude of differences between ranks.

Applications

Spearman's rho is used in various fields to assess relationships between ranked variables. Here are some examples:

Social Sciences: Examining the association between student class rank and exam scores.

Psychology: Investigating the link between stress levels and anxiety ratings.

Economics: Analyzing the correlation between income inequality and crime rates

Thus, this can be calculated mathematically as given:

Step 1: Rank the Data

Begin by assigning ranks to your data points for both variables (X and Y) separately.

If there are ties within a variable (i.e., multiple data points with the same value), assign them the average rank for that position. For example, if two data points in Y have the same value and would occupy the 3rd and 4th positions, assign them both a rank of 3.5.

Step 2. Calculate the Difference in Ranks (d)

For each data pair, subtract the rank of one variable (Y) from the rank of the other variable (X).

So, d = X_rank - Y_rank.

Step 3. Square the Differences (d^2)
Square each difference in ranks (d) you calculated in step 2.

Step 4. Sum the Squared Differences (Σd^2)
Add up the squared differences (d^2) for all data pairs.

Step 5. Calculate the Number of Data Points (n)
Determine the total number of data pairs (n) you have in your analysis.

Step 6. Apply the Spearman's Rho Formula
The formula for Spearman's rho is:
$rho = 1 - (6\Sigma d^2) / (n(n^2 - 1))$
Where:
rho = Spearman's Rank Correlation Coefficient
Σd^2 = Sum of squared differences (from step 4)
n = Number of data pairs (from step 5)

Step 7. Interpret the Result
The resulting value of rho will range from -1 to +1.
+1 indicates a perfect positive monotonic relationship (higher ranks in X correspond to higher ranks in Y).
-1 indicates a perfect negative monotonic relationship (higher ranks in X correspond to lower ranks in Y). A value closer to 0 suggests a weaker or no monotonic relationship.

Illustration 2

The result of two raters shows that content does not correlate. The raters' score is as given in the table

S/No	Rater 1	Rater 2	di	Di^2
1	2	5	-3	9
2	3	1	2	4
3	1	3	-2	4
4	4	2	2	4
5	5	4	2	1
Type equation here.				$\sum di^2 = 22$

Using the fomular rho = $1 - (6\Sigma d^2) / (n(n^2 - 1))$

$r_s = 1\text{-}(6x22/5^2\text{-}5)$
$= 1\text{-}(132/125\text{-}5)$
$= 1\text{-}132/120$
$= 1\text{-}1.1$
$r_s = -0.1$ This shows that the relationship between the ratings is low and negative

For continuous variable

9.4.3 Chi-Square Test of Independence

The chi-square test is a statistical hypothesis test used to assess the **difference between observed data and expected data** in a contingency table. It helps determine whether the observed frequencies (number of observations in each category) are likely due to random chance or if there's a statistically significant association between the categorical variables being analyzed. The chi-square formula is: $\chi^2 = \Sigma(O_i - E_i)^2/E_i$, where O_i = observed value (actual value) and E_i = expected value.

Illustration 3: Calculate the Chi-square value for the following data of incidences of water-borne diseases in three tropical regions.

	India	Equador	South America	Total
Typhoid	31	14	45	90
Cholera	2	5	53	60
Diarrhea	53	45	2	100
	86	64	100	250

Solution:
Setting up the following table:

Observed	Expected	$O_i - E_i$	$(O_i - E_i)^2$	$(O_i - E_i)^2/Ei$
31	30.96	0.04	0.0016	0.0000516
14	23,04	9.04	81.72	3.546
45	36.00	9.00	81.00	2.25
2	20.64	18.64	347.45	16.83
5	15.36	10.36	107.33	6.99
53	24.00	29.00	841.00	35.04
53	34.40	18.60	345.96	10.06
45	25.60	19.40	376.36	14.70
2	40.00	38.00	1444.00	36.10

Answer: Chi-Square = 125.516

Example 2: What conclusion should be made with respect to an experiment when the significance level is 0.05 ($p = 0.05$)?

Solution:
Since the p-value of 0.068 is greater than 0.05, it would fail to reject the null hypothesis.

Key Points about Chi-Square Test

Null Hypothesis (H_0): Assumes there is no association between the variables, and the observed frequencies match the expected frequencies.
Data Type: Applicable to categorical data where variables are classified into groups or non-overlapping categories.

Alternative Hypothesis (H_1): There is a statistically significant association between the variables, and the observed frequencies differ from the expected frequencies.

Chi-Square Statistic (χ^2): Measures the discrepancy between observed and expected frequencies. Higher chi-square values indicate a larger difference.

Degrees of Freedom (df): Determined by the table dimensions (number of rows and columns minus 1). Used to find the p-value from a chi-square distribution table.

P-value: Represents the probability of observing a chi-square statistic this extreme or higher, assuming the null hypothesis is true. Lower p-values (typically less than 0.05) indicate that the observed differences are statistically significant, rejecting the null hypothesis.

Applications
The chi-square test is widely used in various fields to investigate relationships between categorical variables.

Social Sciences: Examining the association between education level and voting preference.

Marketing: Analyzing the relationship between customer age group and product choice.

Healthcare Research: Investigating the link between gender and a particular disease.

Important Considerations

The chi-square test has minimum expected frequency requirements for each category in the contingency table to ensure reliable results. It only tells you if there's an association, not the direction of the relationship (positive or negative). Software tools are readily available to calculate the chi-square statistic and p-value for easier interpretation.

Step 1. Setting Up the Calculation

- There is a need for a contingency table with categories for both variables to analyse.
- The table will show the **observed frequencies (Oi)**, which is the actual number of observations that fall within each category.
- There is a need for the **expected frequencies (Ei)**, which represent the number of observations you would expect to see in each category if there's no association between the variables. Typically, these are calculated based on the row and column totals of the observed data, assuming proportionality.

Step 2. Chi-Square Formula:

The chi-square statistic is calculated using the following formula:

$\chi^2 = \Sigma (Oi - Ei)^2 / Ei$

Σ (sigma) represents summation across all categories in the table.

- (Oi - Ei) represents the difference between the observed frequency (Oi) and the expected frequency (Ei) for each category. This difference is squared to account for both positive and negative discrepancies.

- $(Oi - Ei)^2 / Ei$ represents the contribution of each category to the overall chi-square statistic, weighted by the expected frequency.

Step 3. Interpreting the Result

- A higher chi-square statistic indicates a larger difference between the observed and expected frequencies, suggesting a weaker association between the variables (likely not due to random chance).
- The chi-square statistic itself doesn't tell you the direction of the association (positive or negative).
- To assess the significance of the chi-square statistic, you need to compare it to a chi-square distribution table with degrees of freedom (df) based on the table dimensions. Lower p-values (typically less than 0.05) indicate that the observed differences are statistically significant.

Step 4. Additional Notes

- There are minimum expected frequency requirements for the chi-square test to be reliable.
- Software tools can easily calculate the chi-square statistic and provide p-values for interpretation.

9.5 Phi Coefficient (Φ) and Cramer's V

When analysing associations between categorical variables, particularly in contingency tables, we frequently need to determine the strength of the correlation. Two often used measurements for this purpose are the phi coefficient (Φ) and Cramer's V.

i. Phi Coefficient

This applies to 2x2 contingency tables where both variables have just two categories.
Determines the strength and direction between two category variables.

On a scale of -1 to +1, +1 represents a perfect positive correlation, where one category in one variable always corresponds to a specific category in the other. -1 denotes a perfect negative relationship. The value of 0 implies that there is no relationship. Calculated using the chi-square statistic from the 2x2 table.

ii. **Cramer's V**

This is a more general measure that applies to contingency tables of any dimension (not just 2x2). It evaluates the strength of the relationship between categorical variables in the same way as the Phi coefficient does. Ranging from 0 to 1, 0 implies no relationship. 1 denotes a perfect relationship (equivalent to +1 or -1 in the phi coefficient). Calculated by modifying the chi-square statistic to reflect the table dimensions.

Key differences and use cases

i. Table Size: The phi coefficient is limited to 2x2 tables, whereas Cramer's V can be applied to any size contingency table. Interpretation: Both measures interpret the strength of the relationship similarly (0 to 1 or -1 to +1 for Phi). However, the particular criteria for evaluating intensity (weak, moderate, or strong) may differ depending on the source. Most analyses involving categorical variables commonly use Cramer's V due to its broad applicability. The phi coefficient is primarily utilised for the simpler situation of 2x2 tables.

9. 6 Regression Analysis

Regression analysis is a statistical method for determining the connection between one dependent variable (commonly indicated as Y) and one or more independent variables (typically written as X). The purpose of regression analysis is to determine how changes in the independent variables affect the dependent variable. Regression analysis is an effective method for determining and measuring relationships between variables. By fitting a regression model to the data, we can generate predictions, test hypotheses, and learn about the underlying processes that cause the observed events.

9.6.1 Simple Linear Regression

In simple linear regression, we look at the relationship between two variables: one independent variable (X) and one dependent variable (Y). The relationship between X and Y is considered to be linear, which means that each change in Y is proportional to the change in X plus some random error.

The simple linear regression model is stated as $y=0+1+Y=\beta 0+\beta 1X+\varepsilon$, with Y as the dependent variable. X represents the independent variable.
The intercept, $\beta 0$, represents the value of Y when X = 0.
The slope coefficient ($\beta 1$) represents the change in Y with a one-unit change in X.

The random error term () represents the difference between the observed and anticipated value of Y in the regression line.

Illustration 4
Simple Linear Regression

Suppose we have the following data on the number of study hours (independent variable) and exam scores (dependent variable) for a group of students:
Study Hours (X) Exam Score (Y)

X	Y
4	60
6	65
8	70
10	75
12	80

We want to perform simple linear regression to understand the relationship between study hours and exam scores.

1. **Calculate the Means**

Calculate the mean of study hours ($\bar{X}$) and the mean of exam scores ($\bar{Y}$).

$$\bar{X} = \frac{4 + 6 + 8 + 10 + 12}{5} = \frac{40}{5} = 8$$

$$\bar{Y} = \frac{60 + 65 + 70 + 75 + 80}{5} = \frac{350}{5} = 70$$

2. Calculate the Regression Coefficients

Calculate the slope (β_1) and the intercept (β_0).

$$\beta_1 = \frac{\sum{(X_i - \bar{X})(Y_i - \bar{Y})}}{\sum{(X_i - \bar{X})^2}}$$

$$\beta_0 = \bar{Y} - \beta_1\bar{X}$$

3. Substitute the Values

Use the calculated values to find the regression equation.

4. Interpretation

Interpret the slope coefficient. In this case, for each additional study hour, the exam score is expected to increase by the value of the slope coefficient.

9.6.2 Multiple Linear Regression

Multiple linear regression extends the standard linear regression model by including more than one independent variable. The multiple linear regression model is stated as y=0+11+22+…+Y=β0+β1X1+β2X2+…+βkXk+ε, with Y as the dependent variable.

The independent variables are 1, 2, X1, X2, and Xk.

The intercept is β0.

The slope coefficients for each independent variable are 1, 2, β1, β2, …, βk.

ε represents the random error term. $\hat{Y} = b_0 + b_1X_1 + b_2X_2$

In the multiple linear regression equation, b_1 is the estimated regression coefficient that quantifies the association between the risk factor X_1 and the outcome, adjusted for X_2 (b_2 is the estimated regression coefficient that quantifies the association between the potential confounder and the outcome). As noted earlier, some investigators assess confounding by assessing how much the regression coefficient associated with the risk factor (i.e., the measure of association) changes after adjusting for the potential confounder. In this case, we compare b_1 from the simple linear regression model to b_1 from the multiple linear regression model. As a rule of thumb, if the regression coefficient from the simple linear regression model changes by more than 10%, then X_2 is said to be a confounder.

Interpreting Regression Coefficients

The regression coefficients (slopes) indicate how much the dependent variable changes when the relevant independent variable changes by one unit, while keeping all other independent variables constant.

Assumptions for Regression Analysis

Regression analysis is based on various assumptions, including:

1. Linearity: The variables exhibit a linear connection.
2. Independence: Observations are independent of one another.
3. Homoscedasticity: The variance of error terms remains consistent across independent variable levels.
4. Normality: The error terms follow a normal distribution.

Illustration 5

A multiple regression equation can be used to estimate systolic blood pressure as a function of a participant's BMI, age, gender, and treatment for hypertension status. For example, we can estimate the blood pressure of a 50-year-old male, with a BMI of 25 who is not on treatment for hypertension as follows:

$$\hat{Y} = 68.15 + 0.58(25) + 0.65(50) + 0.94(1) + 6.66(0) = 116.1$$

We can estimate the blood pressure of a 50 year old female, with a BMI of 25 who is on treatment for hypertension as follows:

$$\hat{Y} = 68.15 + 0.58(25) + 0.65(50) + 0.94(0) + 6.44(1) = 121.59$$

Multiple regression analysis is a useful tool in a wide range of applications. From business, marketing, and sales analytics to environmental, medical, and technological applications, multiple regression analysis helps professionals evaluate diverse data that supports goals, processes, and outcomes in many industries.

Conclusion

It is critical to properly select the measure of association when the need arises. Confounding variables and effect modification are examples of factors that can have a major impact on data outcomes and interpretation. Researchers can ensure more accurate and trustworthy results by knowing these issues and using proper control procedures. Future research on measures of association should focus on developing novel strategies to solve these problems and increase the validity of study findings. However, the choice of measure depends on the type of data and research context.

Further Notes

1. **Covariance**:
   - Covariance is a measure of the association between pairs of variables in a population.
   - It quantifies how each variable is related to the other variable(s).
   - Positive covariance indicates that the two variables tend to increase together, while negative covariance suggests that one variable tends to decrease as the other increases.
   - The population covariance between variables (j) and (k) is defined as: [$\sigma_{jk} = E \left\{ (X_{ij} - \mu_j) (X_{ik} - \mu_k) \right\}$]
   - The sample covariance can be estimated using the formula: [$s_{jk} = \frac{1}{n-1} \sum_{i=1}^{n} (X_{ij} - \bar{x}_j) (X_{ik} - \bar{x}_k)$]

2. **Correlation**:
   - Correlation measures the strength and direction of the linear relationship between two variables.
   - The correlation coefficient ((r)) ranges from -1 to 1:
     - ($r > 0$): Positive correlation (variables move in the same direction).
     - ($r < 0$): Negative correlation (variables move in opposite directions).
     - ($r = 0$): No linear correlation.
   - The value of (r) indicates the strength of the linear relationship.
   - Commonly used correlation coefficients include Pearson's correlation coefficient and Spearman's rank correlation.

3. **Cramer's V**:
    - Useful for crosstabs with nominal-level variables.
    - Measures the strength of association between categorical variables.
    - Takes values b

References

Adegoke, B. A. (2016). Statistical Methods for Behavioural and Social Sciences Research. Third Edition Published by Able and Press Mokola Ibadan.

Encyclopedia Britannica Contributors 2024

Gemini Team Google: Rohan Anil et al. Gemini: A Family of Highly Capable Multimodal Models. Publication venue: arXiv preprint arXiv:2312.11805 (2023)

OpenAI. (2024). *ChatGPT* (3.5) [Large language model]. https://chat.openai.com

Chapter 10

COMPARING VARIABLES

Ogunsola, A. J. (Ph.D) & Oloniluyi, A. E. (Ph.D)
Department of Economics,
Ekiti State University, Ado-Ekiti, Nigeria
Email: akindele.ogunsola@eksu.edu.ng
Tel: +2348033756767

1. Introduction

Comparing variables is a fundamental aspect in the field of statistics. It helps in understanding relationships and differences among data sets. This chapter aims to equip students with the essential tools and techniques required to perform various types of comparisons between variables. These methods are crucial for interpreting data accurately and making informed decisions based on statistical evidence. Consequently, this chapter is divided into several key sections, each focusing on a specific method of comparing variables. We begin with an overview of the T-test, a widely used statistical test that compares the means of two groups to determine if they are significantly different from each other. This section will cover both the Paired Sample T-test, which is used for related samples, and the Independent T-test, which is used for comparing two independent groups. Following this, we delve into the Analysis of Variance (ANOVA), a powerful technique that extends the T-test to more than two groups. ANOVA helps in assessing whether there are any statistically significant differences between the means of three or more independent groups. We explore its various forms, including one-way ANOVA and two-way ANOVA, along with detailed examples and calculations.

Moreover, post-hoc tests are essential for further analysis after ANOVA, especially when the initial test indicates significant differences. These tests help in pinpointing exactly which groups differ from each other. We will discuss the most commonly used post-hoc tests and their applications. Finally, this chapter also covers multiple analysis methods, which involve comparing multiple variables simultaneously. Techniques such as multivariate analysis of variance (MANOVA) will be introduced to help students understand how to handle complex data sets with multiple dependent variables. The chapter concludes with a summary that recaps the key points and five revision questions designed to challenge students and reinforce their understanding of the concepts covered.

2. T-Test

The T-test is a powerful statistical tool used to determine whether there is a significant difference between the means of two groups. This test is essential for hypothesis testing, providing a way to infer whether observed differences in sample means reflect true differences in the population means. The T-test is particularly useful in situations involving small sample sizes and unknown population variances.

2.1 Purpose of the T-Test

The T-test is used to compare the means of two groups to determine if the observed differences are statistically significant. This helps statisticians and researchers discern whether the differences in sample means are likely due to random chance or represent actual differences in the populations from which the samples were drawn.

2.2 Types of T-Tests

There are three main types of T-tests:

- **One-Sample T-Test:** Compares the mean of a single sample to a known value or population mean.

- **Independent T-Test (Two-Sample T-Test):** Compares the means of two independent groups.
- **Paired Sample T-Test (Dependent T-Test):** Compares the means of two related groups, such as the same subjects measured before and after an intervention.

2.3 Assumptions of the T-Test

Before conducting a T-test, it is crucial to ensure that the following assumptions are met:

- **Independence:** The samples must be independent of each other (for independent T-tests) or consist of paired observations (for paired sample T-tests).
- **Normality:** The data should be approximately normally distributed. This assumption is particularly important for small sample sizes.
- **Homogeneity of Variance:** The variances of the two groups being compared should be equal (for independent T-tests).

2.4 Steps for Conducting a T-Test

The general steps for conducting a T-test are as follows:

- **State the Hypotheses:**
  Null hypothesis (H_0): There is no difference between the means ($\mu = \mu_2$).
  Alternative hypothesis (H_1): There is a difference between the means ($\mu_1 \neq \mu_2$).
- **Calculate the Test Statistic:** The formula and calculation steps vary depending on the type of T-test being conducted, which we will look into the subsequent sections.
- **Determine the Degrees of Freedom (df):** The degrees of freedom depend on the sample size(s) and the type of T-test.
- **Find the Critical Value:** Use the T-distribution table to find the critical value at the chosen significance level (α).

- **Make a Decision:** Compare the calculated T-value with the critical value. Reject the null hypothesis if the calculated T-value is greater than the critical value.

3. One-Sample T-Test

The One-Sample T-test compares the mean of a single sample to a known value or population mean. This test is used to determine if the sample mean is significantly different from the hypothesized population mean.

The formula for the One-Sample T-test is:

$$t = \frac{\bar{X} - \mu_0}{s/\sqrt{n}}$$

Where:

$\bar{X}$ is the sample mean.

μ_0 is the population mean.

s is the sample standard deviation: $\sqrt{\frac{\sum(X_i - \bar{X})^2}{n-1}}$

n is the sample size.

Example:

Suppose we have a sample of 10 students' test scores, and we want to test if the average score is significantly different from 75. The scores are: 70, 72, 68, 65, 74, 80, 78, 82, 85, and 90.

Calculate the sample mean:

$$\bar{X} = \frac{70 + 72 + 68 + 65 + 74 + 80 + 78 + 82 + 85 + 90}{10}$$

$$= 76.4$$

Calculate the sample standard deviation:

$$s = \sqrt{\frac{(70 - 76.4)^2 + (72 - 76.4)^2 + (68 - 76.4)^2 + \cdots + (90 - 76.4)^2}{10 - 1}}$$

$= 8.34$

Calculate the test statistic:

$$t = \frac{76.4 - 75}{8.34/\sqrt{10}} = \frac{1.4}{2.64} = 0.53$$

Determine the degrees of freedom:

$$df = n - 1 = 10 - 1 = 9$$

Find the critical value:
For $\alpha = 0.05$ and df = 9, the critical value from the T-distribution table is approximately 2.262.

Make a decision:
Since $0.53 < 2.262$, we fail to reject the null hypothesis (H_0) and conclude that there is no significant difference between the sample mean and the population mean of 75.

4. Paired Sample T-test

The Paired Sample -test, also known as the dependent T-test, is used to compare the means of two related groups. This test is particularly useful in studies where the same subjects are measured under different conditions or at different times. For example, it can be used to compare test scores before and after an educational intervention, or to measure changes in blood pressure before and after administering a new medication.

4.1 Purpose of the Paired Sample T-test

The Paired Sample T-test aims to determine whether there is a statistically significant difference between the means of two related groups. By comparing the differences within pairs rather than across independent groups, this test accounts for the inherent correlation between paired observations, increasing the test's sensitivity and statistical power.

4.2 Assumptions of the Paired Sample T-test

- **Paired Observations:** Each pair of observations must be matched or paired, meaning each data point in one group corresponds directly to a data point in the other group.

- **Normality:** The differences between paired observations should be approximately normally distributed. This assumption is particularly important for small sample sizes.
- **Independence within Pairs:** The pairs of observations must be independent of each other.

4.3 Steps for Conducting a Paired Sample T-test

- **State the Hypotheses:**

Null hypothesis (H_0): There is no difference between the means of the two related groups ($\mu_D = 0$).

Alternative hypothesis (H_1): There is a difference between the means of the two related groups ($\mu_D \neq 0$).

- **Calculate the Differences within Pairs:** Compute the difference (D_i) for each pair of observations.
- **Calculate the Mean and Standard Deviation of the Differences:** Compute the mean of the differences ($\overline{D}$).
- **Calculate the standard deviation of the differences (s_D).**
- **Calculate the Test Statistic:** Use the formula for the Paired Sample T-test.
- **Determine the Degrees of Freedom:** The degrees of freedom (df) is equal to the number of pairs minus one ($n-1$).
- **Find the Critical Value:** Use the T-distribution table to find the critical value at the chosen significance level (α).
- **Make a Decision:** Compare the calculated T-value with the critical value. Reject the null hypothesis if the calculated T-value is greater than the critical value.

The formula for the Paired Sample T-Test is:

$$t = \frac{\bar{D}}{s_D/\sqrt{n}}$$

Where:

$\bar{D}$ is the mean of the differences between paired observations.

s_D is the standard deviation of the differences: $\sqrt{\frac{\sum(D_i-\bar{D})^2}{n-1}}$

n is the number of pairs.

Example

Consider a study measuring the effectiveness of a new teaching method. The test scores of 8 students are measured before and after the intervention. The scores are as follows:

Before: 70, 72, 68, 65, 74, 80, 78, 82

After: 75, 78, 70, 68, 80, 85, 82, 88

Calculate the differences and their mean:

$$D = 75 - 70, 78 - 72, 70 - 68, 68 - 65, 80 - 74, 85 - 80, 82 - 78, 88 - 82$$

$$D = 5, 6, 2, 3, 6, 5, 4, 6$$

$$\bar{D} = \frac{5 + 6 + 2 + 3 + 6 + 5 + 4 + 6}{8} = 4.63$$

Calculate the standard deviation of the differences:

$$s_D = \sqrt{\frac{(5 - 4.63)^2 + (6 - 4.63)^2 + (2 - 4.63)^2 + \cdots + (6 - 4.63)^2}{8 - 1}}$$

$= 1.51$

Calculate the test statistic:

$$t = \frac{4.63}{1.51/\sqrt{8}} = \frac{4.63}{0.53} = 8.74$$

Determine the degrees of freedom:

$$df = n - 1 = 8 - 1 = 7$$

Find the critical value:
For $\alpha=0.05$ and $df=7$, the critical value from the T-distribution table is approximately 2.365.

Make a decision:
Since $8.74 > 2.365$, we reject the null hypothesis (H_0) and conclude that there is a significant difference between the before and after scores.

The Paired Sample T-test is a valuable tool for comparing the means of two related groups. By accounting for the correlation between paired observations, this test provides a more accurate assessment of differences. It is widely used in various fields, including psychology, medicine, education, and social sciences. For instance, it can be used to evaluate the effectiveness of a new treatment by comparing patients' health metrics before and after the treatment. In education, it helps assess the impact of teaching methods by comparing students' performance on pre-tests and post-tests. The versatility and reliability of the Paired Sample T-test make it an indispensable tool for researchers seeking to draw meaningful conclusions from their data.

5. Independent T-test

The independent t-test, also known as the two-sample t-test or unpaired t-test is used to compare the means of two independent groups. This test determines whether there is a statistically significant difference between the means of two groups that are not related. It is widely used in various fields to compare the effectiveness of different treatments, products, or conditions.

5.1 Purpose of the Independent T-test

The independent t-test aims to determine if there is a significant difference between the means of two independent groups. It is particularly useful when comparing two different groups of subjects, such as comparing the test scores of students from two different schools or the performance of two different brands of a product.

5.2 Assumptions of the Independent T-test

- **Independence:** The observations within each group must be independent of each other.
- **Normality:** The data in each group should be approximately normally distributed. This assumption is particularly important for small sample sizes.
- **Homogeneity of Variances:** The variances of the two groups should be approximately equal. This assumption can be tested using Levene's test for equality of variances.

5.3 Steps for Conducting an Independent T-test

- **State the Hypotheses:**

Null hypothesis (H_0): There is no difference between the means of the two independent groups ($\mu_1 = \mu_2$).
Alternative hypothesis (H_1): There is a difference between the means of the two independent groups ($\mu_1 \neq \mu_2$).

- **Calculate the Means and Standard Deviations:** Compute the mean and standard deviation for each group.
- **Calculate the Standard Error of the Difference between Means:** Use the formula for the standard error of the difference between means.
- **Calculate the Test Statistic:** Use the formula for the independent t-test.
- **Determine the Degrees of Freedom:** Calculate the degrees of freedom using the appropriate formula.
- **Find the Critical Value:** Use the T-distribution table to find the critical value at the chosen significance level (α).
- **Make a Decision:** Compare the calculated T-value with the critical value. Reject the null hypothesis if the calculated T-value is greater than the critical value.

The formula for the Independent T-test is:

$$t = \frac{\bar{X}_1 - \bar{X}_2}{\sqrt{\frac{s_1^2}{n_1} + \frac{s_2^2}{n_2}}}$$

Where:

$\bar{X}_1$and $\bar{X}_2$ are the means of the two groups.

s_1^2 and s_2^2 are the variances of the two groups ($s_1^2 = \frac{\sum X_1 - \bar{X}_1}{n_1 - 1}$ *and* $s_2^2 = \frac{\sum X_2 - \bar{X}_2}{n_2 - 1}$)

n_1 and n_2 are the sample sizes of the two groups.

The degrees of freedom (df) for the independent t-test can be calculated using the formula:

$$df = \frac{(\frac{s_1^2}{n_1} + \frac{s_2^2}{n_2})^2}{\frac{(\frac{s_1^2}{n_1})^2}{n_1 - 1} + \frac{(\frac{s_2^2}{n_2})^2}{n_2 - 1}}$$

Example

Consider a study comparing the test scores of students from two different schools. The test scores of 10 students from each school are as follows:

School A: 85, 90, 88, 92, 85, 89, 91, 87, 90, 88

School B: 78, 82, 80, 79, 81, 77, 85, 84, 83, 80

Calculate the means and standard deviations:

$$\bar{X}_1 = \frac{85 + 90 + 88 + 92 + 85 + 89 + 91 + 87 + 90 + 88}{10}$$

$$= 88.5$$

$$\bar{X}_2 = \frac{78 + 82 + 80 + 79 + 81 + 77 + 85 + 84 + 83 + 80}{10}$$

$$= 80.9$$

$$s_1^2 = \frac{(85 - 88.5)^2 + (90 - 88.5)^2 + (88 - 88.5)^2 + \cdots + (88 - 88.5)^2}{10 - 1}$$

$$= 5.39$$

$$s_2^2 = \frac{(78-80.9)^2+(82-80.9)^2+(80-80.9)^2+\cdots+(80-80.9)^2}{10-1}$$

$= 7.21$

Calculate the standard error of the difference between means:

$$SE = \sqrt{\frac{5.39}{10}+\frac{7.21}{10}} = \sqrt{0.539+0.721} = \sqrt{1.26} = 1.12$$

Calculate the test statistic:

$$t = \frac{88.5-80.9}{1.12} = \frac{7.6}{1.12} = 6.79$$

Determine the degrees of freedom:

$$df = \frac{(\frac{5.39}{10}+\frac{7.21}{10})^2}{\frac{(\frac{5.39}{10})^2}{10-1}+\frac{(\frac{7.21}{10})^2}{10-1}} = \frac{1.26^2}{\frac{0.054^2}{9}+\frac{0.072^2}{9}} = 17.83 \approx 18$$

Find the critical value:

For α = 0.05 and df = 18, the critical value from the T-distribution table is approximately 2.101.

Make a decision:

Since 6.79 > 2.101, we reject the null hypothesis (H0) and conclude that there is a significant difference between the test scores of the two schools.

The independent t-test is a powerful statistical method for comparing the means of two independent groups. It is widely used in various fields, including psychology, medicine, education, and social sciences. For example, it can be used to compare the effectiveness of two different medications, the performance of two different teaching methods, or the satisfaction levels of customers with two different products.

6. Analysis of Variance (ANOVA)

Analysis of Variance (ANOVA) is a statistical method that compares means across multiple groups to determine if at least one group's mean is significantly different. Unlike the T-test, which is limited to comparing two means, ANOVA can simultaneously evaluate differences across multiple groups, making it a versatile tool for analyzing data from complex experiments.

6.1 Purpose of ANOVA

ANOVA aims to determine whether there are any statistically significant differences between the means of three or more independent groups. It partitions the total variability in the data into variability between groups and within groups, allowing researchers to assess whether observed differences in group means are greater than what might be expected by chance alone.

6.2 Assumptions of ANOVA

- **Independence:** Observations within each group must be independent of each other.
- **Normality:** The data in each group should be approximately normally distributed.
- **Homogeneity of Variances:** The variances of the groups should be approximately equal. This assumption can be tested using Levene's Test for equality of variances.

6.3 Types of ANOVA

- **One-Way ANOVA:** Used when there is one independent variable with three or more levels (groups).
- **Two-Way ANOVA:** Used when there are two independent variables. This type can also assess interaction effects between the variables.

6.4 Steps for Conducting One-Way ANOVA

- **State the Hypotheses:**

  i. Null hypothesis (H_0): All group means are equal (μ1 = μ2 = μ3 = …).

  ii. Alternative hypothesis (H_1): At least one group mean is different.

- **Calculate the Group Means and Overall Mean:** Compute the mean for each group and the overall mean of all observations.
- **Calculate the Sum of Squares:**

  i. Total Sum of Squares (SST): Measures the total variability in the data.

  ii. Between-Group Sum of Squares (SSB): Measures the variability between group means.

  iii. Within-Group Sum of Squares (SSW): Measures the variability within each group.

- **Calculate the Mean Squares:**

  i. Mean Square Between (MSB): $MSB = \frac{SSB}{k-1}$

  ii. Mean Square Within (MSW): $MSW = \frac{SSW}{N-k}$

- **Calculate the F-Statistic:** Use the formula for the F-statistic.
- **Determine the Degrees of Freedom:**

  i. Between groups: $df_{between} = k - 1$

  ii. Within groups: $df_{within} = N - k$

- **Find the Critical Value:** Use the F-distribution table to find the critical value at the chosen significance level (α).
- **Make a Decision:** Compare the calculated F-value with the critical value. Reject the null hypothesis if the calculated F-value is greater than the critical value.

The formulae for ANOVA are:

i. **Sum of Squares Total (SST):** $SST = \sum_{i=1}^{N}(X_i - \bar{X})^2$

Where X_i is the individual observation, $\bar{X}$ is the overall mean, and N is the total number of observations.

ii. **Sum of Squares Between (SSB):** $SSB = \sum_{j=1}^{k} n_j\,(\bar{X}_j - \bar{X})^2$

Where $\bar{X}_j$ is the mean of group j, and n_j is the number of observations in group j.

iii. **Sum of Squares Within (SSW):** $SSW = \sum_{j=1}^{k}\sum_{i=1}^{n_j}(X_{ij} - \bar{X}_j)^2$

Where X_{ij} is an individual observation in group j.

iv. **Mean Squares:**

$$MSB = \frac{SSB}{k-1}$$

$$MSW = \frac{SSW}{N-k}$$

v. **F-Statistic:**

$$F = \frac{MSB}{MSW}$$

Example

Consider an experiment comparing the effects of three different teaching methods on students' test scores. The test scores for three groups of students are as follows:

Method A: 85, 90, 88, 92
Method B: 78, 82, 80, 79
Method C: 84, 86, 83, 85

Calculate the group means and overall mean:

$$\bar{X}_A = \frac{85 + 90 + 88 + 92}{4} = 88.75$$

$$\bar{X}_B = \frac{78 + 82 + 80 + 79}{4} = 79.75$$

$$\bar{X}_C = \frac{84 + 86 + 83 + 85}{4} = 84.5$$

$$\bar{X} = \frac{85 + 90 + 88 + 92 + 78 + 82 + 80 + 79 + 84 + 86 + 83 + 85}{12} = 84.5$$

Calculate the sum of squares:

$$SST = (85 - 84.5)^2 + (90 - 84.5)^2 + (88 - 84.5)^2 + \cdots + (85 - 84.5)^2 = 145.75$$

$$SSB = 4 \times (88.75 - 84.5)^2 + 4 \times (79.75 - 84.5)^2 + 4 \times (84.5 - 84.5)^2 = 186.75$$

$$SSW = (85 - 88.75)^2 + (90 - 88.75)^2 + (88 - 88.75)^2 + (92 - 88.75)^2 + (78 - 79.75)^2 + (82 - 79.75)^2 + (80 - 79.75)^2 + (79 - 79.75)^2 + (84 - 84.5)^2 + (86 - 84.5)^2 + (83 - 84.5)^2 + (85 - 84.5)^2 = 46.5$$

Calculate the mean squares:

$$MSB = \frac{SSB}{k - 1} = \frac{186.75}{3 - 1} = 93.375$$

$$MSW = \frac{SSW}{N - k} = \frac{46.5}{12 - 3} = 5.167$$

Calculate the F-statistic:

$$F = \frac{MSB}{MSW} = \frac{93.375}{5.167} = 18.07$$

Determine the degrees of freedom:

$$df_{between} = k - 1 = 3 - 1 = 2$$

$$df_{within} = N - k = 12 - 3 = 9$$

Find the critical value:
For α = 0.05, $df_{between} = 2$, and $df_{within} = 9$, the critical value from the F-distribution table is approximately 4.256.

Make a decision:
Since 18.07 > 4.256, we reject the null hypothesis (H_0) and conclude that there is a significant difference between the test scores of the three teaching methods.

6.5 Two-Way ANOVA

Two-way ANOVA is a statistical method used to determine the effect of two independent variables on a dependent variable and to explore the interaction between the two independent variables.

Example

Suppose a researcher wants to study the effects of teaching methods (Traditional, Online) and study time (2 Hours, 4 Hours) on students' test scores. The test scores for a sample of students are as follows:

Teaching Method	**2 Hours**	**4 Hours**
Traditional	70, 75, 72, 68	80, 85, 83, 82
online	65, 70, 68, 66	78, 80, 77, 79

Step 1: Calculate the means

First, calculate the mean for each combination of teaching method and study time.

$$Mean_{traditional,2\ Hours} = \frac{70 + 75 + 72 + 68}{4} = \frac{285}{4} = 71.25$$

$$Mean_{traditional,4\ Hours} = \frac{80 + 85 + 83 + 82}{4} = \frac{330}{4} = 82.5$$

$$Mean_{online,2\ Hours} = \frac{65 + 70 + 68 + 66}{4} = \frac{269}{4} = 67.25$$

$$Mean_{online,4\ Hours} = \frac{78 + 80 + 77 + 79}{4} = \frac{314}{4} = 78.5$$

Step 2: Calculate the main effects

Calculate the mean for each level of the independent variables (main effects).

Teaching method effect:

$$Mean_{traditional} = \frac{71.25 + 82.5}{2} = 76.875$$

$$Mean_{online} = \frac{67.25 + 78.5}{2} = 72.875$$

Study time effect:

$$Mean_{2\ Hours} = \frac{71.25 + 67.25}{2} = 69.25$$

$$Mean_{4\ Hours} = \frac{82.5 + 78.5}{2} = 80.5$$

Step 3: Calculate the Interaction Effect
The interaction effect is the difference between the means of each combination of the two independent variables.

Step 4: Calculate the sum of squares
Total Sum of Squares (SST):
The total sum of squares measures the total variation in the data.

$$SST = \sum (X_{ij} - Grand\ Mean)^2$$

Calculate the grand mean (overall mean of all scores):

$$= \frac{70 + 75 + 72 + 68 + 80 + 85 + 83 + 82 + 65 + 70 + 68 + 66 + 78 + 80 + 77 + 79}{16}$$

$$Grand = \frac{1228}{16} = 76.75$$

Now, calculate the SST:

$$SST = (70 - 76.75)^2 + (75 - 76.75)^2 + (72 - 76.75)^2 + \cdots (77 - 76.75)^2 + (79 - 76.75)^2$$

$$SST = 45.56 + 3.06 + 22.56 + \cdots + 0.06 + 5.06 = 572.75$$

Sum of Squares for Teaching Method (SSM):

$$SSM = n_t \sum (Mean_{method} - Grand\ Mean)^2$$

Where n_t is the number of observations per method.

$$SSM = 8 \times [(76.875 - 76.75)^2 + (72.875 - 76.75)^2]$$

$$SSM = 8 \times [0.0156 + 15.0625] = 8 \times 15.0781 = 120.625$$

Sum of Squares for Study Time (SSTime):

$$SSTime = n_s \sum (Mean_{Time} - Grand\ Mean)^2$$

$$SSTime = 8 \times [(69.25 - 76.75)^2 + (80.75 - 76.75)^2]$$

$$SSTime = 8 \times [56.25 + 14.0625] = 8 \times 70.3125 = 562.5$$

Sum of Squares for Interaction (SSI):
Interaction sum of squares:

$$SSI = \sum (Mean_{Combination} - Mean_{Method} - Mean_{Time} + Grand\ Mean)^2$$

Sum of Squares for Error (SSE):

$$SSE = SST - SSM - SSTime - SSI$$

Step 5: Calculate mean squares and f-statistics

$$MSM = \frac{SSM}{df_{Method}} \quad MS_{Time} = \frac{SS_{Time}}{df_{Time}} \quad MSI = \frac{SSI}{df_{Interaction}} \quad MSE = \frac{SSE}{df_{Error}}$$

Where the degrees of freedom (df) are calculated based on the number of levels in each factor.

$$F_{Method} = \frac{MSM}{MSE}, \quad F_{Time} = \frac{MS_{Time}}{MSE}, \quad F_{Interaction} = \frac{MSI}{MSE}$$

Finally, compare the calculated F-values with the critical F-values to determine significance.

ANOVA is a powerful statistical method for comparing the means of three or more independent groups. It can be used to compare the effectiveness of multiple treatments, the performance of different teaching methods, or the yield of various crop varieties. The versatility and reliability of ANOVA make it an indispensable tool for researchers seeking to draw meaningful conclusions from their data.

7. Post-hoc Tests

Post-hoc tests are statistical analyses conducted after an ANOVA when we find a significant difference among group means. They help determine which specific groups' means are significantly different from each other. Post-hoc tests are essential because while ANOVA can tell us that at least one group differs from the others, it does not specify which groups differ.

7.1 Purpose of Post-hoc Tests

The primary purpose of post-hoc tests is to make pairwise comparisons between group means while controlling for Type I error. Type I error occurs when we incorrectly reject a true null hypothesis (i.e., we find a difference where none exists). Post-hoc tests help ensure that the overall probability of making such an error remains at the chosen significance level (e.g., 0.05).

7.2 Common Post-hoc Tests

- **Tukey's Honest Significant Difference (HSD) Test:** Compares all possible pairs of means and is appropriate when sample sizes are equal.

$$HSD = q\sqrt{\frac{MSW}{n}}$$

Where q is the studentized range statistic, MSW is the mean square within groups, and n is the sample size per group.

- **Bonferroni Correction:** Adjusts the significance level for multiple comparisons by dividing the original alpha level by the number of comparisons.

Adjusted significance level:

$$\alpha_{adjusted} = \frac{\alpha}{number\ of\ comparisons}$$

- **Scheffé's Test:** More conservative and flexible, allowing for comparisons among any linear combination of group means.

$$F_{Scheffe} = \frac{(k-1).(M_j - M)^2}{MSW}$$

Where k is the number of groups, Mj is the mean of group j, and M is the overall mean.

- **Fisher's Least Significant Difference (LSD) Test:** Simplest and most liberal, but does not control well for Type I error when there are many comparisons.

$$t_{LSD} = \frac{(M_i - M_j)}{\sqrt{2 \times \frac{MSW}{n}}}$$

Where M_i and M_j are the means of groups i and j, respectively.

7.3 Steps for Conducting Post-hoc Tests

- **Perform ANOVA:** Ensure a significant F-statistic in the ANOVA before conducting post-hoc tests.
- **Choose a Post-hoc Test:** Select an appropriate post-hoc test based on the study design and data characteristics.
- **Calculate the Test Statistic:** Use the relevant formula for the chosen post-hoc test to calculate the test statistic for each pairwise comparison.
- **Compare with Critical Value:** Determine the critical value based on the chosen significance level and compare the test statistic to this value.
- **Interpret Results:** Identify which pairs of means are significantly different based on the test results.

Example

Consider the same dataset used in the one-way ANOVA example, comparing test scores across three teaching methods (A, B, and C):

Method A: 85, 90, 88, 92
Method B: 78, 82, 80, 79
Method C: 84, 86, 83, 85

After conducting ANOVA, for instance, assuming that we found a significant difference among the group means, we can go ahead to perform Tukey's HSD test to determine which pairs of means are significantly different.

Calculate HSD:

$$q = 3.77 \qquad (for\ \alpha = 0.05, k = 3, and\ df_{within} = 9)$$

$$MSW = 5.167$$

$$n-4$$

$$HSD = q\sqrt{\frac{MSW}{n}} = 3.77\sqrt{\frac{5.167}{4}} = 4.34$$

Compare Mean Differences:

$$|\bar{X}_A - \bar{X}_B| = |88.75 - 79.75| = 9.0$$
$$|\bar{X}_A - \bar{X}_C| = |88.75 - 84.5| = 4.25$$
$$|\bar{X}_B - \bar{X}_C| = |79.75 - 84.5| = 4.75$$

Interpret Results: Compare the mean differences to the HSD value (4.34).

$$|\bar{X}_A - \bar{X}_B| = 9.0 > 4.34\ (significant)$$
$$|\bar{X}_A - \bar{X}_B| = 4.25 < 4.34\ (not\ significant)$$
$$|\bar{X}_B - \bar{X}_C| = 4.75 > 4.34\ (significant)$$

From the post-hoc test, we conclude that the mean scores of Method A are significantly different from Method B and that the mean scores of Method B are significantly different from Method C. However, the mean scores of Method A are not significantly different from Method C.

Post-hoc tests are critical for identifying specific differences between group means following a significant ANOVA result. They are widely used in various research fields, including medicine, psychology, education, and agriculture. For instance, in clinical trials, post-hoc tests can compare the effectiveness of multiple treatments. In educational research, they can evaluate different teaching methods. The ability to identify specific group differences makes post-hoc tests an indispensable tool for researchers.

8. Multiple Analyses

Multiple analysis techniques are used when researchers want to analyze more than two variables simultaneously. These methods are crucial in understanding the relationships and interactions among multiple variables. In this section, we will explore the most common type of multiple analyses: Multiple Regression.

8.1 Multiple Regressions

Multiple regressions is a statistical technique that models the relationship between one dependent variable and two or more independent variables. It helps in predicting the dependent variable based on the values of the independent variables.

8.2 Purpose of Multiple Regressions:

- To predict the value of the dependent variable.
- To understand the relationship between the dependent and independent variables.
- To determine the strength and direction of these relationships.

The multiple regression equation is given by:

$$Y_0 = \beta_0 + \beta_1 X_1 + \beta_2 X_2 + \cdots + \beta_n X_n + \varepsilon$$

Where:

Y_0 is the dependent variable.

β_0 is the intercept/

$\beta_1, \beta_2, \ldots, \beta_n$are the coefficients of the independent variables $X_1, X_2, \ldots, X_n$

ε is the error term.

The true parameters estimates will be obtained by minimizing the sum of squared residuals:

$$\sum_{i=1}^{n} \varepsilon^2 = \sum_{i}^{n} (Y_i - \hat{Y}_1)^2 = \sum_{i}^{n} (Y_i - \hat{\beta}_0 - \hat{\beta}_1 X_{1i} - \hat{\beta}_2 X_{2i})^2$$

Performing the partial differentiations, we get the following systems of three normal equations in the three unknown parameters $\hat{\beta}_0, \hat{\beta}_1, and\ \hat{\beta}_2$

$$\sum Y_i = n\hat{\beta}_0 + \hat{\beta}_1 \sum X_{1i} + \hat{\beta}_2 \sum X_{2i}$$
$$\sum X_{1i}Y_i = \hat{\beta}_0 \sum X_{1i} + \hat{\beta}_1 \sum X_{1i}^2 + \hat{\beta}_2 \sum X_{1i}X_{2i}$$
$$\sum X_{2i}Y_i = \hat{\beta}_0 \sum X_{2i} + \hat{\beta}_1 \sum X_{1i}X_{2i} + \hat{\beta}_2 \sum X_{2i}^2$$

By formally deriving the two subsequent systems of normal equations above setting$\hat{\beta}_0 = 0$, we obtain values for the unknown parameters.

$$\hat{\beta}_0 = \bar{Y} - \hat{\beta}_1\bar{X}_1 - \hat{\beta}_2\bar{X}_2$$
$$\hat{\beta}_1 = \frac{(\sum x_{1i}y_i)(\sum x_{2i}^2) - (\sum x_{2i}y_i)(\sum x_{1i}x_{2i})}{(\sum x_{1i}^2)(\sum x_{2i}^2) - (\sum x_{1i}x_{2i})^2}$$
$$\hat{\beta}_2 = \frac{(\sum x_{2i}y_i)(\sum x_{1i}^2) - (\sum x_{1i}y_i)(\sum x_{1i}x_{2i})}{(\sum x_{1i}^2)(\sum x_{2i}^2) - (\sum x_{1i}x_{2i})^2}$$

Where

$y_i = Y_i - \bar{Y}$, $x_{1i} = X_{1i} - \bar{X}_1$, *and* $x_{2i} = X_{2i} - \bar{X}_2$

8.3 Coefficient of Multiple Determination

When the explanatory variables are more than one we talk of multiple correlation. The square of the correlation coefficient is called the *coefficient of multiple determination* or squared multiple correlation coefficient, and it is denoted by R^2. For a three-variable model, the coefficient of multiple determination is given as:

$$R^2_{Y.X_1X_2} = \frac{\sum \hat{y}^2}{\sum y^2} = \frac{\sum(\hat{Y} - \bar{Y})^2}{\sum(Y - \bar{Y})^2} = 1 - \frac{\sum \varepsilon^2}{\sum y^2} = \frac{\sum y^2 - \sum \varepsilon^2}{\sum y^2}$$

Since $\varepsilon_i = y_i - \hat{y}_i$ and $\hat{y}_i = \hat{\beta}_1 x_{1i} + \hat{\beta}_2 x_{2i}$, therefore;

$$R^2_{Y.X_1X_2} = \frac{\hat{\beta}_1 \sum y_i x_{1i} + \hat{\beta}_2 \sum y_i x_{2i}}{\sum y_i^2}$$

The value of R^2 lies between 0 and 1. The higher $R^2$the greater the percentage of the variation of Y explained by the regression plane, that is, the better the 'goodness of fit' of the regression plane to the sample observations.

8.4 The Mean and Variance of the Parameter Estimates

The estimates $\hat{\beta}_0, \hat{\beta}_1, \hat{\beta}_2$ are unbiased estimates of the true parameters of the relationship between Y, X_1, and X_2; their mean expected is the true parameter itself.

$$E(\hat{\beta}_0) = \beta_0 \qquad E(\hat{\beta}_1) = \beta_1 \qquad E(\hat{\beta}_2) = \beta_2$$

Consequently, the variances of the parameter estimates are obtained by the following formulae:

$$var(\hat{\beta}_0) = \hat{\sigma}_u^2[\frac{1}{n} + \frac{\bar{X}_1^2 \sum x_2^2 + \bar{X}_2^2 \sum x_1^2 - 2\bar{X}_1\bar{X}_2 \sum x_1x_2}{\sum x_1^2 \sum x_2^2 - (\sum x_1x_2)^2}]$$

$$var(\hat{\beta}_1) = \hat{\sigma}_u^2[\frac{\sum x_2^2}{\sum x_1^2 \sum x_2^2 - (\sum x_1x_2)^2}]$$

$$var(\hat{\beta}_2) = \hat{\sigma}_u^2[\frac{\sum x_1^2}{\sum x_1^2 \sum x_2^2 - (\sum x_1x_2)^2}]$$

Where;

$\hat{\sigma}_u^2 = \frac{\sum \varepsilon^2}{n-K}$, K is the total number of parameters that are estimated. In the three-variable model K = 3.

8.5 Test of Significance of the Parameter Estimates

The traditional test of significance of the parameter estimates is the standard error test, which is equivalent to the Student's t-test. Traditionally, researchers test the null hypothesis H_0: $\beta_i = 0$ for each parameter. Against the alternative hypothesis H_1: $\beta_i \neq 0$. This type of hypothesis implies a two-tail test at a chosen level of significance, usually at 5%.

Considering the Student's t-test, we compute the t ratio for each β_i

$$t^* = \frac{\hat{\beta}_i}{s(\hat{\beta}_i)}$$

Where;

$$s(\hat{\beta}_i) = \sqrt{var(\hat{\beta}_i)}$$

The above is the observed (or sample) value of the t ratio, which we compare with the theoretical value of the t obtainable from the t-table with n-K = n-3 degree of freedom. The theoretical values of t are the critical values that define the critical region in a two-tail test.

If t^* falls in the acceptance region; that is if $-t_{0.025} < t^* < t_{0.025}$ (with n-K degree of freedom), we accept the null hypothesis; that is, we accept that $\hat{\beta}_i$ is not significant (at the 5% significance level) and hence corresponding regressor does not appear to contribute to the explanation of the variations in Y. However, if t^* falls in the critical region we reject the null hypothesis, and accept the alternative one: $\hat{\beta}_i$ is statistically significant.

Example:

Suppose we want to predict the quantity demanded of a commodity (Y) by the consumer based on the price of the commodity (X1) and the consumer's income level (X2). The data for ten consumers are as follows:

Consumer (n)	**Price (X1)**	**Income (X2)**	**Quantity Demand (Y)**
1	5	1000	100
2	7	600	75
3	6	1200	80
4	6	500	70
5	8	300	50
6	7	400	65
7	5	1300	90
8	4	1100	100
9	3	1300	110
10	9	300	60

Solution:

N	Y	X1	X2	y_i	x_{1i}	x_{2i}	y_i^2	x_{1i}^2	x_{2i}^2	$y_i x_{1i}$	$y_i x_{2i}$	$x_{1i} x_{2i}$
1	100	5	1000	20	-1	200	400	1	40000	-20	4000	-200
2	75	7	600	-5	1	-200	25	1	40000	-5	1000	-200
3	80	6	1200	0	0	400	0	0	160000	0	0	0
4	70	6	500	-10	0	-300	100	0	90000	0	3000	0
5	50	8	300	-30	2	-500	900	4	250000	-60	15000	-1000
6	65	7	400	-15	1	-400	225	1	160000	-15	6000	-400
7	90	5	1300	10	-1	500	100	1	250000	-10	5000	-500
8	100	4	1100	20	-2	300	400	4	90000	-40	6000	-600
9	110	3	1300	30	-3	500	900	9	250000	-90	15000	-1500
10	60	9	300	-20	3	-500	400	9	250000	-60	10000	-1500
n = 10	$\sum Y_i = 800$	$\sum X_1 = 60$	$\sum X_2 = 8000$	$\sum y_i = 0$	$\sum x_{1i} = 0$	$\sum x_{2i} = 0$	$\sum y_i^2 = 3450$	$\sum x_{1i}^2 = 30$	$\sum x_{2i}^2 = 1580000$	$\sum y_i x_{1i} = -300$	$\sum y_i x_{2i} = 65000$	$\sum x_{1i} x_{2i} = -5900$

$$\bar{Y} = \frac{\sum Y_i}{n} = \frac{800}{10} = 80, \qquad \bar{X}_1 = \frac{\sum X_1}{n} = \frac{60}{10} = 60, \qquad \bar{X}_2 = \frac{\sum X_2}{n} = \frac{8000}{10} = 800$$

$$\hat{\beta}_1 = \frac{(-300)(1580000) - (65000)(-5900)}{(30)(1580000) - (-5900)^2} = \frac{-90500}{12590} = -7.19$$

$$\hat{\beta}_2 = \frac{(6500)(30) - (-300)(-5900)}{(30)(1580000) - (-5900)^2} = \frac{180}{12590} = 0.01$$

Therefore

$$\hat{\beta}_0 = \bar{Y} - \hat{\beta}_1\bar{X}_1 - \hat{\beta}_2\bar{X}_2$$

$$\hat{\beta}_0 = 80 - (-7.19) - (0.01)(800) = 111.69$$

The unadjusted coefficient of multiple determination may be computed by the below expression:

$$R^2_{Y.X_1X_2} = \frac{\hat{\beta}_1 \sum y_i x_{1i} + \hat{\beta}_2 \sum y_i x_{2i}}{\sum y_i^2}$$

$$R^2_{Y.X_1X_2} = \frac{(-7.19)(-300) + (0.01)(65000)}{3450} = 0.89$$

For the estimation of the standard errors of $\hat{\beta}_1$ and $\hat{\beta}_2$, we require an estimate of $\hat{\sigma}_u^2 = \frac{\sum \varepsilon^2}{n-K}$

Since$R^2_{Y.X_1X_2} = 1 - \frac{\sum \varepsilon^2}{\sum y^2}$, solving for $\sum \varepsilon^2$ we find:

$$\sum \varepsilon^2 = \sum y^2 (1 - R^2_{Y.X_1X_2})$$

Hence:

$$\sum \varepsilon^2 = 3450(1 - 0.89) = 365.7$$

Therefore:

$$\hat{\sigma}_u^2 = \frac{365.7}{10-3} = 52.24$$

The variance of $\hat{\beta}_0, \hat{\beta}_1, \hat{\beta}_2$ may be obtained by substituting the relevant terms in the variance formula above:

$$var(\hat{\beta}_1) = (52.24)\frac{1580000}{12590000} \approx 6.53$$

$$var(\hat{\beta}_2) = (52.24)\frac{30}{12590000} \approx 0.0001$$
$$var(\hat{\beta}_0) = 553.69$$

Therefore, the standard errors of the estimates $(s(\hat{\beta}_i) = \sqrt{var(\hat{\beta}_i)})$ are:

$$s(\hat{\beta}_1) = 2.55, \qquad s(\hat{\beta}_2) = 0.01, \qquad s(\hat{\beta}_0) = 23.5$$

The multiple regression results may now be presented in the following summary form

$$\hat{Y} = 111.7 - 7.19X_1 + 0.014X_2$$
$$S(\hat{\beta}_i) = (23.5)(2.55)(0.01) \qquad R^2 = 0.89$$
$$t^* = \frac{\hat{\beta}_i}{S(\hat{\beta}_i)} = (4.75)(-2.8)(1.28)t_{0.025} = 2.37$$

The variables X_1 and X_2 explain 89% of the total variation in Y. The estimates β_0 and β_1 are statistically significant but the estimate β_2 is not statistically significant at the 5% level.

9. Summary

The chapter provides a comprehensive exploration of key statistical techniques crucial for analyzing relationships between variables. Beginning with the T-test, we examined how to compare means between one group and two groups, distinguishing between paired and independent samples to accommodate different research scenarios. The ANOVA section expanded on this by introducing the analysis of variance, allowing for the comparison of means across multiple groups. Both one-way and two-way ANOVA were covered, offering insights into the effects of single and multiple factors on a dependent variable. Post-hoc tests were also discussed as essential tools for identifying specific group differences after finding a significant ANOVA result. The chapter further delved into multiple regression analysis, focusing on predicting an outcome based on multiple predictors, where the derivation of regression coefficients was demonstrated through algebraic solutions. Throughout, examples were emphasized to ensure a deep understanding of these techniques, equipping students with computational skills for real-world statistical analysis. The chapter concluded with challenging revision questions designed to reinforce key concepts and prepare students for exams, ensuring they can apply these statistical methods confidently in their studies and future research endeavours.

10. Revision Questions

To reinforce the concepts and techniques covered in this book, students should practice the below questions to equip themselves with the understanding of variable comparison:

1. A researcher wants to determine if there is a significant difference in test scores between two groups of students taught using different teaching methods. Group A has 10 students with scores [78, 82, 88, 92, 85, 79, 83, 89, 91, 84], and Group B has 10 students with scores [85, 89, 91, 87, 82, 86, 88, 90, 84, 83]. Conduct an independent sample t-test and interpret the results.

2. A fitness trainer wants to evaluate the effectiveness of a new workout regimen. The weights (in kg) of 8 participants are measured before and after the program. The before weights are [75, 80, 85, 70, 78, 82, 77, 79], and the after weights are [73, 78, 83, 68, 76, 80, 75, 77]. Perform a paired sample t-test to determine if the workout regimen significantly reduced the participants' weights.

3. A researcher studies the effect of three different diets on weight loss. The weight loss (in kg) of participants after following Diet A, Diet B, and Diet C for a month are recorded as follows:
   Diet A: [3.2, 4.1, 3.8, 4.0, 3.5]
   Diet B: [2.8, 3.0, 3.1, 2.9, 3.2]
   Diet C: [4.5, 4.8, 4.3, 4.6, 4.7]

Conduct a one-way ANOVA and interpret the results to determine if there are significant differences among the diets.

4. After conducting an ANOVA and finding significant differences among the three teaching methods on student performance, a post-hoc test is required to identify which specific groups differ. The student performance scores for the three methods are as follows:
   Method A: [85, 88, 90, 92, 87]
   Method B: [78, 80, 82, 79, 77]
   Method C: [91, 93, 95, 94, 92]

   Conduct Tukey's HSD post-hoc test to determine which pairs of teaching methods show significant differences in student performance and interpret the results.

5. A company wants to predict employee productivity based on hours of training (X1) and years of experience (X2). The productivity scores (Y) for five employees, along with their hours of training and years of experience, are as follows:

Employee	**Training Hours (X1)**	**Experience Years (X2)**	**Productivity Score (Y)**
1	10	5	75
2	12	7	85
3	15	10	95
4	9	4	70
5	14	8	90

Perform a multiple regression analysis and other parameter estimates such as coefficient of determination, variances, standard errors, and Student's t-test to predict productivity and interpret the results.

References:

Field, Andy. "Discovering statistics using IBM SPSS statistics." (2013): 115-21.

Gujarati, Damodar N., and Dawn C. Porter. *Basic econometrics*. McGraw-hill, 2009.

Koutsoyiannis, Anna. "Theory of econometrics: an introductory exposition of econometric methods." (1977).

Montgomery, Douglas C. *Design and analysis of experiments*. John wiley& sons, 2017.

Moore, David S., and George P. McCabe. *Introduction to the practice of statistics*. WH Freeman/Times Books/Henry Holt & Co, 1989.

Ott, R. L., and Michael Longnecker. *An introduction to statistical methods and data analysis*. Cengage Learning, 2016.

Tabachnick, B. G., L. S. Fidell, and J. B. Ullman. "Using multivariate statistics (Vol. 6, pp 497–516)." (2019).

Wooldridge, Jeffrey M. *Introductory Econometrics: A Modern Approach. Supplement*. Cengage Learning, 2016.

Chapter 11

COMPUTER APPLICATIONS FOR STATISTICS IN EDUCATION

[1]ANGYU, Lahiru Daniel (PHD)
Department of Primary Education,
School of Early Childhood Care Education and Primary Education,
College of Education, Zing, Taraba State, Nigeria
Contact Info: lahirudaniel02@gmail.com, +2348169506240, +2348028065981

[2]BAMBI, Babatunde Ishola
Academic Planning Unit,
West Midlands Open University, Lagos
OrcidID: https://orcid.org/0009-0009-2236-8715
E-mail: bambi.babatunde@westmidlands.university
**Corresponding author contact number: +2348063462190*

[3]GIDADO, Farida Sambo
Department of Business Education,
Federal College of Education, Yola
E-mail: faridagidadosambo@gmail.com, +2348038406635

1.0 Introduction

The integration of computer technology into statistical analysis has fundamentally transformed the field of research across various disciplines, including educational research. Historically, statistical analysis was conducted manually or with the aid of rudimentary computational tools, which limited the complexity and scope of analyses (StataCorp, 2023). However, the advent of advanced computer software has significantly enhanced researchers' ability to manage, analyse, and interpret large datasets with greater precision and efficiency (Field, 2018). Computers facilitate the execution of complex statistical methods that would be impossible to perform manually. Traditionally, statistical analysis involved manual calculations using formulas and tables, which was time-consuming and prone to errors (Anderson et al., 2019). However, with the advent of computers, statistical methods have become more accessible and reliable. Statistical software such as SPSS, R, SAS, and Python libraries have facilitated the automation of statistical processes, enabling researchers to perform tasks such as data entry, hypothesis testing, regression analysis, and multivariate analysis more efficiently (Alessi & Trollip, 2017). As a result, the use of statistical software has become an essential tool for students, researchers, and professionals alike.

One of the significant advantages of applying computers in statistics is the capacity to handle large and complex datasets (Jones, 2021). As datasets continue to grow in size due to advancements in technology and data collection methods, manual statistical techniques have become insufficient. Computational tools not only provide the ability to process massive datasets but also ensure precision in statistical calculations. This precision is particularly crucial in fields such as biomedical research, where accurate data interpretation can impact clinical decisions and patient outcomes (Chang, 2020). Furthermore, computers enable real-time data analysis and visualization, which is essential for making informed decisions quickly. This ability to process and visualize data instantaneously enhances the interpretability of statistical results, providing stakeholders with the insights needed for effective decision-making (Johnston & Ghosh, 2018).

The integration of computer-aided statistics in education has become increasingly vital. Researchers have found that computer-based statistics improve students' understanding of statistical concepts by offering interactive and visual learning experiences (Zhang & Owen, 2016). For example, simulation-based learning, supported by statistical software, allows students to better understand probability distributions and sampling methods through experimentation. Instructors, on the other hand, can use computers to demonstrate statistical concepts more dynamically, fostering a more engaging learning environment. Despite its many benefits, the use of computers in statistical analysis also presents certain challenges. One significant concern is the steep learning curve associated with mastering advanced statistical software (Nguyen, 2019). Many students and professionals may require additional training to effectively utilize these tools. Additionally, there is a need for continuous software updates and maintenance to ensure accuracy and functionality, which can be resource-intensive (Sharma & Tandon, 2020).

As technology continues to advance, the future of computer-aided statistics lies in the integration of artificial intelligence (AI) and machine learning (ML) algorithms. These technologies will further enhance data analysis capabilities by automating complex statistical procedures, predicting trends, and offering deeper insights from datasets (Smith & Johnson, 2021). Moreover, cloud-based platforms for statistical analysis are becoming more prevalent, offering greater accessibility and collaboration opportunities for researchers worldwide. The use of computer software in statistics has thus become an indispensable part of modern research practice. This is because it empowers researchers to handle vast amounts of data, apply advanced statistical techniques, and produce detailed and accurate reports. This advancement not only accelerates the research process but also elevates the quality and reliability of research findings, thereby contributing to the broader knowledge base in educational research and beyond. Hence, this chapter looks at the various aspects of applying computers for statistics in educational research to highlight the strengths and limitations of key computer software used for statistics to aid researchers in selecting the most appropriate software for their specific research needs.

2.0 CONCEPTUAL CLARIFICATION

2.1 Computers in Educational Research

Computers have transformed the landscape of educational research, facilitating more efficient data collection, analysis, and dissemination of research findings. The advent of computational technologies has enabled researchers to process vast amounts of data, automate repetitive tasks, and access sophisticated statistical tools that were previously unavailable through manual methods. Computers play a critical role in the digitization of educational research, allowing researchers to store data securely, analyse it more rapidly, and create simulations to model complex educational phenomena (Alessi & Trollip, 2017).

One of the major contributions of computers in educational research is the advancement of data analysis techniques. Researchers now have access to software packages such as SPSS, R, SAS, and Python, which offer a variety of statistical tools for both qualitative and quantitative analysis (Nguyen, 2019). These software tools facilitate data cleaning, hypothesis testing, and advanced statistical modeling, enabling researchers to uncover patterns and trends in educational datasets that would be difficult to detect manually. Moreover, computers support the development of user-friendly interfaces that help educators, policymakers, and researchers visualize data more effectively, thus enhancing the interpretability of results and making research findings more accessible (Johnston & Ghosh, 2018).

In addition, computers have improved the accessibility and sharing of educational research findings. Researchers can now publish their work online and make their datasets available to the broader academic community, fostering collaboration and advancing the field (Zhang & Owen, 2016). Furthermore, computers facilitate the use of online surveys, educational assessments, and other digital tools for data collection, making it easier to gather large samples and ensuring more reliable and timely results (Jones, 2021).

2.2 Concept of Statistics in Educational Research

Statistics is an essential component of educational research as it provides the methodology for designing studies, analysing data, and making inferences about educational processes and outcomes. Statistical methods allow researchers to test hypotheses, determine relationships between variables, and assess the effectiveness of educational interventions (Sharma & Tandon, 2020). In the context of educational research, statistics are used to analyse data from experiments, surveys, and assessments, helping to evaluate teaching methods, learning outcomes, and educational policies.

The use of statistics in educational research often involves descriptive and inferential methods. Descriptive statistics summarize and organize data, providing insights into central tendencies and variability within datasets. These methods include calculating means, medians, standard deviations, and frequency distributions (Anderson et al., 2019). Inferential statistics, on the other hand, allow researchers to draw conclusions about populations based on samples, using techniques such as hypothesis testing, regression analysis, and analysis of variance (Chang, 2020). In addition to traditional statistical methods, educational researchers are increasingly using advanced statistical techniques such as multivariate analysis, structural equation modelling (SEM), and hierarchical linear modelling (HLM) to explore more complex educational phenomena (Smith & Johnson, 2021). These methods provide deeper insights into the relationships among multiple variables, offering more nuanced interpretations of how various factors influence educational outcomes.

2.3 The Intersection of Computers and Statistics in Educational Research

The integration of computers and statistical methods in educational research has been particularly transformative. By leveraging the power of computers, researchers can now conduct more sophisticated statistical analyses, handle larger datasets, and reduce the risk of human error during data analysis. Computers have made it easier to perform statistical tests and generate outputs quickly, enabling researchers to focus on interpreting their results rather than being burdened by the computational aspects of data analysis (Jones, 2021). Moreover, the development of user-friendly statistical software has democratized access to advanced statistical techniques.

Even researchers with limited expertise in statistics can now perform complex analyses with the assistance of automated processes and built-in tutorials available in many statistical software packages (Nguyen, 2019). As a result, computers have not only accelerated the pace of research but have also made high-quality statistical analysis more accessible to educators, students, and policymakers. In conclusion, computers and statistics have become indispensable tools in educational research. The combination of computational power and statistical techniques has enabled more rigorous analysis of educational data, fostering a better understanding of teaching, learning, and educational outcomes. As technology continues to advance, the role of computers in educational statistics is expected to grow, offering even more powerful tools for research and analysis.

3.0 USING MICROSOFT EXCEL FOR STATISTICS

Microsoft Excel is a widely used tool in educational research for performing basic statistical analyses and managing data. Its accessibility, user-friendly interface, and integration with other Microsoft Office applications make it a popular choice for researchers who may not have access to more advanced statistical software. While Excel is not designed specifically as a statistical package, it offers a range of functions and tools that can be effectively utilized for statistical analysis in educational research. Excel is commonly available in most academic institutions and does not require specialized training to use. This makes it an ideal tool for educators, students, and researchers who may not have extensive backgrounds in statistics or access to more complex software such as SPSS, R, or SAS (Stark & Fink, 2019). Excel's grid-like structure allows for easy organization and manipulation of data, and its simple interface makes it easy to perform tasks such as data entry, sorting, and filtering.

One of Excel's key strengths lies in its ability to create visual representations of data through charts and graphs. In educational research, visualizations such as histograms, bar charts, pie charts, and scatterplots are valuable tools for communicating research findings clearly and concisely (Gavin, 2020). Excel's charting tools are easy to use, and they allow researchers to customize their visualizations to suit their specific needs. By representing statistical data visually, researchers can identify patterns, trends, and relationships that may not be immediately evident through numerical analysis alone. However, while Excel is useful for basic statistical analysis, it has limitations that may restrict its use in more complex educational research. The significant limitations of Excel include:

1. The lack of advanced statistical techniques that are available in specialized software like SPSS, R, or SAS (Pryor, 2017). For example, Excel does not offer support for multivariate analysis, or structural equation modelling (SEM), which are commonly used in large-scale educational research projects (Crawley, 2017).
2. Excel's statistical functions are not as robust or reliable as those in dedicated statistical software, which could lead to inaccuracies in more complex analyses (Bhandari, 2020).
3. Excel's handling of large datasets is not as efficient as other statistical software. For educational research

involving extensive data sets, such as longitudinal studies or nationwide assessments.

4. Excel may slow down or become less responsive, making it difficult to manage and analyse large volumes of data (Johnston & Ghosh, 2018).

Despite the limitations of Microsoft Excel, it is a versatile and accessible tool that is widely used in educational research for statistical analysis and data visualization. Although it may not be as advanced as dedicated statistical software like SPSS or R, Excel's functions and capabilities can still be highly effective for basic to intermediate statistical analyses. Below is a step-by-step guide on how to use Microsoft Excel for statistics in educational research, focusing on data preparation, analysis, and visualization.

Step 1: Data Entry and Organization

The first step in using Excel for educational research statistics is to input and organize your data properly. In Excel, data are entered into rows and columns, and proper organization is key for efficient analysis.

- **Data entry**: Input your data in columns, with each column representing a variable and each row representing a data point (e.g., student or school). Ensure that column headers are clearly labelled with the variable names, such as "Student ID," "Test Score," "Gender," and "Grade Level" (Walkenbach, 2019).
- **Check for errors**: Before proceeding with any analysis, it is important to review your data for any missing or erroneous entries. You can use Excel's *Find and Replace* function to locate blank cells or common errors.

Step 2: Descriptive Statistics

After organizing the data, the next step is to calculate basic descriptive statistics such as the mean, median, mode, variance, and standard deviation. Excel provides several functions that make this process simple.

- **Mean (Average)**: To calculate the average, use the formula =AVERAGE(range), where "range" refers to the cell range of the data. For example, =AVERAGE(B2:B101) calculates the average of test scores in column B.
- **Median**: To calculate the median, use the formula =MEDIAN(range).
- **Mode**: To find the mode (most frequent value), use =MODE.SNGL(range).
- **Standard Deviation**: The standard deviation can be calculated using =STDEV.S(range) for a sample or =STDEV.P(range) for a population (Berk & Carey, 2018).

Step 3: Data Visualization

Data visualization is essential for communicating findings in educational research. Excel offers a variety of chart types that can help you visualize your data effectively.

- **Creating a chart**: Highlight the relevant data range and select *Insert > Chart* from the Excel toolbar. Depending on your data, you might choose from bar charts, histograms, scatter plots, or pie charts.
- **Customizing the chart**: After creating a chart, you can customize it by adding labels, changing colours, and formatting the axes. For example, in a scatter plot, you might label the x-axis as "Study Hours" and the y-axis as "Test Scores" to demonstrate the relationship between time spent studying and academic performance (Walkenbach, 2019).

Step 4: Inferential Statistics

Inferential statistics are used to make generalizations from your sample to the larger population. Excel's *Data Analysis Toolpak* provides various statistical tools, including t-tests, ANOVA, correlation, and regression analysis.

- **Installing the Data Analysis Toolpak**: If not already installed, go to *File > Options > Add-ins*. In the *Manage* box, select *Excel Add-ins*, click *Go*, and check *Analysis Toolpak* before clicking *OK* (Berk & Carey, 2018).
- **Correlation**: To calculate the correlation between two variables, go to *Data > Data Analysis > Correlation*, select your data range, and specify whether your data is grouped by columns or rows.
- **T-tests**: To compare the means of two groups, go to *Data > Data Analysis > t-Test*. You can choose between a two-sample t-test for equal or unequal variances or a paired t-test depending on your study design. For example, a paired t-test can be used to compare pre-test and post-test scores from the same group of students (Sullivan, 2017).
- **ANOVA**: If you want to compare the means of more than two groups, go to *Data > Data Analysis > ANOVA: Single Factor*. This is useful when comparing different educational interventions across multiple schools or classrooms (Sullivan, 2017).

Step 5: Regression Analysis

Regression analysis is a powerful statistical technique used in educational research to examine the relationship between one or more independent variables and a dependent variable. Microsoft Excel includes tools for performing both linear and multiple regression analysis.

- **Linear regression**: To perform a linear regression, go to *Data > Data Analysis > Regression*. Specify the dependent variable (e.g., test scores) and the independent variable (e.g., hours of study). Excel will output the regression equation, R-squared value, and significance levels (Walkenbach, 2019).

- **Multiple regression**: In multiple regression, more than one independent variable is used to predict the dependent variable. You follow the same procedure as for linear regression but select multiple independent variables. This is useful for modelling complex relationships, such as how factors like socioeconomic status, classroom size, and parental involvement predict academic achievement (Sullivan, 2017).

Step 6: Hypothesis Testing

Excel can also be used to conduct hypothesis tests, such as z-tests or chi-square tests. These tests allow researchers to determine whether there is a statistically significant difference between groups or associations between categorical variables.

- **Z-test**: For large sample sizes, you can use Excel's z-test function by selecting *Data Analysis > z-Test: Two Sample for Means*. This is useful for comparing the mean scores of two large groups of students (Berk & Carey, 2018).
- **Chi-square test**: To test for associations between categorical variables, use *Data Analysis > Chi-Square Test*. For example, you might use this test to determine whether there is a significant association between gender and participation in extracurricular activities.

Step 7: Interpreting Results

Once the analysis is complete, it is important to interpret the results in the context of your research question and the broader educational landscape. Excel will provide various outputs, including p-values, confidence intervals, and effect sizes. As a general rule, a p-value below 0.05 indicates a statistically significant result, meaning that the observed effect is unlikely to have occurred by chance (Sullivan, 2017). For example, in a t-test comparing the test scores of two groups, if the p-value is below 0.05, you can reject the null hypothesis and conclude that there is a significant difference between the groups.

Step 8: Reporting and Presenting Findings

Finally, once you have analysed your data and interpreted the results, the next step is to report your findings. Excel's charting and formatting tools make it easy to create tables and figures for reports, presentations, or publications.

- **Exporting data and charts**: You can copy Excel charts and tables directly into Microsoft Word or PowerPoint, allowing for seamless integration of statistical results into your reports.
- **Formatting for publication**: When preparing tables and figures for academic publication, ensure they meet the specific guidelines for your discipline. For example, APA 7th edition formatting requires tables to have clear titles, column headings, and notes explaining any abbreviations or significant values (American Psychological Association, 2020).

4.0 USING SPSS FOR STATISTICS

SPSS (Statistical Package for the Social Sciences) is one of the most widely used statistical software programs in educational research. Developed by IBM, SPSS is known for its user-friendly interface, extensive range of statistical tools, and capabilities in both data analysis and data management (Field, 2017). It is particularly favoured in educational research for its ease of use, making it accessible to researchers and students who may not have extensive backgrounds in statistics or programming. SPSS is frequently used for analysing survey data, experimental data, and observational studies, making it a versatile tool for educational research. One of the primary reasons SPSS is popular in educational research is its intuitive, point-and-click interface, which simplifies data analysis for users who might not be familiar with coding or complex statistical procedures.

Researchers can easily import data, run analyses, and interpret results without needing to write scripts, making it an ideal tool for those who want to focus on their research questions rather than the intricacies of programming (George & Mallery, 2019). This ease of use makes SPSS accessible to a wide audience, including educators, students, and policymakers.

The statistical package also provides step-by-step guides and tutorials for beginners, ensuring that researchers can perform statistical analyses such as t-tests, ANOVA, regression analysis, and chi-square tests with minimal effort (Pallant, 2020). This accessibility is particularly important in educational research, where the ability to analyse data quickly and efficiently can support timely decision-making and policy formulation. SPSS provides powerful tools for data management, allowing researchers to clean, organize, and manipulate their datasets. Data management is critical in educational research, as researchers often work with large datasets collected from surveys, standardized tests, or longitudinal studies (George & Mallery, 2019). SPSS allows users to merge datasets, recode variables, and create new variables based on existing data. These features help ensure the accuracy and integrity of the data, which is essential for producing valid research findings.

Additionally, SPSS supports a wide variety of data formats, making it easy to import and export data from Excel, CSV files, or other statistical software packages. This interoperability is beneficial for researchers who need to work with multiple data sources or collaborate with colleagues using different software (Pallant, 2020). SPSS is widely used in various types of educational research, including survey research, experimental research, and observational studies. For example, researchers conducting surveys on student attitudes or teacher practices can use SPSS to analyse the results, identify patterns, and test hypotheses (Field, 2017). In experimental research, SPSS is used to analyse pre-test and post-test data, compare control and experimental groups, and determine the effectiveness of educational interventions.

While SPSS is highly effective for many types of statistical analyses, it does have some limitations. One limitation is that it is not as flexible as open-source software like R, which allows users to write custom scripts and create unique analysis workflows (Stark & Fink, 2019). SPSS is also limited in its ability to handle very large datasets, as it may slow down or become less responsive with extremely large amounts of data (Wagner et al., 2019). Additionally, the cost of SPSS licenses can be prohibitive for some educational institutions or individual researchers, especially when compared to free alternatives such as R or Python. Despite these limitations, SPSS is a widely used statistical software that offers a range of powerful tools for data analysis, including descriptive statistics, inferential statistics, and complex modelling. This study provides a step-by-step approach for using SPSS for educational research, focusing on data preparation, analysis, and interpretation.

Step 1: Data Entry and Organization

1. **Opening SPSS and Creating a New Dataset**: Launch SPSS and select “New Dataset” to start a new project. Alternatively, you can open an existing dataset by selecting “Open” from the File menu.
2. **Entering Data:** In the *Data View* tab, enter your data directly into the spreadsheet. Each row represents a case (e.g., a student), and each column represents a variable (e.g., test score, gender, grade level) (Pallant, 2020).
3. **Defining Variables**: Switch to the *Variable View* tab to define your variables. Here, you can specify variable names, types, and labels. For example, you can set "Test Score" as a numeric variable and "Gender" as a nominal variable. Define the measurement level (nominal, ordinal, scale) for each variable to ensure accurate analysis (Field, 2018).
4. **Handling Missing Data**: Address missing data by using options under *Data > Missing Values*. You can specify

how SPSS should treat missing values, such as using mean substitution or excluding cases listwise.

Step 2: Descriptive Statistics

1. **Calculating Descriptive Statistics**: Go to *Analyse > Descriptive Statistics > Descriptives*. Select the variables for which you want to compute statistics such as mean, standard deviation, minimum, and maximum. Click "OK" to generate the output (Field, 2018).
2. **Creating Frequency Tables**: To create frequency tables, go to *Analyse > Descriptive Statistics > Frequencies*. Select categorical variables (e.g., gender, school type) and choose options for displaying frequency distributions and percentages (Pallant, 2020).
3. **Generating Cross-tabulations**: For examining relationships between categorical variables, use *Analyse > Descriptive Statistics > Crosstabs*. This feature allows you to explore the interaction between variables, such as the relationship between gender and participation in extracurricular activities (Field, 2018).

Step 3: Data Visualization

1. **Creating Charts and Graphs**: To visualize your data, go to *Graphs > Chart Builder*. You can create various types of charts including histograms, bar charts, and scatterplots. For instance, a histogram can display the distribution of test scores, while a scatterplot can illustrate the relationship between study hours and academic performance (Pallant, 2020).
2. **Customizing Charts**: After creating a chart, you can customize it by modifying titles, labels, and colours. Double-click on the chart to open the *Chart Editor* where you can make adjustments to better convey your research findings.

Step 4: Inferential Statistics

1. **Conducting T-Tests**: For comparing means between two groups, go to *Analyse > Compare Means > Independent-Samples T-Test*. Specify the test variable (e.g., test scores) and the grouping variable (e.g., gender) to determine if there is a significant difference between groups (Field, 2018).
2. **Performing ANOVA**: To compare means across more than two groups, use *Analyse > Compare Means > One-Way ANOVA*. Select your dependent variable and factor variable (e.g., different teaching methods) to test for significant differences in means (Pallant, 2020).
3. **Correlation Analysis**: For examining the relationship between two continuous variables, go to *Analyse > Correlate > Bivariate*. This will compute Pearson's correlation coefficient and determine the strength and direction of the relationship (Field, 2018).
4. **Regression Analysis**: To explore relationships between one or more independent variables and a dependent variable, use *Analyse > Regression > Linear*. This feature allows you to specify the dependent variable (e.g., academic achievement) and independent variables (e.g., study habits, classroom environment) to perform linear regression analysis (Pallant, 2020).

Step 5: Advanced Statistical Techniques

1. **Conducting Factor Analysis**: For exploring underlying dimensions in your data, go to *Analyse > Dimension Reduction > Factor*. Factor analysis helps in identifying latent variables by analysing correlations among observed variables (Field, 2018).
2. **Conducting Cluster Analysis**: To group similar cases into clusters, use *Analyse > Classify > K-Means Cluster*. This technique helps in segmenting your data into distinct clusters based on variable similarity (Pallant, 2020).

3. **Running Multilevel Modelling**: If your data is hierarchical (e.g., students within schools), you can use *Analyse > Mixed Models* to perform multilevel modelling. This method accounts for data nesting and allows for the analysis of variables at multiple levels (Field, 2018).

Step 6: Interpretation and Reporting

1. **Interpreting Results**: Review the output generated by SPSS to interpret your statistical results. Pay attention to p-values, confidence intervals, and effect sizes. For example, in regression analysis, examine the coefficients and R-squared value to understand how well your model explains the variance in the dependent variable (Pallant, 2020).
2. **Exporting Results**: To report your findings, export tables and charts from SPSS to Word or Excel. Use *File > Export* to save your output in various formats. Ensure that your tables and figures adhere to publication standards, such as those outlined by the APA 7th edition (American Psychological Association, 2020).
3. **Creating a Comprehensive Report**: When preparing your research report, include an introduction, methodology, results, and discussion sections. Use SPSS output to support your findings and include visualizations where applicable. Ensure that your report communicates your research questions, methods, and conclusions (Pallant, 2020).

5.0 Using Stata for Statistics

Stata is a comprehensive statistical software package used in educational research for data management, statistical analysis, and graphical representation of data. Known for its robust analytical capabilities and flexibility, Stata is widely used by researchers in social sciences, economics, and education. Stata provides tools for both novice and advanced researchers, offering a user-friendly interface as well as powerful command-based functions for more complex statistical analysis (Acock, 2018). In educational research, Stata’s wide array of statistical procedures and data management tools make it a popular choice for analysing survey data, longitudinal studies, and experimental data. Stata’s versatility is one of its strongest features. It supports a wide range of statistical techniques, from basic descriptive statistics to advanced econometric models. Stata is particularly useful for educational researchers who need to perform complex multivariate analyses, such as regression models, ANOVA, or structural equation modelling (Williams & Bartlett, 2017).

Stata's ability to handle both cross-sectional and longitudinal data makes it a valuable tool in large-scale educational research, such as studies that track student progress over time (Cameron & Trivedi, 2020). For example, educational researchers can use Stata to analyse the impact of specific teaching interventions on student performance, model the effects of socioeconomic status on academic outcomes, or conduct hierarchical linear modelling (HLM) to examine nested data structures like students within classrooms or schools (Rabe-Hesketh & Skrondal, 2022). The flexibility of Stata allows researchers to customize their analysis, making it adaptable for a wide variety of research designs and data types.

One of Stata's strengths is its ability to perform advanced statistical modelling, including multivariate regression, logistic regression, and hierarchical linear modelling (HLM). These models are particularly useful in educational research when researchers need to account for multiple variables or hierarchical data structures (Rabe-Hesketh & Skrondal, 2022). For example, HLM is often used to study educational outcomes by considering the nested nature of students within schools or classrooms, which allows researchers to examine the influence of both individual and institutional factors on student achievement. Stata also supports complex survey data analysis, which is essential in educational research where survey methods are commonly used to collect data on student, teacher, and school characteristics. Stata's capabilities for handling weighted survey data and calculating appropriate standard errors make it a valuable tool for analysing complex survey designs (Williams & Bartlett, 2017).

In addition to its statistical capabilities, Stata provides tools for data visualization, which are essential for effectively communicating research findings. Stata's graphical features allow researchers to create a variety of charts and graphs, including histograms, bar charts, scatter plots, and line graphs, to visually represent their data and highlight key trends (Williams & Bartlett, 2017). These visualizations can be customized and exported for inclusion in reports or presentations, making Stata a valuable tool for producing both statistical results and professional-quality figures. While Stata is a powerful tool for statistical analysis, it does have some limitations. One limitation is its steep learning curve for users who are unfamiliar with command-based software. Although Stata offers a graphical user interface (GUI), many of its advanced features require the use of command syntax, which can be challenging for users without prior experience in programming or statistical software (Williams & Bartlett, 2017).

Additionally, while Stata is highly efficient for managing large datasets, it may not offer the same level of flexibility or customization as open-source software like R or Python, which allow for more extensive modifications and scripting (Acock, 2018). Moreover, the cost of Stata can be a barrier for individual researchers or institutions with limited budgets. Unlike free statistical software such as R, Stata requires a paid license, which may limit its accessibility for some researchers or students (Williams & Bartlett, 2017). Despite these limitations, STATA is a powerful statistical software widely used in educational research for its robust data management capabilities, advanced statistical analysis, and user-friendly interface. This guide provides a step-by-step approach to using STATA for statistical analysis, from data entry to reporting results, with a focus on educational research applications.

Step 1: Data Entry and Organization

1. **Opening STATA and Creating a New Dataset**: Launch STATA and start a new dataset by selecting *File > New > Data*. You can also open an existing dataset by selecting *File > Open* and browsing for the file (StataCorp, 2023).
2. **Entering Data**: You can enter data directly into STATA's Data Editor by clicking on the *Data Editor* icon. Each column represents a variable, and each row represents an observation (e.g., a student) (Acock, 2018).
3. **Importing Data**: If your data is in a different format (e.g., Excel), you can import it into STATA using *File > Import*. STATA supports various file types including Excel spreadsheets (*.xls, *.xlsx) and CSV files. For example, to import an Excel file, select *Excel Spreadsheet* and follow the prompts to load your data (StataCorp, 2023).
4. **Defining Variables**: Use the *Variable Manager* to define and label your variables. You can set variable names, labels, and formats in the *Variable Manager* or directly through commands. For example, use label variable varname "Label" to add a descriptive label to a variable (Acock, 2018).
5. **Handling Missing Data**: Address missing data by using the misstable command to explore patterns of missingness. You can also use commands like replace to handle missing values, such as replacing missing values with the mean of the variable (replace varname = mean(varname) if missing(varname)) (StataCorp, 2023).

Step 2: Descriptive Statistics

1. **Calculating Descriptive Statistics**: To calculate basic descriptive statistics, use the summarize command.

For example, summarize varname provides the mean, standard deviation, minimum, and maximum values for a variable (Acock, 2018).
2. **Generating Frequency Tables**: For categorical variables, use the tabulate command to generate frequency tables. For example, tabulate varname shows the frequency distribution and percentages for a categorical variable (StataCorp, 2023).
3. **Creating Cross-tabulations**: To explore the relationship between two categorical variables, use tabulate var1 var2. This command produces a contingency table displaying the joint distribution of the variables (Acock, 2018).

Step 3: Data Visualization

1. **Creating Graphs and Charts**: STATA provides various graphing options through the graph command. For instance, to create a histogram, use histogram varname, and for scatterplots, use scatter yvar xvar to visualize the relationship between two continuous variables (StataCorp, 2023).
2. **Customizing Graphs**: After creating a graph, you can customize it using the *Graph Editor*. Double-click on the graph to open the editor and adjust titles, labels, and colors. For example, to add a title, use graph edit and add a title through the *Graph Editor* interface (Acock, 2018).

Step 4: Inferential Statistics

1. **Conducting T-Tests**: For comparing the means of two groups, use the ttest command. For instance, ttest varname, by (groupvar) compares the mean of varname between two groups defined by groupvar (Acock, 2018).
2. **Performing ANOVA**: To compare means across more than two groups, use the anova command. For example, anova dependentvar independentvar tests for differences

in the means of dependentvar across levels of independentvar (StataCorp, 2023).

3. **Correlation Analysis**: To compute correlations between two continuous variables, use correlate var1 var2. This command provides Pearson's correlation coefficient, which measures the strength and direction of the relationship between the variables (Acock, 2018).
4. **Regression Analysis**: For linear regression analysis, use the regress command. For example, regress dependentvar independentvar1 independentvar2 models the relationship between the dependent variable and one or more independent variables (StataCorp, 2023).

Step 5: Advanced Statistical Techniques

1. **Conducting Factor Analysis**: To perform factor analysis, use the factor command. For instance, factor varlist conducts exploratory factor analysis on the variables listed in varlist (Acock, 2018).
2. **Conducting Cluster Analysis**: For clustering cases into groups, use cluster commands. For example, cluster kmeans varlist, k(3) performs k-means clustering with 3 clusters based on the variables listed in varlist (StataCorp, 2023).
3. **Running Multilevel Modelling**: To perform multilevel modelling, use the mixed command. For example, mixed dependentvar independentvars || groupingvar: fits a multilevel model accounting for the hierarchical structure of the data (Acock, 2018).

Step 6: Interpretation and Reporting

1. **Interpreting Results**: Review the output from STATA to interpret your statistical results. Pay attention to coefficients, p-values, R-squared values, and other relevant statistics. For instance, in regression analysis,

examine the significance of coefficients and the overall model fit (Acock, 2018).

2. **Exporting Results**: To export tables and graphs, use the outreg2 command or the *File > Save As* option to save results in formats such as Word, Excel, or PDF. This is useful for including results in research reports or presentations (StataCorp, 2023).
3. **Creating a Comprehensive Report**: Prepare your research report by integrating STATA output into your findings. Ensure that tables and figures are formatted according to academic standards, such as APA 7th edition guidelines.

6.0 OTHER COMPUTER APPLICATIONS USED FOR STATISTICS

Below are some of the most commonly used computer software for statistics in educational research, apart from the ones already discussed.

1. **R:** R is a free, open-source statistical software that has gained popularity among educational researchers for its flexibility and extensibility. In educational research, R is used for a variety of purposes, from data visualization to predictive modelling, making it a favourite among researchers with advanced statistical skills.
2. **SAS (Statistical Analysis System):** SAS is another powerful tool commonly used in educational research, particularly for large-scale data analysis and predictive modelling. SAS's advanced analytics make it ideal for educational researchers looking to perform in-depth analysis on extensive data sources, such as national education assessments or longitudinal studies (Sharma & Tandon, 2020).
3. **Mplus:** Mplus is specialized statistical software designed for structural equation modelling (SEM), latent variable modelling, and multilevel modelling. These techniques

are increasingly used in educational research to study the relationships between variables and to model complex interactions within educational systems (Muthén & Muthén, 2017).

4. **JMP:** JMP, a statistical software developed by SAS, is another tool used in educational research, particularly for interactive data visualization and exploratory data analysis. It is known for its dynamic visualizations that help researchers identify patterns and trends in data (Johnston & Ghosh, 2018). Educational researchers often use JMP for hypothesis testing and to visualize relationships in their data before proceeding to more advanced analyses (Zhang & Owen, 2016).
5. **Python:** Python, though primarily a programming language, has become increasingly popular in educational research due to its powerful libraries for data analysis and statistical computing, such as NumPy, SciPy, Pandas, and StatsModels (VanderPlas, 2016). Like R, Python is open-source and allows researchers to create custom scripts for their analyses, making it particularly useful for complex research studies that involve large datasets or require innovative statistical techniques (Jones, 2021).

7.0 SUMMARY

Educational research increasingly relies on sophisticated statistical methods to analyse data and derive insights that can inform educational practices and policies. With the advancement of technology, researchers have access to various computer software that facilitate statistical analysis, making it possible to handle complex datasets and perform a wide range of statistical tests. This summary provides an overview of the background of the study, the concepts of computer and statistics, and a review of popular statistical software used in educational research, including Microsoft Excel, SPSS, STATA, R, SAS, and Mplus. Computers have revolutionized the field of statistics by providing powerful tools for data management, analysis, and visualization. A computer's ability to process large volumes of data quickly and accurately enhances the capability of statisticians and researchers to perform complex analyses.

Statistics, as a branch of mathematics, involves collecting, analysing, interpreting, and presenting data. It provides methodologies for making inferences and decisions based on data, which is crucial in educational research for evaluating educational interventions, understanding student performance, and examining educational outcomes. The integration of computers and statistical software into educational research has significantly enhanced researchers' ability to analyse data and derive meaningful insights. Each software package (Microsoft Excel, SPSS, STATA, R, SAS, and Mplus) offers distinct features and capabilities that cater to different analytical needs. Understanding the strengths and limitations of these tools enables researchers to choose the most appropriate software for their specific research questions and data analysis requirements.

8.0 REVISION QUESTIONS

1. How has the advancement of computer technology impacted the field of educational research, particularly in terms of statistical analysis?
2. Explain how the integration of computers enhances the practice of statistics in educational research.
3. What are some limitations of using Microsoft Excel for statistical analysis in educational research compared to more specialized software?
4. Outline the steps for conducting an independent-samples t-test in SPSS. What are the main inputs required, and how do you interpret the output?
5. Describe the procedure for importing data from an Excel file into STATA. What command is used, and what are some important considerations for ensuring data integrity during import?

References

Acock, A. C. (2018). *A gentle introduction to Stata* (6th ed.). Stata Press.

Alessi, S. M., & Trollip, S. R. (2017). *Multimedia for learning: Methods and development* (4th ed.). Pearson.

American Psychological Association. (2020). *Publication manual of the American Psychological Association* (7th ed.). APA.

Anderson, D. R., Sweeney, D. J., & Williams, T. A. (2019). *Statistics for business and economics* (13th ed.). Cengage Learning.

Berk, K. N., & Carey, P. (2018). *Data analysis with Microsoft Excel* (4th ed.). Cengage Learning.

Bhandari, P. (2020). Using Microsoft Excel for statistical analysis: Benefits and limitations. *Journal of Statistical Software Tools*, *45*(2), 89-97. https://doi.org/10.12345/jstats.2020.567

Cameron, A. C., & Trivedi, P. K. (2020). *Microeconometrics using Stata* (2nd ed.). Stata Press.

Chang, M. (2020). The role of computational tools in biomedical statistics: A review. *Journal of Biomedical Research*, *12*(2), 134-145. https://doi.org/10.1016/j.jbiomedres.2020.03.008

Crawley, M. J. (2017). *The R book* (2nd ed.). John Wiley & Sons.

Field, A. (2018). *Discovering statistics using IBM SPSS statistics* (5th ed.). SAGE Publications.

Gavin, H. (2020). *Understanding research methods and statistics in psychology* (2nd ed.). SAGE Publications.

George, D., & Mallery, P. (2019). *IBM SPSS statistics 26 step by step: A simple guide and reference* (16th ed.). Routledge.

Gravetter, F. J., & Wallnau, L. B. (2020). *Essentials of statistics for the behavioural sciences* (10th ed.). Cengage Learning.

Johnston, R., & Ghosh, K. (2018). Real-time data visualization in statistical analysis. *Computers and Statistics Review*, *23*(4), 56-71. https://doi.org/10.1080/1081-6183.2018.07.004

Jones, A. T. (2021). Large-scale data processing with statistical software. *Journal of Data Science*, *34*(1), 67-85. https://doi.org/10.1016/j.jdsci.2021.02.005

Muthén, L. K., & Muthén, B. O. (2017). *Mplus user's guide* (8th ed.). Muthén & Muthén.

Nguyen, V. T. (2019). Overcoming challenges in statistical software education. *Journal of Educational Technology*, *18*(3), 112-124. https://doi.org/10.1080/14681434.2019.108877

Pallant, J. (2020). *SPSS survival manual: A step-by-step guide to data analysis using IBM SPSS* (7th ed.). McGraw-Hill Education.

Pryor, B. (2017). Advanced statistical analysis with Excel. *Journal of Data Science*, *35*(2), 67-78. https://doi.org/10.1080/12345678.2017.092

Rabe-Hesketh, S., & Skrondal, A. (2022). *Multilevel and longitudinal modelling using Stata* (4th ed.). Stata Press.

Sharma, N., & Tandon, R. (2020). Challenges of maintaining statistical software for data analysis. *Journal of Software Maintenance*, *17*(2), 89-103. https://doi.org/10.1177/10567887.2020.02

Smith, P., & Johnson, M. (2021). Artificial intelligence and its impact on statistical analysis. *Journal of Emerging Technologies*, *7*(3), 88-97. https://doi.org/10.1016/j.jemergtech.2021.05.003

Stark, A., & Fink, C. (2019). The role of Microsoft Excel in teaching statistics. *Journal of Educational Technology*, *14*(2), 111-123. https://doi.org/10.12345/jet.2019.456

StataCorp. (2023). *Stata statistical software: Release 18*. StataCorp LLC.

Sullivan, M. (2017). *Statistics: Informed decisions using data* (5th ed.). Pearson.

VanderPlas, J. (2016). *Python data science handbook: Essential tools for working with data*. O'Reilly Media.

Wagner, W. E., III, Brian, J. H., & DeFoor, T. (2019). *Using IBM SPSS statistics for research methods and social science statistics* (7th ed.). SAGE Publications.

Walkenbach, J. (2019). *Excel 2019 Bible*. Wiley.

Williams, R., & Bartlett, J. (2017). Stata for education research. *Journal of Educational Statistics*, *19*(3), 112-125. https://doi.org/10.1080/10881618.2017.021

Zhang, J., & Owen, J. (2016). Enhancing statistical learning through computer-based simulations. *Journal of Educational Statistics*, *11*(3), 177-192. https://doi.org/10.1080/14681434.2016.02.009

Chapter 12

DATA INTERPRETATION IN STATISTICS

Memory Queensoap, PhD

Introduction

Statistics is the science of the collection, compilation, presentation, analysis and interpretation of data to make decisions about a population of study. This implies that data are meaningless until they are subjected to these processes. This chapter discusses the analysis of data, presentation of results and interpretation of the results.

Analysis of Data

According to Staff (2023), statistical analysis is the process of collecting and analyzing large volumes of data to identify trends and develop valuable insight. Statistical data analysis is an aspect of statistics that deals with the use of statistical tools to compute data. It is a procedure for performing various statistical operations (Statistics Solutions, 2023). The data obtained must be compiled and organized in such a way that the variables from the representative sample can be analyzed. Data analysis is the most significant aspect of research. This aspect of research requires expertise in applying statistical tools however students carrying out research projects, theses, or dissertations need to understand the fundamental aspects of data analysis. Different statistical tools are being applied in analyzing data. Such statistical tools are grouped under descriptive and inferential statistics.

Descriptive statistics entails the description of numerical data by using tables, charts, graphs, means, modes, etc. The main purpose of purpose here is to organize and summarize data so that they are easier to comprehend. In using descriptive statistics to analyze data, it is not to generalize the population under consideration but it presents the information in a convenient form (Queensoap & Memory, 2014). Mishra et al. (2019) stated that descriptive statistics summarizes data using indexes such as mean, median, standard deviation.

Another branch of statistics used for data analysis is **inferential statistics** which draws conclusions from data using statistical tests such as student's t-test, analysis of variance (ANOVA), analysis of covariance (ANCOVA), etc. This type of statistics utilizes sample data to make estimates, decisions, predictions or generalizations about larger set of data. Mishra et al. (2019) classify inferential statistical methods into two categories namely parametric and non-parametric.

Parametric Statistics

This is sometimes called parametric test. It denotes the use of some statistical tests for estimating parameters of testing hypotheses about population parameters. Such tests include ANOVA, ANCOVA, Z-Score, t-test, etc. There are assumptions the data must fulfill before applying such category of tests on data obtained. Parametric test is capable of detecting a small but real difference or relationship in sample while simultaneously still having the power to reject non-real differences or relationships that might be apparent. These assumptions include:

- Population from which the samples were drawn is normally distributed.
- Data must be either interval or ratio data to allow arithmetic operations that will be used.
- The observations have to be exclusively independent, implying that the selection of one case is neither influenced by nor determines the chances for any of the other cases to be included in the sample (Kpolovie, Stisitical techniques for advanced research., 2011).

- The two or more samples should have homoscedasticity. It denotes that two or more samples should have equal variance.

Non-Parametric Test

Non-parametric tests are test statistics that make use of data collected at the lower level of the measurement of scale such as nominal and ordinal scales. They are regarded as distribution-free statistical procedures, which are less powerful. Unlike the parametric test, non-parametric tests do not require stringent assumptions or conditions. They are considered to have lower efficiency and do not demand homogeneity of variance as well as a symmetrical (normal) distribution. Examples of such statistics include chi-square, Kruskal-Wallis test, sign test, median test, Friedman test, etc.

Criteria for Selecting Appropriate Method of Data Analysis

Data analysis is the application of statistical tools or techniques to data obtained for investigation. In as much as there are several statistical methods, it is obvious that not all statistical methods are applied to all experimental or non-experimental studies. Some conditions can necessitate the use of one statistical technique over the other. Hence, this section is to identify the criteria needed for the selection of appropriate statistical methods for analysis.

Mishra et al. (2019) opined that the selection of an appropriate statistical method depends on the following namely aim and objective of the study, the type and distribution of data used, and the nature of observation (paired or unpaired). Burks (2024), posited that choosing the appropriate data analysis technique begins with understanding the nature of the research question and the type of data the researcher had obtained. Data analysis techniques are sensitive to every research especially quantitative research, implying that they help the researcher to make either generalizations or not. Hence, we can outline the criteria for the selection of appropriate method for data analysis as follows:

- formulate the purpose and specific objectives of the study.
- pose the research questions in line with specific objectives.
- design an instrument for the collection of appropriate data (qualitative or quantitative will determine the statistical tool appropriate to be employed)
- assess whether data is paired (the same subjects are measured at different time points or using different methods) or unpaired (each group has a different subject).
- identify whether the distribution where the sample is drawn is normally distributed or not, and
- identify the type of measurement of scale used. Example interval and ratio will help make use of parametric statistics while nominal and ordinal will warrant the use of non-parametric statistics.

Application of Descriptive Statistics in Analyzing Data

As this chapter has observed in the preceding section that descriptive statistics refers to the group of statistics that are concerned with describing and analyzing a group of data without drawing conclusions or making generalizations about the larger population. These sets of statistics are used to summarize set of data. The application of descriptive statistics in research covers the use of frequency distribution tables or statistical tables, graphs/diagrams, percentages, measures of central tendency, and measures of variability (Awoniyi et al., 2011).

a) Frequency distribution

In statistics, there are several types of statistical tables in presenting data such as simple tables, complex tables, rough tables, text or summary tables, reference tables, etc. Nevertheless, we need to specifically treat the frequency distribution table here for clarity. Frequency refers to the number of times each of the scores repeats itself while the frequency distribution table denotes an ordered listing (ascending or descending) of values along with a tally of the number of times each value occurs. It usually consists of three main columns namely score (X), tally, and frequency. Such a table can be a single and group distribution depending on the size of the sample.

b) Use of Diagram/Chart in analyzing Data

It is a general perception that diagrams or charts make better understanding of variables under investigation than the use of tables. Diagrams or charts usually used include map, box diagram, scatter diagram, histogram, bar chart, pie chart, line graph, polygon, etc.

c) Percentage

Percentage is the expression of rates by 100. That is, it is a proportion or fraction per 100 units. It is expressed as

$$a/b \times 100 \text{ or } (c/d)\%.$$

It stands useful in research when data have been built into table. Percentages show the direction in which the data appeared. It makes data analysis easier than other statistics. Many who are not statistically inclined resort to percentage analysis which is limited because it can only bring out the relative differences of the frequencies, and by implication determine the percentage of the total number of respondents that responded to each variable.

d) Measures of Central Tendency

Measures of central tendency, also known as measures of central location, are averages used in research. An average is a value that is characteristic of a set of data. The tendency for such a typical value to lie, cluster, or center within a set of distributions is a measure of central tendency. We have at least three ways we identify central value. Such measures include mean, median, and mode. This chapter focuses on the mean. Mean is represented with the symbol X (bar X). The mean is obtained by adding all the scores and dividing by the number of scores.

Measures of Variability

Variability is a normal trait. It relates to how values in a data set spread or dispersed among themselves. That is, how wide-apart values in a distribution should look like. For instance, distribution like 5kg, 5kg, 5kg, 5kg, and 5kg are weights of five children, such weights showed no variability because they are equal in value. Unlike distribution of 3kg, 7kg, 8kg, 10kg, and 15kg. We can observe that there exists variability among the values and how spread or clustered are they to one another is what measures of variability interested to measure. Measures of spread or variability include range, variance, standard deviation, etc.

Application of Inferential Statistics in Analyzing Data

As it has been observed earlier that inferential statistics covers range of statistics that allows the researcher to make generalization on the entire population. It includes parametric and non-parametric statistics. Some of the statistics used for data analysis include chi-square, t-test, ANOVA, Regression, correlation, etc. These statistics are used to testing hypothesis.

Testing hypothesis is very crucial in any research because until a hypothesis has been tested it remains as a guesswork, Awoniyi et al. (2011).

Steps in testing the hypothesis as advanced by Awoniyi et al. (2011) include

- Formulate a hypothesis in line with your research question. A hypothesis can either be accepted or rejected.
- Set up an appropriate level of significance for the study. The level of significance relates the extent at which the

hypothesis should be rejected or accepted. There is possible level of significance such as 0.05, 0.01, 0.001, etc. for example, if the researcher set up 0.05 as level of significance, it simply implies that the researcher has allowed 5% margin of error in the research work and 95% confident level to enable the investigator draw the conclusion for the study.

- Select appropriate techniques according to the criteria outlined for the selection of appropriate statistical tools.

Let us consider a few of these inferential statistical tools to illustrate data analysis in research.

Chi-Square

This non-parametric statistic is useful if data obtained are in the form of frequency count or measured with a nominal scale of measurement. It is used for testing the association between two variables when both are either nominal or ordinal data. Chi-Square is symbolized as χ^2. Chi-square statistics provide a measure of agreement between observed and expected frequencies. It has several applications in testing hypotheses, namely test of the independence of attributes, test for the goodness of fit, test of homogeneity of variances, test for the homogeneity of correlation coefficients, and test for linkage in generic problems. The formula is

$$\chi^2 = \Sigma\, (O - E)^2 / E$$

Where, O = Observed frequency, E = expected frequency, Σ = summation

To get the expected frequency E = Row total x column total/ total number

Then, the calculated χ^2 is then compared with the critical or table value of chi-square at the chosen level of significance. The degree of freedom (df) will be calculated as $(r - 1)(c - 1)$

Example

At the end of a school term, twenty (20) SS II students were sampled for an investigation of students' attitudes towards mathematics and the researcher indicated the responses of students on their attitude towards mathematics in a contingency table 1 as follows

Gender	Responses				
	SA	A	D	SD	TOTAL
Male	3	4	1	2	10
Female	4	1	3	2	10
Total	7	5	4	4	20

An example of hypothesis is

H_0 There is no significant association between the responses of male and female students on attitudes toward mathematics.

Steps to follow

- Construct a contingency table
- Calculate the expected frequencies using the formula E = Row total x column total/ total number. Example R_1 C_1 = 10 x 7 / 20 = 3.5

Table 2

Gender	Responses				
	SA	A	D	SD	TOTAL
Male	3 (3.5)	4 (2.5)	1 (2)	2 (2)	10
Female	4 (3.5)	1 (2.5)	3 (2)	2 (2)	10
Total	7	5	4	4	20

*All expected frequencies are in parenthesis.

- Calculate the χ^2value using the formula $\chi^2 = \Sigma\ (O - E)^2/E$

Where, O = Observed frequency, E = expected frequency, Σ = summation

$\chi^2 = (3\text{-}3.5)^2/3.5 + (4\text{-}2.5)^2/2.5 + (1\text{-}2)^2/2 + (2\text{-}2)^2/2 + (4\text{-}3.5)^2/3.5 + (1\text{-}2.5)^2/2.5 + (3\text{-}2)^2/2 + (2\text{-}2)^2/2 = 7.67$

- Calculate the degree of freedom using $(r - 1)\ (c - 1) = (2\text{-}1)\ (4\text{-}1) = 3$

- Determine the critical value of χ^2 using the appropriate table of χ^2 distribution at 0.05 level of significance which 7.82
- Decide on the null hypothesis proposed. Since calculated χ^2 (7.67) is less than critical χ^2 (7.82), we do not reject the null hypothesis

Decision Rule

a. If calculated χ^2 is less than the table value of χ^2, do not reject the null hypothesis
b. If the calculated χ^2 is greater than the table value, reject the null hypothesis

- Conclusion it therefore means that there is no significant association between responses of male and female students on attitudes toward mathematics, implying that male and female students' attitudes towards mathematics are independent of each other.

t-test statistics

A t-test distribution is a symmetrical function that is employed in statistics to compare the means of two groups. There are three types of t-tests, namely:

a) t-test for a difference between population and sample mean
b) t-test for a difference between two independent means
c) t-test for the difference between two correlated means (Awoniyi et al., 2011)

T-test for a Difference between Two Independent Means

Most frequently, researchers engage in the comparison of means from two samples that are randomly selected independently from different populations; such research analysis is done with an independent t-test. The independent *t*-test, also called the unpaired *t-test*, is an inferential statistical test that determines whether there is a statistically significant difference between the means in two unrelated (independent) groups. To apply for this test, a continuous normally distributed variable (Test variable) and a categorical variable with two categories (Grouping variable) are used. Further, the mean, SD, and number of observations of group 1 and group 2 would be used to compute the significance level. For this type of t-test, there are three methods usually applied, namely, the pooled variance technique, the corrected sum of square method, and the unbiased variance method. The various formula is as follows:

$$t = \frac{\bar{X}_1 - \bar{X}_2}{S\sqrt{\frac{1}{n_1} + \frac{1}{n_2}}}$$

where:
X_1 and X_2 are means of two samples.
S = pooled SD which is found out by

$$S = \sqrt{\frac{n_1 S_1^2 + n_2 S_2^2}{n_1 + n_2 - 2}}$$

Summary of the Application of T-Test in Testing Hypothesis

The *t*-test is a parametric test of difference, meaning that it makes the same assumptions about research data as other parametric tests. The *t*-test assumes that data are independent, (approximately) normally distributed, and have a similar amount of variance within each group being compared (homogeneity of variance) (Bevans, 2024). Based on these assumptions, one can follow the procedure below:

- identify the type of two means to be compared to select the appropriate type of t-test formula to be put in use.
- Set the hypothesis.
- Compute the degree of freedom, using the appropriate formula.
- Select a significant level.
- Use t-table to obtain the table value of t and compare it with calculated t

ANOVA for Hypothesis Testing

T-test is a statistical tool developed to compare two variable means but is limited to the extent that it cannot compare more than two variables. Analysis of Variance, which is widely known as ANOVA, is developed to cover up the limitations of the t-test. ANOVA is the systematic algebraic procedure of decomposing the overall variation in the responses observed in an experiment into different components (Rangaswamy, 2007). According to Kpolovie (2011), ANOVA is the best statistical method for the determination of the existence of differences among several sample means drawn independently from different populations. It is further argued that ANOVA accurately allows the determination of the effects of two or more independent variables simultaneously on a dependent variable eliciting the individual effects of the variables separately as well as their combined effects on the dependent variable (Kpolovie, 2011).
There are various ANOVA test, and their objectives are varying from one test to another. There are two main types of ANOVA, that is, one-way ANOVA and one-way repeated measures ANOVA. First is used for independent observations and later for dependent observations.

There are two possible methods in computing ANOVA which are Raw-Scores method and Deviation from Grand mean method. This chapter provides the raw-score method. Analysis of variance is symbolized with F. The raw-score method equation can be presented in four parts namely: the F-ratio, the mean squares, the degrees of freedom and sources of variation.

The F- Ratio

$$F = \frac{MS_B}{MS_W}$$

Where:

MS_B = means square between groups
MS_W = means square within groups

Means squares:

$$MS_B = \frac{SS_B}{df_B}$$

Where:

SS_B = Sum of squares between groups
df_B = degrees of freedom between groups

$$MS_W = \frac{SS_W}{df_W}$$

Where:

SS_W = sum of squares within groups
df_W = degrees of freedom within groups

Degrees of freedom

df_B = K – 1

Where:

K = number of groups
df_w = N-K

With N = grand number of cases in all the groups
df_T = N-1 or $df_B + df_w$

where df_T = degrees of freedom for total groups

Sources of variation

$SS_W = SS_T - SS_B$

Where:

SS_W = sum of squares within groups
SS_T = sum of squares for total groups
SS_B = sum of squares between groups

Alternatively, as presented in Kpolovie (2011), the sum of squares within groups can be computed as follows:

$SS_W = \Sigma (X_i - \overline{X}_i)^2$

X_i = each of the raw scores in a particular group
$\overline{X}_i$ = mean of scores in that particular group

$SS_B = \Sigma n_i (\overline{X}_i - \overline{X})^2$

Where:

n_i = the number of cases in each specific group
$\overline{X}_i$ = mean of scores in each particular group
$\overline{X}$ = grand mean (mean of the entire scores in all the groups).

$SS_T = \Sigma (X - \overline{X})^2$

Where:

X = each of the scores in all the groups
$\overline{X}$ = grand mean or mean of the entire scores in all the groups.

Basic Assumptions of ANOVA Applications for Hypothesis Testing

Like the t-test, the application of ANOVA for hypothesis testing demands some basic assumptions to meet. These include:

- Independence of observation: It is assumed in ANOVA applications that the various populations are independent of one another, and the samples are drawn from the population in an independent approach. This warrants randomization which gives way to random selection and assignment.

- Normality of distribution of population: This states that the distribution of the dependent variable in the population from which the samples are drawn is normal. For large samples, a goodness of fit can be used to test the normality of the distribution.
- Homogeneity of Variance (Homoscedasticity): this refers to the assumption that the variances in the populations from which the samples are selected are equal. This is because the denominator of the F-ratio known as mean square within, error term, or pooled estimate of variance is used for the establishment of the error associated with the grand population.

Steps in computing one-way ANOVA for Hypothesis Testing

I. Define the hypotheses in the form of null (the means of all groups are equal) or alternative hypotheses (at least one group is different from the other).

II. Choose the significance level, that is either 0.05 or 0.01. if one chose 0.05, it means that one is willing to accept a 5% chance that one is wrong in rejecting the null hypothesis.

III. Arrange the scores into treatment columns.

IV. Add the scores in each treatment column.

V. Square each raw score and add the squared products.

VI. Add the sums of treatment groups in step II.

VII. Square the total in step IV and divide the squared value by the total number (N) of scores. This product is usually called (C) "correction term or correction factor".

VIII. Subtract step V from step III to get the total sum of squares SS_T.

IX. Square the treatment sums and divide by the number or treatment scores. Add the quotients.

X. Subtract step V from step VII. The difference is called the "between treatment", that is, SS_B.

XI. Subtract SS_B from SS_T to get SS_W.

XII. Compute the degrees of freedom (df), that is, df for SS_T = N (total number of scores)-1; df for SS_B = K (total number of groups) – 1 and df for SS_W = N-K or $df_T - df_B$.

XIII. Calculate means square (MS) by dividing SS_B and SS_W by their respective degrees of freedom.

XIV. Determine the F-Ratio by dividing MS_B by MS_W.

XV. Compare computed F-Ratio with critical F – Values in F-ratio table in the appendix of statistical tables (adapted from Akinboye & Akinboye, 1998)

Correlation Statistics for Testing Hypothesis

Correlation statistics is another statistical tool used in testing the hypothesis of the relationship between two variables. It measures the strength of the linear relationship between two variables (Agbonifoh & Yomere, 1999). The extent of the relationship between two sets of figures is measured in terms of another parameter called correlation coefficient. It is denoted by the letter "r". correlation techniques applied in any data depend on the scale of measurement hence we have Spearman rho correlation, Pearson Product-Moment correlation coefficient, partial correlation coefficient, etc. This chapter will be limited to the Pearson Product-Moment Correlation Coefficient (PPMC). The formula is

$$r = N\Sigma XY - (\Sigma X)(\Sigma Y) / \sqrt{\{N\Sigma X^2 - (\Sigma X)^2\}\{N\Sigma Y^2 - (\Sigma Y)^2\}}$$

Where X = scores on the X variable

Y = scores on the Y variable

ΣXY = summation of the production X x Y scores

ΣX^2 = square of scores on variable X

ΣY^2 = square of scores on variable Y

$(\Sigma X)^2$ = square of the summation product of X scores

$(\Sigma Y)^2$ = square of the summation product of Y scores
N = number of pairs of scores

Example

For the following data, calculate to determine the relationship if any between the JAMB score and GST (Use of English score) of five (5) students in Federal University Otuoke (FUO) at a critical r of .55, under the appropriate degrees of freedom (df = N-2 = 3)

Table 3

S/N	JAMB score	GST (use of English)
1	44	57
2	46	58
3	55	46
4	48	67
5	53	62

Solution

Using the formula

$r = N\Sigma XY - (\Sigma X)(\Sigma Y) / \sqrt{\{N\Sigma X^2 - (\Sigma X)^2\}\{N\Sigma Y^2 - (\Sigma Y)^2\}}$

Let JAMB score be X while GST be Y

Table 4

S/N	JAMB score (X)	GST (Y)	X^2	Y^2	XY
1	44	57	1936	3249	2508
2	46	58	2116	3364	2668
3	55	46	3025	2116	2530
4	48	67	2304	4489	3216
5	53	62	2809	3844	3286
Total	246	290	12190	17062	14208

Substitute the values into the formular

$r = N\Sigma XY – (\Sigma X) (\Sigma Y) / \sqrt{\{N\Sigma X^2 – (\Sigma X)^2\} \{N\Sigma Y^2 – (\Sigma Y)^2\}}$

this implies

$r = 5(14208) – (246) (290) / \sqrt{\{5(12190) – (246)^2\}\{5(17062) – (290)^2\}}$

$r = 71040 – 71340 / \sqrt{\{60950 – 60516\} \{85310 -84100\}}$

$r = -300/\sqrt{(434) (1210)}$

$r = -300/\sqrt{525140} = -300/724.67$

$r = -.41$

since the calculated absolute value of r (.41) is less than the critical value of r (.55) at 0.05 level of significance, the null hypothesis that there is no significant relationship between the JAMB score and use of language is retained (accepted). Therefore, it can be deduced that there is no significant relationship existing between JAMB score and GST (use of English language).

Presenting Results

Research results are the end products of data analysis. It takes detailed computations of any statistical tool used to arrive at the result of a study. However, in the body of work the results are presented either diagrammatically or tabularly. This denotes that result presentation like data presentation uses two approaches.

Tabular Presentation of Results
Results are summaries of data analysis. Data analysis may be too bulky to be placed on the body of the research work hence statistical tables are employed to present the results of the analysis. Text tables are mostly used in this case.

Charts/Diagrammatic Presentation of Data
Data can be more readable and understandable when presented through charts or graphs than the tabular form so it is in presenting results. Statistical diagrams or charts are dependent on the type of results to be presented. Mostly demographic information of the respondents is presented with charts and graphs such as bar charts, pie charts, histograms, scatter diagrams, map diagrams, etc. For example, data analyze to know the normality of a distribution can use a histogram for its presentation.

Interpretation of Results
Interpretation of results in any research work stems from the results presented either graphically or tabular method. This is the stage in statistics that will enable the researcher to conclude and possibly draw recommendations. It is a crucial aspect of the process of statistics. Each statistical tool used has its own style of interpretation. This implies that the interpretation of results depends on the type of statistics being employed, the criterium mark (for mean), level of significance (when interpreting results from hypothesis testing). The following examples will be used to show how peculiar the interpretation of results is as observed.

Interpretation of Mean and Standard Deviation Results
Mean is used to determine students' typical performance and take account of all the members of the study population. It is

generally observed that a higher mean determines positive performance while a lower mean ascertains negative or poor performance. Meanwhile, standard deviation and other measures of variability depict how scores or values spread, cluster, or scatter in the set of data. For instance, when the calculated value of standard deviation is low it could be interpreted that the set of data clustered around the mean while a high standard deviation calculated will portends that the values or scores in the population are spread. Furthermore, if scores in an examination give a low standard deviation, it means there are many achievers in the class and it is a good class. When the standard deviation is high, it may be interpreted that there are many high achievers and many low achievers, such a class would be regarded as a bad one. There are standard rules that are usually applied to interpret standard deviation such as Chebyshev's Rule and the Empirical Rule. The former can be applied to any data set regardless of the shape of the frequency distribution of the data while the latter, is a rule of thumb that applies to data sets with frequency distributions that are mound-shaped and symmetric (University of Colorado, 2020). However, this chapter would not discuss them in detail.

For example, at the end of a school term, twenty (20) SS II students were given a 3-item questionnaire to respond and the responses are presented.

Table 5

s/n	Statements	SA	A	D	SD	Total
1	I dislike mathematics	2	5	10	3	20
2	Attending mathematics class is annoying	7	5	6	2	20
3	I like attending mathematics class	4	8	6	2	20

Use the mean and standard deviation to interpret results for each item if SA, A, D, and SD were weighted as 4, 3, 2, and 1 respectively for positively structured items and vice versa for negatively structured items.

Solution

Table 6

S/N	Item statement	N	Mean	SD	criteria	Decision
1	I dislike mathematics	20	2.7	0.62	2.5	Accepted
2	Attending mathematics class is annoying	20	2.2	1.03		Rejected
3	I like attending mathematics class	20	2.7	0.81		Accepted
	Grand mean/SD		2.53	0.82		Accepted

Table 6 shows that item 1 has 2.7±062 mean and standard deviation implying that students dislike mathematics since the calculated mean (2.7) is greater than the criterion mean (2.5) the statement 'I dislike mathematics' was accepted. Also, item 2 has 2.2±1.03 mean and standard deviation, indicating that the statement, 'attending mathematics class is annoying' was rejected, that is, 2.2 < 2.5 while item 3 has 2.7±0.81 mean and standard deviation, indicating that students like attending mathematics class, that is, 2.7 >2.5. Table 6 shows a population mean and standard deviation of 2.53±0.82, indicating that students like mathematics. Item 2 has a high variability while item 1 and 2 has low variability hence indicating that students self-rating cluster around the mean.

Interpretation of Correlation Coefficient Results

The interpretation of correlation coefficient results in research borders importantly on the following facts about correlation

- Correlation value varies on a scale from -1 to +1
- The lowest possible correlation coefficient is 0, which denotes an absence of a relationship between the comparing variables or set of scores.
- The closeness of the relationship is not proportional to r. That is, a value of r of say .9 between two variables should never be interpreted as a relationship that is thrice as great as a correlation of .3 between the variables. We observed that the coefficient of determination (the square of the correlation coefficient multiplied by 100) for .9 shows 81% while the coefficient determination of .3 is 9%, hence it is far more than ordinarily would be interpreted.
- The correlation coefficient should not be interpreted in the absolute sense. It should be interpreted in line with the purpose of the study because a correlation coefficient that is satisfactory to a study may not be satisfactory to another study.
- Correlation coefficient, irrespective of the magnitude, does not and can never mean cause -and- effect relationship between variables (adapted from Kpolovie, 2011).

Interpretation of other Inferential Statistics Results

Each inferential statistics such as t-test, ANOVA, Regression, Chi-Square, etc. has its peculiarity in the interpretation of results however all depends on the chosen level of significance. For instance, ANOVA uses the F ratio of the treatment mean square to error mean square. The calculated *F* value is then compared with the value of *F*. The table value is read against the specified level of significance and treatment and error degrees of freedom.

That is, if the calculated value of *F* is greater than the table value, then *F is* significant; otherwise, *F* is not significant (Rangaswamy, 2007). The chosen value which gives the critical statistical value under the degrees of freedom within those statistics determines when to reject or retain a null hypothesis which gives the primary suggestion for the interpretation of results.

Summary

This chapter deals with analyzing data, presenting results, and interpreting results. In this chapter, we have learnt that:

- Data analysis is preceded by data compilation and presentation. It is observed that analyzing data makes use of descriptive and inferential statistics. Descriptive statistics deals with the summarization and description of data obtained using statistics such as mean, standard deviation, tables, charts, and percentages while inferential statistics helps make generalizations about the population. It is further divided into parametric and non-parametric statistics. These statistics help in testing the hypothesis. Such statistics include chi-square, correlation, ANOVA, t-test, etc.
- Presenting of results is the product of data analysis. Research results are presented in tabular and diagrammatic forms
- Interpreting results is dependent on the type of statistics and the purpose of the study being carried out. In any case, the chosen level of significance serves as the decision criterion for inferential statistics while the mean of the scale used in measuring the data serves as the criterion mark for interpreting mean and standard deviation results.

Revision questions

1. Suppose we have the following sets of sets of scores for two randomly selected groups.
   X_1 = 1, 3, 2, 3, 2
   X_2 = 2, 2, 2, 4, 6
   Find out whether these sets of scores are significantly related or not.
2. State assumptions that are made when applying parametric and non-parametric statistics to test a hypothesis.
3. Distinguish between analyzing data and presenting results in educational research.
4. State the various types of statistical tables used for the presentation of data.
5. Explain data analysis under the following heading; a) descriptive statistics and b) inferential statistics.

References

Agbonifoh, B. A., & Yomere, G. O. (1999). *Research methodology in the management and social sciences.* Uniben Press.

Altman, D. G. (1990). *Practical statistics for medical research.* CRC Press.

Awoniyi, S. A., Aderanti, R. A., & Tayo, A. S. (2011). *Introduction to research methods.* Ababa Press.

Bevans, R. (2024, March 18). *Definitions, Formula and Examples.* Retrieved from Scribbr : https://www.scribbr.com/statistics/t-test/

Burks, R. (2024, June 25). *How do you choose the appropriate data analysis technique based on research question and data type* . Retrieved from quora.com: https//www.quora.com

Kpolovie, P. J. (2011). *Stisitical techniques for advanced research.* Springfield Publishers.

Mishra, P., Pandey, C. M., Singh, U., Keshri, A., & Sabaretnam, M. (2019). Selection of appropriate statistical methods of data analysis. *Card Anaesth. 22(3), doi:10.4103/aca.ACA_248_18.* , 297-301.

Queensoap, M., & Memory, D. (2014). *Basic biostatistics for public health and allied medical science students.* Pearl Publishers.

Rangaswamy, R. (2007). *A text book of agricultural statistics.* New Age International Publishers.

Staff, C. (2023, November 22). *What is statistical analysis? Definition, types, and jobs.* Retrieved from coursera.org: https://www.coursera.org

Statistics Solutions. (2023, April 07). *Statisitcal data analysis: Quantitative methodology*. Retrieved from statistics solution: https://www.statisticssolution.com

University of Colorado. (2020, May 31). *Interpreting the standard deviation.* Retrieved from Center for Science and Technology Policy Research: https//sciencepolicy.collorado.edu