QUANTITATIVE ANALYSIS

FOR

ACCOUNTING TECHNICIANS SCHEME OF WEST AFRICA

(ATSWA)

STUDY PACK

DISCLAIMER

CONTENTS

PREFACE

INTRODUCTION

The Council of the Association of Accountancy Bodies in West Africa (ABWA) recognized the difficulty of students when preparing for the Accounting Technicians Scheme West Africa examinations. One of the major difficulties has been the non-availability of study materials purposely written for the Scheme. Consequently, students relied on text books written in economic and socio-cultural environments quite different from the West African environment.

AIM OF THE STUDY PACK

In view of the above, the quest for good study materials for the subjects of the examinations and the committee of the ABWA Council to bridge the gap in technical accounting training in West Africa led to the production of this Study Pack.

The Study Pack assumes a minimum prior knowledge and every chapter reappraises basic methods and ideas in line with the syllabus.

READERSHIP

The Study Pack is primarily intended to provide comprehensive study materials for students preparing to write the ATSWA examination.

Other beneficiaries of the Study Pack include candidates of other Professional Institutes, students of Universities and Polytechnics pursuing first degree and post graduate studies in Accounting, advanced degrees in Accounting as well as Professional Accountants who may use the Study Pack as reference materials.

APPROACH

The Study Pack has been designed for independent study by students and as concepts have been developed methodically or as a text to be used in conjunction with tuition at schools and colleges. The Study Pack can effectively be used as course text and for revision. It is recommended that readers have their own copies.

STRUCTURE OF THE STUDY PACK

The layout of the chapters has been standardised so as to present information in a simple form that is easy to assimilate.

The Study Pack is organised into chapters. Each chapter deals with a particular area of the subject, starting with learning objective and a summary of sections contained therein.

The introduction also gives specific guidance to the reader based on the contents of the current syllabus and the current trends in examinations. The main body of the chapter is subdivided into sections to make for easy and coherent reading. However, in some chapters, the emphasis is on the principles or applications while others emphasise methods and procedures.

At the end of each chapter is found the following
- Summary
- Points to note (these are used for purpose of emphasis or clarification)
- Examination type questions and;
- Suggested answers

HOW TO USE THE STUDY PACK

Students are advised to read the Study Pack and attempt the questions before checking the suggested answers.

FOREWARD

The ABWA Council, in order to actualise its desire and ensure the success of students at the examinations of the Accounting Technicians Scheme West Africa (ATSWA), put in place a Harmonisation Committee, to among other things, facilitate the production of Study Packs for student. Hitherto, the major obstacle faced by students was the dearth of study text which they needed to prepare for the examinations.

The Committee took up the challenge and commenced the task in earnest. To start off the process, the existing syllabus in use by some member Institutes were harmonised and reviewed. Renowned professionals in private and public sectors, the academia, as well as eminent scholars who had previously written books on the relevant subjects and distinguished themselves in the profession, were commissioned to produce Study packs for the twelve subjects of the examination.

A minimum of two Writers and a Reviewer were tasked with the preparation of a Study Pack for each subject. Their output was subjected to a comprehensive review by experienced imprimaturs. The Study Packs cover the following subjects:

PART I
1. Basic Accounting Processes and Systems
2. Economics
3. Business Law
4. Communication Skills

PART II
1. Principles and Practice of Financial Accounting
2. Public Sector Accounting
3. Quantitative Analysis
4. Information Technology

PART III
1. Principles of Auditing
2. Cost Accounting
3. Preparation Tax Computation and Returns
4. Management

Although, these Study packs have been specially designed to assist candidates preparing for the technicians examinations of ABWA, they should be used in conjunction with other materials listed in the bibliography and recommended text.

PRESIDENT, ABWA

NOTE: This is a work in progress. All topics in the syllabus are covered but editing for necessary corrections is in progress.
Thanks.

ACKNOWLEDGMENTS

PAPER 7: QUANTITATIVE ANALYSIS

AIMS:

- To provide candidates with a sound foundation in Quantitative Techniques which will assist understanding and competence in business decision-making processes that are encountered in practice.

- To develop a thorough understanding in statistical, business mathematical and operations research techniques which will help in the day-to-day performance of duties of a typical Accounting Technician.

- To examine candidates' competence in the collection, collation, manipulation and presentation of statistical data for decision-making.

- To examine the candidates' ability to employ suitable mathematical models and techniques to solve problems involving optimization and rational choice among competing alternatives.

OBJECTIVES:

On completion of this paper, candidates should be able to:

a. discuss the role and limitations of statistics in government, business and economics;

b. identify sources of statistical and financial data;

c. collect, collate, process, analyse, present and interpret numeric and statistical data;

d. analyse statistical and financial data for planning and decision-making purposes;

e. Use mathematical techniques of the Operative Research to allocate resources judiciously; and

f. Apply mathematical models to real life situations and to solve problems involving choice among alternatives.

STRUCTURE OF PAPER:

The paper will be a three-hour paper divided into two sections.

Section A (50 Marks): This shall consist of 50 compulsory questions made up of 30 multiple-choice Questions and 20 short answer questions covering the entire syllabus.

Section B (50 marks) six questions out of which candidates are expected to answer only four, each at 12½ marks.

CONTENTS:
1. **STATISTICS   40%**
(a)   **Handling Statistical Data   8%**
    (i)   Collection of Statistical Data
        -     primary and secondary data
        -     discrete and continuous data
        -     sources of secondary data: advantages and disadvantages
        -     internal and external sources of data
        -     mail questionnaire, interview, observation, telephone: advantages and disadvantages of each method.
    (ii) Sampling Methods
        -     purpose of sampling
        -     methods of sampling: simple, random, stratified, systematic, quota, multistage, cluster
        -     advantages and disadvantages of each method
    (iii) Errors and approximations
        -     errors, level of accuracy and approximations
        -     types of errors: absolute, relative, biased and unbiased.
        -     laws of error including simple calculations of errors in sum, difference, product and quotient
    (iv) Tabulation and Classification of Data
        -     tabulation of data including guidelines for constructing tables
    (v)   Data Presentation

- frequency table construction and cross tabulation
- charts: bar charts (simple, component, percentage component and multiple), pie chart, Z- chart and Gantt chart
- graphs: histogram, polygon, Ogives, Lorenz curve

(b) **Measures of Location**   3%
    (i)   Measures of Central Tendency
         - arithmetic mean, median, mode, geometric and harmonic means
         - characteristic features of each measure
    (ii) Measures of partition
        - percentiles, deciles and quartiles

(c) **Measures of Variation/Spread/Dispersion**   2%
        - range, mean deviation, variation, standard deviation, coefficient of variation, quartile deviation and skewness (all both grouped and ungrouped data)
        - estimation of quartiles and percentiles from Ogives

(d) **Measures of Relationships**   3%
    (i)   Correlation (Linear)
         - Meaning and usefulness of correlation
         - scatter diagrams, nature of correlation (positive, zero, Negative)
         - meaning of correlation coefficient and its determination and interpretation
         - rank correlation such as spearman"s rank correlation coefficient, pearson product moment correlation.
    (ii) Regression Analysis (Linear)
         - normal equations/least squares method and the determination of the regression line
         - interpretation of regression constant and regression coefficients
         - use of regression line for estimation purposes

(e) **Time Series**   6%
    (i)   Meaning of time series
    (ii) Basic components and the two models
    (iii) Methods for measuring trend i.e. graphical, moving averages, least squares, semi-averages
    (iv) Methods of determining seasonal indices i.e. average percentage, moving average, link relative, ratio to trend and smoothening

(f) **Index Numbers**   6%
    (i)   meaning
    (ii) problems associated with the construction of index numbers.
    (iii) unweighted index i.e. sample aggregative index, mean of price. relatives.
    (iv) Weighted index numbers e.g. use Laspeyre, Paasche, Fisher and Marshall Edgeworth.

(g) **Probability**   4%
    (i)   Definition of probability
    (ii) Measurement (addition and multiplication laws applied to mutually exclusive, independent and conditional events)
    (iii) Mathematics expectation

(h) **Estimation and Significance Testing**   8%
    (i)   Interval Estimation
         - confidence interval concept and meaning
         - confidence interval for single population mean and single population proportions.
         - point estimation for mean, proportion and standard error
    (ii) Hypothesis
         - Concept and meaning
         - types (Null and alternative)
    (iii) Type I and type II errors; level of significance
    (iv) Testing of hypothesis about single population mean and single proportions for small and large samples
    (v) Sampling distribution of sample means and single proportions including their standard errors.

2.   **BUSINESS MATHEMATICS**   27%
    (a) **Functional Relationships**   **10%**
        (i)   definition of a function
        (ii) types of functions: linear, quadratic, logarithmic, exponential and their solutions including graphical treatment

      (iii) applications involving cost, revenue and profit functions
      (iv) break-even analysis
      (v)   determination of break-even point in quantity and value, significance of break-even point.
      (vi) simple linear inequalities not more than two variables including graphical approach

(b) **Mathematics of Finance**   **8%**
      (i)   Sequences and series (limited to arithmetic and geometric progressions), sum to infinity of a geometric progression (business applications)
      (ii)   simple and compound interests
         -    present value of simple amount
         -    present value of a compound amount
      (iii) Annuities
         -    types of annuities e.g. ordinary and annuity due
         -    sum of an ordinary annuity (sinking funds)
         -    present value of an annuity
      (iv) Net Present Value (NPV)
      (v)   Internal Rate of Return (IRR)

(c) **Differentiation**   **6%**
      (i)   meaning of slope or gradient or derivative
      (ii)   rules for differentiating the following functions: power (e.g. $y=ax_n$), product, quotient, function of function, exponential, implicit and logarithmic functions
      (iii) applications of a differential e.g. funding marginals, elasticity, maximum and minimum values
      (iv) simple partial differentiation

(d) **Integration**   **3%**
      (i)   rules for integrating simple functions only
      (ii)   applications of integration in business e.g. finding total functions from marginal functions, determination of consumers and producers surpluses

3.   **OPERATIONS RESEARCH**   **33%**

(a) **Introduction**   **3%**
      (i)   main stages of an Operation Research (OR) project
      (ii)   relevance of Operations Research in business

(b) **Linear Programming**   **8%**
      (i)   concept and meaning (as a resource allocation tool)
      (ii)   underlying basic assumptions
      (iii) problem formulation in linear programming
      (iv) methods of solution
         -    graphical methods (for 2 decision variables)
      (v)   interpretation of results
         -    Results from tableau
         -    Results from simplex method, shadow price, marginal value, worth of resources
         -    Determination of dual/shadow costs

(c) **Inventory and Production Control**   **8%**
      (i)   Meaning of an inventory
      (ii)   Functions of inventory
      (iii) Inventory costs e.g. holding cost, ordering costs, shortage costs, cost of materials.
      (iv) General inventory models e.g. deterministic and stochastic model: periodic review system and re-order level system
      (v)   Basic Economic Order Quantity (EOQ) model including assumptions of the model

(d) **Network Analysis**   **6%**
      (i)   Critical Path Analysis (CPA) and Programme Evaluation and Review Technique (PERT)
      (ii)   Drawing the network diagram
      (iii) Meaning of critical path and how to determine it and its duration
      (iv) Calculation of floats or spare times

(e) **Replacement Analysis**   **3%**
      (i)   Replacement of items that wear gradually
      (ii)   Replacement of items that fail suddenly

(f) **Transportation Model**   **5%**
      (i)   Nature of transportation models

(ii) Balanced and unbalanced transportation problems
(iii) Methods for funding initial basic feasible transportation cost: North West Corner Method (NWCM), Least Cost Method (LCM), and Vogel"s Approximation Method (VAM)

RECOMMENDED TEXTS:
1. ATSWA Study Pack on Quantitative Analysis

2. Adamu, S. O. and Johnson T. L.: Statistics for Beginners, Evans Nigeria

OTHER REFERENCE BOOK

Donald H. Saders: Statistics, A Fresh Approach, McGraw-Hill

SECTION A

STATISTICS

CHAPTER ONE

STATISTICS, DATA COLLECTION AND SUMMARY

1.0   Learning objectives

At the end of the chapter, students and readers should be able to

- know the meaning of statistics

- know types of data, how to collect and classify them

- understand the concept of data presentation in form of charts and graphs

- understand various methods of sampling

- distinguish between error and approximation.

1.1   Introduction

Statistics is a scientific method that concerns data collection, presentation, analysis, interpretation or inference about data when issues of uncertainties are involved. Definitely, statistics is useful and needed in any human area of endeavour where decision making is of vital importance. Hence, it is useful in Accountancy, engineering, Education, Business, social Sciences, Law, Agriculture, to mention a few.

Statistics can be broadly classified into two:

i.   Descriptive   and   ii.   Inferential.

Descriptive statistics deals with data collection, summarizing and comparing numerical data, while inferential statistics deals with techniques and tools for collection of data from a population. These data are then studied by taking a sample from the population. By doing this, knowledge of the population characteristics is gained and vital decisions are made about the sampled population.

1.2   Types of data

Data are the basic raw facts needed for statistical investigations. These investigations may be needed for planning, policy implementation, and other purposes.

Data can be quantitative or qualitative. It is quantitative when the data set have numerical values. Examples of quantitative data include:

i.	heights of objects,	ii.	Weights of objects

iii.   prices of goods and services,

iv.   The family expenditure on monthly basis.

The quantitative data, which are also termed numeric data, can be further classified into two:

a.   **Discrete:** The data values are integers (whole numbers). Examples include: (i) number of books sold in bookshop, (ii) number of registered ABWA students.

b.   **Continuous:** the data values are real numbers which can be integer, fraction or decimal. Examples include: (i) heights of objects; (ii) weights of objects; (iii) price of goods.

For the qualitative data, it is the set of data which have no numerical values and can not be adequately represented by numbers. They are also called non-numerical data. They are further classified into two:

a.   **Ordinal Non-numeric:** The data set is on ordinal scale when ranking for is allowed. Example can be seen in the beauty contest judgment.

b.   **Nominal Non-numeric:** The data set is in categorical form where the facts collected are based on classification by group or category. Examples include: (i) level of education, (ii) sex, (iii) occupation.

1.2.1 Types of data by generation source

The data source can be broadly classified into: primary and secondary

Primary data: These data originate or are obtained directly from respondents for the purpose of an initial investigation.

Secondary data: These are data obtained, collected, or extracted from already existing records or sources. They are derived from existing published or unpublished records of government agencies, trade associations, research bureaus, magazines and individual research work.

It will be of great benefit to know that data source can be either internal or external. By internal source, the data collected are within the organization. Such data are used within the organization. Examples are(i) the sales receipts, (ii) invoice, (iii) work schedule.

In the external source, the data are collected outside the organization. The data are generated or collected from any other places which are external to the user.

1.3   Methods of data collection

The major practising methods of data collection are: The interview, Questionnaire, Observation, Documents/Reports and Experimental methods.

1.3.1   Interview method

This is a method in which the medium of data collection is through conversation by face-to-face contact, through a medium like telephone.

In general, the interview method can be accomplished by:

i.   Schedule;   ii.   Telephone;   and   iii.   Group discussion.

i.   **Interview schedule:** This is the use of schedule by the interviewer to obtain necessary facts/data. A schedule is a form or document which consists of set a of questions to be completed by interviewer as he/she asks the respondent questions.

ii.   **Telephone interview:** Here the questions are asked through the use of telephone in order to get the needed data. The respondents and interviewers must have access to telephone.

iii.   **Interview by Group discussion:** The interview is conducted with more than one person with a focus on a particular event.

1.3.2   Questionnaire Method

The interviewers send the copies of their questionnaires to the respondents to be filled without necessarily being present there. A questionnaire is a document that consists of a set of leading questions which are logically arranged and are to be filled by the respondent himself.

The types of questions in questionnaire can be classified into two; namely:

a.   Close – ended or coded question is the type of questions in which alternative answers are given for the respondent to pick one.

b.   Open-ended or uncoded questions is the type of questions in which the respondent is free to give his own answer and is not restricted to some particular answers.

In drafting a questionnaire, the following qualities are essentials to note:

i.   A questionnaire must be well structured so that major sections are available. The first section usually deals with personal data while other sections consider necessary and relevant questions on the subject matter of investigation.

ii.   A questionnaire must be clear in language, not ambiguous and not lengthy.

iii.   A questionnaire question should not be a leading question.

1.3.3 Observation Method

This method enables one to collect data on behaviour, skills etc. of persons, objects that can be observed in their natural ways. Observation method can either be of a controlled or uncontrolled type. It is controlled when issues to be observed are predetermined by rules and procedures; while otherwise, it is termed uncontrolled type.

1.3.4 Experimental Method

Here, experiments are carried out in order to get the necessary data for the desired research. There are occasions when some factors are to be controlled and some not controlled in this method. The method is mostly used in the sciences including engineering.

1.3.5 Documents/Reports

This method allows one to check the existing documents/reports from an organization.

1.4 Presentation of Data

Statistical data are organized and classified into groups before they are presented for analysis. Four important bases of classification are:

i.   Qualitative – By type or quality of items under consideration.

ii. Quantitative – By range specified in quantities.

iii. Chronological – Time series – Monthly or Yearly: An analysis of time series involving a consideration of trend, cyclical, periodic and irregular movements.

iv. Geographical – By location.

The classified data are then presented in one of the following three methods: (a) Text presentation, (b) Tabular presentation and (c) Diagrammatic presentation:

1.4.1 Text presentation:

This is a procedure by which text and figures are combined. It is usually a report in which much emphasis is placed on the figures being discussed.

For instance, a text presentation can be presented as follows:

The populations of science and management students are 3000 and 5000 respectively for year 2006 in Osun State Polytechnic, Iree.

1.4.2 Tabular presentation:

A table is more detailed than the information in the text presentation. It is brief and self explanatory. A number of tables dealt with in statistical analysis are – general reference table, summary table, Time series table, frequency table. Tables may be simple or complex. A simple table relates a single set of items such as the dependent variable against another single set of items – the independent variable. A complex table on the other hand has a number of items presented and often shows sub-divisions.

Essential features of a table are:

- A title to give adequate information about it.

- Heading for identification of the rows and columns.

- Source about the origin of the figures; and

- Footnote to give some detailed information on some figures in the table.

Example 1.4.2.1: A typical example of simple table.

Classification of two hundred Polytechnic students on departmental basis.

Department	No of Students
Accountancy	60

Business Administration	50
Marketing	40
Banking / Finance	50
Total	**200**

Example 1.4.2.2: A Typical Example of Complex Table

Departmental classification of 200 University students on the basis of gender.

Department	No of Students		Total
	Male	**Female**	
Accountancy	40	20	60
Business Administration	36	14	50
Marketing	30	10	40
Banking / Finance	24	26	50
Total	**130**	**70**	**200**

Example 1.4.2.3: Formulation of Frequency Table

The following marks are obtained by 30 students in accountancy department out of a total marks of 100:

6,   10,   16,   60,   50,   30,   70,   14,   32,   80,   98,   80,   28,

32,   60,   46,   42,   44,   58,   54,   50,   64,   30,   22,   18,   16,

86,   56,   64,   28.

Prepare a frequency table with class intervals of 1 – 10, 11 – 20, 21 – 30, etc.

Solution

Class Intervals	Tally Bars	Frequency
1 – 10	||	2
11 – 20	||||	4
21 – 30	⊬⊬	5
31 – 40	||	2

41 – 50	┼┼┼┼	5
51 – 60	┼┼┼┼	5
61 – 70	|||	3
71 – 80	||	2
81 – 90			1
91 – 100			1
		N = 30

1.4.3 Diagrammatic Presentation:

Diagrams are used to reflect the relationship, trends and comparisons among variables presented on a table. The diagrams are in form of charts and graphs.

Examples are:

i.   The Bar charts

-        Simple bar chart; histogram; component bar chart

-        Percentage component bar charts;

-        Multiple bar charts

ii.   Pie Charts.

iii.   Pictograms

iv.   Statistical maps and area diagrams

v.   Graphs – a graph shows relationship between variables concerned by means of a curve or a straight line. A graph will, for example, show the relationship between output and cost, or the amount of sales to the time the sales were made.

Typical graphs used in business are the cumulative frequency curve, histogram based on a time series data; the z – charts and the net balance charts; break-even chart and progress chart.

Example 1.4.3.1: Simple Bar Chart
Draw simple bar chart for the table in Example 1.4.2.1. i.e:

Department	No of Students

Accountancy	60
Business Administration	50
Marketing	40
Banking / Finance	50
Total	**200**
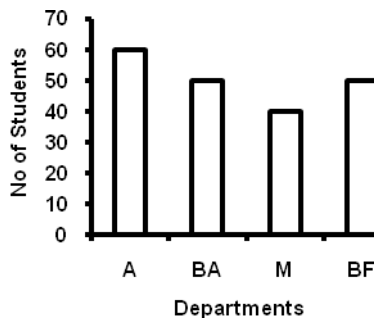
Legend

Departments	Key
Accountancy	A
Business Admin	BA
Marketing	M
Banking/Finance	B/F

Example 1.4.3.2: Multiple Bar Chart

Draw Multiple bar chart for the table in Example 1.4.2.2. i.e:

Department	No of Students		Total
	Male	**Female**	
Accountancy	40	20	60
Business Administration	36	14	50
Marketing	30	10	40
Banking / Finance	24	26	50
Total	**130**	**70**	**200**

Example 1.4.3.3: Pie Chart

Draw Pie chart for the table in Example 1.4.2.1.

Solution

Calculation of angles

Total no of students is 200

∴ For Accountancy, the corresponding angle is $= \dfrac{60}{200} \times \dfrac{360}{1} = 108^0$

For Business Administration, the corresponding angle is $= \dfrac{50}{200} \times \dfrac{360}{1} = 90^0$

For Marketing, the corresponding angle is $= \dfrac{40}{200} \times \dfrac{360}{1} = 72^0$

For Banking / Finance, the corresponding angle is $= \dfrac{50}{200} \times \dfrac{360}{1} = 90^0$
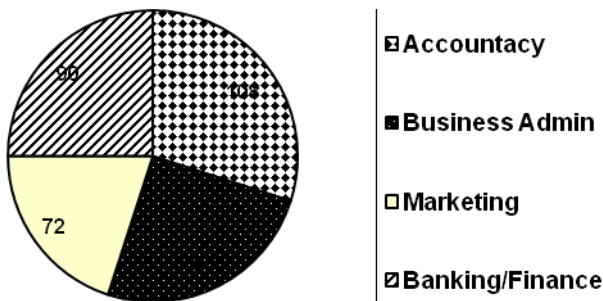
Check: $108^0 + 90^0 + 72^0 + 90^0 = 360^0$

Pie chart

Example 1.4.3.4: Component Bar Chart

Draw the percentage component bar chart from the table below.

	Accountancy	Banking/Finance	Total
Male	40	20	60
Female	30	10	40
	70	30	100

Solution

	Accountancy	%	Banking/Finance	%	Total
Male	40	57%	20	67%	60
Female	30	43%	10	33%	40
	70	100	30	100	100
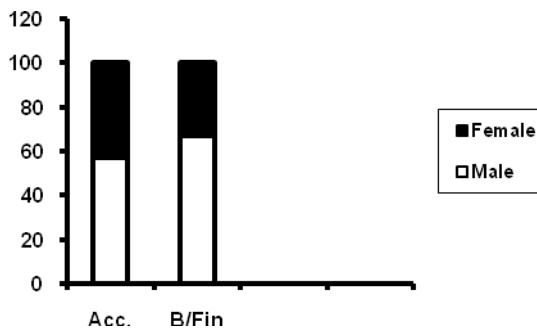
Example 1.3.4.5:

From Example 1.4.2.3 draw the Histogram for the frequency distribution.

Class Interval	Frequency
1 – 10	2
11 – 20	4
21 – 30	5
31 – 40	2
41 – 50	5
51 – 60	5
61 – 70	3
71 – 80	2
81 – 90	1
91 – 100	1
	30

Solution

Class Interval	Class Boundaries	Frequency
1 – 10	0.5 – 10.5	2
11 – 20	10.5 – 20.5	4
21 – 30	20.5 – 30.5	5
31 – 40	30.5 – 40.5	2
41 – 50	40.5 – 50.5	5
51 – 60	50.5 – 60.5	5
61 – 70	60.5 – 70.5	3
71 – 80	70.5 – 80.5	2
81 – 90	80.5 – 90.5	1
91 – 100	90.5 – 100.5	1

The Histogram
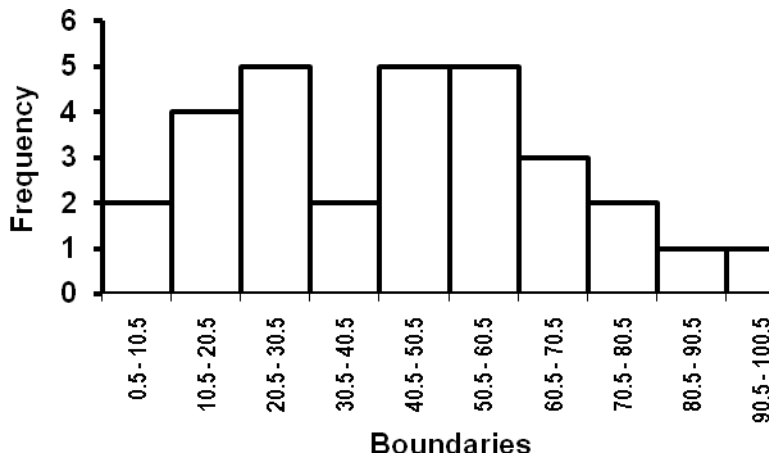
Example 1.4.3.6:

From Example 1.4.2.3 construct the cumulative frequency (Ogive).

Boundaries	Frequency	C. F
0.5 – 10.5	2	2
10.5 – 20.5	4	6
20.5 – 30.5	5	11
30.5 – 40.5	2	13

40.5 – 50.5	5	18
50.5 – 60.5	5	23
60.5 – 70.5	3	26
70.5 – 80.5	2	28
80.5 – 90.5	1	29
90.5 – 100.5	1	30
	30	
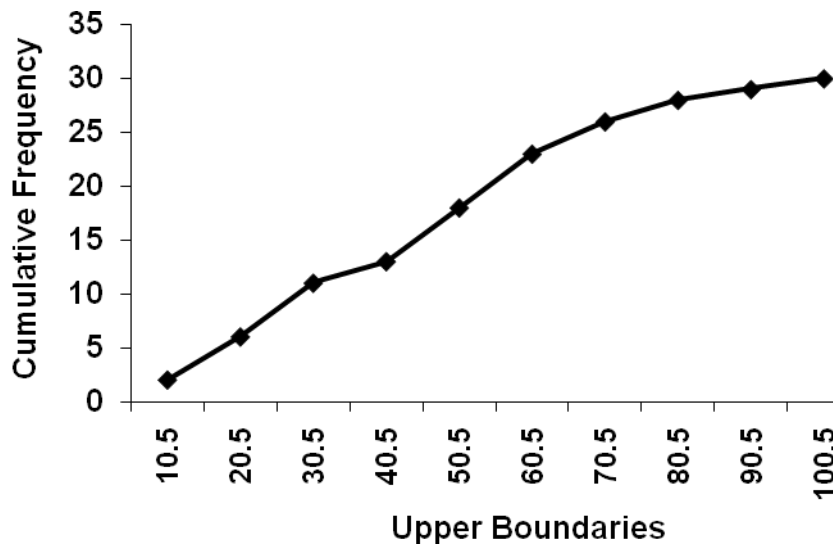
1.5 Sampling Techniques

Sampling technique is an important statistical method that entails the use of fractional part of population to study and make decisions about the characteristics of the given population. As an introduction to the sampling techniques, the following concept shall be discussed:

1.5.1 Definition of some important concepts on sampling techniques

i. **Population:** a population is a collection or group of items satisfying a definite characteristic. Items in the definition can be referred to

    a.   human beings (as population census);

    b.   animate objects as cattle, sheep, goats etc

    c.   inanimate objects such as chairs, tables, etc.

d.   a given population like students of ATS.

There are two types of population, namely: finite and infinite populations:

By finite population, we mean the population which has countable number of items; while the items in infinite population are uncountable. Population of students in a college is an example of finite population while sand particles under beach are that of infinite population.

ii.   **Sample:** A sample is the fractional part of a population selected in order to observe or study the population for the purpose of making scientific statement or decision about the said population.

iii.   **Census:** This is a complete enumeration of all units in the entire concerned population with the aim of collecting vital information on each unit. Year 2006 head count in Nigeria is a good example of census.

iv.   **Sample Survey:** A sample survey is a statistical method of collecting information by the use of fractional part of population as a representative sample.

v.   **Sampling frame:** This is a list containing all the items or units in the target population and it is normally used as basis for the selection of sample. Examples are: list of church members, telephone directory, etc.

vi.   Some sampling notations and terms:

a.   Let "N" represent the whole population, then the sample size is denoted by "n".

b. $f = {}^n/_N$, where f is called the sampling fraction, N and n are as defined in vi(a).

c.   The expansion or raising factor is given as $g = {}^N/_n$, where g is the expansion factor, N and n are as defined in vi(a).

1.5.2   Types of sampling techniques

Sampling techniques or methods can be broadly classified into two types; namely:

i.   probability sampling methods;

ii.   purposive sampling method.

1.5.2.1 The probability sampling method

The probability sampling technique is a method involving random selection. It is done in such a way that each unit of the population is given a definite chance or probability of being included in the sample. Among these sampling techniques are simple random sampling, systematic sampling, stratified sampling, multi-stage sampling and cluster sampling.

i. **Simple Random Sampling (SRS):** This is a method of selecting a sample from the population with every member of the population having equal chance of being selected. The sampling frame is required in this technique.

The method can be achieved by the use of

a. table of random numbers;

b. lottery or raffle draw approach.

We are to note that SRS can be with replacement or without replacement. When it is with replacement, a sample can be repeatedly taken or selected; while in the SRS without replacement, a sample can not be repeatedly taken or selected.

Advantages of Simple Random Sampling

a. SRS gives equal chance to each unit in the population which can be termed a fair deal.

b. SRS method is simple and straight forward.

c. SRS estimates are unbiased.

Disadvantages of Simple Random Sampling

a. The method is not good for a survey if sampling frame is not available.

b. There are occasions when heavy drawings are made from one part of the population, the idea of fairness is defeated.

ii. **Systematic Sampling (SS):** This is a method that demands for the availability of the sampling frame with each unit numbered serially. Then from the frame, selection of the samples is done on regular interval K after the first sample is selected randomly within the first given integer number K; where

K $=$ N/n, and must be a whole number, and any decimal part of it should be cut off or cancelled. N and n are as earlier defined. That is, after the first selection within the

K interval, other units or samples of selection will be successfully equi-distant from the first sample (with an interval of K).

For instance, in a population N = 120 and 15 samples are needed, the SS method will give us

$$K \quad = \quad {}^{120}/_{15} \quad = \quad 8$$

It means that the first sample number will be between 1 and 8. Suppose serial number 5 is picked, then the subsequent numbers will 13, 21, 29, 37, ---. (Note that the numbers are obtained as follows: $5 + 8 = 13, \ 13 + 8 = 21, \ 21 + 8 = 29, \ 29 + 8 = 37$ etc).

Advantages of Systematic Sampling

    a.   It is an easy method to handle.

    b.   The method gives a good representative if the sampling frame is available.

Disadvantages
    a.   If there is no sampling frame, the method will not be appropriate.

iii.   **Stratified Sampling (STS):** This is a method that is commonly used in a situation where the population is heterogeneous. The principle of breaking the heterogeneous population into a number of homogeneous groups is called stratification. Each group of unique characteristics is known as stratum and there must not be any overlapping in the groups or strata. Some of the common factors that are used to determine the division or breaking into strata are income levels, employment status, etc.

The next step in stratification is the selection of samples from each group (stratum) by simple random sample. An example of where to apply STS can be seen in the survey concerning standard of living or housing pattern of a town.

Advantages of STS
    a.   By pooling the samples from the strata, a more and better representative of the population is considered for the survey or investigation.

    b.   Breaking the population into homogeneous goes a long way to have more precise and accurate results.

Disadvantages of STS

a.   There are difficulties in deciding the basis for stratification into homogeneous groups.

b.   STS suffers from the problems of assigning weights to different groups (strata) when the condition of the population demands it.

iv.   **Multi – stage sampling:** This is a sampling method involving two or more stages. The first stage consists of breaking down the population into first set of distinct groups and then select some groups randomly.

The list of selected groups is termed the primary sampling units. Next, each group selected is further broken down into smaller units from which   samples are taken to form a frame of the second-stage sampling units. If we stop at this stage, we have a two-stage sampling.

Further stages may be added and the number of stages involved is used to indicate the name of the sampling. For instance, five-stage sampling indicates five stages are involved.

An example where the sampling method [MSS] can be applied is the survey of students' activities in high institutions of a country. Here, first get the list   of all higher institutions in the country and then select some randomly. The next stage is to break the selected institutions into faculties and select some. One   can go further to break the selected faculties into departments and select some departments to have the third-stage sampling.

Advantages Multi-stage sampling(MSS)

1. The method is simple if the sampling frame is available at all stages.
2. It involves little cost of implementation because of the ready-made availability of sampling frame.

Disadvantages of MSS

1. If it is difficult to obtain the sampling frame, the method may be tedious.
2. Estimation of variance and other statistical parameters may be very complicated.

v.   **Cluster sampling [CS]:**   Some populations are characterized by having their units existing in natural clusters. Examples of these can be seen in farmers' settlement (farm

settlements are clusters); students in schools (schools are clusters). Also, there could be artificial clusters when high institutions are divided to faculties.

The cluster method involves random selection of A clusters from M clusters in the population which represents the samples.

1.5.2.2 Non-Probability Sampling

The sampling techniques which are not probabilistic are **Purposive Sampling (PS)**
In this sampling, no particular probability for each element is included in sample. It is at times called the judgment sampling.

In purposive sampling, selection of the sample depends on the discretion or judgment of the investigator. For instance, if an investigation is to be carried out on students' expenditures in the university hostels, the investigator may select those students who are neither wiser nor extravagant in order to have a good representation of the targeted population.

Advantage of PS
i.   It is very cheap to handle.

ii.   No need of sampling frame.

Disadvantage of PS
i.   The method is affected by the prejudices of the investigator.

ii.   Estimates of sampling errors cannot be possible by this method.

a.   Quota Sampling (QS)

This is a sampling technique in which the investigator aimed at obtaining some balance among the different categories of units in the population as selected samples. A good example of this can be seen in the quota selection of students to Federal Colleges in Nigeria.

Advantages of QS
i.   It has a fair representation of various categories without probability basis.

ii.   No need for sampling frame.

Disadvantages of QS
i.   Sampling error cannot be estimated.

1.6   Errors and Approximation

Approximation and errors are inevitable phenomena in data. Hence, the two concepts shall be examined in the following subsequent sections.

1.6.1 Approximation:

The results of the measurements, counting, recordings and observation in data collection, and even in computations are subjected to approximations due to a certain degree of accuracy desired. In the light of this, the concept of approximation is relevant and necessary to know or determine the degree of accuracy to be maintained in data collection and various computations.

Formally, approximation can be defined as a technique of rounding off a number so that the last digit(s) is(are) affected in order to make the number clearer and more understandable.

For instance, let's compare these two statements:

a.   the total number of people in a town is 515,176;

b.   the total number of people in a town is 515,000

There is no doubt that statement (b) is more comprehensive than statement (a) even though statement (a) is more accurate.

Basically, the approximation concept is achieved by rounding off statistical figures or numbers to:

i.   specific nearest units e.g to the nearest 100,

ii.   specific significant figures e.g to three significant figures.

Furthermore, the principle of rounding off can be done by either

i.   rounding up; where any figure which is up to half of the required specification unit is rounded up to a whole unit; or

ii.   rounding down, where any figure or number is not up to half of the required specification unit is rounded down to zero unit.

The following examples shall be used to demonstrate the principle of approximation.

Example 1.6.1.1

Approximate 514,371 to: (a)   to the nearest 1,000

                                             (b)   to the nearest 100

                                             (c)   to 3 significant figures

                                             (d)   to 4 significant figures.

Solution:

a.   514,000 (because 371 is not up to half of 1000)

b.   514,400 (because 71 is up to half of 100)

c.   514,000 (because 3 of hundred unit is the fourth digit and is not up to 5)

d.   514,400 (because of ten unit is the fifth digit and is up to 5)

Example 1.6.1.2

Approximate the following:

a.   154.235 to 2 decimal places and 2 significant figures.

b.   0.02567 to 3 decimal places and 3 significant figures.

Solution

a.   i.   To 2 decimal places   = 154.24

      ii.   To 2 significant figures   = 150.000

b.   i.   To 3 decimal places   = 0.026

      ii.   To 3 significant figures   = 0.0257

1.6.2   Errors

1.6.2.1 In statistical thinking, the word "error" is used in a very restricted sense. It denotes the difference between the estimation (through measurements or computations as obtained by the investigator) and the actual size of the object or item under investigation.

Sources of some errors can be traced to:

i.   Origin: where for example, precision in measuring is not possible due to limitations of the measuring instruments;

ii.   Inadequacy: which can be due to small samples or insufficient coverage;

iii.   Manipulation, which can be seen as errors unconsciously committed by enumerators in measuring the object;

iv.   Calculation: with all arithmetic mistakes and other calculation errors.

Errors can be broadly classified into two; namely:

a.   Sampling error and;   b.   Non-sampling error.

Sampling Error, there are series of errors that can be committed due to sampling principles. When these accumulated errors are combined, they form the sampling errors. Suppose in a class of 30, a sample of five is taken in order to obtain necessary information about the class.   The difference between the estimate based on the sample (of five students) and the actual class value is termed the sampling error. At times, sampling error is called "sampling variability".

For non-sampling errors, these are errors committed which are not due to sampling principles. Therefore, the term "non-sampling errors" refers to errors in measurements, counting, assessments, judgments, calculations (including arithmetic mistakes), interpretations, instructions and even in machines. For instance, there exists only non-sampling error in census while both sampling and non-sampling errors are present in sample survey.

1.6.2.2 Measurement of errors:

Errors can be measured as absolute, relative and percentage.

Let   x be the observed (estimated) value and

        x + e be the true value, then

error   = $(x + e) - x = e$ ---1.6.2.2.1

This error(e) in equation (1.6.2.2.1), which is the difference between observed and true value is known as the <u>absolute Error</u>; and it is dependent upon the x unit of measurement.

The <u>Relative Error</u> is the absolute error divided by the true value.

i.e   relative Error = $^e/_{(x + e)}$ ------------1.6.2.2.2

While the <u>Percentage Error</u> = Relative Error x 100

The relative and percentage errors are both independent of the unit of measurement. Therefore, they are good and true indices of measurement or calculation accuracy.

Example 1.6.2.2.1
A population with 518,413 as actual number has an estimated figure of 518,000. Determine the:

a.   Absolute Error;   b.   Relative Error

c.   Percentage Error.

Solution
a.   Using equation (1.6.2.2.1), the

Absolute error   =   518,413 – 518,000   = 413.

b.   Using equation (1.6.2.2.2), the

Relative error   = $\underline{413}$ =   0.0007973
                        518000

c.   Percentage error   =   0.0007973 x 100%

                                =   0.07973%

1.7   Summary

Data are defined as raw facts in numerical form. Its classification by types, method of collection and the forms of presentation are discussed. Some sampling techniques with their advantages and disadvantages are presented. The issues of approximation and errors are also presented.

1.8   Multiple choice and short – answer Questions

1.   Which of the following is a non-numeric ordinal data?

a.   income   b.   price of commodity   c.   occupation

d.   rating in beauty contest   e.   students number in a class

2.   A schedule in statistics refers to_____

a.   examination time table

b.   a set of questions used to gather pieces of information which is filled by an informant or respondent.

c.   a set of questions used to gather information and filled by the investigator him/herself.

d.   a set of past examination questions.

e.   paper used by bankers to carry out investigation.

3.   Which of the following sampling methods does not need sampling frame?

a.   simple random sampling.   b.   purposive sampling

c.   systematic sampling   d.   cluster sampling

e.   stratified sampling.

4.   The following are qualities of a good questionnaire except

a.   that each question in the questionnaire must be precise and unambiguous.

b.   avoidance of leading question in the questionnaire.

c.   that a questionnaire must be lengthy in order to accommodate many questions.

d.   that a questionnaire must be well structured into sections such that questions in each section are related.

e.   avoidance of double barrelled questions in the questionnaire.

5.   In a sampling technique, a list consisting of all units in a target population is

known as__

6.   A small or fractional part of a population selected to meet some set objective is known as_____

7.   Data presentation in form of charts and graphs is known as_____

8.   A procedure in the form of report involving a combination of text and figures is known as_____

9.   The act of rounding off data is known as_____

10.   The difference between the actual and estimated value is known as_____

Answers to Chapter 1

1. d,   2. b,   3. b,   4. c,   5. Sampling frame.

6. Sample,   7. Diagrammatic presentation

8.   Text presentation,   9.   Approximation

10.   Absolute error

CHAPTER TWO

MEASURES OF LOCATIONS

2.0 Learning objectives

At the end of the chapter, students or readers should be able to

- know the meaning of "Measure of Central Tendency".
- understand and solve problems on various means such as Arithmetic mean, Geometric mean and Harmonic mean.
- know and handle problems on the mode and median.
- understand the concept of measures of partition.
- handle problems on quintiles, deciles and percentiles.

2.1 Introduction

Measure of location is a summary statistic which is concerned with a figure which represents a series of values. Measures of location can be further classified into

a. Measures of Average or Measures of Central Tendency.
b. Measures of Partition.

In average or measures of central tendency, there is an average value which is a representative of all the values in a group of data. It is used to give these typical values of averages which usually tend to lie centrally within the set of data arranged in an array, hence they are called measures of central tendency.

The common and usable types of averages or measures of central tendency are

a. Mean; consisting of the following types: Arithmetic Mean (AM), Geometric Mean (GM) and Harmonic Mean (HM).
b. Mode
c. Median

Each of the above averages will be discussed in the subsequent sections.

2.2 Mean

2.2.1 Arithmetical Mean (A.M):-

This is the sum of all the values $(x_1, x_2, ----, x_n)$ in the data group divided by the total number of the values. In symbolic form, for individual series (as case 1) the mean can be expressed as

$$\bar{x} = \frac{\Sigma x_i}{n} = \frac{x_1 + x_2 + ... + x_n}{n} \qquad ....2.2.1$$

where $\bar{x}$ represents the arithmetical mean,

Σ is the summation usually called "sigma" and

n is the number of the values

In the case of discrete or ungrouped values with frequency distribution, the above formula can be modified as:

$$\bar{x} = \frac{\Sigma f_i x_i}{\Sigma f_i} \qquad \text{or simply} \qquad \bar{x} = \frac{\Sigma fx}{\Sigma f} \qquad ...2.2.2$$

where, $\Sigma f = n$ (as defined as 2.2.1) and f represents frequency.

Also for grouped frequency distribution table, formula 2.2.2 will still be equally applied.

However, the x_i values in the formula are obtained by the use of Mid-Value. This is also called class mark. The class mark is obtained by adding the lower and upper boundary limits of the class interval and then divide by two.

i.e   Class Mark $= \dfrac{L_B + U_B}{2}$ $\qquad ...(2.2.3)$

where L_B = Lower boundary / limit.

U_B = Upper boundary / limit.

Example 2.2.1.1(a)

From the underlisted data generated on the number of sales of cement (in bags): 12, 7, 2, 6, 13, 17, 14, 5, 9, 4, determine the mean

$$\bar{x} = \frac{\Sigma x_i}{n} = \frac{x_1 + x_2 + ... + x_{10}}{n}$$

$$\bar{x} = \frac{12 + 7 + 2 + 6 + 13 + 17 + 14 + 5 + 9 + 4}{10}$$

$$\bar{x} = \frac{89}{10} = 8.9$$

Example 2.2.1.1(b)

From the underlisted data generated on the volume (litres) of water: 2.7, 6.3, 4.1, 6.4, 3.5, 4.7, 13.8, 7.9, 12.1, 10.4, determine the mean

$$\bar{x} = \frac{\Sigma x_i}{n} = \frac{x_1 + x_2 + ... + x_{10}}{n}$$

$$\bar{x} = \frac{2.7 + 6.3 + 4.1 + 6.4 + 3.5 + 4.7 + 13.8 + 7.9 + 12.1 + 10.4}{10}$$

$$\bar{x} = \frac{71.9}{10} = 7.19$$

Example 2.2.1.2(a)

These are the data collected on the number of students that register for STA 415 course, obtain the mean for them.

X	1	2	3	4	5
F	8	7	10	6	2

$$\text{Mean } (\bar{x}) = \frac{\Sigma f_i x_i}{\Sigma f_i}$$

X	F	fx
1	8	8
2	7	14
3	10	30
4	6	24
5	2	10
	33	86

NOTE: This is a work in progress. All topics in the syllabus are covered but editing for necessary corrections is in progress.

Thanks.

$$\text{Mean } (\bar{x}) \quad = \quad \frac{86}{33} \quad = \quad 2.6061$$

Example 2.2.1.2(b)

These are the data generated on the weight (kilogram) of soap produced in the production section, obtain the mean for them.

Weight (x)	1.0	2.0	3.0	4.0	5.0
Frequency (f)	4	3	7	5	6

Solution

$$\text{Mean } (\bar{x}) \quad = \quad \frac{\Sigma f_i x_i}{\Sigma f_i}$$

x	f	fx
1	4	4
2	3	6
3	7	21
4	5	20
5	6	30
	25	81

$$\text{Mean } (\bar{x}) \quad = \quad \frac{81}{25} \quad = \quad 3.24$$

Example 2.2.1.3

The following are data collected on the ages of the people living in Opomulero village.

Class	0 – 10	10 – 20	20 – 30	30 – 40	40 – 50
Frequency	12	8	7	7	6

Obtain the mean age of the data.

Solution

$$\text{Mean} \ (\bar{x}) \quad = \quad \frac{\Sigma f_i x_i}{\Sigma f_i}$$

Class	X	f	fx
0 – 10	5.0	12	60
10 – 20	15.0	8	120
20 – 30	25.0	7	175
30 – 40	35.0	7	245
40 – 50	45.0	6	270
		40	870

$$\text{Mean} \ (\bar{x}) \quad = \quad \frac{870}{40} \quad = \quad 21.75$$

Example 2.2.1.4

From the following data generated on the ages of people living in an Estate.

Class	1 – 10	11 – 20	21 – 30	31 – 40	41 – 50
Frequency	6	17	7	12	8

Obtain the mean age of the data.

Solution

$$\text{Mean} \ (\bar{x}) \quad = \quad \frac{\Sigma f_i x_i}{\Sigma f_i}$$

Class	X	f	fx
1 – 10	5.5	6	33.0
11 – 20	15.5	17	263.5
21 – 30	25.5	7	178.5
31 – 40	35.5	12	426.0
41 – 50	45.5	6	273
		50	1174

Mean $(\bar{x})$ = $\dfrac{1174}{50}$ = 23.48

Remark I: It is possible to reduce the volume of figures involved in the computation of the mean by the use of "assumed mean" method. In this method, it requires that one of the observed values of x_i is taken. It is advisable that the middle value of x_i is taken.

If A denotes the assumed mean, a new variable d is introduced by the expression $d_i = x_i - A$, then the actual mean is obtained by

$\bar{x}$	=	$A + \dfrac{\Sigma d_i}{N}$	(for ungrouped data)	2.2.4
			OR	
$\bar{x}$	=	$A + \dfrac{\Sigma f_i d_i}{\Sigma f_i}$	(for grouped data)	2.2.5

Example 1: Using the data in Example 2.2.1.1(a), obtain the arithmetic mean by the assumed mean method.

Solution

x_i	$d_i = x_i - 9$	
12	3	
7	-2	
2	-7	
6	-3	
13	4	$\bar{x} = A + \dfrac{\Sigma di}{N}$
17	8	
14	5	$= 9 + \dfrac{(-1)}{10}$
5	-4	
9	0	$= 9 - 0.1$
4	-5	$= 8.9$
$\Sigma d_i = -1$		≈ 9

A = 9 is the assumed mean

Note that it gives the same answer as Example 2.2.1.1(a)

Example 2: Using the data in Example 2.2.1.3, obtain the mean age by the assumed mean method.

Solution

Class	x	f	If A = 25 $d_i = x - 25$	fd_i
0 – 10	5	12	-20	-240
10 – 20	15	8	-10	-80
20 – 30	25	7	0	0
30 – 40	35	7	10	70
40 – 50	45	6	20	120
		40		$\Sigma fd_i = -130$

$$\bar{x} = A + \frac{\Sigma fd_i}{\Sigma f_i}$$

$$= 25 + \frac{-130}{40}$$

$$= 25 - 3.25$$

$$= 21.75$$

which is the same as the answer obtained in Example 2.2.1.3.

Remark II: Assumed mean and common factor approach can equally be used to obtain the mean.
Here, it is only applicable to group data with equal class interval.

By this approach, equation 2.2.5 becomes:

$$\bar{x} = A + C\left(\frac{\Sigma f_i d_i}{\Sigma f_i}\right) \qquad\qquad 2.2.6$$

where C is common factor or the class interval.

Example: Using the data in Example 2.2.1.3, obtain the actual mean.

Solution

Class	X	F	If A = 25 $d_i = x - 25$	If C = 10 $d_i' = d_i/10$	fd_i'
0 – 10	5	12	-20	-2	-24
10 – 20	15	8	-10	-1	-8
20 – 30	25	7	0	0	0
30 – 40	35	7	10	1	7
40 – 50	45	6	20	2	12
		40			$\Sigma fd_i' = -13$

$$\bar{x} = A + \frac{C\Sigma f_i d_i'}{\Sigma f_i}$$

$$= 25 + \frac{10(-13)}{40}$$

$$= 25 - 3.25$$

$$= 21.75$$

2.2.2 Geometric Mean (G.M)

This is the nth root of the product of the values of variable x_i, where i = 1, 2, --- , n

i.e $GM(x) = \sqrt[n]{x_i \times x_2 ----x_n}$

$$= (\Pi\, x_i)^{1/n} \qquad\qquad ---2.2.2.1$$

where Π stands for the product of the data values

Example 2.2.2.1
The following are the data generated on the number of sales of cement (in bags): 12, 7, 2, 6, 13, 17, 14, 5, 9, 4. Compute the Geometric Mean.

Solution
$GM(x) = \sqrt[n]{x_i}$, i = 1, 2, --- , n

$GM(x) = \sqrt[10]{12\times7\times2\times6\times13\times17\times14\times5\times9\times4}$

$$= \sqrt[10]{561375360}$$

$$= 7.4977$$

Remark:

i. The GM is more appropriate for data in the form of rates, ratios and percentages.

ii.   In the computation of GM, logarithm approach can be applied. Here, one takes the logarithm of both sides of the equation 2.2.2.1 to give

$$\text{Log GM}(x) \quad = \quad \frac{\text{Log } x}{n} = \quad y$$

The antilogarithm of y gives the GM(x)

i.e $GM(x) = \quad 10^y$

Logarithm to base 10 or e i.e Napierian logarithm can be used, but one must be consistent with the base used.

2.2.3   Harmonic Mean (HM)

This is an average often used in engineering and related disciplines. It is good for calculation of average speed of machines.

For the values of variable x_i, i = 1, 2, --- , n, the HM of these values is given by

$$HM(x) = \quad \frac{n}{\sum \dfrac{1}{x_i}} \qquad \text{or} \qquad n\left(\sum \frac{1}{x_i}\right)^{-1}$$

Example 2.2.3.1
Using the data in example 2.2.2.1, obtain the Harmonic mean.

$$HM(x) = \quad \frac{n}{\sum \dfrac{1}{x_i}}$$

$$= \quad \frac{10}{\frac{1}{12} + \frac{1}{7} + \frac{1}{2} + \frac{1}{6} + \frac{1}{13} + \frac{1}{17} + \frac{1}{14} + \frac{1}{5} + \frac{1}{9} + \frac{1}{4}}$$

$$= \quad \frac{10}{0.0833 + 0.1429 + 0.5 + 0.1667 + 0.0769 + 0.0588 + 0.0714 + 0.2 + 0.1111 + 0.25}$$

$$= \quad \frac{10}{1.6611}$$

$$\therefore HM(x) \quad = \quad 6.0201$$

2.3   Mode and Median

The MODE is the value which occurs most frequently in a set of data. For a data set in which no measured values are repeated, there is no mode. Otherwise, a distribution can be Unimodal, Bimodal or Multimodal according to the data set.

From a grouped data with frequency distribution, the mode is determined by either

a.   graphical method through the use of Histogram; and

b.   the use of formula.

a.   In the graphical method, the following steps are to be taken:

i.   Draw the histogram of the given distribution.

ii.   Identify the highest bar (the modal class).

iii.   Identify the two linking bars to the modal class, i.e the bar before and after the modal class.

iv.   Use the two flanking bars to draw diagonals on the modal class.

v.   Locate the point of intersection of the diagonals drawn in step (iv) above, and vertically draw a line from the point of intersection to the boundary axis which gives the desired mode.

b.   By formula, the mode is defined by

$$\text{Mode} = L_1 + \left(\frac{\Delta_1}{\Delta_1 + \Delta_2} \right) C$$

where

L_1 =   Lower class boundary of the modal class

Δ_1 =   Modal class frequency – Frequency of the class before the modal class.

Δ_2 =   Modal class frequency – Frequency of the class after the modal class.

C =   Modal class size

The <u>MEDIAN</u> of a data set is the value of the middle item of the data when all the items in the data set are arranged in an array form (either ascending or descending order).

From an ungrouped data, the position (Median) is located by $^{(N+1)}/_2$, where N is the total number of items in the data set.

In a grouped data, the position of the Median is located by $^N/_2$. Here, the specific value of the Median is determined by either

li

    a.   graphical method through the use of Ogive (Cumulative frequency curve)or

    b.   the use of formula

a.   In the graphical method, the following procedure is used:

    i.   Draw the Ogive of the given distribution.

    ii.   Locate the point $^N/_2$ on the cumulative frequency axis of the Ogive.

    iii.   Draw a parallel line through the value of $^N/_2$ to the curve and then draw a perpendicular line from curve intercept to the boundary axis in order to get the Median value.

b.   The Median is obtained by the use of the following formula:

$$\text{Median} = L_1 + \left(\frac{^N/_2 - (F_i)}{f_i}\right)C$$

where L_1 =   Lower class boundary of median class.

        N =   Total number of items in the data set.

        F_i =   Summation of all frequencies before the median class.

        f_i =   Frequency of the median class.

        C =   Median class size or interval.

Worked Examples:

Example 2.3.1: Case I (n is odd and discrete variable is involved)
The following are the data generated from the sales of oranges within eleven days: 4, 7, 12, 3, 5, 3, 6, 2, 3, 1, 3, calculate both the mode and median.

Solution
Mode =   Most occurring frequency

Mode =   3

Median   = $\dfrac{\{n+1\}^{th}}{2}$ position

i.e 1, 2, 3, 3, 3, 3, 4, 5, 6, 7, 12

Median (n = 11)   = $\dfrac{11+1}{2}$ = $\dfrac{12}{2}$ = 6th position

    $\therefore$ The median = 3

Example 2.3.2: (n is odd and continuous variable is involved)
The following are the data generated from the measurement of an iron-rod; 2.5, 2.0, 2.1, 3.5, 2.5,

2.5, 1.0, 2.5, 2.2, calculate both the mode and median.

Solution

1.0, 2.0, 2.1, 2.2, 2.5, 2.5, 2.5, 2.5, 3.5

Mode = 2.5

Median (n = 9) = $\dfrac{n+1}{2}$ = $\dfrac{9+1}{2}$ = $\dfrac{10}{2}$ = 5th position

∴ The median = 2.5

Example 2.3.3: (n is even and discrete variable is involved)

The following are the data generated from the sales of orange within 12 days

4, 7, 12, 3, 5, 3, 6, 2, 3, 1, 3, 4, calculate the mode and median.

Solution

Mode = Most occurrence frequency

Mode = 3

Median = $\dfrac{(n+1)^{th}}{2}$ position

If n = 12, $\dfrac{12+1}{2}$ = $\dfrac{13}{2}$ = 6.5^{th} position

Rearrange the data first and take the middle value.

1, 2, 3, 3, 3, 3, 4, 4, 5, 6, 7, 12

The 6.5^{th} position gives $\dfrac{3+4}{2}$ = 3.5

Example 2.3.4 (n is even and continuous variable is involved):

The following are the data generated from the measurement of an iron-rod; 2.5, 2.0, 2.1, 3.5, 2.5, 2.5, 1.0, 2.5, 2.2, 2.6, calculate the mode and median

Solution

Mode = Most occurring frequency

Mode = 2.5

Median = $\dfrac{(n+1)^{th}}{2}$ position

If n = 12, $\dfrac{12+1}{2}$ = $\dfrac{13}{2}$ = 6.5^{th} position

Rearrange the data first and take the middle value.

1.0, 2.0, 2.1, 2.2, 2.5, 2.5, 2.5, 2.5, 2.6, 3.5

The 6.5th position gives $\dfrac{2.5 + 2.5}{2}$ $=$ 2.5

$\therefore$ Median $= 2.5$

Example 2.3.5: Case III: Grouped data by graphical method

Use the graphical method to determine the mode of the following data

Tyre range	Frequency	C. F	Class Boundaries
2 – 4	3	3	1.5 – 4.5
5 – 7	4	7	4.5 – 7.5
8 – 10	6	13	7.5 – 10.5
11 – 13	2	15	10.5 – 13.5
14 – 16	1	20	13.5 – 16.5
	20		

Solution

-       Draw the histogram of the given table

-       Interpolate on the modal class as discussed earlier in order to obtain the mode.
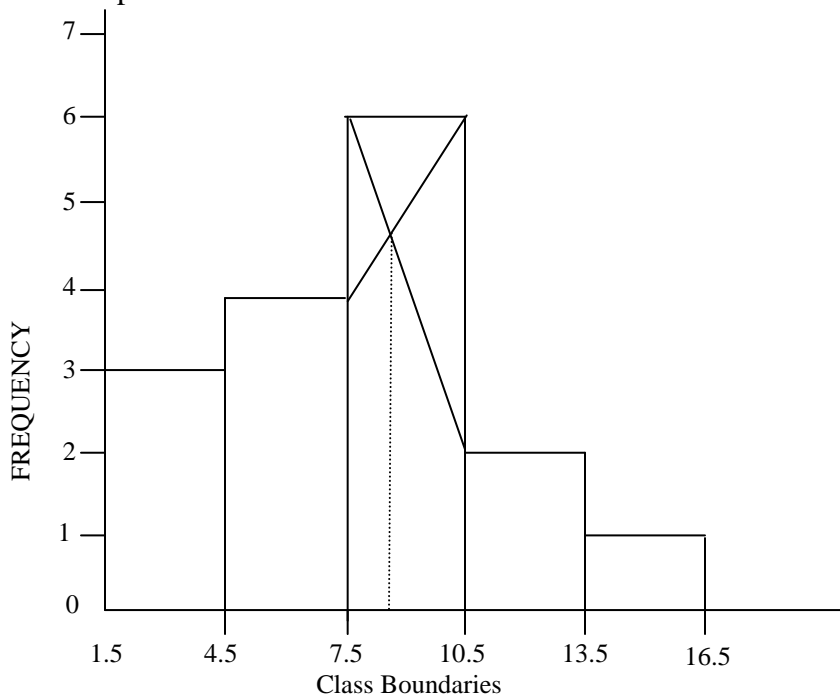
Mode   =   8.5

Example 2.3.6: Case IV: Graphical method

Use graphical method to find the median using case III data

Solution:

Tyre range	Frequency	C. F	Less than Class Boundary
2 – 4	3	3	4.5
5 – 7	4	7	7.5
8 – 10	6	13	10.5
11 – 13	2	15	13.5
14 – 16	1	20	16.5
	20		
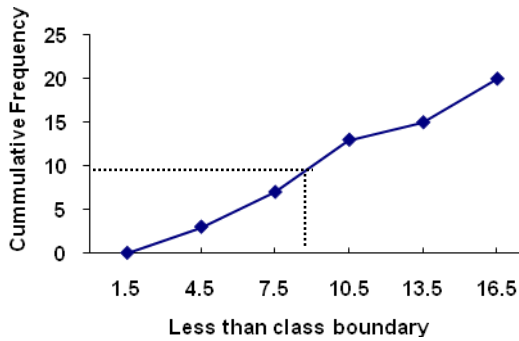
Median frequency $= \dfrac{N}{2} = \dfrac{20}{2} = 10$

Median $= 9.0$

Example 2.3.7: Graphical method to find mode and median.

Determine the mode and median by graphical method from the following data.

Class interval	Frequency(f)	C. F
2 – 4	3	3
4 – 6	4	7
6 – 8	6	13
8 – 10	7	20
10 – 12	2	22
	22	

Solution
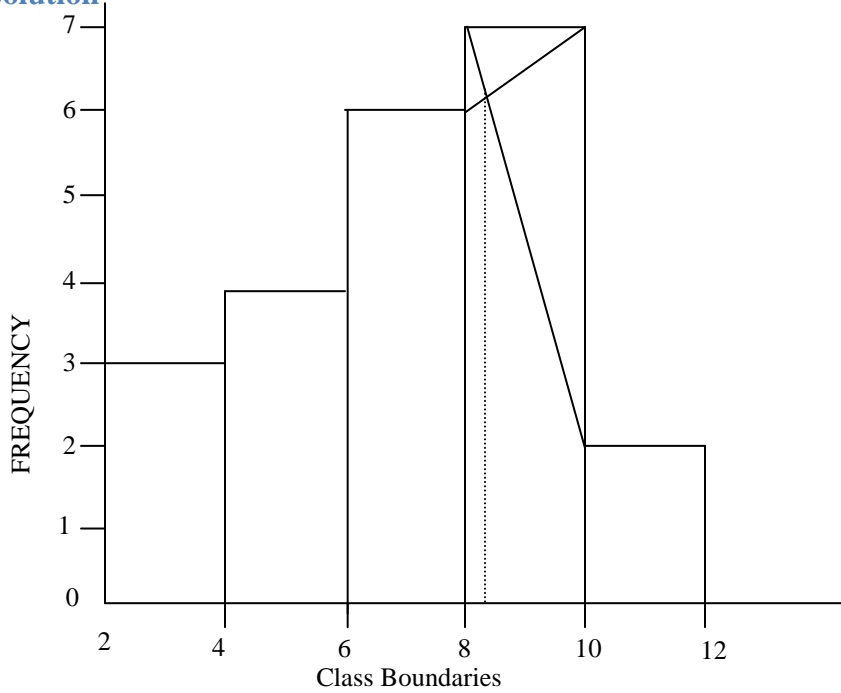
Mode = 8.50

To find the Median, we proceed as follows:

Table	Frequency (f)	C. F	Less than Boundary
2 – 4	3	3	4
4 – 6	4	7	6
6 – 8	6	13	8
8 – 10	7	20	10
10 – 12	2	22	12
	22		
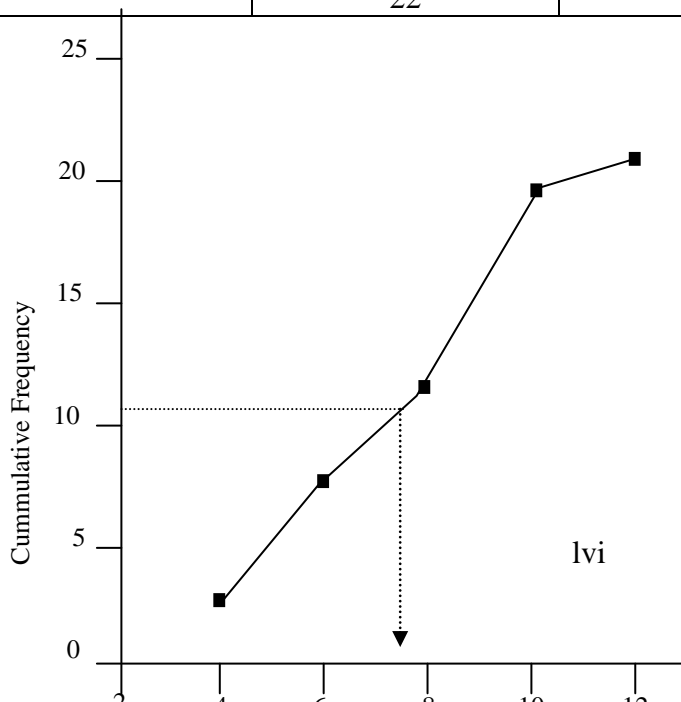
lvi

Median frequency $=$ $N/2$ $=$ $22/2$ $=$ 11

　　Median $=$ 　7.8

Example 2.3.8: Case III: Grouped data using formula

The following are the data generated on the sales of electric data

Class interval	Frequency
2 – 4	3
4 – 6	4
6 – 8	6
8 – 10	7
10 – 12	2

Determine the mode and median using the formula.

Solution

Class	Frequency	Cumulative Frequency
2 – 4	3	3
4 – 6	4	7
6 – 8	6	13 →Median class
8 – 10	7	20 → Modal class
10 – 12	2	22
	22	

N $=$ 22; $N/2$ $=$ $22/2$ $=$ 11

Mode $=$ L_1 $+$ $\left(\dfrac{\Delta_1}{\Delta_1 + \Delta_2}\right)C$

where

　　L_1 $=$ Lower class boundary of the modal class

　　Δ_1 $=$ Modal class frequency – Frequency of the class before the modal class.

　　Δ_2 $=$ Modal class frequency – Frequency of the class after the modal class.

　　C $=$ Modal class size

　　$L_1 = 8,$ $\Delta_1 = 7 - 6 = 1,$ $\Delta_2 = 7 - 2 = 5,$ $C = 2$

Mode $=$ $8 + \left(\dfrac{1}{\quad}\right)(2)$

$$1 + 5$$

$$= \quad 8 + \left(\frac{1}{6}\right)(2)$$

$$= \quad 8 + {}^1/_3$$

$$= \quad 8 + 0.33 \quad = \quad 8.33$$

$\therefore$ Mode $\quad = \quad 8.33$

Median $\quad = \quad L_1 \quad + \quad \left(\frac{N/_2 - (F_i)}{f_i}\right)C$

where $L_1 \quad = \quad$ Lower class boundary of the median class

$\quad f_i \quad = \quad$ Frequency of the median class

$\quad C \quad = \quad$ Median class size

$\quad F_i \quad = \quad$ Summation of all frequencies before the median class

$\quad L_1 = 6, \quad F_{med} = 6, \quad C = 2, \quad \Sigma f_i = 3 + 4 = 7, \quad N = 22$

Median $\quad = \quad 6 \quad + \quad \left(\frac{22/_2 - 7}{6}\right)(2)$

Median $\quad = \quad 6 \quad + \quad \left(\frac{11 - 7}{6}\right)(2)$

Median $\quad = \quad 6 \quad + \quad \left(\frac{4}{6}\right)(2)$

Median $\quad = \quad 6 \quad + \quad \frac{4}{3}$

Median $\quad = \quad 6 \quad + \quad 1.3$

Median $\quad = \quad 7.3$

Example 2.3.9

The following are the data generated on the sales of tyres:

Tyre range	2 – 4	4 – 6	6 – 8	8 – 10	10 – 12
Frequency	3	4	6	2	5

Calculate the mode and median.

Solution

Tyre range	Frequency	Cumulative Frequency
2 – 4	3	3
4 – 6	4	7
6 – 8	6	13 →Median/Modal class
8 – 10	2	15
10 – 12	5	20
	20	

$N = 20;$ $^N/_2 = {}^{20}/_2 = 10$

$$\text{Mode} = L_1 + \left(\frac{\Delta_1}{\Delta_1 + \Delta_2}\right)C$$

$L_1 = 6.0,$ $\Delta_1 = 6 - 4 = 2,$ $\Delta_2 = 6 - 2 = 4,$ $C = 3$ (i.e $4.5 - 1.5$)

$$\text{Mode} = 6.0 + \left(\frac{2}{2 + 4}\right)(3)$$

$$= 6.0 + 1$$

$$\text{Mode} = 7.0$$

$$\text{Median} = L_1 + \left(\frac{^N/_2 - (F_i)}{f_i}\right)C$$

$L_1 = 6.0,$ $F_{med} = 6,$ $C = 3,$ $F_{i} = \Sigma f_i = 3 + 4 = 7,$ $N = 20$

$$\text{Median} = 6.0 + \left(\frac{^{20}/_2 - 7}{6}\right)(3)$$

$$\text{Median} = 6.0 + \left(\frac{10 - 7}{6}\right)(3)$$

$$\text{Median} = 6.0 + \left(\frac{3}{6}\right)(3)$$

$$\text{Median} = 6.0 + \frac{3}{2}$$

$$\text{Median} = 6.0 + 1.5$$

Median = 7.5

2.4 Measures of Partition

These are other means of measuring location. We will recall that the median characterises a series of values by its midway position. By extension of this idea, there are other measures which divide a series into a number of equal parts but which are not measures of central tendency. These are the measures of partition. The common ones are the quartiles, deciles and the percentiles. They are collectively called the QUANTILES.

The methods of computation for the quantiles follow the same procedure for the Median according to what is being measured. For instance, the position of the first quartile from a grouped data is located by $^N/_4$ and the value is determined graphically from Ogive or by use of formula as done in the Median.

Also the position for locating third quartile is $^{3N}/_4$, for the seventh decile is $^{7N}/_{10}$ and tenth percentile is $^{10N}/_{100}$. We should note that the second quartile, fifth decile and fiftieth percentile coincide the Median.

The use of formulae to compute the quantiles are summarized below:

$$\text{Quartile } (Q_i) = L_i + \left(\frac{^N/_4 - F_i}{f_i} \right) C \quad , \quad i = 1, 2, 3$$

$$\text{Decile } (D_i) = L_i + \left(\frac{^N/_{10} - F_i}{f_i} \right) C \quad , \quad i = 1, 2, 3, 4, 5, 6, 7, 8, 9$$

$$\text{Percentile } (P_i) = L_i + \left(\frac{^N/_{100} - F_i}{f_i} \right) C, \quad i = 1, 2, 3, 4, 5, -----,98, 99.$$

where

L_i = Lower Class boundary of the ith quantile class.

f_i = frequency of ith quantile class.

$F_i = \Sigma f_i$ = Sum of the frequencies of all classes lower than the ith quantile class.

C = Class interval of the ith quantile class.

N = Total number of items in the distribution.

Example 2.4.1:

Find Q_1, Q_3, D_7, P_{20} for the following data:

Interval	X	F	c. f	Less than Class Boundaries
0 – 2	1	2	2	2.5
3 – 5	4	4	6	5.5
6 – 8	7	8	14	8.5
9 – 11	10	4	18	11.5
12 – 14	13	2	20	14.5
		20		

Position of Q_1 = $(N/4)^{th}$

Frequency of Q_1 = $^{20}/_4$ = 5

Q_1 = 3 – 5 class

$$Q_1 = L_1 + \left(\frac{\frac{N}{4} - F_i}{f}\right)C$$

$L_1 = 2.5,$ $^N/_4 = 5,$ $F_i = 2,\ f_i = 4, C = 3$

$$Q_1 = 2.5 + \left(\frac{5-2}{f}\right)(3)$$

Q_1 = 2.5 + (¾)3

Q_1 = 2.5 + 2.25

Q_1 = 4.75

Q_3's position; $\dfrac{3N}{4} = \dfrac{3 X 20}{4} = 15^{th}$

$$Q_3 = L_1 + \left(\frac{\frac{3N}{4} - F_i}{f_i}\right)C$$

Q_3's class = 9 – 11

$L_1 = 8.5,$ $F_i = 2 + 4 + 8,$ $f_i = 4, C = 3$

$$Q_3 = 8.5 + \left(\frac{15-14}{4_i}\right)(3)$$

Q_3 = 8.5 + (1/4)3

$Q_3 = 8.5 + 0.75$

$Q_3 = 9.25$

Decile (D_i) $=$ L_i $+$ $\left(\dfrac{N/10 - F_i}{f_i}\right) C$, $i = 1, 2, 3, ---, 9$

$\therefore$ Decile (D_7) $=$ L_7 $+$ $\left(\dfrac{7N/10 - F_i}{f_i}\right) C$, $i = 1, 2, 3, ----, 9$

$\therefore$ its frequency $\dfrac{N(i)}{10}$ $=$ $\dfrac{20(7)}{10}$ $=$ $\dfrac{140}{10}$ $=$ 14

$L_7 = 5.5;$ $f_7 = 8,$ $F_7 = 6,$ $C = 8.5 - 5.5 = 3$

D_7 $=$ $5.5 + \dfrac{3(14-6)}{8}$

$=$ $5.5 + 0.375(8)$

$=$ $5.5 + 3$

$=$ 8.5

Percentile $(P_i) =$ L_i $+$ $\left(\dfrac{N/100 - F_i}{f_i}\right) C$, $i = 1, 2, 3, ---, 99$

$\therefore N(i) =$ $\dfrac{20(20)}{100}$ $=$ $\dfrac{400}{100}$ $=$ 4

P_{20} $=$ L_{20} $+$ $\dfrac{C}{f_{20}}\left(\dfrac{N(20) - F_{20}}{100}\right)$, $i = 1, 2, 3, ---, 99$

$L_{20} = 2.5,$ $f_{20} = 4, F_{20} = 2,$ $C = 5.5 - 2.5 = 3$

P_{20} $=$ $2.5 + \frac{3}{4}(4-2)$ $=$ $2.5 + 0.75(2)$ $=$ $2.5 + 1.5$

$=$ 4

2.5 Summary

In this chapter, the measures of central tendency, which consist of mean, mode and median, were considered for both grouped and ungrouped data. The measures of location such as Quartile, Decile and Percentile were also discussed in the chapter.

2.6 Questions (Multiple-Choice and Short-Answer Questions)

1. Which of the following means is often used in engineering and related disciplines?

    a. Arithmetic mean.   b. Geometric mean

    c. Harmonic mean   d. Arithmetic and Geometric mean.

    e. Hyperbolic mean.

2. Which of the following is not a measure of central tendency?

   a. Arithmetic mean   b. Mode   c. Median

   d. Decile   e. Geometric mean

3. Which of the following is not a measure of partition?

   a. Median   b. Mode   c. Percentile

   d. Quantile   e. Quintile.

4. Which of the following formulae is used for the computation of quintile?

   a. $L_i + C\left(\dfrac{N(i)/2 - F_i}{f_i}\right)$ b. $L_i + C\left(\dfrac{N(i)/4 - F_i}{f_i}\right)$

   c. $L_i + C\left(\dfrac{N(i)/5 - F_i}{f_i}\right)$ d. $L_i + C\left(\dfrac{N(i)/10 - F_i}{f_i}\right)$

   e. $L_i + C\left(\dfrac{N(i)/100 - F_i}{f_i}\right)$

5. In the graphical method of obtaining the quartiles, which of the following diagrams is used?

   a. Bar chart   b. Histogram   c. Pie chart

   d. Ogive   e. component bar chart.

Use the following data to answer questions 6 to 10:

6, 3, 8, 8, 5

6. Calculate the Arithmetic mean.

7. Calculate the Geometric mean.

8. Calculate the Harmonic mean.

9. Determine the Mode.

10. Determine the Median.

Answers to Chapter Two

1. c,   2. d;   3. b;   4. c;   5. d

6. $\bar{x}$ = $\dfrac{6 + 3 + 8 + 8 + 5}{5}$ = $\dfrac{30}{5}$ =   6

7. G_{xi} = $\sqrt[n]{\prod x_i}$ = $\sqrt[5]{6x3x8x8x5}$

      = $\sqrt[5]{5760}$ =   5.65

8. $Hm(\bar{x})$ = $\dfrac{n}{\sum \dfrac{1}{x_i}}$

      = $\dfrac{5}{\dfrac{1}{6}+\dfrac{1}{3}+\dfrac{1}{8}+\dfrac{1}{8}+\dfrac{1}{5}}$ = $\dfrac{5}{0.1667+0.3333+0.125+0.125+0.2}$

      = $^5/_{0.95}$ =   5.263 $\approx$ 5.3

9.   Mode is 8

10.   6, 3, 8, 8, 5

      Rearrange ➜   3, 5, 6, 8, 8

      $\therefore$ Median = 6.

CHAPTER THREE
MEASURES OF VARIATION

3.0 Learning objectives

At the end of the chapter, readers and students should be able to

- know the meaning of measure of variation;

- determine various measures of variation such as mean deviation, variance, standard deviation and quantile deviation for both grouped and ungrouped data;

- compute the coefficient of variation and of skewness.

3.1 Introduction

The degree to which numerical data tend to spread about an average value is referred to as measure of variation or dispersion. At times, it is called measure of spread. The purpose and significant uses of these measures are of paramount value in statistics.

The popular ones among these measures of variation are the

i.   Range;   ii.   Mean Deviation;

iii.   Variance/standard Deviation;   iv.   Semi – interquantile range

3.2 Measures of spread

3.2.1 Range

The range for a set of data, is the difference between the highest number and the lowest number of the data. That is,

Range = Highest number – Lowest number (for ungrouped data); OR

Range = Upper bound of the highest class – Least bound of the lowest class

(for grouped data)

Example 3.2.1.1

Determine the range for following set of numbers:

a.   32, 6, 10, 27, 30, 5, 45

b.   09, 14, 16, 13, 14, 21

Solution:

The data here is ungrouped type

a.   Range = Highest – Lowest

         = 45 – 05   = 40

b.   Range =   21 – 09   = 12

Example 3.2.1.2

Obtain the range for the following table

Class   Frequency

Class	Frequency
1 – 10	4
11 – 20	10
21 – 30	12
31 – 40	12
41 – 50	9

Solution

The data in this question is grouped.

Range =   50 – 1 = 49

3.2.2   Mean Deviation (MD)

The Mean Deviation is the arithmetic mean of the absolute deviation values from the mean. For ungrouped data (of x_1, x_2, --- ,x_n), the M.D is given as

$$M.D = \frac{\Sigma|x_i - \overline{x}|}{N} = \frac{\Sigma|d_i|}{N} \qquad \text{---3.2.2.1}$$

where $d_i = x_i - \overline{x}$, the symbol $||$ stands for modulus or absolute value. For grouped data, it is written as:

$$M.D = \frac{\Sigma f|x_i - \overline{x}|}{\Sigma f} = \frac{\Sigma f|d_i|}{\Sigma f} \qquad \text{---3.2.2.2}$$

where $f = N$

Examples 3.2.2.1

Calculate the mean deviation from number of books sold in a small bookshop as given below: 15, 17, 14, 16, 18

Solution

$$\text{The mean } (\overline{x}) = \frac{\Sigma x}{N} = \frac{15 + 17 + 14 + 16 + 18}{5} = \frac{80}{5} = 16$$

The following format shall be used for MD:

X	$d_i = x_i - \overline{x}$	$|x_i - \overline{x}|$

15	-1	1
17	1	1
14	-2	2
16	0	0
18	2	2
80		6

Thus, MD $=\dfrac{\Sigma |d_i|}{N}=\dfrac{6}{5}=$ 1.20

Example 3.2.2.2

The distribution of book sales (in hundreds) in large bookshop is given in the table below.
Determine the M. D from this table

Class	Frequency
1 – 10	10
11 – 20	15
21 – 30	17
31 – 40	13
41 – 50	05

Solution:

Class	Frequency	x	fx	$d_i = |x_i - \overline{x}|$	$f|d_i|$
1 – 10	10	5.5	55	18	180
11 – 20	15	15.5	232.5	8	120
21 – 30	17	25.5	433.5	2	34
31 – 40	13	35.5	461.5	12	156
41 – 50	05	45.5	227.5	22	110
	60		1410		600

Mean (x) $=\dfrac{\Sigma fx}{\Sigma f}$

$=\dfrac{1410}{60}=$ 23.5

$\therefore$ Mean deviation (MD) $=\dfrac{\Sigma f_i|d_i|}{\Sigma f_i}$

$=\dfrac{600}{}$

$$\frac{}{60} = 10$$

3.2.3 Variance and Standard Deviation

The variance is similar to the mean deviation (M.D) in the sense that it is based on the difference between each value in a data set and the mean of the group. It differs in that these differences are squared before being summed.

It is important to know further that the standard deviation is the positive square root of the variance.

In an ungrouped data, the population variance is given as

$$\sigma^2 = \frac{\Sigma(x - \mu)^2}{N} \qquad \text{---3.2.3.1}$$

Where σ^2 is the population variance.

μ is the population mean.

N is the total number of items.

From the above, the population standard deviation is

$$\sigma = \sqrt{\frac{\Sigma(x - \mu)^2}{N}} \qquad \text{---3.2.3.2}$$

The variance for a sample is not exactly equal to the population variance. It is given as:

$$S^2 = \frac{\Sigma(x_i - \bar{x})^2}{n - 1} \qquad \text{---3.2.3.3}$$

where $\bar{x}$ is the sample mean and n is the sample size.

Also for grouped data, the population variance is:

$$\sigma^2 = \frac{\Sigma f(x - \mu)^2}{\Sigma f - 1} \qquad \text{---3.2.3.4}$$

and for sample variance, we have

$$S^2 = \frac{\Sigma f(x_i - \bar{x})^2}{\Sigma f - 1} = \frac{\Sigma f(x_i - \bar{x})^2}{n - 1} \qquad \text{---3.2.3.5}$$

where f stands for frequency and $\Sigma f = n$

Example 3.2.3.1

The figures 9, 5, 9, 7, 10, 14, 12, 10, 6, 17 represent the volumes of sales (₦'000) of Adebayo Spring Water within the first ten days of operation. Calculate the population variance and hence the standard deviation of sales.

Solution

Let the sales be represented by x, then,

$$\sigma^2 \;=\; V(x) \;=\; \frac{\Sigma(x-\mu)^2}{N} \qquad \text{where } \mu \;=\; \frac{\Sigma x_i}{N}$$

$$\mu \;=\; \frac{9 + 5 + 9 + 7 + 10 + 14 + 12 + 10 + 6 + 17}{10}$$

$$=\; \frac{99}{10} \;=\; 9.9$$

$$\therefore \sigma^2 \;=\; \{(9-9.9)^2 + (5-9.9)^2 + (9-9.9)^2 + (7-9.9)^2 + (10-9.9)^2 + (14-9.9)^2$$
$$+ (12-9.9)^2 + (10-9.9)^2 + (16-9.9)^2 + (17-9.9)^2\} / (10)$$

$$\therefore \sigma^2 \;=\; \frac{(0.81 + 24.01 + 0.81 + 8.41 + 0.01 + 16.81 + 4.41 + 0.01 + 37.21 + 50.41)}{10}$$

$$=\; \frac{142.9}{10}$$

$$\sigma^2 \;=\; 14.29$$

$$\sigma \;=\; \text{S.D} \;=\; \sqrt{V(x)}$$

$$\text{S.D} \;=\; \sqrt{14.29} \;=\; 3.78$$

Example 3.2.3.2

The table below shows the frequency table of the age distribution of 25 children in a family.

Age (x)	5	8	11	14	17
Frequency (f)	6	7	5	4	3

Determine the sample variance and its standard deviation.

Solution

X	F	Fx	$\lvert d_i \rvert = x_i - \overline{x}$	d_i^2	fd_i^2
5	6	30	-4.92	24.2064	145.2384
8	7	56	-1.9	3.61	25.27
11	5	55	1.1	1.21	6.05
14	4	56	4.1	16.81	67.24
17	3	51	7.1	50.41	151.23
	25	248			395.0284

$$\bar{x} \;=\; \frac{\Sigma fx}{\Sigma f} \;=\; \frac{248}{25}$$

$$d_i = x_i - 9.92$$

$$=\; 9.92$$

$$S^2 \;=\; \frac{\Sigma f_i(x_i - 9.92)^2}{25 - 1} \;=\; \frac{395.0284}{24}$$

$$=\; 16.4595$$

Hence,

$$S \;=\; \sqrt{16.4595} \;=\; 4.0570$$

Remark:

From the above computations, it is clearly seen that the set of formulae used above is tedious. Hence, the call or demand for a short cut and easier method is necessary. The following short cut formulae will be used:

For ungrouped data:

$$S^2 \;=\; \frac{\Sigma x_i^2 - n(\bar{x})^2}{n - 1} \quad \text{or} \quad \frac{\Sigma x_i^2 - (\Sigma x)^2/n}{n - 1} \qquad\qquad \text{---3.2.3.6}$$

Also for grouped data:

$$S^2 \;=\; \frac{\Sigma fx_i^2 - n(\bar{x})^2}{\Sigma f - 1} \quad \text{or} \quad \frac{\Sigma fx_i^2 - (\Sigma fx)^2/n}{n - 1} \qquad\qquad \text{---3.2.3.7}$$

$$\text{where } n = \Sigma f$$

Example 3.2.3.3: Another method for ungrouped data

The values 9, 5, 9, 7, 10, 14, 12, 10, 6, 17 represent the volume of sales (₦'000) of Stephen Pure Water within the first ten days of operation. Calculate the population variance and hence the standard deviation of sales using the shortcut method.

Solution

Let the sales be represented by x, then

$$S^2 \;=\; \frac{\Sigma x_i^2 - n(\bar{x})^2}{n - 1}$$

$$\bar{x} \;=\; \frac{9 + 5 + 9 + 7 + 10 + 14 + 12 + 10 + 6 + 17}{10}$$

$$=\; \frac{99}{10} \;=\; 9.9$$

$$\Sigma x^2 \;=\; 9^2 + 5^2 + 9^2 + 7^2 + 10^2 + 14^2 + 12^2 + 10^2 + 6^2 + 17^2$$

$$=\; 1101$$

$$\therefore S^2 \;=\; \frac{1101 - 10 \times (9.9)^2}{10 - 1}$$

$$=\; \frac{1101 - 10 \times 98.01}{9} \;=\; \frac{1101 - 980.1}{9}$$

$$= \frac{120.9}{9} = 13.433$$

$$S = \sqrt{\frac{\Sigma x_i^2 - n(\overline{x})^2}{n-1}} = \sqrt{13.433} = 3.6651$$

Example 3.2.3.4: Grouped Data

X	F	x^2	fx_i^2	fx
5	6	25	150	30
8	7	64	448	56
11	5	121	605	55
14	4	196	784	56
17	3	289	867	51
	25	695	2854	248

$$S^2 = \frac{\Sigma fx_i^2 - n(\overline{x})^2}{\Sigma f - 1} = \frac{\Sigma fx_i^2 - (\Sigma fx)^2/n}{n-1} \qquad \text{---3.2.3.7}$$

$$= \frac{2854 - (248)^2/25}{25-1}$$

$$= \frac{2854 - 61504/25}{24}$$

$$= \frac{2854 - 2460.16}{24}$$

$$= \frac{393.84}{24} = 16.41$$

$$S = \sqrt{\frac{\Sigma fx_i^2 - (\Sigma fx)^2/n}{n-1}}$$

$$=$$

$$= \sqrt{16.41} = 4.0509$$

3.2.4 Semi – Interquartile Range (SIR)

Semi – Interquantile Range is obtained from quantile range and it is denoted and computed by:

$$\text{SIR} = \frac{Q_3 - Q_1}{2} \qquad \text{--- 3.2.4.1}$$

where Q_1 = First Quantile, Q_3 = Third Quantile and the numerator $(Q_3 - Q_1)$ is known as the quantile range.

Example 3.2.4.1

Determine the semi – interquartile range for the income distribution of SAO company employees as given in the following table:

Income class in N'000	10 – 20	20 – 30	30 – 40	40 – 50
Frequency (f)	7	10	8	5

Solution

Income Class	Frequency	CF
10 – 20	7	7
20 – 30	10	17
30 – 40	8	25
40 – 50	5	30
	30	

The first quartile Q_1 is in the class containing the $30/4 = 7.5^{th}$ item. This is the class 20 – 30

$$\therefore L_1 = 20, \qquad F_1 = 7; \qquad f_1 = 10, \qquad C = 20 - 30 = 10, \qquad {}^N/_4 = 7.5$$

$$
\begin{aligned}
Q_1 \;=\;& L_1 + \frac{C({}^N/_4 - F_1)}{f} \\
=\;& 20 + {}^{10}/_{10}({}^{30}/_4 - 7) \\
=\;& 20 + 1.0(7.5 - 7) \\
=\;& 20 + 0.50 \quad = \quad 20.50
\end{aligned}
$$

The third quartile (Q_3) is in the class containing the ${}^{3(30)th}/_4$ i.e 22.5^{th} item. This is the class 30 – 40.

$$\therefore L_3 = 30, \quad f_3 = 8, \quad F_3 = 17, \qquad C = 30 - 40 = 10, \quad {}^{3N}/_4 = 22.5$$

$$
\begin{aligned}
Q_3 \;=\;& L_3 + \frac{C({}^{3N}/_4 - F_3)}{f} \\
=\;& 30 + {}^{10}/_8(22.5 - 17) \\
=\;& 30 + 1.25(5.5)
\end{aligned}
$$

NOTE: This is a work in progress. All topics in the syllabus are covered but editing for necessary corrections is in progress.

Thanks.

$$= \quad 30 + 6.875 \quad = \quad 36.875$$

$$\therefore \text{Semi} - \text{Interquartile range} = \quad \tfrac{1}{2}(Q_3 - Q_1)$$

$$= \quad \tfrac{1}{2}(36.875 - 20.50)$$

$$= \quad \tfrac{1}{2}(16.375)$$

$$= \quad 8.1875$$

3.3 Coefficient of Variation and of Skewness

3.3.1 Coefficient of Variation (CV)

This is a useful statistical tool. It shows degree of variation between two sets of data. It is a dimensionless quantity and defined as:

$$CV \quad = \quad \frac{\text{Standard Deviation}}{\text{Mean}} \times 100\%$$

$$= \quad \frac{\sigma}{\overline{x}} \times 100\%$$

By a way of interpretation, the smaller the C.V of a data set, the higher the precision of that set. Also in comparison, the data set with least C.V has a better reliability than the other.

Example 3.3.1.1:

The following data consists of the ages of twelve banks in year (last birthday) during their last anniversary in Oyo state. Calculate the coefficient of variation: 15, 14, 8, 12, 11, 10, 16, 14, 12, 11, 13, 8

Solution

$$\text{Coefficient of Variation} \quad = \quad \frac{S}{\overline{x}} \times 100\% \quad = \quad \frac{\text{Standard Deviation}}{\text{Mean}} \times 100\%$$

$$\overline{x} \quad = \quad \frac{15 + 14 + 8 + 12 + 11 + 10 + 16 + 14 + 12 + 11 + 13 + 8}{12}$$

$$= \quad \frac{144}{12} \quad = \quad 12$$

$$S_x^2 \quad = \quad \frac{\Sigma f x_i^2 - n\overline{x}^2}{\Sigma f - 1} \quad , \text{ where}$$

$$\Sigma f x_i^2 = 15^2 + 14^2 + 8^2 + 12^2 + 11^2 + 10^2 + 16^2 + 14^2 + 12^2 + 11^2 + 13^2 + 8^2 = 1800$$

$$\therefore S_x^2 \quad = \quad \frac{1800 - 12(12)^2}{12 - 1}$$

$$= \frac{1800 - 12 \times 144}{11}$$

$$= \frac{1800 - 1728}{11}$$

$$= \frac{72}{11} = 6.5455$$

$$S = \sqrt{6.5455}$$

$$= 2.5584$$

$$\therefore CVs = \frac{S}{X} = \frac{2.5584}{12}$$

$$= 0.2132$$

3.3.2 Coefficient of Skewness

If a frequency curve is not symmetrical, its skewness is the degree of asymmetry of the distribution.

The Pearsonian coefficient of skewness is given as:

$$\text{Coefficient of skewness} = \frac{\text{Mean} - \text{Mode}}{\text{Standard Deviation}}$$

OR it can also be obtained as

$$\text{Coefficient of skewness} = \frac{3(\text{Mean} - \text{Median})}{\text{Standard Deviation}}$$

The coefficient of skewness is zero for a symmetry or normal distribution. For a positively skewed distribution, the mean is larger than the mode, while for a negatively skewed distribution the mean is smaller than the mode. The median takes a value between the mean and the mode.

Example 3.3.2.1:

Given the following arithmetic means, the medians and the standard deviations of two distributions:

i.   Arithmetic mean = 20,   Median = 22   and   Standard Deviation = 8
ii.   Arithmetic mean = 20,   Median = 23   and   Standard deviation = 10

Determine which of the distributions is more skewed.

Solution

By the Pearson measure of skewness, we have

Skewness $=$ $\dfrac{3(\text{Arithmetic mean} - \text{Median})}{\text{Standard Deviation}}$

i. Skewness $=$ $\dfrac{3(20 - 22)}{8}$ $=$ $\dfrac{-6}{8}$ $=$ -0.75

ii. Skewness $=$ $\dfrac{3(20 - 23)}{10}$ $=$ $\dfrac{-9}{10}$ $=$ -0.90

since $|-0.90| > |-0.75|$, the second (ii) distribution is more skewed.

3.4   Summary

Measure of variation has been described as measure of spread. It is commonly used to obtain the spread about the average and to make comparison of spreads for two sets of data. Among these measures of dispersion discussed, we have the range, mean deviation, variance/standard deviation and semi – interquantile range. Here, both ungrouped and grouped data are considered.

Also, the concepts of coefficient of variation and of skewness are discussed.

3.5   Revision Questions

1. For a set of data, the difference between the highest and the lowest number is known as_____

   a. Mean   b. Variance   c. Range

   d. Interquatile range   e. Mean deviation

2. The main difference between the mean deviation and variance is that_____

   a. differences between the data set and mean are zero.

   b. the differences between the data set and the mean are squared before being summarized in variance.

   c. the square roots of differences between the data set and the mean are obtained

    d. the differences are in the order of arrangement.

    e. the difference between the data set and the mean are in geometric order.

3.   Semi – Interquantile ange is determined by

    a. $\dfrac{Q_3 - Q_2}{2}$
    b. $\dfrac{Q_3 - Q1}{2}$
    c. $\dfrac{Q_3 + Q_2}{2}$

    d.. $\dfrac{Q_3 + Q_1}{2}$
    e. $\dfrac{Q_4 - Q1}{2}$

4.   For a normal distribution, the coefficient of skewness is

    a. 1   b.. −1,   c. 0   d. 2,   e. −2

Use the following information to answer questions 5 to 10. given the following set of data: 3, 7, 2, 8, 5, 6, 4

5.   Determine the range.

6.   Determine the mean deviation.

7.   Determine the variance.

8.   Determine Q_1.

9.   Determine Q_3.

10.   Determine Semi-interquantile range.

Answers to Chapter Three

1. c,   2. b,   3. b,   4. c,   5. 5,   6. 2

7. $^{40}/_7$ = 5.714   8. Q_1 = 2,   9. Q_3 = 6

10. $\dfrac{Q_3 - Q1}{2}$ = $\dfrac{6 - 2}{2}$ = 2

CHAPTER FOUR

MEASURE OF RELATIONSHIP AND REGRESSION

4.0   Learning objectives

At the end of the chapter, readers or students should be able to

- understand the concepts of univariate and bivariate;

- understand the concept of scatter diagram;

- know the various types of correlation coefficients;

- compute Pearson product moment correlation coefficient and make necessary interpretations;

- compute Spearman's rank correlation coefficient and make necessary interpretations;

- understand the concept of regression;

- fit a simple linear regression model to data;

- fit a simple linear regression line using the standard formulas for the slope and intercept;

- make necessary interpretation of the regression constant and regression coefficient;

- forecast (estimation) by the use of fitted regression line;

- know the meaning and computation of coefficient of determination.

4.1   Introduction

The type of data we have been using so far is the univariate type, i.e one variable. But in this chapter we shall be dealing with bivariate data, i.e two variables.

The statistical analysis that requires the use of bivariate is generally termed the measure of relationship and regression analysis which is the focus of this chapter. In this analysis, the interest is usually on relationship or pattern of the relationship. Here, one can look at how students' performance in one subject (Mathematics) affects another subject (Accountancy). Another good example is the consideration of how an increase in the income affects spending habits or savings.

In the measure of relationship, the usual practice is to quantify or qualify and represent the bivariate by letters. Taking for instance, the marks scored by a set of students in Mathematics and Accountancy can be represented by x and y variables respectively and

can be written as point (x,y), where x, and y stand for independent and dependent variables respectively.

A graphical presentation of bivariate data on a two-axis coordinate graph is known as the SCATTER DIAGRAM. Here, the bivariate data are plotted on a rectangular coordinate system in order to see the existing relationship between the two variables under study. The following bivariate data will be used as an illustration:

Example 4.1.1
Draw the scatter diagram for the following data:

Mathematics Scores (X_i) in %	30	40	35	40	20	25	50
Accountancy scores (Y_i) in %	50	70	65	68	40	60	80

SCATTER DIAGRAM
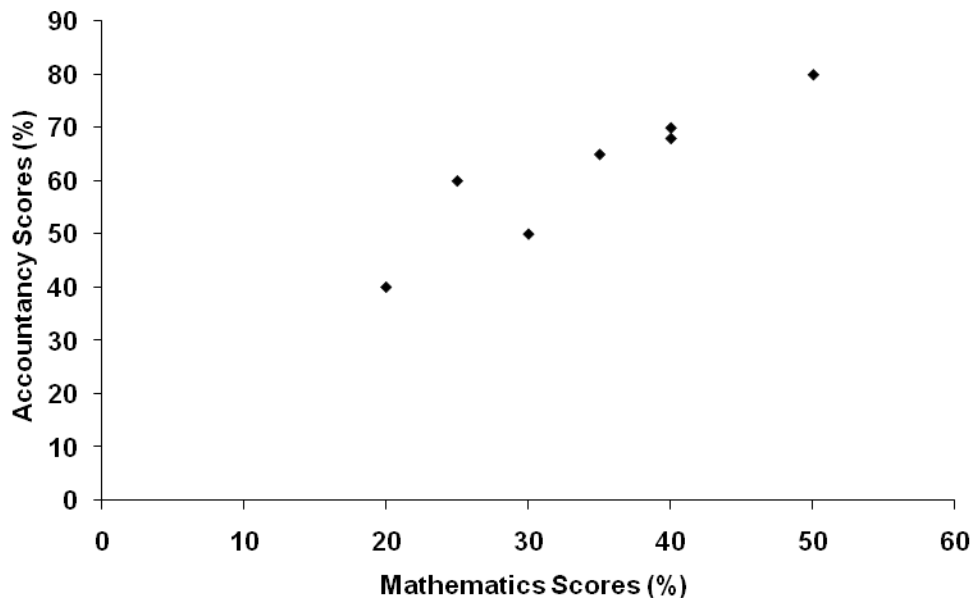
4.2   Types of Correlation

Correlation measures the degree of association (relationship) between two variables. Therefore, two variables are said to be correlated or related when one variable changes in value, it also changes the value of the other variable.   The degree of correlation (r) between two variables x and y is expressed by a real number which lies between −1.0 and +1.0 (i.e −1 ≤ r ≤ 1) and it is called correlation coefficient or coefficient of correlation.

Simple correlation is basically classified according to the value of its coefficient. Hence, there are:

i.   Positive correlation;

ii.   Negative correlation; and

iii.   Zero correlation.

The following scatter diagrams (fig. 4.2.1) depict the above types of correlation:
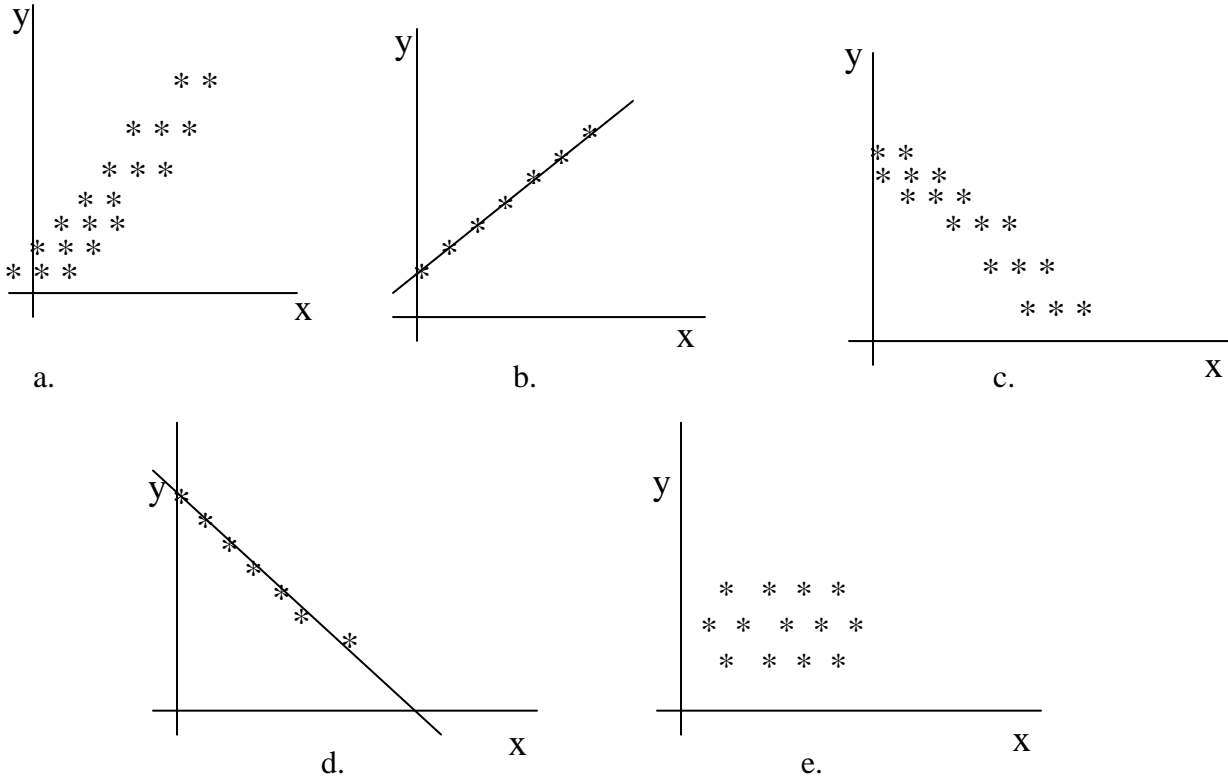
Figure 4.2.1

Figure 4.2.1(a) depicts <u>positive correlation</u>, where the two variables involved are directly proportional to each other. It means that when one variable increases (decreases), the other variable also increases (decreases). Therefore, there is positive correlation when two variables are increasing or decreasing together. Examples of paired variables that give positive correlation include:

i.   Price and supply of a commodity.

ii.   Income and expenditure of a family

Figure 4.2.1(b) depicts a <u>perfect positive correlation</u> where the changes in the two related variables are exactly proportional to each other and in the same direction.

In figure 4.2.1(c), there is <u>Negative correlation</u>. Here is a situation where one variable increases (decreases), the other variable decreases (increases). Therefore, the two variables move in opposite directions. Examples of paired variable that give negative correlation include:

i.   Number of labourers and time required to complete field work.

ii.   Demand and price for a commodity.

Figure 4.2.1(d) depicts a <u>perfect negative correlation</u> where the changes in the two related variables are exactly proportional to each other but at reverse (opposite) directions.

Figure 4.2.1(e) depicts the <u>zero correlation,</u> where there is no correlation existing between the two variables. Here, no fixed pattern of movement can be established.

4.3   Measure of Correlation

4.3.1   The measure of correlation can be determined by using the following methods:

a.   Product moment correlation coefficient method (Karl Pearson's Co-efficient of correlation):

The degree of correlation between variables x and y is measured by the product moment correlation co-efficient r, defined as

$$r = \frac{\sum(x - \bar{x})(y - \bar{y})}{\sqrt{\sum(x - \bar{x})^2 \left(\sum(y - \bar{y})^2\right)}}$$

$$r = \frac{\Sigma xy - (\Sigma x)(\Sigma y)/n}{\sqrt{\left(\Sigma x^2 - (\Sigma x)^2/n\right)\left(\Sigma y^2 - (\Sigma y)^2/n\right)}}$$

$$r = \frac{n\Sigma xy - \Sigma x \Sigma y}{\sqrt{[n\Sigma x^2 - (\Sigma x)^2][n\Sigma y^2 - (\Sigma y)^2]}} \qquad \text{---4.3.1}$$

where

n = the number of items (in pairs)

The numerator of equation 4.3.1 is called <u>co-variance</u> and the denominator is the square root of the product of the variances of variables x and y

$$r = \frac{Co\,var\,iance(x, y)}{\sqrt{Var(x)Var(y)}}$$

Example 4.3.1:

The marks scored by seven students in Mathematics (x) and Accounts (y) are given below. If the maximum scores obtainable in mathematics and accounts are respectively 50 and 100, determine the Pearson correlation coefficient for these scores.

X	30	40	35	40	20	25	50
Y	50	70	65	68	40	60	80

Solution:

x	y	xy	x^2	y^2
30	50	1500	900	2500
40	70	2800	1600	4900
35	65	2275	1225	4225
40	68	2720	1600	4624
20	40	800	400	1600
25	60	1500	625	3600
50	80	4000	2500	6400
240	433	15595	8850	27849

Now, Pearson Correlation Coefficient (r)

$$r = \frac{n\Sigma xy - \Sigma x \Sigma y}{\sqrt{[n\Sigma x^2 - (\Sigma x)^2][n\Sigma y^2 - (\Sigma y)^2]}}$$

$$r = \frac{(7)(15595) - (240)(433)}{\sqrt{[7(8850) - (240)^2][7(27849) - (433)^2]}}$$

$$r = \frac{109165 - 103920}{\sqrt{[4350][7454]}}$$

$$r = \frac{5245}{\sqrt{32424900}}$$

$$r = \frac{5245}{5694.28661} \qquad = \qquad 0.9211$$

Example 4.3.2:
The costs on advertisement (x) and revenues (y) generated by a company for 10 months are given below:

Advertisement (x) (₦'000)	45	70	32	24	75	16	28	43	60	15
Revenue (y) (Million)	42	51	38	39	44	20	22	46	47	36

Determine the product moment correlation coefficient for the table.

Solution:

Let x and y represent the cost of advertisement and revenue respectively.

X	y	xy	x^2	y^2
45	42	1890	2025	1764
70	51	3570	4900	2601
32	38	1216	1024	1444
24	39	936	576	1521
75	44	3300	5625	1936
16	20	320	256	400
28	22	616	784	484
43	46	1978	1849	2116
60	47	2820	3600	2209
15	35	525	225	1225
408	384	17171	20864	15700

Now, Pearson Correlation Coefficient (γ_k)

$$r = \frac{n\Sigma xy - \Sigma x\Sigma y}{\sqrt{[n\Sigma x^2 - (\Sigma x)^2][n\Sigma y^2 - (\Sigma y)^2]}}$$

$$r = \frac{(10)(17171) - (408)(384)}{\sqrt{[10(20864) - (408)^2][10(15700) - (384)^2]}}$$

$$r = \frac{171710 - 156672}{\sqrt{[208640 - 166464][157000 - 147456]}}$$

$$r = \frac{15038}{\sqrt{[42176][9544]}}$$

$$r = \frac{15038}{}$$

$$r = \frac{\sqrt{402527744}}{15038}$$
$$20063.0941$$

r = 0.7495

(b.) Rank correlation (Spearman's Rank Correlation)

There are occasions when we wish to put some object in an ordinal scale without giving actual marks to them. For instance, during a beauty contest the judges place the contestants in some order like first, second, third etc. without actually giving specific marks to them. This process is called ranking and the ordinal scale is used.

Ranking occurs where either for lack of time, money or suitable measurements may be impossible to quantify the items. In dealing with a correlation problem where the values are in ranks, rank correlation methods are used.

The best known of such methods is the Spearman's Rank correlation co-efficient which is defined as:

$$R = 1 - \frac{6\Sigma d_i^2}{n(n^2 - 1)}$$

where d_i = difference in each pair of ranks

n = number of objects being ranked.

R is defined in such a way that when the ranks are in perfect agreement, R equals +1 and when they are in perfect disagreement R equals −1.

Example 4.3.3: If two judges P and Q ranked 10 contestants in a beauty show as follows:

Contestant	H	I	J	K	L	M	N	O	P	Q
Rank by P	4	2	5	6	1	8	3	7	10	9
Rank by Q	3	4	2	8	1	9	5	6	10	7

Obtain the Spearman's Rank correlation coefficient from the given data.

Solution

Contestant	H	I	J	K	L	M	N	O	P	Q
Rank by P(x)	4	2	5	6	1	8	3	7	10	9
Rank by Q(y)	3	4	2	8	1	9	5	6	10	7
d = x − y	1	-2	3	-2	0	-1	-2	1	0	2
d^2	1	4	9	4	0	1	4	1	0	4

Σd^2 = 28

$$R = 1 - \frac{6\Sigma d_i^2}{n(n^2 - 1)}$$

$$= 1 - \frac{6\,(28)}{10(99)} = 1 - \frac{168}{990}$$

$$= 1 - 0.169696969 = 0.83030303$$

Example 4.3.4: The following are the marks obtained by five students in Mathematics (x) and Accounts (y). Determine the rank correlation coefficient of the data.

x	11	12	13	15	19
y	16	12	14	20	18

Solution

Since the marks are not given in ranks, we need to first rank the marks

x	y	Ranking x	Ranking y	d = x − y	d^2
11	16	5	3	2	4
12	12	4	5	-1	1
13	14	3	4	-1	1
15	20	2	1	1	1
19	18	1	2	-1	1
					8

Note that ranking can be done in ascending order or descending order. What is important is consistency in the order chosen to use.

Σd^2 = 8

$$R = 1 - \frac{6\Sigma d_i^2}{n(n^2-1)}$$

$$= 1 - \frac{6\times 8}{5\times 24} = 1 - \frac{(8)}{5(4)} = 1 - \frac{8}{20}$$

$$= 1 - 0.4 = 0.6$$

Remark 1:

Note that when ties of ranks occur (i.e when two or more of the ranks are the same), each value of the ties is given the mean of that ranks. For example, if two places are ranked second then each place takes the value of $1/2$ $(2+3) = 5/2$, since for the ties one place would have been second while the other one would have been third. Similarly if three places are ranked sixth then each one is ranked $1/3(6+7+8) = 21/3$. Having done this, the procedure for getting R is the same as outlined before.

Example 4.3.5: The following are marks obtained in Mathematics (x) and Accounts (y) by 10 students. Calculate the rank correlation coefficient of the data.

Students	H	I	J	K	L	M	N	O	P	Q
X	14	12	18	14	17	16	20	11	13	19
Y	13	11	15	18	15	19	15	12	14	20

Solution

X	y	Ranking x	Ranking y	$d = x - y$	d^2
14	13	6.5	8	-1.5	2.25
12	11	9	10	-1	1
18	15	3	5	-2	4
14	18	6.5	3	3.5	12.25
17	15	4	5	-1	1
16	19	5	2	3	9
20	15	1	5	-4	16
11	12	10	9	1	1
13	14	8	7	1	1
19	20	2	1	1	1
					48.5

Note that in variable x, there are two persons with the same mark 14 and the positions to be taken by the two of them are 6 and 7th positions. Hence, the average position is $= (6 + 7)/2 = 6.5$

For the variable y, there are three persons with the same mark of 15 and the positions to be taken

by the three of them are 4th, 5th and 6th positions. Hence, the average position is 15/3 i.e 5.

Σd^2 = 48.5

R = $1 - \dfrac{6\Sigma d_i^2}{n(n^2 - 1)}$

= $1 - \dfrac{6(48.5)}{10(10^2-1)}$ = $1 - \dfrac{6(48.5)}{10(99)}$ = $1 - \dfrac{291}{990}$

= $1 - 0.2939394$ = 0.7060606

Example 4.3.6: Using the table in example 4.3.2, compute its rank correlation coefficient.

x	y	Rank of x	Rank of y	$d_i = R_x - R_y$	d_i^2
45	42	7	6	1	1
70	51	9	10	-1	1
32	38	5	4	1	1
24	39	3	5	-2	4
75	44	10	7	3	9
16	20	2	1	1	1
28	22	4	2	2	4
43	46	6	8	-2	4
60	47	8	9	-1	1
15	35	1	3	-2	4
					30

R = $1 - \dfrac{6\Sigma d_i^2}{n(n^2 - 1)}$

R = $1 - \dfrac{6(30)}{10(10^2 - 1)}$

R = $1 - \dfrac{180}{990}$

R = $\dfrac{990 - 180}{990}$

R = $\dfrac{810}{990}$

R = 0.8182

Remark: By comparing the two coefficients in examples 4.3.2 and 4.3.6, it could be seen that the Spearman's Rank correlation coefficient is on the high side compared with the product moment correlation coefficient. It therefore, implies that the Spearman's Rank correlation Coefficient is not as accurate as the product moment correlation.

4.3.2 Interpretation of correlation

1. Interpretation of product moment correlation is as follows:

   (a.) there is a perfect relationship between the two variables x and y if r is close to ± 1. That is there is good correlation.

   (b.) when r is close to zero(0), there is no relationship, that is, there is poor or non – existence of correlation.

   For example:

r	=	+0.92 }	Good correlation
r	=	-0.96 }	Good correlation
r	=	-0.12 }	Poor or non-existent correlation
r	=	0 }	Non-existent correlation or no correlation at all
r	=	+0.26 }	Poor or non- existent correlation

2. **Interpretation of Rank correlation:** In rank correlation, when R approaches (+1) there is a perfect agreement but when R tends to (-1), this indicates a perfect disagreement.

4.4 Simple Regression Line

In the previous section, we discussed how to measure the degree of association. In this section, the nature of the relationship that exists between two variables x and y will be looked at, the interest may be about the students' performance who offered mathematics and accounts. We want to know whether their performance in a subject like accounts is affected by how well they do in mathematics. In this case, our variable y represents the accounts (the dependent or response variable), while mathematics represents the independent or explanatory variable. The relationship between these two variables is characterized by mathematical model called a Regression Equation.

For a simple linear regression and for regression y on x, we have:

$$y = a + bx$$

where y is the dependent variable.

x is the independent variable

a is the intercept of the line on the y-axis (the point where the line cuts the y-axis).

b is known as the regression coefficient. This is, the slope or gradient of the regression line.

4.4.1   Methods of obtaining or fitting regression line

We have two methods of fitting the regression line. These are (i) graphical and (ii) calculation.

(a.)   **Graphical method:** The following steps are to be taken:

i.   Draw the scatter diagram for the data.

ii.   Look at two points that a straight line will pass through on the diagram. One of the points ought to be ($\bar{x}, \bar{y}$).

iii.   Estimate constants a and b from the graph.

a. =   intercept on the y – axis of the drawn straight line.

b. =   slope or gradient of the line drawn i.e slope   =   vertical length
horizontal length.

iv.   Regression line y = a + bx is stated.

Example 4.4.1.1: Suppose we are required to find the relationship between two variables y and x on which we have the following data:

x	1	2	3	4	5
y	3	5	7	9	11

Solution
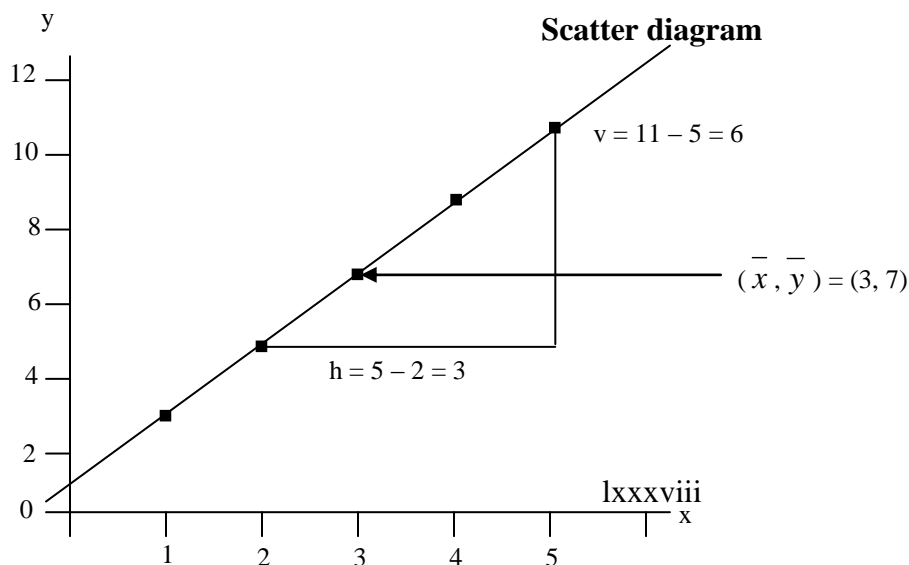
Scatter diagram

$v = 11 - 5 = 6$

$(\bar{x}, \bar{y}) = (3, 7)$

$h = 5 - 2 = 3$

lxxxviii

Slope (b)   = $\dfrac{\text{Vertical length}}{\text{Horizontal length}}$

= $\dfrac{v}{h}$

$a = 1 , \quad b = {}^{v}/_{h} = {}^{6}/_{3} = 2$

$\therefore \ y = a + bx \qquad => \qquad y = 1 + 2x$

(b.)   **Calculation method**: In the calculation method, we use the "normal equation" which is derived by the least squares method.   The said normal equations are:

$na + b\Sigma x \quad = \quad \Sigma y \quad \}$ \qquad\qquad ------4.4.1.1(a)

$a\Sigma x + b\Sigma x^2 = \quad \Sigma xy \quad \}$ \qquad\qquad ------4.4.1.1(b)

which are used to fit the regression line of y on x as $y = a + bx$

It should be noted that when equations 4.4.1.1 are solved simultaneously, we have the following estimates of a and b:

$b \quad = \quad \dfrac{n\Sigma xy - \Sigma x\Sigma y}{n\Sigma x^2 - (\Sigma x)^2}$

$a \quad = \quad \bar{y} - b\bar{x}$

Example 4.4.1.2: Using the last question (under the graphical approach), calculate or fit the regression line of y on x.

Solution:

x	y	xy	x^2
1	3	3	1
2	5	10	4
3	7	21	9
4	9	36	16
5	11	55	25
15	35	125	55

$5a + b(15) \quad = \quad 35 \qquad\qquad ---1$

$a(15) + b(55) = \quad 125 \qquad\qquad ---2$

$5(a) \quad = 35 - b(15)$

$a \quad = \dfrac{35 - b(15)}{5} \qquad\qquad$ OR

a $= 7 - b(3)$

$15(7 - b(3)) + b(55) = 125$

$105 - b(45) + b(55) = 125$

$105 - 45b + 55b = 125$

$-45b + 55b = 125 - 105$

$10b = 20$

$b = 20/10 = 2$

$5a + b(15) = 35$

$5a + 2(15) = 35$

$5a + 30 = 35$

$5a = 35 - 30$

$5a = 5$

$a = 5/5 = 1$

$\therefore y = 1 + 2x$ is the regression line of y on x

Note that when one compares the results of examples 4.4.1 and 4.4.2, the calculation method gives a better result.

Example 4.4.1.3:

The table below shows the income and expenditure (in ₦'000) of a man for 10 months.

Income (x)	8	18	52	38	26	60	40	50	82	75
Expenditure (y)	2	4	5	7	9	11	13	15	20	23

Fit simple linear regression line $y = a + bx$ to the data.

Solution

x	y	xy	x^2	y^2
8	2	16	64	4
18	4	72	324	16
52	5	260	2704	25
38	7	266	1444	49
26	9	234	676	81
60	11	660	3600	121
40	13	520	1600	169
50	15	750	2500	225
82	20	1640	6724	400
75	23	1725	5625	529
449	109	6143	25261	1619

xc

From the model

$$y = a + bx$$

$$b = \frac{n\Sigma xy - \Sigma x\Sigma y}{n\Sigma x^2 - (\Sigma x)^2}$$

$$a = \bar{y} - b\bar{x}$$

$$b = \frac{(10)(6143) - (449)(109)}{(10)(25261) - (449)^2}$$

$$b = \frac{12489}{51009}$$

$$b = 0.2448$$

Also,

$$a = \bar{y} - b\bar{x}$$

$$\bar{y} = \frac{\Sigma y}{n} \,\, \Sigma y = \frac{109}{} = 10.9$$

$$\bar{x} = \frac{\Sigma x}{n} = \frac{449}{10} = 44.9$$

Now,

$$a = 10.9 - (0.2448)(44.9)$$

$$a = 10.9 - 10.9915$$

$$a = -0.0915$$

Hence,

The model is

$$y = a + bx$$

$$y = -0.0915 + 0.2448x$$

4.4.2 Computation or fitting regression line of variable x on variable y.

In section 4.4.1, fitting the regression of variable y on variable x was discussed. But in this section, we present the regression of variable x on variable y which gives the regression line.

$$x = a' + b'y$$

where $$b' = \frac{n\Sigma xy - \Sigma x\Sigma y}{n\Sigma y^2 - (\Sigma y)^2}$$

$$a' = \bar{x} - b'\bar{y}$$

We are to note that variables x and y are now dependent and independent variables respectively.

Example 4.4.2.1:

Using the data in example 4.4.1.3, determine the regression line of variable x on variable y.

Income (x)	8	18	52	38	26	60	40	50	82	75
Expenditure (y)	2	4	5	7	9	11	13	15	20	23

Fit simple linear regression line to the table.

Solution

x	y	xy	x^2	y^2
8	2	16	64	4
18	4	72	324	16
52	5	260	2704	25
38	7	266	1444	49
26	9	234	676	81
60	11	660	3600	121
40	13	520	1600	169
50	15	750	2500	225
82	20	1640	6724	400
75	23	1725	5625	529
449	109	6143	25261	1619

$$b' = \frac{n\Sigma xy - \Sigma x \Sigma y}{n\Sigma y^2 - (\Sigma y)^2}$$

$$= \frac{10 \times 6143 - 449 \times 109}{10 \times 1619 - (109)^2}$$

$$= \frac{61430 - 48941}{16190 - 11881}$$

$$= \frac{12489}{4309}$$

b' = 2.8984

$$a' = \frac{\Sigma x}{n} - b'\frac{\Sigma y}{n}$$

$$= \frac{449}{10} - \frac{2.8984(109)}{10}$$

$$= 44.9 - 2.8984(10.9)$$

$$= 44.9 - 31.5926$$

$$= 13.3074$$

∴   The model is x = 13.3074 + 2.8984y

- If this result is compared with the one obtained in example 4.4.1.3, you can see that the two results have nothing in common; it is not a question of change of subjects. Thus, regression y on x is DEFINITELY different from regression x on y.

4.4.3 Application and Interpretation of Linear Regression

1.   Since regression coefficient (b) is the slope of the regression line, its magnitude gives an indication of the steepness of the line and it can be interpreted as thus.

   i.   If b = 0, it means the line is parallel to x axis.

   ii.   If b is high and positive, it gives very steep and upward slopping regression.

   iii.   If b is negative, it gives downward sloping of the regression line.

2.   The regression line can be used for prediction. When the regression equation or regression line is used to obtain y value corresponding to a given x value, we say that x is used to predict y. We can equally use y to predict x.

Example 4.4.3.1: If y = 16.94 + 0.96x, what is y when x = 50.?

Solution

   y = 16.94 + 0.96(50)

   y = 16.94 + 48

   y = 64.94

4.5   Coefficient of Determination.

   The concept of "coefficient of determination" demonstrates the relationship between the coefficient of correlation and regression coefficients. Coefficient of determination, usually denoted by r^2, is defined by:

   r^2 =   b x b'

   where r, b and b' are earlier defined.

Example 4.5.1

Using the results in examples 4.4.1.3 and 4.4.2.1, determine

a.   coefficient of determination.

b.   correlation coefficient of the data.

Solution:

a.   Since r^2 = $b \times b'$

                    b = 0.2448
                    b' = 2.8984
                 $\therefore r^2$ = 0.2448 x 2.8984
                    = 0.7095

b.   Correlation coefficient of the data is

        r^2 =   0.7095

        r = $\sqrt{0.7095}$

        r =   0.8423

4.6   Summary

Treatment of bivariate data by the use of correlation and regression analyses is presented. Correlation measures the degree of association while regression gives the pattern of the relationship between the two variables. Two types of correlation coefficients, namely: the Pearson's product moment correlation and Spearman's rank correlation are considered. Regression line fitting by graphical and calculation methods are considered. The use of regression to predict correlation coefficient and coefficient of determination is also presented.

4.7   Multiple-choice and short – answer questions

1.   Correlation measures

        a.   pattern of relationship   b.   degree of association

        c.   degree of statistical analysis   d.   pattern of statistical analysis

        e.   nature of statistical analysis

2.   When the rank correlation is -1, it means:

        a.   perfect agreement   b.   directly proportional relationship

        c.   perfect disagreement   d.   indirect relationship

        e.   unstable relationship

3.   Which of the following is the regression coefficient in $y = a + bx$, where x and y are

variables?

   a. a+b,   b. a − b,   c. ab,   d. a   e. b

Use of the following table to answer questions 4 and 5

x	1	2	3	4
y	2	3.5	1	3.5

4. Assuming variables x and y are marks, the product moment correlation coefficient is____

5. Assuming variables x and y are ranks, the rank correlation coefficient is_____

Use the following table to answer questions 6, 7 and 8

x	2	3	4
y	6	4	2

6.   Using the regression line $y = a + bx$, the value of b is _____

7.   Using the regression line $y = a + bx$, the value of a is_____

8.   If correlation coefficient is 0.95, the value of the regression coefficient of regression line x on y is _____

Use the following information to answer questions 9 and 10. Given the regression line $y = 5 + 1.5x$,

9.   the value of y when x = 6 is _____.

10.   the value of x when y = 11 is _____.

Answers to Chapter Four

1. b,   2. c,   3. e

4.   r $= \dfrac{n\sum xy - \sum x \sum y}{\sqrt{\left[n\sum x^2 - (\sum x)^2 \right] \left[n\sum y^2 - (\sum y)^2 \right]}}$

    r $= \dfrac{4(26) - 10(10)}{\sqrt{[4(30) - (10)^2][4(29.5) - (10)^2]}} = \dfrac{4}{\sqrt{(20)(18)}} = 0.2108$

5.   R $= 1 - \dfrac{6\sum d_i^2}{n(n^2 - 1)} = 1 - \dfrac{6(7.5)}{4(4^2 - 1)} = 1 - {}^{45}/_{60}$

          $= 1 - 0.75 = 0.25$

6. $b = \dfrac{n\Sigma xy - \Sigma x\Sigma y}{n\Sigma x^2 - (\Sigma x)^2} = \dfrac{3(32) - 9(12)}{3(29) - 9^2} = \dfrac{96 - 108}{87 - 81} = -2$

7. $a = \bar{y} - b\bar{x} = 4 - (-2)3 = 10$

8. Coefficient of determination $r^2 = (0.95)^2 = 0.9025$

    i.e $0.9025 = b \cdot b' = -2b'$

    $\Rightarrow b' = \dfrac{0.9025}{-2} = -0.4513$

    where b' is the regression coefficient of x on y.

9. $y = 5 + 1.5(6) = 14$

10. $y = 11 = 5 + 1.5x \Rightarrow 1.5x = 6 \Rightarrow x = 4$

CHAPTER FIVE

TIME SERIES ANALYSIS

5.0   Learning objectives

At the end of the chapter, students or readers should be able to

- know the meaning of a time series
- know the basic components of a time series;
- estimate the trend by the use of moving averages and least squares methods;
- carry-out exponential smoothening; and
- estimate seasonal index.

5.1   Introduction

A time series is a set of data that are successively collected at regular intervals of its time of occurrence. The regular interval time can be daily, weekly, monthly, quarterly or yearly. Examples of some time series data include:

- Monthly production of a company.
- Daily sales at a medical store.
- Amount of annual rainfall over a period of time.
- Money deposited in a bank on various working days.

Furthermore, it is essential to know that when a time series is analysed, it has the following benefits:

i.   Understanding the past behaviour of a variable and be able to

- determine direction of periodic fluctuations;

- predict future tendencies of the variable.

ii.   Determining the impact of the various forces influencing different variables which then facilitate their comparison, such as

- the differences that may have to do with price of commodities;

- the physical quantity of goods produced, marketed or consumed in order to make a comparison between periods of time, schools, places and etc.

iii. Knowing the behaviour of the variables in order to iron out intra-year variations as control events.

5.2 Basic components of a time series

In time series, the existence of fluctuations from time to time is caused by composite forces which are at work constantly;

These factors have four components viz:

i. Secular Trend or Secular Variation (T)

ii. Seasonal Fluctuations or Variation (S)

iii. Cyclical Variation (C)

iv. Irregular or Random Variations (I).

5.2.1 Secular Trend or Secular Variation (T)

The pattern of time series trend may be linear or non-linear. It is linear when the series values are concentrated along a straight line on the time – plot. Time – plot is known as the graph of time series values against different time points. The sketches of time plot showing typical examples of a secular trend are given below.

(a) upward movement   (b) Downward movement
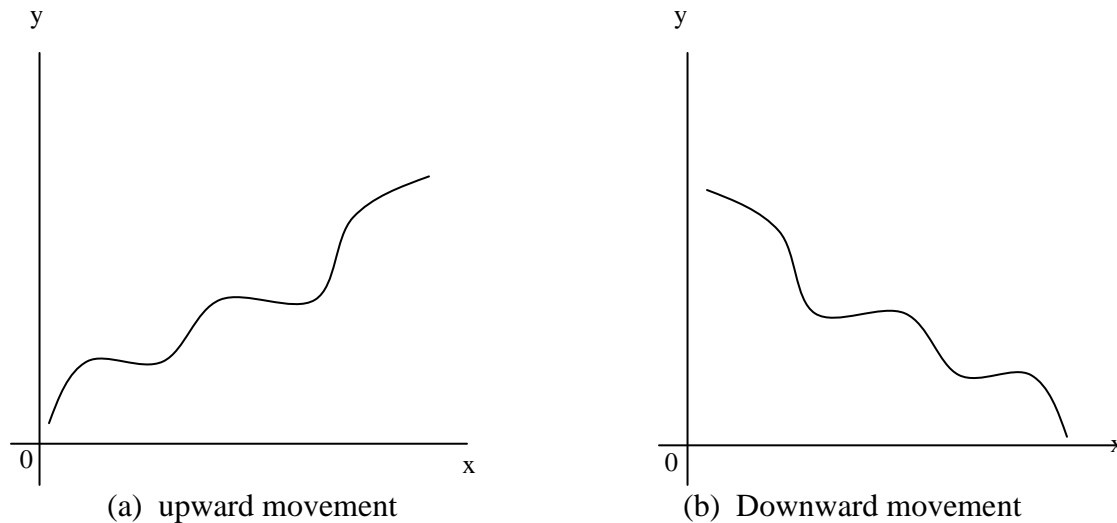
Fig 5.2.1

Example 5.2.1

The following table shows the number of cartons of Gulder beer sold by a retailer in Lagos over twelve consecutive months in a year.

Table 1

Month	1	2	3	4	5	6	7	8	9	10	11	12
Sales	14	24	42	19	29	38	43	58	48	43	64	74

Draw the time – plot of the table above

Solution   **TIME PLOT OF GULDER BEER SOLD IN TWELVE MONTHS**
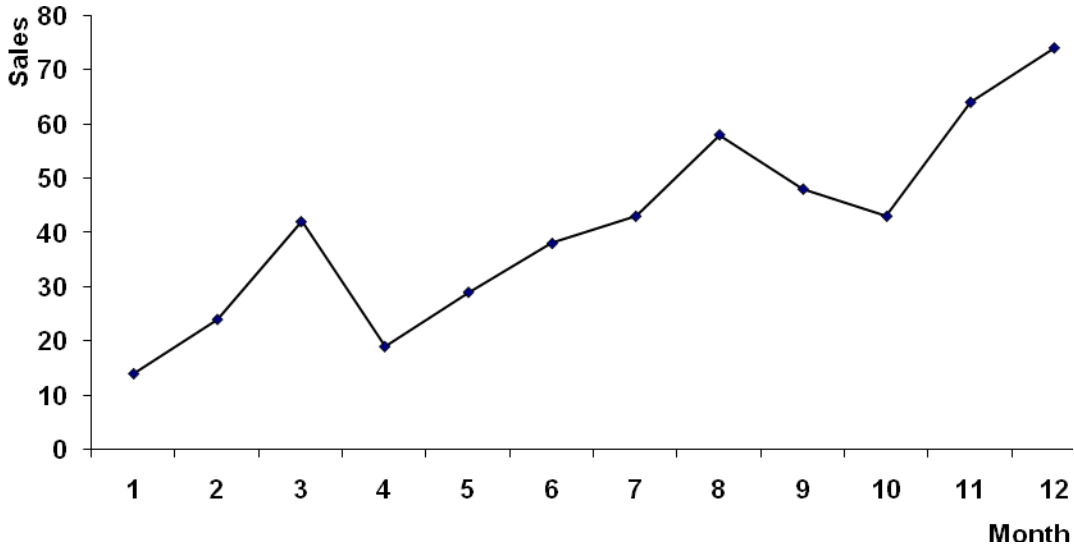
5.2.2   Seasonal Fluctuations or Variations

This is a variation that repeatedly occurs during a corresponding month or period of successive years. It is an annual reoccurring event which a time series appears to follow in a particular period or time of the year.

Data on climate such as rainy season, sales of goods during Christmas are good examples of time series with seasonal variations. The figure below depicts a time-plot showing the presence of seasonal variation.
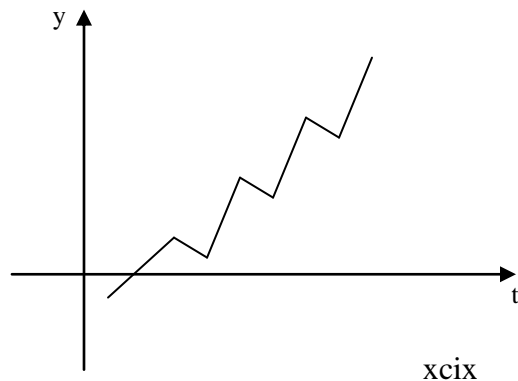
xcix

5.2.3 Cyclical Variation (C)

This is a long-term oscillation or wavelike fluctuation about the trend line of a time series. It is similar to seasonal variation with a difference of reoccurring in more than one year period. Cyclical variations are called business cycles in the sense that periods of prosperity followed by recession or depression and then recovery are caused by aggregate economic conditions rather than seasonal effects. The length of the cycle varies between four and seven years.

The cyclical variation is less predictable. The figure depicted below shows the pattern of cyclical variation.
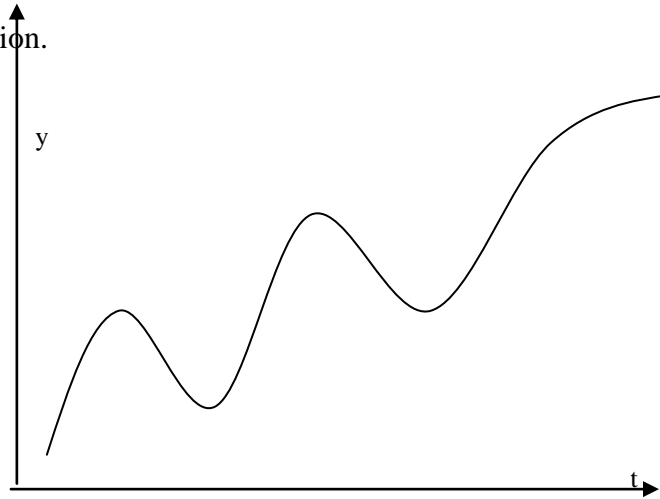
5.2.4 Irregular or Random Variation (I)

This is a variation caused by sporadic events (unpredictable events) such as floods, strikes, disasters, wars.

Random variation represents the residual variation in time series which can not be accounted for by the three other components (i.e Trend, Seasonal, Cyclical variations). See the figure below for a typical random variation.
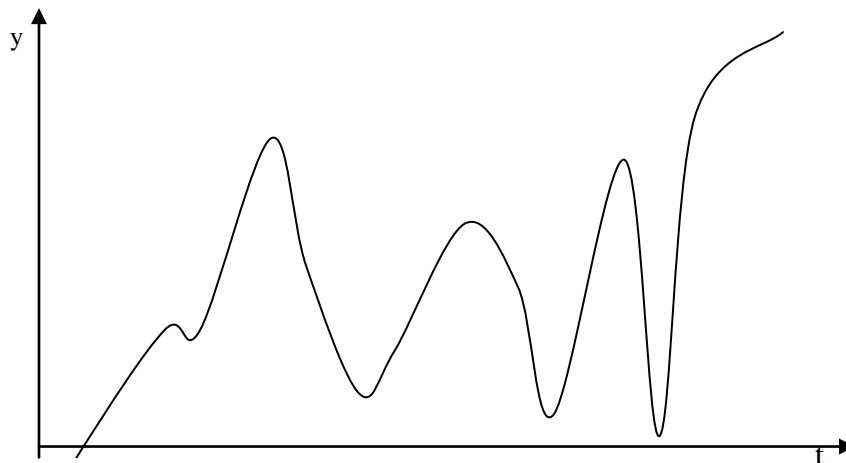
c

5.3   Time series analysis

This is an act of analysing and interpreting time series data. It can be said to be an investigative method into the time series components; which atimes is referred to as decomposition of a time series.

The first step in the time series analysis is the drawing of the "time-plot". This will clearly show the pattern of the series movement.

Let Y represent the series by convention and T, S, C, I as components earlier indicated in section 5.2. The two models of time series are additive and multiplicative models. For the additive, it is written as

$$Y = T + S + C + I \qquad\qquad 5.3.1$$

while in multiplicative model, it is written as;

$$Y = TSCI \qquad\qquad 5.3.2$$

5.4   Estimation of Trend

There are various methods of estimating trend. However, this study pack will consider only two methods, namely:

i.   Moving average method, and

ii.   Least Squares method.

5.4.1   Moving Average Method

This is a popular method of trend estimation. The trend is obtained in this method by smoothing out the fluctuation. The procedure of computing the trend by the Moving Average (M.A) method depends on whether the period of moving average desired, atimes called the order, is even or odd:

For a given time series Y_1, Y_2, Y_3, ---, Y_n, moving averages of order n are given by.

$$\frac{Y_1 + Y_2 + Y_3 + ---- + Y_n}{n}, \qquad\qquad \frac{Y_2 + Y_3 + Y_4 + ---- + Y_{n+1}}{n}$$

$$\frac{Y_3 + Y_4 + Y_5 + ---- + Y_{n+2}}{n}, \qquad\qquad e.t.c \qquad\qquad (5.4.1.1)$$

where the moving averages are expected to be written against the middle items considered for each average and the numerators of equation 5.4.1.1 are called the moving totals.

The following are procedural steps of computing trend by moving average (M.A):

1. Determine the order to be used whether odd or even.

2. When it is odd, directly apply equation 5.4.1.1 to get the desired moving average (M.A); and

3. If it is even, first obtain the moving totals of order n from the given series and then obtain a 2 combined n moving totals of earlier obtained moving totals. The moving totals lastly obtained will be divided by 2n to give the desired M.A. The purpose of using two moving totals is to overcome the problem of placing the M. A against middle items considered. By convention, M. A. must be written against middle item of the items considered. Hence, moving totals make this possible.

The following examples illustrate the trend estimation by Moving Average (M.A) method:

Example 5.4.1.1

Using the data in Table 1 of Example 5.2.1 i.e

Month(x)	1	2	3	4	5	6	7	8	9	10	11	12
Sales(y)	14	24	42	19	29	38	43	58	48	43	64	74

Determine (a) the moving average of order 3.

(b) the moving average of order 4

Solution

a.

Month(x)	Sales(y)	3 – Month moving total	3 – month moving average
1	14		
2	24	80	26.67
3	42	85	28.33
4	19	90	30.00
5	29	86	28.67
6	38	110	36.67
7	43	139	46.33
8	58	149	49.67
9	48	149	49.67
10	43	155	51.67

11	64	181	60.33
12	74		

b.

C_1 Month(x)	C_2 Sales(y)	C_3 4 – Month moving total	C_4 2 of 4 – month moving total	C_5 4 – month average $= C_4 \div 8$
1	14			
2	24			
		99		
3	42		213	26.63
		114		
4	19		242	30.25
		128		
5	29		257	31.13
		129		
6	38		297	37.13
		168		
7	43		355	44.38
		187		
8	58		379	47.38
		192		
9	48		405	50.63
		213		
10	43		442	55.25
		229		
11	64			
12	74			

Merits and Demerits of Moving Average (M.A) Method

Merits:

i. The method is simple if compared with least squares method.

ii. The effect of cyclical fluctuations is completely removed if the period of Moving Average (M.A) is equal to the average period of cycles.

iii. It is good for a time series that reveals linear trends.

Demerits

i.   The extreme values are always lost by Moving Average (M.A) method.

ii.   The method is not suitable for forecasting.

iii.   It is not good for non-linear trend.

5.4.2   Least Squares Method (L.S.M)

Recall that by least squares method, fitting a linear regression of variable (y) on variable (x) gives $y = a + bx$, where a and b are constants and are the intercept and regression coefficient respectively. This method can equally be extended to a time series data by taking the time period t as variable x and the time series value as variable y. Hence, the trend line by L.S.M is given as :

$$y = a + bx \qquad\qquad\qquad 5.4.2.1$$

$$\text{where } b = \frac{\Sigma xy - \Sigma x\Sigma y/n}{\Sigma x^2 - (\Sigma x)^2/n} \qquad\qquad 5.4.2.2$$

$$\text{and } \bar{a} = \bar{y} - b\bar{x} \qquad\qquad\qquad 5.4.2.3$$

The following examples give illustrations of Trend estimation by L.S.M.

Example 5.4.2.1

Using the data in table in Example 5.2.1, estimate the trend and fit the trend by least squares approach.

Month(x)	1	2	3	4	5	6	7	8	9	10	11	12
Sales(y)	14	24	42	19	29	38	43	58	48	43	64	74

Solution

Month(x)	Sales(y)	x^2	xy	Trend $\tilde{Y} = 62.0604 - 1.7657x$
1	14	1	14	60.2947
2	24	4	48	58.5290
3	42	9	126	56.7633
4	19	16	76	54.9976
5	29	25	145	53.2319
6	38	36	228	51.4662
7	43	49	301	49.7005
8	58	64	301	47.9348
9	48	81	432	46.1691
10	43	100	430	44.4034
11	64	121	704	42.6377
12	74	144	888	40,8720
78	607	650	3693	

From

$$Y = a + bx$$

where

$$b = \frac{n\Sigma xy - \Sigma x\Sigma y}{n\Sigma x^2 - (\Sigma x)^2}$$

$$= \frac{(12)(3639) - (78)(607)}{(12)(650) - (78^2)}$$

$$= \frac{44316 - 47346}{7800 - 6084}$$

$$= \frac{-3030}{1716}$$

$$\therefore \quad b = -1.7657$$

Also

$$a = \overline{y} - b\overline{x}$$

$$\overline{y} = \frac{\Sigma y}{n}$$

$$\overline{y} = \frac{607}{12} = 50.5833$$

$$\overline{x} = \frac{\Sigma x}{n}$$

$$\overline{x} = \frac{78}{12} = 6.5$$

Now,

$$a = \overline{y} - b\overline{x}$$

$$a = 50.5833 - (-1.7657)(6.5)$$

$$= 50.5833 + 11.47705$$

$$= 62.06035$$

$$\therefore a = 62.0604$$

Recall

$$y = a + bx$$

$$y = 62.0604 + (-1.7657)x$$

$$y = 62.0604 - 1.7657x$$

REMARK: As the data in time series involved large number of period and data, the use of L.S.M will require us to reduce the computational tediousness by coding the period aspect of the data.

The principle of the coding is achieved by the change of variable x to t using the relationship: $t_i = x_i - x_m$, $i = 1, 2, --, n$

Where x_m is the median of x_i. By this idea, the new variable t_i will give $\Sigma t_i = 0$. If this condition ($\Sigma t_i = 0$) is imposed on any time series data, then the normal equations of equations 5.4.2.2 and 5.4.2.3 will reduce to:

$$b \quad = \quad \frac{\Sigma y_i t_i}{\Sigma t_i^2} \qquad \qquad \qquad ......5.4.2.4$$

$$a \quad = \quad \frac{\Sigma y_i}{n} \quad = \quad Y \qquad \qquad \qquad .....5.4.2.5$$

for the Trend $T = a + bt$

Example 5.4.2.2
Use the Least Square Method of coding to fit the regression line to the following table:

Year	Q_1	Q_2	Q_3	Q_4
2000	20	40	25	45
2001	30	50	37	60
2002	42	60	45	65

Solution
Let's consider the following arrangement of the data

Year	Quarter	X	$t = x - 5.5$	y	ty	t^2
2000	1	0	-5.5	20	-110	30.25
	2	1	-4.5	40	-180	20.25
	3	2	-3.5	25	-87.5	12.25
	4	3	-2.5	45	-112.5	6.25
2001	1	4	-1.5	30	-40	2.25
	2	5	-0.5	50	-25	0.25
	3	6	0.5	37	18.5	0.25
	4	7	1.5	60	90	2.25
2002	1	8	2.5	42	105	6.25
	2	9	3.5	60	210	12.25
	3	10	4.5	45	202.5	20.25
	4	11	5.5	65	357.5	30.25
			0	519	428.5	143

* Note that the median of variable x = 5.5

Least squares equation is $y = a + bx$, where

$$b \quad = \quad \frac{\Sigma ty}{\Sigma t^2} \quad = \quad \frac{428.5}{143} \quad = \quad 2.9965$$

and $a = \bar{y} - bt$

$$\bar{y} = \frac{519}{12} = 43.25$$

Hence,   a= $43.25 - 2.9965(5.5)$

$= 43.25 - 16.4808$

$= 26.7692$

$\therefore Y = 26.7092 + 2.9965(x - 5.5)$

$= 26.7692 + 2.9965x - 16.4808$

$Y = 10.2884 + 2.9965x$

Merits and Demerits of L. S. M

Merits

i.   No extreme values are lost in this method as in the case of the M. A.

ii.   The method is free from subjective error.

iii.   The method can be used for forecasting.

Demerits

i.   It requires more time for computation.

5.4.3   Exponential Trend (Exponential Smoothening)

When the trend shows an exponential function, where the x and y variables are in arithmetic and geometric progressions respectively, smoothening the trend is the appropriate approach. This is achieved as follows:

For an exponential function

$Y = ab^x$ …5.4.3.1

where a and b are constants, taking the logarithm of both sides of equation 5.4.3.1 yields

$Log\ Y = Log\ a + x\ Log\ b$ …..5.4.3.2

If $Log\ Y = z$, $Log\ a = A$ and $Log\ b = B$, equation 5.4.3.2 becomes

$Z = A + Bx$ ….5.4.3.3

Which is now in linear form of variables Z and x. The normal equations generated for the

estimation of A and B are given as:

$$\Sigma Z = nA + B\Sigma x \qquad\} \qquad ....5.4.3.4$$

$$\Sigma Zx = A\Sigma x + B\Sigma x^2 \qquad\} \qquad ....5.4.3.5$$

Solving equations 5.4.3.4, and 5.4.3.5 we finally get a = Anti-Log A; b = antilog B.

The estimated exponential trend is obtained by putting the estimated values of a and b in equation 5.4.3.1.

Example 5.4.3.1

Given the population censuses of a country for certain number of periods as:

Census Year (x)	1950	1960	1970	1980	1990
Population in Million (y)	25.0	26.1	27.9	31.9	36.1

Fit an exponential trend $Y = ab^x$ to the above data using the Least Squares Method (L.S.M).

Solution.

Census (x)	Population (y)	$u = \dfrac{(x - 1970)}{10}$	$v = \log y$	u^2	uv
1950	25.0	-2	1.3979	4	-2.7958
1960	26.1	-1	1.4166	1	-1.4166
1970	27.9	0	1.4456	0	0
1980	31.9	1	1.5038	1	1.5038
1990	36.1	2	1.5375	4	3.1150
		$\Sigma u = 0$	7.2314	10	0.4064

$Y = ab^x$

$\text{Log } y = \log ab^x$

$\text{Log } y = \text{Log } a + x\text{Log } b$

$v = A + BX$

where $\text{Log } y = V$, $\text{Log } a = A$, $\text{Log } b = B$.

$\therefore u = (x - 1970)/10$

So that $\Sigma u = 0$

$\therefore A = \dfrac{\Sigma v}{} = \dfrac{7.3214}{} = 1.4643 \rightarrow a = \text{antilog } A$

$$\frac{n}{} \qquad \frac{5}{} \qquad\qquad = 29.1259$$

$$B \;=\; \frac{\Sigma uv}{\Sigma u^2} \;=\; \frac{0.4064}{10} \;=\; 0.04064 \qquad\qquad \rightarrow\; b \;=\; \text{Antilog B}$$
$$= 1.0981$$

From $Y = ab^x$

$$Y \;=\; 29.1259 x 1.0981^{\left(\frac{x-1970}{10}\right)} \;=\; \qquad 29.1259(1.0981)^{\left(\frac{x-1970}{10}\right)}$$

To obtain the trend values Y for different x, we use the linear trend

$$V \;=\; A \;+\; Bu$$

$$V \;=\; 1.4643 \;+\; 0.04064u$$

5.5 Estimation of Seasonal Variation

Estimation of Seasonal variation involves the use of original time series data and the trend obtained. Here, the approach depends on the model assumed. The model can be additive or multiplicative as stated in equations 5.3.1 and 5.3.2.

For the <u>additive model</u>, the following steps are to be adhered to:

1. Arrange the original given time series(Y) in column;

2. The Trend T to be placed as next column;

3. Obtain the seasonal variation (S) in another column by subtracting T from Y; i.e

   $S = Y - T - (C + I)$, from equation 5.3.1. Here, it is assumed that

   $C + I = 0$, hence $S = Y - T$

   (Note that T(trend) can be obtained by either M. A. or L. S.M);

4. Obtain the average seasonal variation for each period/season (weeks, months, quarters, half-yearly, etc). This is known as seasonal index (S.I).

5. Check whether the (S.I) obtained in step 4 is a balanced one. Add the indices obtained and if the total is equal to zero, it implies balance; otherwise adjust them to balance.

6. For the adjustment of S. I., divide the <u>difference</u> (to zero) by the number of period/season and then add or subtract to S. I. obtained in step 4 as dictated by the sign of the difference in order to make the sum total zero.

For the <u>multiplicative model</u>, the following steps are to be followed:

1. The same as step 1 in the additive model.

2. The same as step 2 in the additive model.

3. Obtain the seasonal variation (S) in another column by dividing Y by T; i.e

S $=$ $\dfrac{Y}{T}$, from equation 5.3.2, It is also assumed that $CI = 1$,

hence $S = {}^{Y}/_{T}$.

4. The same as step 4 in the additive model.

5. Check whether the S. I. obtained in step 4 is balanced or not by having the total of the indices as number of period/season. For example, a quarterly period will have total of indices as 4, for half-yearly 2, etc. atimes when indices are expressed in percentages, it will be 400 for quarter, 200 for half-yearly, etc. If it is not balanced, there is need to adjust the indices to the balanced form.

6. The same as step 6 in the additive model with modification of <u>difference</u> to 4 or 2 as the period indicates.

Example 5.5.1:

A company secretary preparing for his retirement within the next six years decided to invest quarterly in the purchase of shares either private placement or through public offers. The table below shows his quarterly investments in 2000, 2001, 2002 and 2003

Year	**Quarter 1**	**Quarter 2**	**Quarter 3**	**Quarter 4**
2000	15	35	40	20
2001	25	45	55	30
2002	37	55	60	40
2003	47	62	72	58

Estimate the trend by least squares method. Hence, compute seasonal variation.

Solution

Define x as follows

Quarter 1	2000,	$x = 0$
Quarter 2	2000,	$x = 1$
Quarter 3	2000,	$x = 2$
Quarter 4	2000,	$x = 3$
Quarter 1	2001,	$x = 4$ e.t.c

Computation of least square trend line (Directed method)

x	**y**	**x^2**	**xy**	**Trend y = a – bx** **= 23.6692 + 2.6441x**
0	15	0	0	23.6692
1	35	1	35	26.0133

cx

2	40	4	80	28.6574
3	20	9	60	31.3015
4	25	16	100	33.9456
5	45	25	225	36.5897
6	55	36	330	39.2338
7	30	49	210	41.8779
8	37	64	296	44.5220
9	55	81	495	47.1661
10	60	100	600	49.8102
11	40	121	440	52.4543
12	47	144	564	55.0984
13	62	169	806	57.7425
14	72	196	1008	60.3866
15	58	225	870	63.0307
120	696	1240	6119	

From

$$y = a + bx$$

$$b = \frac{n\Sigma xy - \Sigma x\Sigma y}{n\Sigma x^2 - (\Sigma y)^2}$$

$n = 16, \Sigma xy = 6119, \Sigma x = 120, \Sigma y = 696$

$$b = \frac{(16)(6119) - (120)(696)}{(16)(1240) - (120)^2}$$

$$= \frac{97904 - 83520}{19840 - 14400}$$

$$= \frac{14384}{5440}$$

$$= 2.6441$$

Also,

$$a = \overline{y} - b\overline{x}$$

$$\overline{y} = \frac{\Sigma y}{n} = \frac{696}{16} = 43.5$$

$$\overline{x} = \frac{\Sigma x}{n} = \frac{120}{16} = 7.5$$

$a = 43.5 - (2.6441)(7.5)$

$a = 43.5 - 19.8308$

$a = 23.6692$

Hence,

$$y \quad = \quad a + bx$$

$$\text{Trend(T)}= y \quad = \quad 23.6692 + 2.6441x$$

x	y	Trend T = a + bx	Variation by multiplicative model (Y/T)
0	15	23.6692	0.6337
1	35	26.0133	1.3455
2	40	28.6574	1.3958
3	20	31.3015	0.6389
4	25	33.9456	0.7475
5	45	36.5897	1.2299
6	55	39.2338	1.4019
7	30	41.8779	0.7164
8	37	44.5220	0.8310
9	55	47.1661	1.1661
10	60	49.8102	1.2046
11	40	52.4543	0.7626
12	47	55.0984	0.8530
13	62	57.7425	1.0737
14	72	60.3866	1.1923
15	58	63.0307	0.9202

Seasonal Index Table

Year	Quarter 1(Q_1)	Quarter 1(Q_2)	Quarter 1(Q_3)	Quarter 1(Q_4)
2000	0.6337	1.3455	1.3958	0.6389
2001	0.7475	1.2299	1.4019	0.7164
2002	0.8310	1.1661	1.2046	0.7626
2003	0.8530	1.0737	1.1923	0.9202
Total	**3.0652**	**4.8152**	**5.1946**	**3.0381**
(Index) Average	0.7663	1.2038	1.2987	0.7595
	-0.007075	-0.007075	-0.007075	-0.007075

Adjusted Index	0.759225	1.196725	1.291625	0.752425

Total of Averages $=$ $0.7663 + 1.2038 + 1.2987 + 0.7595$

$=$ 4.0283

Since the total of averages is supposed to be 4, we need to adjust.

Adjustment by $\dfrac{4 - 4.0283}{4}$ $=$ $\dfrac{-0.0283}{4}$ $=$ -0.007075

5.6 Summary

The time series is described as a set of data collected at regular intervals of time. Analysis of time series reveals its principal components as Trend, Seasonal, Cyclic and Random variations. The methods of obtaining trend, as discussed in this book, are moving averages and least squares methods. The concept of seasonal index is considered from the perspective of additive and multiplicative models. Exponential smoothening is also discussed with numerical example.

5.7 Multiple-Choice and Short-Answer Questions

1. Which of the following is not a time series data?

    a. Monthly production of a company.

    b. Daily sales at a medicine store.

    c. Expenditures at home.

    d. Daily deposits in a bank.

    e. Annual rainfall.

2. For monthly time series data, the order of the moving average is_____

    a. 2; b. 4; c. 6; d. 10; e. 12

3. Using the conventional symbol of time series components, which of the following is the additive model?

    a. $P = T + C + I + S$; b. $Y = T + S + C + I$

    c. $Y = TSCI$ d. $P = TSCI$

    e. $Y = ab^x$

4. Which of the following is not true of moving average method?

    a. The method is simple if compared with Least Squares Method.

    b. The effect of cyclical fluctuations is completely removed by the method.

    c.   The extreme values are always lost.

    d.   The method is suitable for forecasting.

    e.   It is not good for non-linear trend.

Use the following table to answer questions 5 – 8.

Time (t)	Value of Series (Y)	Moving Average of order 3
1	24	
2	30	a
3	25	b
4	35	c
5	33	d
6	38	

5.   Find **a** from the table.

6.   Find **b** from the table.

7.   Find **c** from the table.

8.   Find **d** from the table.

Use the following information to answer questions 9 and 10

Time (t)	Series (Y)	Trend by LSM $Y = 30 + 3t$	Seasonal Variation assuming additive model
1	30		p
2	32		
3	37		q

9.   Find **p**.

10.   Find **q**.

Answers to Chapter Five

1. c

2. e

3. b

4. d

5. $a = \dfrac{24 + 30 + 25}{3} = 26.33$

6. $b = \dfrac{30 + 25 + 25}{3} = 30$

7. $c = \dfrac{25 + 35 + 33}{3} = 31$

8. $d = \dfrac{35 + 33 + 38}{3} = 35.33$

9. $P = 30 - \{30 + 3(1)\} = -3$

10. $q = 37 - (30 + 3(3)) = -2$

CHAPTER SIX

INDEX NUMBERS

6.0   Learning objectives

At the end of the chapter, readers and students are expected to

- understand the concept of Index numbers;

- differentiate between price indices and price relatives;

- know the differences between unweighted index numbers and weighted index numbers;

- compute and handle problems on index numbers by the use of weighted methods such as Laspeyre, Paasche, Fisher and Marshall Edgeworth methods.

6.1   Introduction

It is a usual practice in business, economy and other areas of life to find the average changes in price, quantity or value of related group of items or variables over a certain period of time, or for geographical locations. Index number is the statistical concept or devise that is usually used to measure the changes,

Spiegel defined Index number as "A statistical measure designed to show changes in a variable or a group of variables with respect to time, geographical location or other characteristics".

By the principle of index number, the statistical device measures

- the differences in the marginal of a group of related variables;

- the differences that may have to do with price of commodities;

- the physical quantity of goods produced, marketed or consumed in order to make a comparison between periods of time, schools, places and etc.

Based on the above principle of the index numbers, it can be classified in terms of the variables that it tends to measure. Hence, it is broadly categorized into:

- Price Index Number consisting of retail price.

- Quantity or Volume Index numbers.

- Value Index Numbers.

6.1.1 Uses / Applications of Index Numbers

It is important and of great benefits to state the following uses of Index numbers:

- To deflate a value series in order to convert it into physical terms.
- To keep abreast of current business condition i.e it acts as business or economic barometer.
- To give the trend movement in business or economy.
- To forecast by using series of the indices.
- To assess the worth of purchasing power of money.
- To compare the standard of living in various areas/countries or geographical locations.
- To compare students' intelligence in various schools or countries.

6.1.2 Problems associated with Constructions of Index Numbers

The following are among the problems usually encountered in the construction of index numbers:

i.   Definition of the purpose for which index number is being compiled or constructed.

ii.   Selection of commodities/items to include in the index. Here, one has to decide what type of item, what quantity and quality are to be selected.

iii.   Selection of sources of data. The data source must be reliable, hence, utmost care must be taken in selecting the source.

iv.   Method of collecting data: Once the source of data has been determined, the next line of action is to decide on an efficient and effective method of data collection. This will facilitate accurate and reliable results. Here, the method depends on the source whether it is of primary or secondary type of data.

v.   Selection or choice of base year: The base year is the reference point or period that other data are being compared with. It is essentially important that the period of choice as the base year, is economically stable and free from abnormalities of economy such as inflation, depression, famines, boom etc. The period should not be too far from the current period.

vi.   Methods of Combining data: The use of appropriate method of combining data depends on the purpose of the index and the data available. The choice of appropriate method dictates the formula to be used.

vii.   Choice of Weight: The weights in index number refer to relative importance attached to various items. It is therefore necessary to take into account the varied relative importance of items in the construction of index numbers in order to have a fair or an accurate index number.

6.2   Construction Methods of Price Index Numbers

The Price Index number measures the changes in the general level of prices for a given number or group of commodities. It could be wholesale price index or retail price index, or index for prices of manufactured products.

The construction methods of price index numbers can be broadly classified into two:

(i)   the use of unweighted price index numbers and

(ii)   use of weighted price index numbers.

6.3   Unweighted Price Index Number Method

In the unweighted index number, it is observed that equal importance is attached to all items in the index.

The following unweighted indices shall be considered in this section:

(a)   Simple Price Relative Index Number.

(b)   Simple Aggregate Price Index.

(c)   Simple Average of Relative Method.

6.3.1   Simple Price Relative Index Numbers (SPRI)

This Index number is the simplest form of all index numbers. It can be simply defined as the ratio of the price of a single commodity in a current or given year to the price of the same commodity in the base year.

This can be expressed mathematically as follows:

$$\text{SPRI} = \frac{P_{ti}}{P_{oi}} \times 100$$

where P_{ti} = current or given year price for item i.

P_{oi} = Base year price for item i.

6.3.2 Simple Aggregate Price Index (SAPI)

This index measures the changes in price level over time,using only the arithmetic mean and ignoring differences in the relative importance of the commodities. It is expressed as the total prices of commodities in a current or given year as a percentage of the total prices of the base year of the same commodities. Symbolically, it is given as

$$\text{SAPI} = \frac{\Sigma P_{ti}}{\Sigma P_{oi}} \times 100$$

where ΣP_{ti} = the sum of prices for item i at period t.

ΣP_{oi} = the sum of prices for item i at the base year.

6.3.3 Simple Average of Relative Price Index (SARPI)

This approach is to remove the shortcomings of the method of SAPI. The index is computed by the following formula:

$$\text{SARPI} = \frac{\sum \left(\dfrac{P_{ti}}{P_{oi}} \times 100 \right)}{n} \qquad \text{or} \qquad \frac{\sum \left(\dfrac{P_{ti}}{P_{oi}} \right) \times 100}{n}$$

where Σ stands for summation

n = the total number of items/commodities

P_{ti} and P_{oi} are as defined in SPRI.

Example 6.3.3.1
Determine the SPRI, SAPI and SARPI for the following table using 2004 as base year.

Table 6.1

Items /	Price per unit (Naira/Cedi)

commodities	Year 2004	Year 2005	Year 2006
A	20	30	60
B	30	42	55
C	10	15	20
D	6	8	15

Solution:

i.   SPRI $=$ $\dfrac{P_{ti}}{P_{oi}}$ x 100

SPRI$_{2005}$ $=$ $\dfrac{P_1}{P_0}$ x 100   SPRI$_{2006}$ $=$ $\dfrac{P_2}{P_0}$ x 100

$P_{0(2004)}$	$P_{1(2005)}$	$P_{2(2006)}$	SPRI$_{2005}$ P_1/P_0 x 100	SPRI$_{2006}$ P_2/P_0 x 100
20	30	60	150	300
30	42	55	140	183
10	15	20	150	200
6	8	15	130	250

ii.   SAPI

Items / commodities	Price per unit (Naira/Cedi)		
	Year 2004	Year 2005	Year 2006
A	20	30	60
B	30	42	55
C	10	15	20
D	6	8	15
	66	95	150

SAPI $=$ $\dfrac{\Sigma P_{ti}}{\Sigma P_{oi}}$ x 100

using 2004 as base year i.e $\Sigma Poi = 66$

SAPI$_{2005}$ $=$ $\dfrac{\Sigma P_{ti}}{\Sigma P_{oi}}$ x 100

$$= \quad \frac{95}{66} \times 100$$

$$= \quad 143.9394$$

$$\approx \quad ₦144$$

$$SAPI_{2006} \quad = \quad \frac{\Sigma P_{ti}}{\Sigma P_{oi}} \times 100$$

$$= \quad \frac{150}{66} \times 100$$

$$= \quad 227.2727$$

$$\approx \quad ₦227$$

iii. $SARPI = \sum \left[\left(\dfrac{P_{ti}}{P_{oi}} \right) X100 \right] / n$

Items	P_0	P_1	P_2	$\underline{P_1}$	$\underline{P_2}$
	Year 2004	Year 2005	Year 2006	P_0	P_0
A	20	30	60	1.5	3.00
B	30	42	55	1.4	1.83
C	10	15	20	1.5	2.00
D	6	8	15	1.3	2.50
	66	95	150	5.70	9.33

$$SARPI_{2005} \quad = \quad \frac{\Sigma(P_1/P_0 \times 100)}{n}$$

$$= \quad \frac{5.7 \times 100}{4}$$

$$= \quad 142.5 \qquad \approx \qquad ₦143$$

$$SARPI_{2006} \quad = \quad \frac{\Sigma(P_2/P_0 \times 100)}{n}$$

$$= \quad \frac{9.33 \times 100}{4}$$

$$= \quad 233.25 \qquad \approx \qquad ₦233$$

6.4   Weighted Index Numbers

In this index, weights are attached to each item or commodity on the assumption that such weights denote the relative importance of the items. The weighted index number can be broadly categorized into

i.   Weighted Aggregative Indices

ii.   Weighted Average of Relative Indices.

6.4.1 Weighted Aggregative Indices

As weights give relative importance to the group of items, various methods of the weighted aggregative index involve different weighing techniques. Some of the methods usually encountered, include:

a.   Laspeyre

b.   Paasche

c.   Fisher

d.   Marshall Edgeworth

(a)   Laspeyre's Index

In 1864, Laspeyre invented this method of weighted aggregative index by making use of the base year quantities as weights in the Laspeyre's Price Index. The method is commonly and widely used because it is easy to compute being based on fixed weights of the base year. It is computed using the formula

$$\text{La-}P_{oi} = \frac{\Sigma p_i q_o}{\Sigma p_o q_o} \times 100$$

where Σ stands for summation

P_i = price of the current or given year.

q_o = quantity of the base year.

The method of the index has an upward bias.

(b)   Paasche Index

The German statistician, Paasche, introduced this method in the year 1874. Its focus is on the usage of current or given year quantities as weights in the Paasche Price index. The formula for its construction is given by

$$\text{Pa-}P_{oi} = \frac{\Sigma p_i q_i}{\Sigma p_o q_i} \times 100$$

where p_i = current or given year price

p_o = base year price.

q_i = current or given year quantity.

The method is tedious to compute and of downward bias.

NOTE: This is a work in progress. All topics in the syllabus are covered but editing for necessary corrections is in progress.

Thanks.

(c) **Fisher Ideal Index**

This method is based on the geometric mean of Laspeyre's and Paasche indices. The method is theoretically better than other methods because it overcomes the shortcomings of these other methods. It is computed by

$$\text{Fis-P}_{oi} = \sqrt{\frac{\Sigma p_i q_o}{\Sigma p_o q_o} \times \frac{\Sigma p_i q_i}{\Sigma p_o q_i} \times 100}$$

OR

$$= \sqrt{(\text{La-P}_{oi}) \times (\text{Pa-P}_{oi})}$$

where pi, qi, po, qo are as defined before.

(d) **Marshall – Edgeworth Index**

The method uses the average of base and current/given year quantities as weights. It constructs the index by the formula:

$$\text{MaxEd-P}_{oi} = \frac{\Sigma p_i\{(q_o + q_i)/2\}}{\Sigma p_o\{(q_o + q_i)/2\}} \times 100$$

OR

$$= \frac{\Sigma(p_i q_o + p_i q_i)}{\Sigma(p_o q_o + p_o q_i)} \times 100$$

where pi, qi, po, qo are as defined before.

Example 6.4.1.1

Use the following table to calculate price index numbers for the year 2006 by taking 2001 as the base year and using the following methods:

i.   Laspeyre Index.

ii.   Paasche Index

iii.   Fisher Index

iv.   Marshall – Edgeworth Index

Table 6.2

Commodity	2001		2006	
Item	Price	Quantity	Price	Quantity
A	14	45	20	35
B	13	15	19	20
C	12	10	14	10
D	10	5	12	8

Solution

Commodity	2001		2006	
Item	p_o Price	q_o Quantity	p_i Price	q_i Quantity
A	14	45	20	35
B	13	15	19	20
C	12	10	14	10
D	10	5	12	8

$$\text{La-P}_{oi} = \frac{\Sigma p_i q_o}{\Sigma p_o q_o} \times 100$$

Items	p_o	q_o	p_i	q_i	$p_o q_o$	$p_i q_o$	$p_i q_i$	$p_o q_i$	$p_i q_o + p_i q_i$	$p_o q_o + p_o q_i$
A	14	45	20	35	630	900	700	490	1600	1120
B	13	15	19	20	195	285	380	300	665	495
C	12	10	14	10	120	140	140	100	280	220
D	10	5	12	8	50	60	96	40	156	90
					995	1385	1280	930	2701	1925

i. $\quad \text{La-P}_{oi} = \dfrac{\Sigma p_i q_o}{\Sigma p_o q_o} \times 100$

$$\text{Laspeyre Index}_{(2006)} = \frac{\Sigma p_i q_o}{\Sigma p_o q_o} \times 100$$

$$= \frac{1385}{995} \times 100$$

$$= 1.3920 \times 100$$

$$= 139.1960$$

$$\approx 139$$

ii. $\quad \text{Paasche Index}_{(2006)} = \dfrac{\Sigma p_i q_i}{\Sigma p_o q_i} \times 100$

$$= \frac{1280}{930} \times 100$$

$$= 1.3763 \times 100$$

$$= 137.6344$$

$$\approx \quad 138$$

iii. Fisher Index $= \sqrt{\dfrac{\Sigma p_i q_o}{\Sigma p_o q_o} \times \dfrac{\Sigma p_i q_i}{\Sigma p_o q_i} \times 100}$

$\therefore$ Fisher Index$_{(2006)}$ $= \sqrt{\dfrac{\Sigma p_i q_o}{\Sigma p_o q_o} \times \dfrac{\Sigma p_i q_i}{\Sigma p_o q_i} \times 100}$

$= \sqrt{\dfrac{1385}{995} \times \dfrac{1280}{930} \times 100}$

$= \sqrt{\{1.9158 \times 100\}}$

$= 1.3841 \times 100$

$= 138.4128$

≈ 138

iv. Marshall-Edgeworth Index $= \dfrac{\Sigma p_i\{(q_o + q_i)/2\}}{\Sigma p_o\{(q_o + q_i)/2\}} \times 100$

$= \dfrac{\Sigma(p_i q_o + p_i q_i)}{\Sigma(p_o q_o + p_o q_i)} \times 100$

OR

Marshall-Edgeworth Index$_{(2006)}$ $= \dfrac{\Sigma(p_i q_o + p_i q_i)}{\Sigma(p_o q_o + p_o q_i)} \times 100$

$= \dfrac{2701}{1925} \times 100$

$= 1.4031 \times 100$

$= 140.3117$

$\approx 140.$

6.4.2 Weighted Average of Relative Index

This index overcomes the drawback of simple average of relative index where equal importance is given to the items or commodities in the group. Here, weighted average can be based on arithmetic mean (A.M) or geometric mean (G.M). Based on weighted A.M, the price index is given by

$P_{oi} = \dfrac{\Sigma(p_i/p_o)w_i}{\Sigma w_i} \times 100$

where $w_i = p_o q_o$

Example 6.4.2.1
Obtain the Weighted Average Relative Price Index for the table used in Example 6.4.1.1, i.e table 6.2

Items	2001		2006		W_i	p_i/p_o	$(p_i/p_o)w_i$	p_iq_0
	Price p_o	Quantity q_o	Price p_i	Quantity q_i	$=$ p_oq_o			
A	14	45	20	35	630	$20/14 = 1.4286$	900.0180	900
B	13	15	19	20	195	$19/13 = 1.4615$	284.9925	285
C	12	10	14	10	120	$14/12 = 1.1667$	140.0040	140
D	10	5	12	8	50	$12/10 = 1.2000$	60.0000	60
					995		1385.0145	1385

$$P_{oi} = \frac{\Sigma(p_i/p_o)w_i}{\Sigma w_i} \times 100 = \frac{\Sigma p_iq_0}{\Sigma p_oq_0} \times 100$$

$$= \frac{1385.0145}{995} \times 100 = \frac{1385}{995} = 1.39$$

6.5 Construction of Quantity Index Number

The procedures for the construction of price index numbers discussed above are also applicable for the construction of quantity index numbers. All that need to be done is to replace price(p) by quantity(q) and q by p in the formula. For instance, in the formulae for price index numbers (p_{oi}), the corresponding quantity index number is Q_{oi}. That is,

If $\quad P_{oi} = \dfrac{\Sigma p_i}{\Sigma p_o} \times 100$, $\quad Q_{oi} = \dfrac{\Sigma q_i}{\Sigma q_o} \times 100$

If $\quad P_{oi} = \dfrac{\Sigma p_iq_o}{\Sigma p_oq_o} \times 100$, $\quad Q_{oi} = \dfrac{\Sigma q_ip_o}{\Sigma q_op_o} \times 100$

6.6 Chapter Summary

The concept of Index Number is an important statistical concept which is used to measure changes in a variable or group of variables with respect to time and other characteristics. The chapter covered the unweighted (price relative) and weighted indices. Among the weighted indices considered are Laspeyre, Paasche, Fisher Ideal and Marshall-Edgeworth indices.

6.7 Exercise Questions

MULTIPLE-CHOICE QUESTIONS AND SHORT – ANSWER QUESTIONS

1.   Which of the following is not one of the uses of an index number?

   a.   To deflate a value series in order to convert it into real physical terms

   b.   To compare student's intelligence in various schools or countries.

   c.   To select the commodities sources.

   d.   To forecast by using series of the indices.

   e.   To assess the worth of purchasing power of money.

2.   Which of the following is not a problem in the construction of an index number?

   a.   Selection of sources of data.

   b.   Definition of the purpose for which index is needed.

   c.   Method of collecting data for index.

   d.   Method of combining the data.

   e.   Unweighted average price index.

3.   Which of the following is not a weighted price index number?

   a.   Laspeyre Index.

   b.   Simple Aggregate Price Index

   c.   Fisher Ideal Index

   d.   Marshall-Edgeworth Index

   e.   Paasche Index

4.   The Weighted Aggregative Index with a backward bias is _____

   a.   Laspeyre Index

   b.   Simple Aggregate Price Index

   c.   Fisher Ideal Index

   d.   Marshall-Edgeworth Index

   e.   Paasche Index

5.   Which of the following is Unweighted Price Index?

   a. $\Sigma q_i/q_o \times 100$

   b. $(\Sigma p_i/p_o)/2 \times 100$

   c. $\dfrac{\Sigma p_i q_o}{\Sigma p_o q_o} \times 100$

   d. $\dfrac{\Sigma p_i q_i}{\Sigma p_o q_i} \times 100$

   e. $\dfrac{\Sigma p_i\{(q_o + q_i)/2\}}{\Sigma p_o\{(q_o + q_i)/2\}} \times 100$

Use the following table to answer Questions 6 and 7

Items	2000 price	2005 price
A	50	80
B	70	100
C	90	100
D	100	120

6.   The Simple Aggregate Price Index of the above table using 2000 as base year is _____

7.   The simple Average of Relative Price Index with year 2000 as base year is _____

Use the following table to answer Questions 8 to 10

Items	2001		2006	
	Price	Quantity	Price	Quantity
A	80	50	100	60
B	90	60	100	70
C	100	70	120	90

8.   Using year 2001 as base year, the Laspeyre price index for year 2006 is _____

9.   Using year 2001 as base year, the Paasche price Index for year 2006 is _____

10.   Fisher's Ideal price index of the above table is _____

ANSWERS

1.   c;   2.   e;   3.   b;

4.   e;   5.   b;

6. $\text{SAPI} = \dfrac{\Sigma p_i}{\Sigma p_0} \times 100 = \dfrac{400}{310} \times 100 = 129.03$

7. $\text{SARPI} = \dfrac{\{\Sigma(p_i/p_o)\}}{n} \times 100 = \dfrac{1.6 + 1.429 + 1.1111 + 1.2}{4} \times 100$

            $= \dfrac{5.3401}{4} \times 100 = 133.50$

Items	p_o	q_o	p_i	q_i	p_iq_i	p_iq_o	p_iq_o	p_oq_o
A	80	50	100	60	6000	4800	5000	4000
B	90	60	100	70	7000	6300	6000	5400
C	100	70	120	90	10800	9000	8400	7000
					23800	20100	19400	16400

8. $\text{Lasp.} = \dfrac{\Sigma p_i q_o}{\Sigma p_o q_o} \times 100 = \dfrac{194000}{16400} \times 100 = 119.29$

9. $\text{Paasche} = \dfrac{23800}{20100} \times 100 = 118.41$

10. $\text{Fisher} = \sqrt{\{(\text{Lasp.}) \times (\text{Paasch})\}}$

          $= \sqrt{(118.29 \times 118.41)} = 118.35$

CHAPTER SEVEN

PROBABILITY

7.0 Learning objectives

At the end of the chapter, students or readers should be able to

- understand the concept of probability and be able to handle simple probability problems;
- compute probabilities by using the additive and multiplicative laws;
- use mathematical expectations in discrete probability distribution;
- compute probabilities using Binomial, Poisson and Normal Distributions.

7.1 Introduction

As various activities of business and life in general depend on chance and risks, the study of probability theory is an essential tool to make correct and right decisions.

7.2 Concept of probability theory

As earlier said, probability theory is majorly concerned with chance and calculated risks in the face of uncertainty. This is achieved by building a mathematical concept, based on the study of some samples of the theoretical or imaginary population ,which leads to formulation of some simple rules for the purpose of inference about the corresponding population.

7.2.1 Importance and uses of probability

Without any doubt, the importance of probability is felt in large percentage of human endeavour, most especially in business, commerce and other related areas where problems of risks are usually involved.

However, some of the uses of probability are itemized below:

1. Probability theory is used as quantitative analysis of some problems arising from business and other areas.
2. Probability theory is used as the basis of statistical inference.
3. Probability theory plays vital roles in insurance and statistical quality control.

7.2.2   Definitions of some useful terms

a.   **An experiment** means performing an act which involves unpredictable outcome. For example, we have tossing of coins, throwing a die, etc.

b.   **An outcome** is one of the possible results that can happen in a trial of an experiment.

Examples 7.2.2.1:

i.   The outcomes of a fair coin = {H, T} , where H and T stand for Head and Tail respectively.

ii.   The outcomes of a die = {1, 2, 3, 4, 5, 6}.

c.   **Sample space** is the list or set of all possible outcomes of an experiment, while each outcome is called SAMPLE POINT.

For example, the sample space of a die is {1, 2, 3, 4, 5, 6} , while 3 is a sample point.

d.   **An event** is a collection of sample points which have certain quality or characteristics in common.

In the set theory, it can be defined as a subset A of the points in the sample space S.

7.2.3   Definition of Probability

Probability is a statistical concept that measures the likelihood of an event happening or not. As a measure of chance, which an event E is likely to occur, it is convenient to assign a number between 0 and 1. Therefore, the probability of an Event E written as P(E) is a positive number which can be formally expressed as $0 \leq P(E) \leq 1$. Note that

i.   If P(E) = 0, then it is sure or certain that the event E will not occur or happen;

ii.   If (P(E) = 1, then it is sure or certain that the event E will occur or happen.

There are two schools of taught about the procedures of getting the estimates for the probability of an event. These are the

1.   **Classical or a Priori approach:** In this approach, it is assumed that all the n trials of an experiment are <u>equally likely</u> and the outcomes are mutually exclusive (can not happen

together). If an event E can occur in r different ways out of the sample space, then the probability of E is written as

$$P(E) = {}^r/_n$$

Example 7.2.3.1: If a fair coin is tossed once, what is the probability of obtaining Head?

Solution:

By a fair coin, it means the coin is not loaded in any way and it is equally unbiased.

Let E be the event of a Head, then $n(E) = 1$ and $n(S) = 2$.

$P(E) = \dfrac{\text{Number of times head is obtained}}{\text{Number of possible outcomes}}$

$\quad = \quad ½ \quad = \quad 0.5$

Example 7.2.3.2: If a fair die is cast once, what is the probability of obtaining an even number?

Solution:

Let the sample space(s) $\quad = \quad \{1, 2, 3, 4, 5, 6\}$

$\qquad n(s) \quad = \quad 6$

Let E be set of even numbers $= \quad \{2, 4, 6\}$

$\qquad n(E) \quad = \quad 3$

$P(E) \quad = \quad \dfrac{n(E)}{n(s)} \quad = \quad \dfrac{3}{6} \quad = \quad 0.5$

2. **Frequency or Posteriori approach:** It is assumed in this approach that after n repetitions of an experiment, where n is <u>very large</u>, an event E is observed to occur in r of these, then the probability of E is

$$P(E) = {}^r/_n$$

Example 7.2.3.3: SAO Bank Plc gave out loans to 50 customers and later found that some were defaulters while only 10 repaid as scheduled. Determine the probability of repayment of loan in that bank.

Solution:

$P(\text{repayment}) = \dfrac{\text{Number that repaid}}{\text{Total number of people that took loan.}}$

$$= \quad \frac{10}{50} \quad = \quad 0.2$$

Remark:

However, the two above approaches are found to have difficulties because of the two underlined terms: (equally likely and very large) which are believed to be vague and even being relative terms to individual. Due to these difficulties, a new approach tagged "Axiomatic Approach" was developed. This axiomatic approach is based on the use of set theory.

Axiomatic definition of Probability

Suppose there is a sample space S and A is an event in the sample space. Then to every event A, there exists a corresponding real value occurrence of the event A which satisfies the following three axioms:

i. $0 \le P(A) \le 1$

ii. $P(S) = 1$

iii.   If $A_1, A_2, \text{---}, A_k$ are mutually exclusive events,

$$P(A_1 \text{ or } A_2 \text{ or --- or } A_k) = P(A_1 \cup A_2 \cup \text{---} \cup A_k)$$

i.e $P(A) \quad = \quad$ sum of the probabilities of the simple events, comprising the event A

$$= \quad \frac{\text{Number of sample points in A}}{\text{Total number of sample points in S}}$$

$$= \quad \frac{n(A)}{n(S)}$$

where n is defined as the number of sample points or elements.

Example 7.2.3.4: If a fair die is cast, determine the probability of

a.   each sample point.

b.   the sum total of all the sample points.

Solution:

a.   Since the die is fair, each sample point is equally chanced.

Sample space (S) $=$ $\{1, 2, 3, 4, 5, 6\}$

$\therefore P(1) = P(2) = P(3) = P(4) = P(5) = P(6) = \frac{1}{6}$

(This confirms the axiom (i) above)

b. $P(S) \quad = \quad P(1) + P(2) + P(3) + P(4) + P(5) + P(6)$

$$= \quad \frac{1}{6} + \frac{1}{6} + \frac{1}{6} + \frac{1}{6} + \frac{1}{6} + \frac{1}{6} \quad = \quad 1$$

(This also confirms axiom (ii) above)

Example 7.2.3.5: A ball is drawn at random from a box containing 8 green balls, 5 yellow balls and 7 red balls. Determine the probability that it is

a.   green,   b.   yellow,   c.   red,   d.   not green

e.   green or yellow.

Solution

Let G, Y and R represent green, yellow, and red balls respectively.

Then, the sample points in each event is

$n(G) = 8$, $n(Y) = 5$ and $n(R) = 7$

Also, the sample space (S) consists of $8 + 5 + 7 = 20$ sample points

i.e $n(S) = 20$

$\therefore$ the probability of:

a. $P(G)$ = $\dfrac{n(G)}{n(S)}$ = $\dfrac{8}{20}$ = 0.4

b. $P(Y)$ = $\dfrac{n(G)}{n(S)}$ = $\dfrac{5}{20}$ = 0.25

c. $P(R)$ = $\dfrac{n(G)}{n(S)}$ = $\dfrac{7}{20}$ = 0.35

d. $P(\text{not green})$ = $P(G')$ = $1 - P(G)$

                                                  = $1 - 0.4$ $= 0.6$

e. $P(\text{green or yellow})$ = $P(G \cup Y)$ $= P(G) + P(Y)$

                                                  $= 0.4 + 0.25 = 0.65$

Alternate method

$P(G \cup Y)$ = $P(S) - P(R)$ = $1 - 0.35 = 0.65$

7.2.4   Use of Tree diagram

There is a simple way of determining the probability of issues involving experiments. This can be achieved by the use of Tree diagram. A tree diagram can be described as a set or group of connected lines with each line looking like a branch of a tree. It represents the trials or stages of an experiment in a diagram. Here, the weight attached to the outcome of trial or stage is written on the branch, while the corresponding outcome is at end of the branch.

Example 7.2.4.1: Draw a tree diagram to determine the probability of obtaining Head (H) in a coin tossed once.

Solution

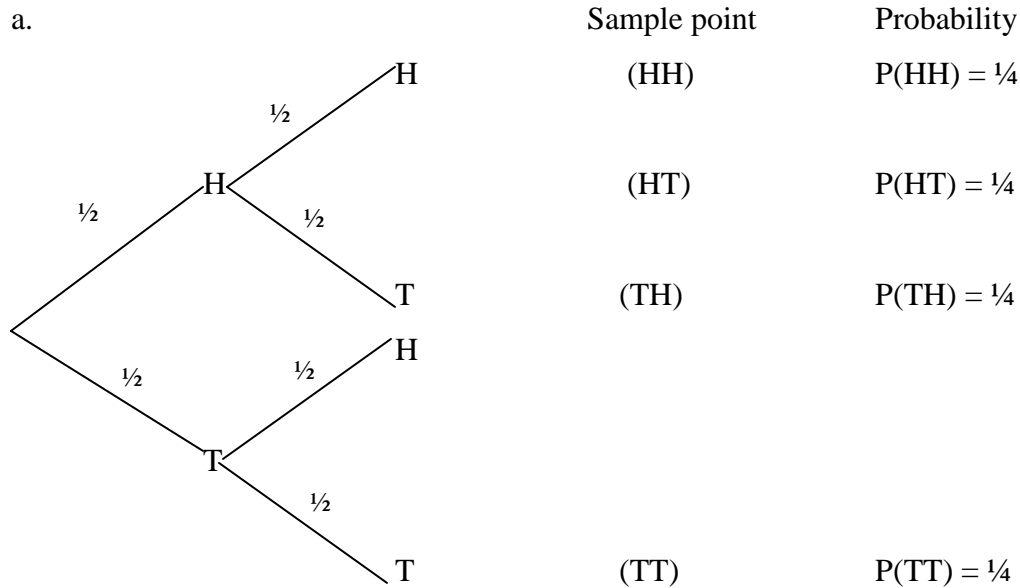|	Sample point	Probability
H	H	$P(H) = \frac{1}{2}$
T	T	$P(T) = \frac{1}{2}$

(tree: $\frac{1}{2}$ → H, $\frac{1}{2}$ → T)
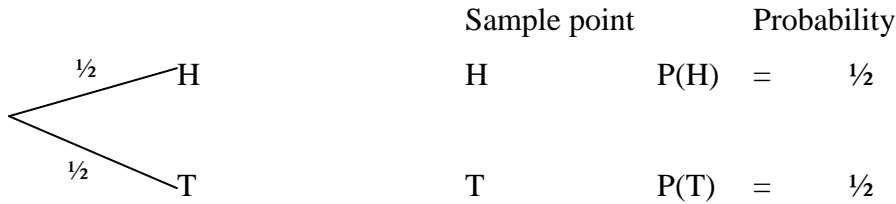
Example 7.2.4.2: Draw a tree diagram and obtain

a.   the probabilities in the case of tossing a fair coin twice;

b.   the probability of having a Tail (T) in the outcome.

Solution

a.

	Sample point	Probability
H	(HH)	$P(HH) = \frac{1}{4}$
	(HT)	$P(HT) = \frac{1}{4}$
	(TH)	$P(TH) = \frac{1}{4}$
	(TT)	$P(TT) = \frac{1}{4}$

(tree diagram with branches of probability $\frac{1}{2}$ each)

b.   Pr(of having T)   = P(HT) + P(TH)   = $\frac{1}{4} + \frac{1}{4} = \frac{1}{2}$

7.3   Addition Law of Probability

If E_1, E_2 are two events in a sample space S, then the probability of E_1 or E_2 occurring is given by

$P(E_1 \text{ or } E_2)$ = $P(E_1 \cup E_2)$ = $P(E_1) + P(E_2) - P(E_1 \cap E_2)$ by set theory

$$= \frac{n(E_1)}{n(S)} + \frac{n(E_2)}{n(S)} - \frac{n(E_1 \cap E_2)}{n(S)}$$

This is an addition rule of probability for any two events. We should note further that in set theory that "$\cap$" is interpreted as "and" while "$\cup$" is as "or".

By extension, the addition rule for three events E_1, E_2 and E_3 can be written as:

$$P(E_1 \cup E_2 \cup E_3) = P(E_1) + P(E_2) + P(E_3) - P(E_1 \cap E_2) - P(E_1 \cap E_3) - P(E_2 \cap E_3)$$
$$+ P(E_1 \cap E_2 \cap E_3)$$

Lastly, when events E_1, E_2 are mutually exclusive, $P(E_1 \cap E_2) = 0$ and the addition rule becomes:

$$P(E_1 \cup E_2) = P(E_1) + P(E_2)$$

Similarly for three mutually exclusive events E_1, E_2 and E_3, the addition rule also becomes:

$$P(E_1 \cup E_2 \cup E_3) = P(E_1) + P(E_2) + P(E_3)$$

Example 7.3.1:

From the following data 1, 2, 3, 4, 5, 6, 7, 8, 9, 10; obtain the probability of the following

    i.   An even number   ii.   A prime number

    iii.   A number not greater than 5.

    iv.   An even number or a prime number.

Solution:

Let E_1, E_2 and E_3 represent even number, prime number and numbers not greater than 5 respectively

Sample Space = {1, 2, 3, 4, 5, 6, 7, 8, 9, 10}

i.   Let E_1 = {2, 4, 6, 8, 10}   = Even numbers

        Then $p(E_1)$ = 5/10 = ½ or 0.5.

ii.   Let E_2 = {2, 3, 5, 7} = Prime numbers

        Then, $p(E_2)$ = 4/10 = 2/5 or 0.4

iii.   Let E_3 = {1, 2, 3, 4, 5} number not greater than 5

        Then $p(E_3)$ = 5/10 = ½

iv. $p(E_1 \cup E_2)$ = $p(E_1) + p(E_2) - p(E_1 \cap E_2)$

        where $E_1 \cap E_2 = \{2\} \Rightarrow p(E_1 \cap E_2) = \dfrac{1}{10}$

$\therefore p(E_1 \cup E_2) = p(E_1) + p(E_2) - p(E_1 \cap E_2) = \{^5/_{10} + {}^2/_5 - {}^1/_{10}\}$

$$= \frac{5}{10} + \frac{2}{5} - \frac{1}{10}$$

$$= \frac{5 + 4}{10} - \frac{1}{10}$$

$$= \frac{9}{10} - \frac{1}{10}$$

$$= \frac{9 - 1}{10} = \frac{8}{10} = 0.8$$

OR

$E_1 \cup E_2 = \{2,3,4,5,6,7,8,10\}$

$\therefore \quad p(E_1 \cup E_2) = 8/10 = 0.8$

Example 7.3.2: There are 24 female students in a class of 60 students. Determine the probability of selecting a student who is either a male or female.

Solution:

Let E_1 and E_2 represent female and male respectively.

Then $\quad P(E_1 \cup E_2) \quad = \quad P(E_1) + P(E_2), \quad$ because E_1 and E_2 are mutually

$$= \quad ^{24}/_{60} + ^{36}/_{60} \quad \quad \text{exclusive i.e } p(E_1 \cap E_2) = 0$$

$$= \quad 1$$

7.4 Conditional Probability and Independence

Let E_1 and E_2 be two events such that there is an intersection between E_1 and E_2, and that $P(E_1) > 0$. then, we denote the conditional probability by $P(E_2/E_1)$; and it can be read "the probability of E_2 given that E_1 has occurred or happened . The symbol (/) represents given. Here, event E_1 is known or assumed to have occurred or happened and it affects the sample space. By definition, the conditional probability is expressed as

$$P(E_2/E_1) \quad = \quad \frac{P(E_1 \cap E_2)}{P(E_1)} \quad , \text{ or}$$

$$P(E_1 \cap E_2) \quad = \quad P(E_1)P(E_2/E_1)$$

A practical example of conditional probability is seen in situations where two events are involved and must be satisfied; e.g probability of a student offering physics given that he must be a male.

The concepts of dependent and independent events go a long way to explain how the occurrence of an event affects another one. Here, if the events E_1 and E_2 occurrence or

non-occurrence does not affect each other, the events are said to be <u>independent</u>; otherwise it is <u>dependent</u> event. Examples of independent events are:

    i.   obtaining Head in coin tossing,

    ii.   passing of QA and Economics in ATS examination..

              Also, examples of dependent are:

    i.   Weather condition and sales of minerals;

    ii.   drug taken and rate of recovery from illness.

By applying the idea of independent event to our conditional probability $P(E_2/E_1)$, it becomes $P(E_2/E_1) = P(E_2)$ because the probability of E_2 occurring or happening is not affected by the occurrence or non-occurrence of E_1. Hence

$P(E_1 \cap E_2)$ = $P(E_1) \cdot P(E_2)$

By extension for independent events E_1, E_2 and E_3,

$P(E_1 \cap E_2 \cap E_3)$ = $P(E_1) \cdot P(E_2) \cdot P(E_3)$

Example 7.4.1:

If the probability that Dayo will be alive in 20 years is 0.8 and the probability that Toyin will be alive in 20 years is 0.3, what is then the probability that they will both be alive in 20 years?

Solution

    p(that Dayo will be alive in 20 years)   =   0.8

    p(that Toyin will be alive in 20 years)   =   0.3

    p(that they will both be alive in 20 years)   =   0.8 x 0.3   =   0.24

    Note: Both are independent events.

Example 7.4.2:

Suppose that a box contains 5 white balls and 4 black balls. If two balls are to be drawn randomly one after the other without replacement, what is the probability of

    i.   selecting a black ball given that a black ball was selected first.

    ii.   both balls drawn are black.

Solution:

Let E_1 be the event "first ball drawn is black" and E_2 be the event "second ball drawn is black", where the balls are not replaced after being drawn. Here E_1 and E_2 are dependent events and

conditional probability approach is an appropriate method.

The probability that the first ball drawn is black is

$$pr(E_1) = \frac{4}{5+4} = \frac{4}{9}$$

The probability that the second ball drawn is black, given that the first ball drawn was black, is $pr(E_2/E_1)$

$$pr(E_2/E_1) = \frac{3}{5+3} = \frac{3}{8}$$

Thus, the probability that both balls drawn are black is

$$p\{E_1E_2\} = pr(E_1) \cdot pr(E_2/E_1) = {}^4/_9 \cdot {}^3/_8 = {}^{12}/_{72} = {}^1/_6$$

7.5 Multiplication Law of Probability

Let us recall that when combining events in the addition rule, we dealt with the case of "OR" where we use the set symbol $\cup$ (union) to represent it. In a similar vein, the multiplication rule of probability is dealing with "AND" when combining events; and this is represented by the set symbol $\cap$ (intersection).

From conditional probability of events E_1, E_2;

$$P(E_2/E_1) = \frac{P(E_1 \cap E_2)}{P(E_1)},$$

or $\quad P(E_1 \cap E_2) = P(E_2/E_1) \times P(E_1)$

it is also true that $\quad P(E_1 \cap E_2) = P(E_1/E_2) \times P(E_2)$

This expression represents the multiplication rule of the probability where two events E_1 and E_2 are involved.

Example 7.5.1:

The probability that an employee comes late to work in any given day is 0.14. The probability that employee is a female in the company is 0.275. Obtain the probability that an employee selected from the company is a female late comer?

Solution:

Let $\quad E_1 \quad = \quad$ event of employee comes late

$\quad E_2 \quad = \quad$ event of employee is a female

Then, $\quad p(E_1 \cap E_2)$ is required

$$p(E_1 \cap E_2) \quad = \quad p(E_2) \cdot p(E_1/E_2)$$

$$= \quad 0.275 \times 0.14$$

$$= \quad 0.0385$$

Note: E_1, E_2 are independent events.

Example 7.5.2: A box contains 6 black and 8 yellow balls. It two successive draws of one ball are made, determine the probability that the first drawn ball is black and the second drawn ball is yellow if (i) the first drawn ball is replaced before the second draw; (ii) the first drawn ball is not replaced.

Solution:

Let B and Y represent black and yellow balls respectively.

i.   This is a case of selection with replacement, and the events involved are independent.

$$Pr(BY) = P(B) \times P(Y)$$

where $P(B) \quad = \quad \dfrac{6}{6+8} \quad = \quad \dfrac{6}{14} \quad = \quad \dfrac{3}{7}$, and

$P(Y) \quad = \quad \dfrac{8}{6+8} \quad = \quad \dfrac{8}{14} \quad = \quad \dfrac{4}{7}$, and

∴ $P(BY) = \quad {}^3/_7 \times {}^4/_7 \quad = \quad {}^{12}/_{49}$

ii.   This is selection without replacement;

$$Pr(BY) = \quad P(B) \times P(Y/B)$$

where $P(B) \quad = \quad \dfrac{6}{6+8} \quad = \quad \dfrac{6}{14} \quad = \quad \dfrac{3}{7}$, and

$P(Y/B) = \quad \dfrac{8}{5+8} \quad = \quad \dfrac{8}{13}$

∴ $P(BY) = \quad {}^3/_7 \times {}^8/_{13} \quad = \quad {}^{24}/_{91}$

7.6   Mathematical Expectation

Mathematical expectation is an important concept in the study of statistics and probability in the sense that it is mostly used in the determination of moment. Such moments include the statistics like mean, variance, moment generating functions etc.

Let X be a discrete random variable having the possible values x_1, x_2, x_3, --- , x_n with the probabilities P_1, P_2, P_3, --- , P_n respectively. Then the mathematical expectation, at times called "expected value" or "expectation", of X is denoted by E(X) and is defined as:

$$E(X) \quad = \quad x_1P_1 + x_2P_2 + ---- + x_nP_n$$

n

$i = 1$ cxl

$$= \quad \Sigma x_i P_i$$

or $\quad E(X) \quad = \quad \sum_{i=1}^{n} x_i f(x_i)$ or simply as $\quad \Sigma x f(x) \quad$ --- 7.8.1

where $P_i = f(x_i)$

A special class of equation 7.8.1 is where the probabilities are all equal to give

$$E(X) \quad = \quad \frac{x_1 + x_2 + \text{----} + x_n}{n} \quad , \text{which is the arithmetic mean.}$$

Similarly for a continuous random variable X having density function f(x), the expected value is defined by:

$$E(X) \quad = \quad \int_{-\infty}^{\infty} x f(x) \, dx$$

We have to note the following expectation as defined below:

i. $E(X) \quad = \quad \Sigma x f(x) = \quad \mu, \text{ the mean.}$

ii. $E(X^2) \quad = \quad \Sigma x^2 f(x)$

iii. $\text{Variance } (X) \quad = \quad \sigma^2 \quad = \quad \Sigma(x - \mu)^2 f(x)$

where (i) is the expectation of variable X and (ii) is the expectation of the square of X.

The following are some of the expectation properties:

a.   If C is any constant, then

$$E(CX) = \quad CE(X)$$

b.   IF X and Y are any random variables, then

$$E(X + Y) \quad = \quad E(X) + E(Y)$$

c.   If X and Y are independent random variables,

$$E(XY) = \quad E(X) \, E(Y).$$

Example 7.6.1:

Find: a. $E(x)$,   b. $E(x^2)$ and   c. $E[(x - x)^2$ for the probability distribution shown in the following table:

X	6	10	14	18	22
p(x)	1/6	1/5	3/7	1/3	1/11

Solution:

a. $E(x) \quad = \quad \Sigma x \cdot P(x)$

$= \quad 6 \, x \, ^1/_6 + 10 \, x \, ^1/_5 + 14 \, x \, ^3/_7 + 18 \, x \, ^1/_3 + 22 \, x \, ^1/_{11}$

$= \quad 1 + 2 + 6 + 6 + 2$

$= \quad 17. \text{ This represents the mean of the distribution.}$

cxli

b. $E(x^2)$ = $\Sigma x^2 . P(x)$

= $6^2 \times \frac{1}{6} + 10^2 \times \frac{1}{5} + 14^2 \times \frac{3}{7} + 18^2 \times \frac{1}{3} + 22^2 \times \frac{1}{11}$

= $6 + 20 + 84 + 108 + 44$

= 262. This represents the second moment about the origin zero.

c. $E[(x - \overline{x})^2]$ = $\Sigma(x - \overline{x})^2 . P(x)$

= $(6 - 17)^2 \times \frac{1}{6} + (10 - 17)^2 \times \frac{1}{5} + (14 - 17)^2 \times \frac{3}{7} +$

$(18 - 17)^2 \times \frac{1}{3} + (22 - 17)^2 \times \frac{1}{11}$

= $20.1667 + 9.8 + 3.8571 + 0.3333 + 2.2727$

= 36.4298

≈ 36; this represents the variance of the distribution.

Example 7.6.2: A company organized a lottery where there are 200 prizes of ₦50, 20 prizes of ₦250 and 5 prizes of ₦1000. If the company is ready to issue and sell 10,000 tickets, determine a fair price to pay for a ticket.

Solution: Let X and f(x) represent the amount of money to be won and density function respectively. Then the following table depicts the distribution:

X(₦)	50	250	1000	0
Frequency	200	20	5	10,000 − 225 = 9775
f(x)	200 / 10,000 = 0.02	20 / 10,000 = 0.002	0.0005	0.9775

Note: $f(x) = \underline{\text{Number of prizes in a category}}$
            Total Number of available tickets

Then, the fair price = E(X) = $50(0.02) + 250(0.002) + 1000(0.0005) + 0(0.9775)$

= ₦2

Example 7.6.3: Determine the expected value of the sum of points in tossing a pair of fair dice.

Solution: Let X and Y represent the points showing on the two dice. Then

E(X) = E(Y) = $1(\frac{1}{6}) + 2(\frac{1}{6}) + 3(\frac{1}{6}) + --- + 6(\frac{1}{6}) = \frac{7}{2}$

By the expectation property (b), $E(X + Y) = E(X) + E(Y) = \frac{7}{2} + \frac{7}{2} = 7$

7.7 Some special probability distributions

The following distributions are special in the sense that they have practical applications and being used frequently in computations of probability.

7.7.1 Binomial distribution

Binomial probability distribution or Bernoulli distribution is the most fundamental and important discrete probability distribution in statistics and is defined by the probability function:

$$P(X) = \frac{n!}{x!(n-x)!} p^x q^{n-x} , x = 0, 1, 2, \text{---} , n \qquad \text{--- 7.9.1}$$

where $\dfrac{n!}{x!(n-x)!} = \binom{n}{x} = {}^nC_x$, (expansion principle in permutation and Combination which is read as n combination x)

$n! = n(n-1)(n-2)\text{-----}(n\text{-}n)$ (read as n factorial); $(n\text{-}n)! = 0!$

$x! = x(x-1)(x-2)\text{----}(x\text{-}x)$ (read as x factorial) ; $(x\text{-}x)! = 0!$

$0! = 1$,

p is the constant probability of success,

q is the constant probability of failure (i.e $1 - p = q$) and

x is the number of successes in a series of n

independent trials $(x \le n)$

with the distribution (Binomial) parameters n and p, the following are some of its properties:

i.   Mean (μ) =   np

ii.   Variance (σ) =   npq

iii.   Standard deviation = $\sqrt{npq}$

iv.   Coefficient of skewness   = $\dfrac{q-p}{\sqrt{npq}}$

Example 7.7.1.1:

If 30% of the bolts produced by a machine are defective, determine the probability that out of 4 bolts chosen at random,

    (a.)   1;   (b.)   0   and   (c.) at most 2

bolts will be defective

Solution:

The probability of a defective bolt is $p = 0.3$, and the probability of a non-defective bolt is $q = 1 - p = 1 - 0.3 = 0.7$

(a.) Pr{1 defective bolt out of 4} $= \binom{4}{1}(0.3)^1(0.7)^3$

$= 4 \times 0.3 \times 0.343$

$= 0.4116$

(b.) Pr{0 defective bolt out of 4} $= \binom{4}{0}(0.3)^0(0.7)^4$

$= 1 \times 1 \times 0.2401$

$= 0.2401$

(c.) Pr(2 defective bolts out of 4} $= \binom{4}{2}(0.3)^2(0.7)^2$

$= 6 \times 0.09 \times 0.49$

$= 0.2646$

Thus,

Pr{at most 2 defective bolts} $=$ Pr{o defective bolt} $+$ Pr{1 defective bolt} $+$

Pr{2 defective bolts}

$= 0.2401 + 0.4116 + 0.2646$

$= 0.9163$

Example 7.7.1.2: A bank conducted a survey on its customers and found out that 35% use savings account. If ten customers who deployed cash are observed,

(a) determine the probability that

(i). none will deposit in savings account.

(ii) only one will deposit in savings account

(iii) only two will deposit in savings account.

(iv) not more than two will deposit in savings account.

(b) determine the mean and variance of the customers involved in savings account.

Solution: Let X represent the number of customers who deposit in savings account. Then we have:

$$P(X = x) = \frac{n!}{x!(n-x)!} p^x q^{n-x} = \frac{10!}{x!(10-x)!}(0.35)^x(0.65)^{10-x}$$

a. $P(X = 0)$ $= \dfrac{10!}{0!(10-0)!} (0.35)^0 (0.65)^{10-0} = 0.013463$

b. $P(X = 1)$ $= \dfrac{10!}{1!(10-1)!} (0.35)^1 (0.65)^{10-1} = 0.07249169$

c. $P(X = 2)$ $= \dfrac{10!}{2!(10-2)!} (0.35)^2 (0.65)^{10-2} = 0.175653$

d. $P(X \le 2)$ $=$ $P(X=0) + P(X=1) + P(X=2)$

   $=$ $0.013463 + 0.0724916 + 0.175653$

   $=$ 0.2616

e.   Mean $=$ $np = 10(0.35) = 3.5$ and

   Variance $= npq = 10(0.35)(0.65) = 2.275$

7.7.2   Poisson Distribution

Siemon Poisson (1781 – 1840) derived the Poisson distribution as a limiting case of binomial distribution in 1837 in sequel to the following conditions:

i.   p, the constant probability of success for each trial is very closed and tends to zero (i.e $p \to 0$).

ii.   n, the number of trials is very large and even tends to infinity (i.e. $n \to \infty$)

iii. $np = m$(say), is the mean, finite and constant.

Therefore, the probability function of the Poisson distribution is defined by:

$P(X) = \dfrac{e^{-m}m^x}{x!}$ $x = 0, 1, 2, ---$

With the Poisson parameter m, we have the following properties:

i.   Mean $(\mu) = \lambda$ ii.   Variance $(\sigma^2) = m$

iii.   Standard deviation (σ) $= \sqrt{m}$

iv.   Coefficient of skewness $= {}^1/_{\sqrt{m}}$

Example 7.7.2.1

If 3% of the electric bulbs manufactured by a company are defective, find the probability that in a sample of 100 bulbs:

a.   0;   b.   1;   c.   2;   d.   3;

bulbs will be defective

Solution:

$$P(X=x) \quad = \quad \frac{e^{-\lambda}\lambda^x}{x!} \qquad x = 0, 1, \text{---}$$

$$P \quad = \quad 3\% \ = \ 0.03$$

$$n \quad = \quad 100$$

$$\lambda \quad = \quad np \quad = \quad 0.03 \ \text{x} \ 100 = 3$$

a. $P(X=0) \quad = \quad \dfrac{e^{-3}3^0}{0!} \quad = \quad e^{-3} \ \text{x} \ 1 = \quad e^{-3} \quad = 0.0498$

b. $P(X=1) \quad = \quad \dfrac{e^{-3}3^1}{1!} \quad = \quad e^{-3} \ \text{x} \ 3 = \ 0.0498 \ \text{x} \ 3 \ = 0.1494$

c. $P(X=2) \quad = \quad \dfrac{e^{-3}3^2}{2!} \quad = \quad \dfrac{e^{-3} \ \text{x} \ 9}{2} = \dfrac{0.0498 \ \text{x} \ 9}{2} \ = 0.2240$

d. $P(X=3) \quad = \quad \dfrac{e^{-3}3^3}{3!} \quad = \quad \dfrac{e^{-3} \ \text{x} \ 27}{6} = \dfrac{0.0498 \ \text{x} \ 9}{2} \ = 0.2240$

7.7.3   Normal distribution

Normal distribution is a continuous probability distribution and it is commonly used. It is also known as Gaussian distribution.

If x is a continuous random variable following the normal distribution with mean μ and standard deviation σ, then probability density function of Normal distribution is given by

$$P(X) \quad = \quad f(x) \quad = \quad \frac{1}{\sigma\sqrt{2\pi}} e^{-(x-\mu)^2/2\sigma^2}$$

where constants $\quad \pi = 3.1416 \quad$ and $\quad e = 2.7183$

If we compute $\displaystyle\int_a^b f(x)dx$ or $\displaystyle\int_{-\infty}^{\infty} f(x)dx$, it gives us the area under the normal curve. For instance, $\displaystyle\int_{-\infty}^{\infty} f(x)dx = 1$. We are to note further that the area under the normal curve can be obtained within certain range. Some of the properties of Normal distribution are given below:

a.   Mean = Mode = Median; and this makes it not to be skewed.

b.   Its graph is symmetry about mean and this makes to be bell shaped.

c.   Normal distribution is a continuous probability distribution with parameters μ and σ.

cxlvi

d.   The total area under the normal curve is equal to 1.

e.   The sub – areas under the normal curve correspond to some probabilities. This is achieved by the use of standard Normal variate (standard score) and the Normal Z table. The standard score Z is obtained by the expression
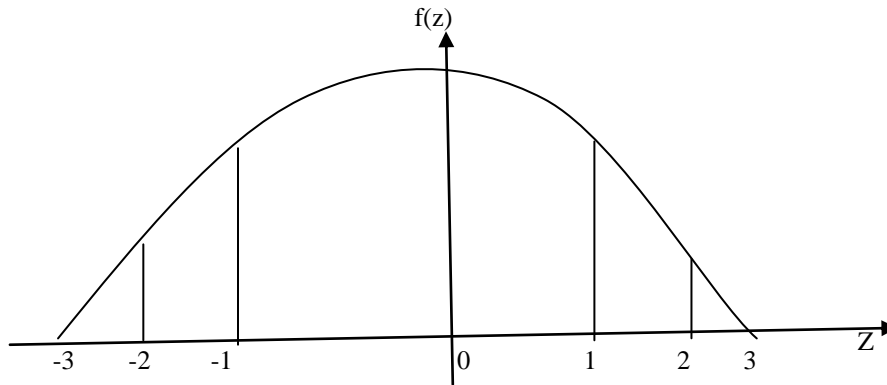
$$Z = \frac{x - \mu}{\sigma}$$

and to evaluate the sub-areas on the normal curve, the following steps are to be adhered to:

i.   The area required is sketched on the X – graph.

ii.   Determine the Z – value for each end point of the required area.

iii.   Sketch the corresponding area of the required one on the Z – graph.

A typical graph of Z – score is given below

Fig. 7.7.3.1

iv.   Use the Normal Z table in the appendix to evaluate the required area. For instance,

P(-1 ≤ Z ≤ 1) = 0.6827,   P(-2 ≤ Z ≤ 2) = 0.9545.

We are to note further that the Normal Z table can be read from the following forms and types.
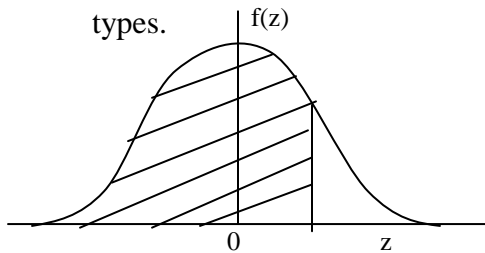
In this type of graph, the table is read directly for the shaded area.

In this type of graph, the table is read by adding 0.5 to the table value for the shaded area.
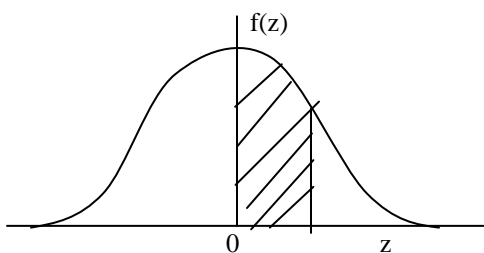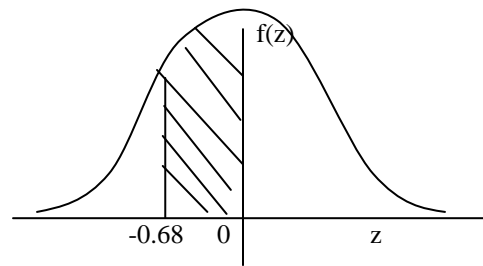
f(z)

0

Z
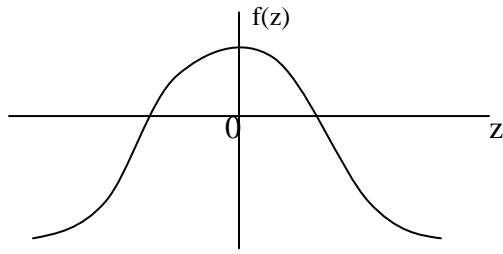
In this type of graph, the table is read by subtracting the table value from 1 for the shaded area.

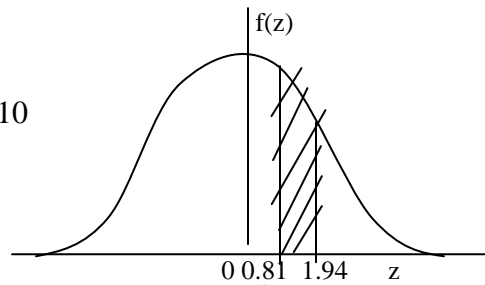**Example 7.7.3.1:** Determine the area under the standard normal curve:

a.   between Z = -0.68 and Z = 0.   b.   between Z = 0.81 and Z = 1.94

c.   to the right of Z = -1.28

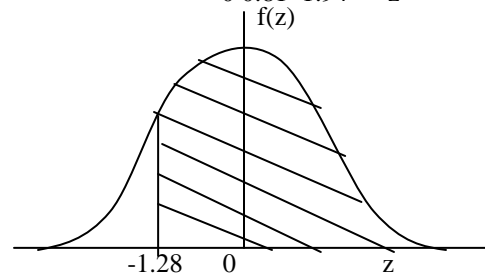**Solution**

a.   P(-0.68 ≤ Z ≤ 0)

                =   0.2517 – 0.0

                =   0.2517

f(z)

-0.68   0   z

b.   P(0.81 ≤ Z ≤ 1.94)

                =   0.4738 – 0.2910

                =   0.1828

f(z)

0 0.81 1.94   z

c.   P(Z ≥ -1.28)

                =   0.3997 + 0.5

                =   0.8997

f(z)

-1.28   0   z

Example 7.7.3.2:

The mean inside diameter of a sample of 200 washers produced by a machine is 0.502 inches(in), and the standard deviation is 0.005inch. The purpose for which these washers is intended is to allow a maximum tolerance in the diameter of 0.480 to 0.516; otherwise, the washers are considered defective. Determine the percentage of defective washers produced by the machine, assuming that the diameters are normally distributed.
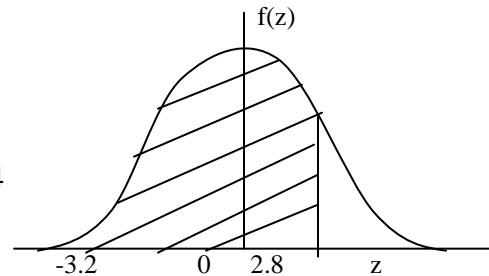
Solution:

Standard score Z $= \dfrac{x - \mu}{\sigma}$

0.476 in standard units (Z) $= \dfrac{0.486 - 0.502}{0.005}$

$= \dfrac{-0.016}{0.005}$ $=$ -3.2

0.516 in standard units $= \dfrac{0.516 - 0.502}{0.005}$

$= \dfrac{0.014}{0.005}$ $=$ 2.8

P(-3.2 < x < 2.8)

$=$ P(-3.2 < z < 2.8)

$=$ $1 - \varphi(3.2) - \varphi(2.8)$

$=$ $1 - 0.4993 - 0.4974$

$=$ 0.0033

7.7.4   Approximations to Binomial Distribution

7.7.4.1 Approximation to the Binomial distribution

Recall that the probability function of a Binomial distribution is

P(X) $= \dfrac{n!}{x!(n-x)!} \; p^{x}q^{n-x}$

where n and p are parameters specifying the distribution.

Computations with binomial when n is large usually become tedious and at times, they create problems. In order to overcome these problems, the idea of approximations to binomial was established.

a.   Poisson approximation to the Binomial distribution

Poisson approximation to Binomial distribution is made possible on the following conditions:

i.   the sample size n must be large (n>30).

ii.   p, the probability of success, must be small and close to zero so that np is less than 5 (np<5).

iii.   the mean of Poisson (m) is obtained by m = np.

(Note: some authors use λ for the mean).

b.   Normal approximation to the Binomial distribution

Normal approximation to Binomial distribution is made possible on the following conditions:

i.   the sample size n is large.

ii.   the probabilities of success (p) and failure (q) are not small so that np and nq are greater than 5.

iii.   that mean and variance of Normal distribution are obtained by $\mu = np$ and $\sigma^2 = npq$.

iv.   that by Central Limit Theorem, the standardized random variable

$$ Z \quad = \quad \frac{q - p}{\sqrt{npq}} $$

is obtained for the Normal approximation.

The following examples illustrate the said approximations:

Example 7.7.4.1.1: Five percent of the goods produced in a manufacturing company turn out to be defective. Determine the probability that in a sample of 20 goods selected at random exactly two will be defective by using:

a.   the Binomial distribution.

b.   the Poisson approximation to the Binomial distribution.

Solution

a.   Let X represent the number of defective goods with the given probability of defective p = 0.05.

. $P(X = 2) \quad = \quad \dfrac{20!}{2!(18)!} (0.05)^2 (0.95)^{18} \quad = \quad 0.1888$

b.   Mean of Poisson m(or λ) = np = 20 x 0.05 = 1

$P(X=x) \quad = \quad \dfrac{e^{-m}m^x}{x!} \qquad m = 0, 1, \text{---}$

. $\therefore P(X=2) \quad = \quad \dfrac{e^{-1}1^2}{2!} \quad = \quad 0.1839$

Note that the answers of (a) and (b) very close.

cl

Example 7.7.4.1.2: If a fair coin is tossed 10 times, determine the probability of getting between 4 and 6 heads inclusive by using the

a.   binomial distribution;

b.   normal approximation to the binomial distribution

Solution

a.   Let X represent the random variable of a number of heads.

. $P(X = 4)$ $= \dfrac{10!}{4!6!} (0.5)^4(0.5)^6 = \dfrac{105}{512} = 0.2051$

. $P(X = 5)$ $= \dfrac{10!}{5!5!} (0.5)^5(0.5)^5 = \dfrac{63}{256} = 0.2461$

. $P(X = 6)$ $= \dfrac{10!}{6!4!} (0.5)^6(0.5)^4 = \dfrac{105}{512} = 0.2051$

     $\therefore P(4 \leq x \leq 6) = 0.2051 + 0.2461 + 0.2051 = 0.6563$

b.   By normal approximation, where $\mu = np = 10(0.5)^5 = 0.3125$

     and $\sigma = \sqrt{npq} = \sqrt{10(0.5)(0.5)} = 1.58$, the standard scores are.

     (Note that the data is now treated as continuous variable type, hence 4 to 6 heads shall be treated as 3.5 to 6.5 heads.)

     3.5 in standard score = $(3.5 - 5)/1.58 = -0.95$

     6.5 in standard score = $(6.5 - 5)/1.58 = 0.95$

     Probability of $(-0.95 \leq Z \leq 0.95)$ $= 0.8289 - 0.1711$

                               $= 0.6578$

     $\therefore P(4 \leq x \leq 6)$ by normal $= 0.6578$

     Also, the answers of (a) and (b) are very close.

7.7.4.2 Normal approximation to the Poisson distribution

     Without any doubt, there is a relationship between the Binomial and Normal distributions as treated in section 7.9.4.1. Hence, it is expected that there will also be a relationship between the Poisson and Normal distributions.

cli

If X is the Poisson random variable, then the standardized random variable is

$$Z = \frac{x - m}{\sqrt{m}}$$

where m is the Poisson mean and the Poisson distribution approaches the normal distribution as m $\rightarrow \infty$ i.e as m is getting larger.

7.8   Summary

The principle of set theory is treated with its applications to elementary probability and general problem solving. The additive and multiplicative laws, and even the conditional probabilities are considered.

Treatment of some special distributions such as Binomial, Poisson and Normal are considered with their approximations.

7.9 Multiple choice and short – answer questions

Use the following information to answer the next three questions. Given a set of numbers

S = {2, 4, 7, 10, 13, 16, 22, 27, 81, 102}

If a number is picked at random, determine the probability that the number is

1.   even

a.   0.4   b.   0.2   c.   0.6

d.   0.8   e.   0.9

2.   odd

a.   0.4   b.   0.8   c.   0.2

d.   0.9   e.   0.6

3.   prime

a.   0.2   b.   0.5   c.   0.3

d.   0.6   e.   0.4

Use the following information to answer the next 2 questions:

Supposing there is a lottery game in which 2000 tickets are issued and that there are 5 major and 50 minor prizes to be won. If I buy a ticket, the probability that it will win

4.   a major prize is

    a.   0.25   b.   0.025   c.   0.0025   d.   0.5

    e.   0.05

5.   a minor prize is

    a.   0.25   b.   0.025   c.   0.0025   d.   0.5

    e.   0.05

6.   Using the item of question 4, the probability of winning a prize in the lottery game is ....

7.   In a three diagram of method of calculating the probability of an even, the weight attached to the outcome of trial or stage is written on the branch while the corresponding outcome is at the .... of the branch

Use the following statement to answer questions 8 to 10: If a ball is selected randomly from a box containing 6 white balls, 4 blue balls and 5 red balls, obtain the probability that:

8.   White ball is selected_____

9.   Not white ball is selected _____

10.   Blue or red ball is selected_____

Answers to Chapter Seven

1. c 2. a 3. c 4. c 5. b

6. 0.0275 i.e. $\dfrac{5}{2000} + \dfrac{50}{2000} = 0.0275$ 7. end

8. $Pr(\text{white}) = {}^{6}/_{15} = 0.4,$

9. $Pr(\text{Not white}) = 1 - Pr(\text{white}) = 1 - 0.4$
$= 0.6$

10. $Pr(\text{Blue or Red}) = {}^{4}/_{15} + {}^{5}/_{15} = {}^{9}/_{15} = 0.6$

CHAPTER EIGHT

STATISTICAL INFERENCE

8.0 Learning objectives

At the end of the chapter, students or readers should be able to

- know that statistical inference consists of estimation and significance test;

- demonstrate good understanding of both point and interval estimates;

- estimate the population mean and population proportion;

- understand test of hypothesis on population mean and population proportion for both large and small samples;

8.1 Introduction

An important study of statistical theory, which is commonly used in decision making, is the concept of statistical inference. Here, estimates and inference (both inductive and deductive) are made in order to make decisions. These decision makings are carried out on the basis of samples taken from the population. For instance, a research worker performs an experiment and obtains some data. On the basis of the data, some estimates and certain decisions or conclusions are made. These decisions are then generalized to similar situations of that experiment. In order words, the researcher may generalize from a particular experiment to all the classes of all similar experiments.

Succinctly, statistical inference can be broadly classified into two; viz:

i.   Estimation theory and;

ii.   Test of Hypothesis.

8.2 Basic concepts of estimation

a. Population parameter

This is a quantity whose numerical value is known for a population. It is this specified quantity that allows one to make probability statement about the distribution of the given population. Therefore, a population is completely known when we know the probability distribution f(x) with its value(s) of parameter(s) for a given random variable X. For

instance, in binomial distribution, the value of P(probability of success) must be known before any meaningful statement can be made about the populations distribution it is representing.

b.   Sample statistics

In practice, the population parameter is not usually known. Therefore, efforts are to be made to have an approximate value called estimate. All we need to do is to take random samples from the target population and then use these samples to obtain values which serve to estimate the population parameter. This set of estimates serves as a representative of population parameter. Therefore, any quantity obtained from a sample for the purpose of estimating a population parameter is called a sample statistic, or briefly statistic.

The following table shows the usual symbols used for population parameter and sample statistics.

Description of Measure	Population parameter symbol	Sample statistics symbol
Mean	μ	$\bar{x}$
Variance	σ^2	S^2
Number of items	N	n
Proportion	P	p

c.   Estimate and Estimator

An approximate value obtained from a sample is called an estimate; while an estimator is the expression used for computing or determining the estimate.   For example, in computation of mean for variable X, we have

$$\bar{x} \quad = \quad \frac{\Sigma x}{n}$$

$\bar{x}$ is the estimate while $\Sigma x/n$ is the estimator.

Further, an estimate can be point or interval. It is point estimate when the population parameter is given by a single number; while an estimate of population parameter is given between the range of two numbers.

d.   Sampling distribution

This is an important concept which makes inference from sample to population possible. By sampling distribution, we mean the probability distribution of a

sample statistic. This can be achieved by considering all possible samples of size n from N(population size); and for each sample, calculate or determine the statistic estimate. Next obtain the frequency distribution of the statistic estimates in order to have what is called "sampling distribution".

For sampling distribution, we have various types such as: mean, variance, proportion, moment but to mention a few.

8.3   Point Estimation of the Population Mean

The point estimation of the population mean can be obtained by adhering to the following steps:

i.   First obtain the sampling distribution of means for all the possible samples of size n.

ii.   Next, compute the sample mean $\bar{x}$ by finding the average of set of means obtained in (i) i.e finding mean of means to give the desired sample mean $\bar{x}$.

The following example will illustrate the procedures stated above.

Example 8.3.1:

If all possible samples of size two are drawn with replacement from a population consisting of numbers 7, 8, 12, 13, determine the:

a.   mean of the population (μ);

b.   standard deviation of the population;

c.   the mean of the sampling distribution of means ($\bar{x}$)

d.   the standard deviation of the sampling distribution of means.

Solution

a.   Population mean μ $=$ $\dfrac{7 + 8 + 12 + 13}{4}$ $=$ $\dfrac{40}{4}$ $=$ 10

b.   Standard deviation of population (σ),

$$\sigma^2 \quad = \quad \frac{(7 - 10)^2 + (8 - 10)^2 + (12 - 10)^2 + (13 - 10)^2}{4}$$

$$= \quad \frac{9 + 4 + 4 + 9}{4} \quad = \quad \frac{26}{4} \quad = \quad \frac{13}{2} \quad = \quad 6.5$$

$$\therefore \sigma \quad = \quad \sqrt{\sigma^2} \quad = \quad \sqrt{6.5} \quad = 2.55$$

c.   Selection of 2 samples with replacement:

NOTE: This is a work in progress. All topics in the syllabus are covered but editing for necessary corrections is in progress.

Thanks.

(7, 7)	(8, 7)	(12, 7)	(13, 7)
(7, 8)	(8, 8)	(12, 8)	(13, 8)
(7, 12)	(8, 12)	(12, 12)	(13, 12)
(7, 13)	(8, 13)	(12,13)	(13, 13)

The corresponding sample means:

7	7.5	9.5	10.0
7.5	8.0	10.0	10.5
9.5	10.0	12	12.5
10.0	10.5	12.5	13.0

∴ The sampling distribution of means can be stated as:

Number (x)	7	7.5	8.0	9.5	10.0	10.5	12.0	12.5	13.0
Frequency(f)	1	2	1	2	4	2	1	2	1

The mean of sampling distribution of mean ($\bar{x}$) is equal to $\bar{x} = \dfrac{\sum fx}{\sum f}$

$$= \frac{7 + 2(7.5) + 8.0 + 2(9.5) + 4(10) + 2(10.5) + 12 + 2(12.5) + 13}{16}$$

$$= \frac{160}{16} = 10$$

d. $S^2 = \dfrac{(7 - 10)^2 + (7.5 - 10)^2 + \text{---} + (12.5 - 10)^2 + (13 - 10)^2}{16}$

$$= \frac{52}{16} = 3.25$$

$\therefore S = \sqrt{S^2} = \sqrt{3.25} = 1.8$

Remark: From the above computations, we could note that:

i. the results of (a) and (c) are equal. This illustrates the fact that $E(\bar{x}) = \mu = 10$; i.e $\bar{x}$ is an unbiased estimator of μ.

ii. the result of (b) and (d) demonstrates a relationship between the variances of the

population mean and sample mean.

Here, it is established that $S^2 = \sigma^2/n$ for infinite population or sampling with replacement.

i.e $S^2 = 3.25$ and $\sigma^2 = 6.5$ where $n = 2$ (as the sample size)

8.4 Point Estimation of population proportion.

Here, it is assumed that the population is infinite and binomially distributed. With these assumptions, the same procedures used in point estimation of population mean shall be equally applied. That is,

i.   to first obtain the sampling distribution of proportions and next;

ii.   to compute the mean of this set of proportions.

For instance, a consideration of all possible samples of size n drawn from the population and for each sample, the proportion statistic (p) is determined leading to the sampling distribution of proportions with mean (μ_p)

$$\mu_p \quad = \quad p \quad \text{and} \quad \sigma_p \quad = \quad \sqrt{\frac{pq}{n}} = \sqrt{\frac{p(1-p)}{n}}$$

Example 8.4.1
If each person of a group of 400 people tosses a fair die 90 times, how many people should be expected to report that between 20% and 30% of their tosses resulted in having die face 5?

Solution:

The probability P of die face 5 $= \; ^1/_6 \; = \; 0.1667$ while $q = 1 - p$

$q = 1 - \frac{1}{6} = \frac{5}{6} = 0.8333$

$\mu \; = \; np \; = \; 90(^1/_6) \quad = 15 \quad \text{and} \; \sigma \; = \; \sqrt{npq} = \sqrt{90(1/6)(5/6)} = 3.54$

Normal approximation to Binomial shall be used for this problem

The probability that face 5 in 90 tosses will be between 20% and 30% is desired, 20% of $90 = 18$ and 30% of $90 = 27$. Since these numbers (18 and 27) are discrete, we need to adjust them to continuous variable for 17.5 and 27.5.

$\therefore$ The standard scores of these adjusted numbers are:

$$17.5 \text{ in standard units (z)} = \frac{17.5 - 15}{3.54} = z(0.7062)$$

$$26.5 \text{ in standard units (z)} = \frac{27.5 - 15}{3.54} = z(3.53)$$

$$Pr(0.7062 \leq 3.53) = 0.9998 - 0.7612$$

$$= 0.2386$$

$\therefore$ 23.86% of 400 $= 95.44 \approx 95$

About 95 people would report that their experiment resulted in between 20% and 30% of die face 5.

8.5 Useful concepts in Hypothesis testing

The following terms are useful in the principle of Hypothesis testing.

a. **Hypothesis:** The concept of hypothesis in statistics aids decision making. Hypothesis can be defined as the assumption or guess about the population parameters involved.

There are two types of hypothesis; namely, the

i. Null hypothesis

ii. Alternative hypothesis

An hypothesis which states that there is no difference between the procedures, results from samples and other phenomenon is called Null Hypothesis and it is denoted by H_0, while any hypothesis which differs from the null hypothesis or given hypothesis is known as Alternative Hypothesis and it is denoted by H_1. For instance, if the population mean μ = 5.0 and the sample mean $\bar{x}$ = 6.0, we set our hypothesis as thus:

$H_0 : \mu = \bar{x}$ i.e $H_0 : 5.0 = 6.0$

$H_1 : \mu \neq \bar{x}$ i.e $H_1 : 5.0 \neq 6.0$ or

$H_1 : \mu < \bar{x}$ i.e $H_1 : 5.0 < 6.0$

b. **Errors:**

There are two types of errors in hypothesis testing, namely,

(a) Type I error

(b) Type II error.

<div align="center">clx</div>

When a hypothesis which is supposed to be accepted is rejected, we call that error a type I error. On the other hand, if we accept a hypothesis which is supposed to be rejected, a type II error is committed.

The above statements can be summarized in the following table:

Decisions	H_0 true	H_0 False
Accept H_0	Correct decision	Type I error
Reject H_0	Type II error	Correct decision

c.   Level of significance

In hypothesis testing, the maximum probability of the willingness to risk a type I error is called the level of significance or simply significance level and it is denoted by α. It therefore implies that the investigator has $(1 - \alpha)$ confidence that he/she is making right decision

In practice, this must be stated first in any hypothesis testing. The common levels of usage in practice are 1%(or 0.01) and 5% (or 0.05).

d.   Test statistic:

A test statistic is the computed value, which, when compared with the tabulated value, enables one to decide whether to accept or reject hypothesis and hence determine whether the figures of the observed samples differ significantly from that of the population. It is also called *test of hypothesis*, test of significance or rules of decision.

e.   Critical region:

This is a region or critical value of one side of the distribution of the case study, with area equal to the level of significance. It is used to decide whether to accept a hypothesis or not.

f.   One-tailed and Two-tailed tests

The type of tailed test depends on the stated alternative hypothesis (H_1). It is one-tailed test or one-sided test when the alternative hypothesis is one directional. For example, we have the following hypothesis for one-tailed test:

$H_0: \mu = \bar{x}$

$H_1: \mu > \bar{x}$

$H_1: \mu \geq \bar{x}$

$H_1 : \mu < \bar{x}$

$H_1: \mu \leq \bar{x}$

For a two-tailed test or two-sided test, the hypothesis is of two directions e.g

$H_0: \mu = \bar{x}$

$H_1: \mu \neq \bar{x}$

The resultant effect of the type of tailed test is on the significance level to be used in obtaining the critical value from the table. If the significance level is α, then critical values for α and $\alpha/2$ will be obtained respectively for one-tailed and two-tailed tests.

For instance, if $\alpha = 0.05$, we check critical values at 0.05 and $0.05/2 = 0.025$ respectively for the one-tailed and two-tailed tests.

8.6   Testing Hypothesis about a population parameters

The procedure for carrying out significance test about a population parameter is given as follows:

i.   Set up the hypothesis about the parameter of interest and state its significance level.

ii.   Set the hypothesis by using appropriate test statistic.

iii.   Draw conclusion by making decision on the result of your computed value in step (ii) above. Here, table value at a significance level is compared with computed value.

8.6.1   Testing hypothesis about the population mean

This section will present testing hypothesis under two conditions namely; (i) when we have large samples and (ii) when we have small samples.

For large samples ($n \geq 30$) it is usually assumed that the sampling distribution of the desired statistics is normally distributed or approximately normal. It is for this reason that tests concerning large samples assumed normal and used the Z test. Therefore the test statistic for large sample is

$$Z_{cal} = \frac{\bar{x} - \mu_0}{\sigma / \sqrt{n}}$$

where $\bar{x}$ =   sample mean;

μ_0 =   hypothesis value of μ.

σ =   the given value of the standard deviation; but when this not given, one

can use sample standard deviation (S) for large samples.

n   =   sample size

When we have small samples (n < 30), the test – statistic is the t – test and it is defined as thus:

$$t_{cal} = \frac{\bar{x} - \mu_0}{S / \sqrt{n}}$$

where $\bar{x}$, μ_0, n are as defined above and

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

The final decision on the tests is made by comparing the computed values (Z_{cal} and t_{cal}) with the table values.   The normal table is used for the Z – test (for large samples) while t – distribution table at n – 1 degrees of freedom is used for small samples.

When $Z_{cal} > Z_{table}$, and $t_{cal} > t_{table}$, we reject the null hypothesis (H_0); otherwise, accept H_0.

Example 8.6.1

In a University, a sample of 225 male students was taken in order to find the average height. From the samples, the computed average height was 184.0cms while the mean of actual population height was 178.5cms with a standard deviation of 120cms. You are required to show if the sample mean height is significantly different from the population mean at 5% significant level.

Solution:

Hypothesis

$H_0: \mu = \bar{x} \Rightarrow H_0:\ 178.5 = 184.0$

$H_1 : \mu \neq \bar{x} \Rightarrow H_1 : 178.5 \neq 184.0$

$$Z_{cal} = \frac{\overline{x} - \mu_0}{\sigma / \sqrt{n}} = \frac{184.0 - 178.5}{120 / \sqrt{225}} = \frac{5.5}{8}$$

$$= 0.6875$$

Table value of Z at 5% level of significance (for two tailed test) = 1.96

Decision: since $Z_{cal} < Z_{table}$ (0.6875 <1.96), we accept H_0 and concluded that no significance difference between the sample mean and population mean.

Example 8.6.2

SAO is a beer manufacturing company with 5000 workers. The company is interested to know the average number of beer bottles taken per week per person by the company's workers. While the quality control manager thinks that the average number of bottles taken per worker per week is 11, the company secretary thinks that the true value should be more. The quality control manager subsequently selected 11 workers at random and got the following results as beer bottles taken in a week, 13, 4, 17, 9, 3, 20, 16, 12, 8, 18,12.

You are required to set up a suitable hypothesis and test it at 5% level of significance.

Solution:

$H_0 : \mu = 11$

$H_1 : \mu > 11$

$\alpha = 0.05$

Since the sample size (11) is small, we use the t – test

$$t_{cal} = \frac{\overline{x} - \mu_0}{S / \sqrt{n}} \qquad \mu_0 = 11, n = 11$$

$$\overline{x} = \frac{\Sigma x}{n} \quad \text{and} \quad S^2 = \frac{\Sigma(x - \overline{x})^2}{n - 1}$$

mean = $\overline{x}$ = $\dfrac{\Sigma x}{n}$ = $\dfrac{13 + 4 + 17 + 9 + 3 + 20 + 16 + 12 + 8 + 18 + 12}{10}$

$$= \frac{132}{11} = 12$$

The computation of S^2 is set out in the table below

x	x − x̄	$(x - \overline{x})^2$
13	1	1

4	-8	64
17	5	25
9	-3	9
3	-9	81
20	8	64
16	4	16
12	0	0
8	-4	16
18	6	36
12	0	0
		312

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

$$= \frac{312}{11 - 1} = \frac{312}{10} = 31.2$$

$$\therefore \quad S^2 = 31.2, \quad \rightarrow \quad S = \sqrt{31.2}$$

$$= 5.5857$$

$$\approx 5.6$$

using all these pieces of information, our test statistic becomes

$$t_{cal} = \frac{\bar{x} - \mu_0}{S/\sqrt{n}} = \frac{12 - 11}{5.6/\sqrt{11}} = \frac{1}{5.6/3.3166} = \frac{1}{1.6885}$$

$$= 0.5923$$

$$\approx 0.59$$

Table value of t at 5% significance level (for one – tailed test) = 1.81

Decision: the problem is of one – tail test (see the alternative hypothesis), and $t_{cal} < t_{table}$ i.e $0.59 < 1.81$, we accept H_0 and conclude that no significance difference between the sample mean and the population mean.

8.7 Testing Hypothesis about the Population Proportion

In the significance test for a proportion, there is the need to know the population proportion (P_0) and then compute the sample proportion (P) from the sample drawn. The sample proportion is obtained by taking a number of observations (x) (possessing an attribute) out of the total number n so that $p = x/n$

As we have done for the mean, we set up the hypothesis as thus:

$H_0 : P_0 = P$

$H_1 : P_0 \neq P$ (or $P_0 > P$ or $P_0 < P$)

and the significance level given is α.

Then, the test statistic is

$$Z_{cal} = \frac{P - P_0}{\sqrt{\dfrac{P(1-P)}{n}}} \quad \text{or} \quad \frac{x/n - P_0}{\sqrt{\dfrac{x/n(1 - x/n)}{n}}}$$

and we decide by comparing the Z_{cal} with the table value Z_{table} at the given α

Example 8.7.1

A demographer claims that pupils in all the primary schools in Osun State constitute 30% of the total population of Osun state. A random sample of 400 pupils from all primary schools in the Local Government Areas of Osun State shows that 25% of them are pupils of primary school.

Test at 5% level of significance the validity or otherwise of the demographer's claim.

Solution:

Let p be the population proportion of pupils in the primary schools in Osun State and P the estimated proportion.

$H_0 : P_0 = P$ i.e $H_0 : 0.30 = 0.25$

$H_1 : P_0 \neq P$ i.e $H_1 : 0.30 \neq 0.25$

$$\therefore Z_{cal} = \frac{P - P_0}{\sqrt{\dfrac{P(1-P)}{n}}}$$

$$= \frac{0.25 - 0.30}{\sqrt{\dfrac{0.30(1 - 0.30)}{400}}} = \frac{-0.05}{\sqrt{\dfrac{0.30(0.70)}{400}}} = \frac{-0.05}{0.0229} = -2.1834$$

It is a two-sided test

$\therefore$ for $\alpha = 0.05$, then $\alpha/2 = 0.025$ and table value of $Z = 1.96$

Decision: since $|Z_{cal}| > Z_{table}$ i.e $2.1834 > 1.96$,

H_0 is rejected and conclude that the data can not support the demographer's claim.

8.8 *Summary*

Statistical inference has been discussed under two broad headings, namely: the estimation theory and test of hypothesis.

In the estimation theory, point and interval estimates are defined. The principle of sampling distribution is also presented as a basis for obtaining point estimates of mean and proportion. Important concept such as hypothesis, errors, significance level, test statistics, one-tailed and two-tail tests are discussed in the hypothesis testing. Tests concerning the mean and proportion are finally presented.

8.9   Revision Questions

1.   In estimation theory, a statistic of a distribution is a quantity computed from_____

    a.   a population,   b.   the normal distribution

    c.   a sample,   d.   the poisson distribution

    e.   average of the data from both sample and population.

2.   Significance level is referred to as the risk of _____

    a.   sampling error,   b.   non-sampling error,

    c.   bias error   d.   type I error,

    e.   type II error.

3.   Which of the following hypothesis is not for one – tailed test?

    a. $H_1 : \mu > 0$,   b. $H_1 : \mu \geq 0$,   c. $H_1 : \mu < 0$

    d. $H_1 : \mu \leq 0$,   e. $H_1 : \mu \neq 0$.

4.   The test – statistic for a large sample in the hypothesis testing of mean is _____

5.   The degree of freedom for t – test with n samples is _____

6.   The test – statistic for proportion is _____

Use the following information to answer questions 7 to 10

    A sample of 4 items are taken randomly from a population sample, their weights (in Kgs) are recorded as follows: 6, 8, 12, and 14.

7. Determine the mean weight of the sample.

8. Determine the standard deviation of sample.

9. Compute the standard error of the mean.

10. If the population mean $\mu = 9$, compute the test – statistic for the dat

Answers to Chapter Eight

1(c), 2(d), 3(e), 4. $Z_{cal} = \dfrac{\bar{x} - \mu_0}{\sigma/\sqrt{n}}$, 5. n – 1

6. $Z_{cal} = \dfrac{P - P_0}{\sqrt{\dfrac{P(1-P)}{n}}}$, 7. 10, 8. 10,

9. $10/\sqrt{4} = 5$

10. $t_{cal} = \dfrac{\bar{x} - \mu_0}{S/\sqrt{n}} = \dfrac{10 - 9}{5} = 0.2$

SECTION B

BUSINESS MATHEMATICS

CHAPTER 9

FUNCTIONS

At the end of this chapter, students and readers should be able to

- understand the concept of functions
- identify different types of functions
- understand equations
- solve different types of equations by algebraic and graphical methods
- understand the concept of linear inequalities and their solutions
- apply all the concepts above to business and economic problems

9.1   Definition of a Function

A function is a mathematical way of describing a relationship between two or more variables. In other words, it is a mathematical expression involving one or more variables. Functions are useful in business and commerce where fragments of information can be connected together by functional relationships.

For example, a function of x can be written as $f(x) = 4x + 3$; $f(x)$ is read as 'f of x'.

Example 9.1

If   (a) $f(x) = 4x + 3$,   find   (i) $f(0)$,   (ii) $f(1)$,   (iii) $f(14)$.

(b) $f(x) = 2x^2 + 5x + 7$,   find   (i) $f(0)$,   (ii) $f(2)$,   (iii) $f(15)$.

Solutions

(a) $f(x)$ $= 4x + 3$

   (i) $f(0)$ $= 4 \times 0 + 3$ $= 3$

   (ii) $f(1)$ $= 4 \times 1 + 3$ $= 7$

   (iii) $f(14)$ $= 4 \times 14 + 3$ $= 59$

(b) $f(x)$ $= 2x^2 + 5x + 7$

   (i) $f(0)$ $= 2 \times 0^2 + 5 \times 0 + 7$ $= 7$

   (ii) $f(2)$ $= 2 \times 2^2 + 5 \times 2 + 7$ $= 25$

   (iii) $f(15)$ $= 2 \times 15^2 + 5 \times 15 + 7$ $= 532$

Some of the time, y is used to represent $f(x)$

i.e. $f(x) = 7x^2 + 2$ can be written as $y = 7x^2 + 2$; in such a case, y is said to be a function of x.

9.1.1 Explicit and Implicit Functions

An explicit function is a function where one variable is directly expressed in terms of the other variable(s).

For example, $y = 5x + 9$; and $y = 3x^2 + 8$ are explicit functions.

An implicit function is a function where the relationship between the variables is expressed as an equation involving all the variables.

For example, $7x + 2y - 8 = 0$; and $2x^2 + 3xy + 3y^2 + 10 = 0$ are implicit functions.

9.2 Types of Function
9.2.1 Linear Functions

A linear function is one in which the variables are of first degree and is generally of the from

$f(x) = a + bx$ or $y = a + bx$, where a and b are constants.

e.g. $y = 18 - 2x$; $y = 6x + 9$

9.2.1.1 Graph of a Linear Function

The graph of a linear function ($y = a + bx$) is a straight line. 'a' is the intercept on the y-axis (i.e. the value of y when $x = 0$) and 'b' is the gradient or slope of the line.

The gradient of a line shows the increase in y for a unit increase in x. It may be positive or negative and is unique to that line (i.e. a line has only one gradient).

e.g. for the line $y = 5x + 3$, the intercept on the y-axis is 3 and the gradient is 5.

Example 9.2

Draw the graph of each of the following functions
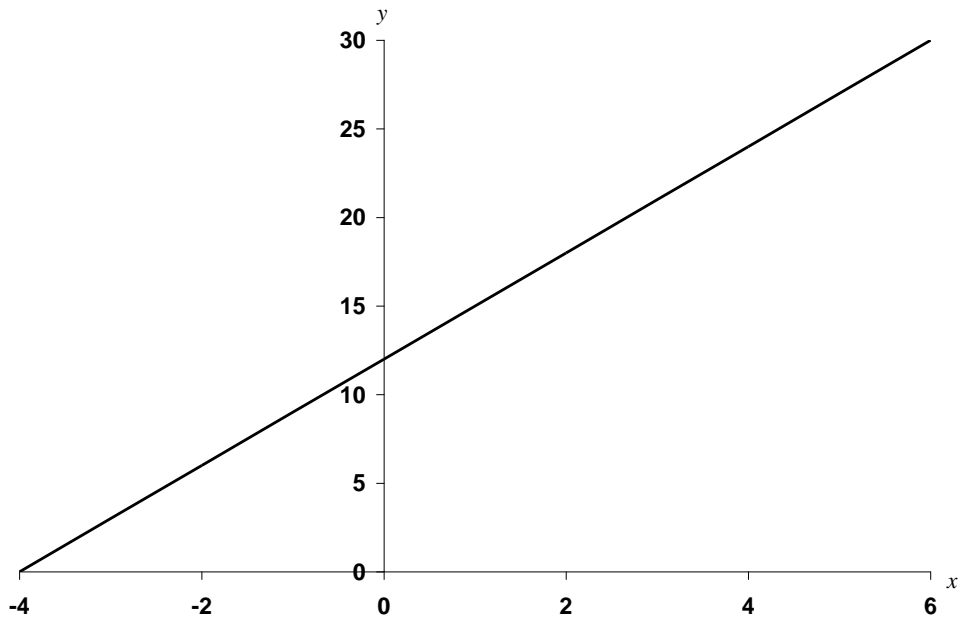
(a) $y = 3x + 12$

(b) $y = 46 - 5x$

Solution

Any two points are enough to draw the graph of a straight line and these are usually the intercept on the y-axis, where $x = 0$; and the intercept on the x-axis, where $y = 0$.

Once the two points are obtained, the line can be drawn and then extended as desirable.

(a) $y = 3x + 12$

When $x = 0$, $y = 12$, *i.e.* $(0, 12)$ is a point on the line.

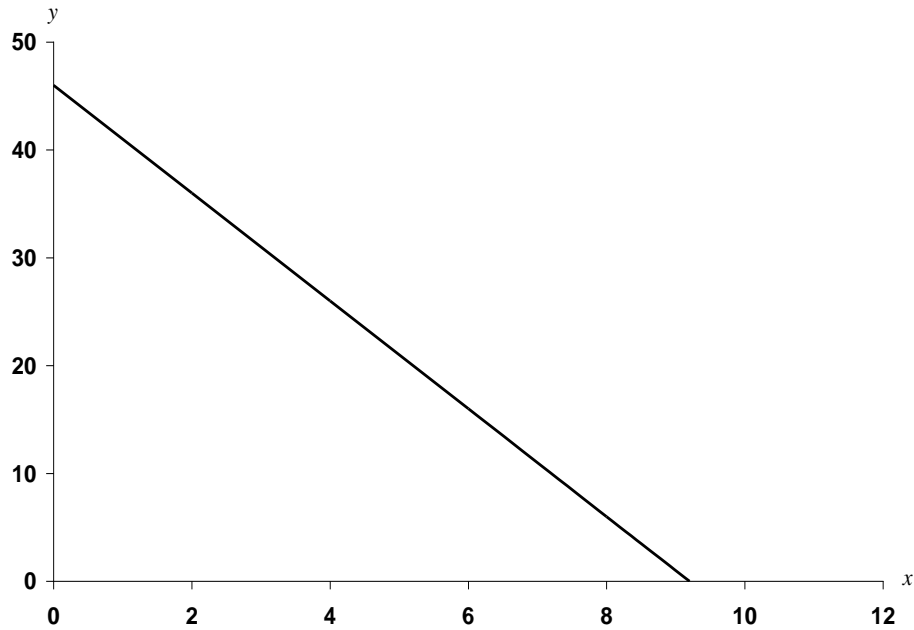When $y = 0$, $x = -4$, *i.e.* $(-4, 0)$ is a point on the line.

(b) $y = 46 - 5x$

When $x = 0$, $y = 46$, *i.e.* $(0, 46)$ is a point on the line.

When $y = 0$, $x = 9.2$, *i.e.* $(9.2, 0)$ is a point on the line.

9.2.2 Quadratic Functions

A quadratic function is one in which the variable x is of second degree and is generally of the form $y = ax^2 + bx + c$, where a, b and c are constants and a must not be zero.

Examples of quadratic functions are:

$$y = x^2 + 5x + 6;$$
$$y = 16x - 3x^2 + 1; \text{ and}$$
$$y = 9x^2$$

are quadratic functions.

9.2.2.1 Graph of Quadratic Functions

The graph of a quadratic function $y = ax^2 + bx + c$ is either cup-shaped ($\leq$) when 'a' is positive or cap-shaped ($\geq$) when 'a' is negative. Usually, the range of values of x is known.

Example 9.3
Draw the graph of the following for $-4 \leq x \leq 3$
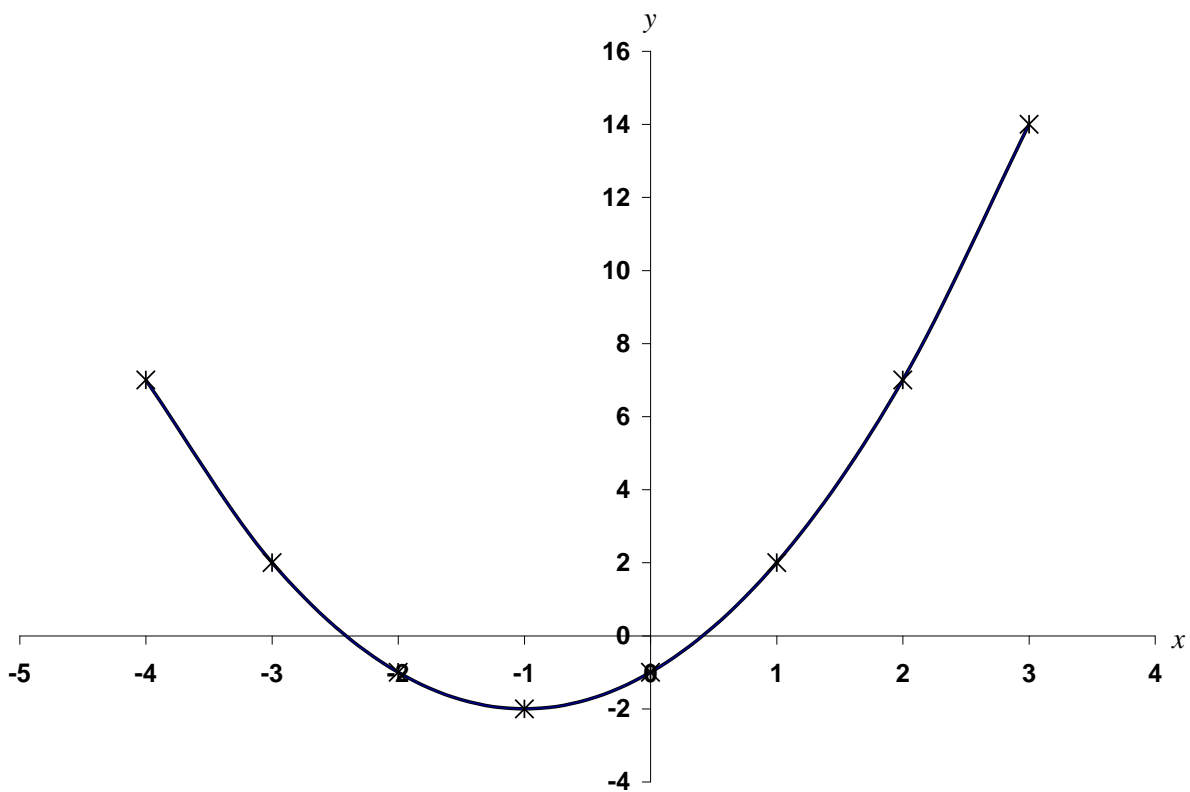(a) $y = x^2 + 2x - 1$
(b) $y = 5 - 2x^2 - 3x$

Solutions

Substitute values of x within the given range in the quadratic function to obtain the corresponding values of y. This gives a table of values, which are plotted on the graph.

(a) $y = x^2 + 2x - 1$;

x	-4	-3	2	-1	0	1	2	3
x^2	16	9	4	1	0	1	4	9
$2x$	-8	-6	-4	-2	0	2	4	6
-1	-1	-1	-1	-1	-1	-1	-1	-1
y	7	2	-1	-2	-1	2	7	14

b)　　$y = 5 - 2x^2 - 3x$;

x	-4	-3	-2	-1	0	1	2	3
5	5	5	5	5	5	5	5	5
$-2x^2$	-32	-18	-8	-2	0	-2	-8	-18
$-3x$	12	9	6	3	0	-3	-6	-9
y	15	-4	3	6	5	0	-9	-22
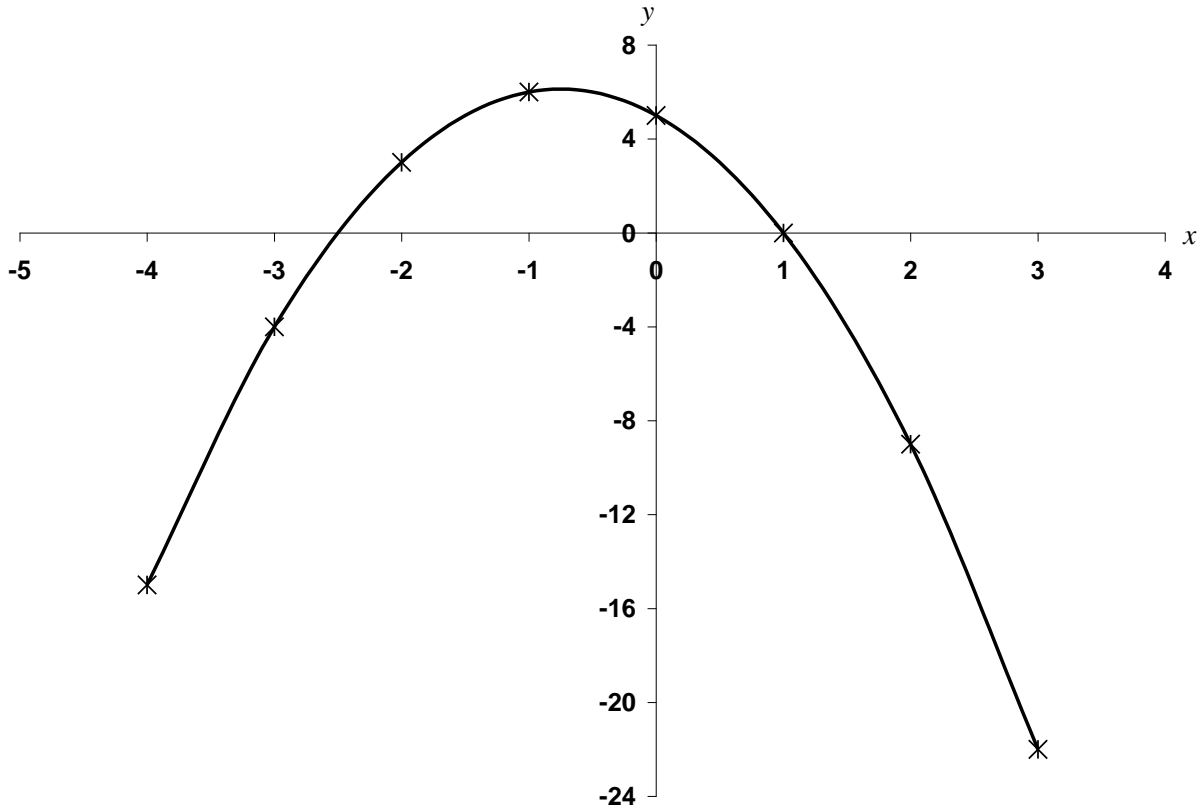
9.2.3　Exponential functions

An exponential function is a function, which has a constant base and a variable exponent. For example if $y = a^x$, then y is said to be an exponential function of x; 'a' is the base and x the exponent.

If 'a' exceeds one, there is an exponential growth but if 'a' is less than one there is an exponential decay.

Most exponential functions used in economic theory have the base 'e' (*i.e.* $y = e^x$). Values for exponential functions with base 'e' are easily obtainable from statistical tables or by the use of a calculator.

9.2.3.1 Graphs of exponential functions

The shape of the graph of an exponential function depends largely on the value of the base.

Example 9.4

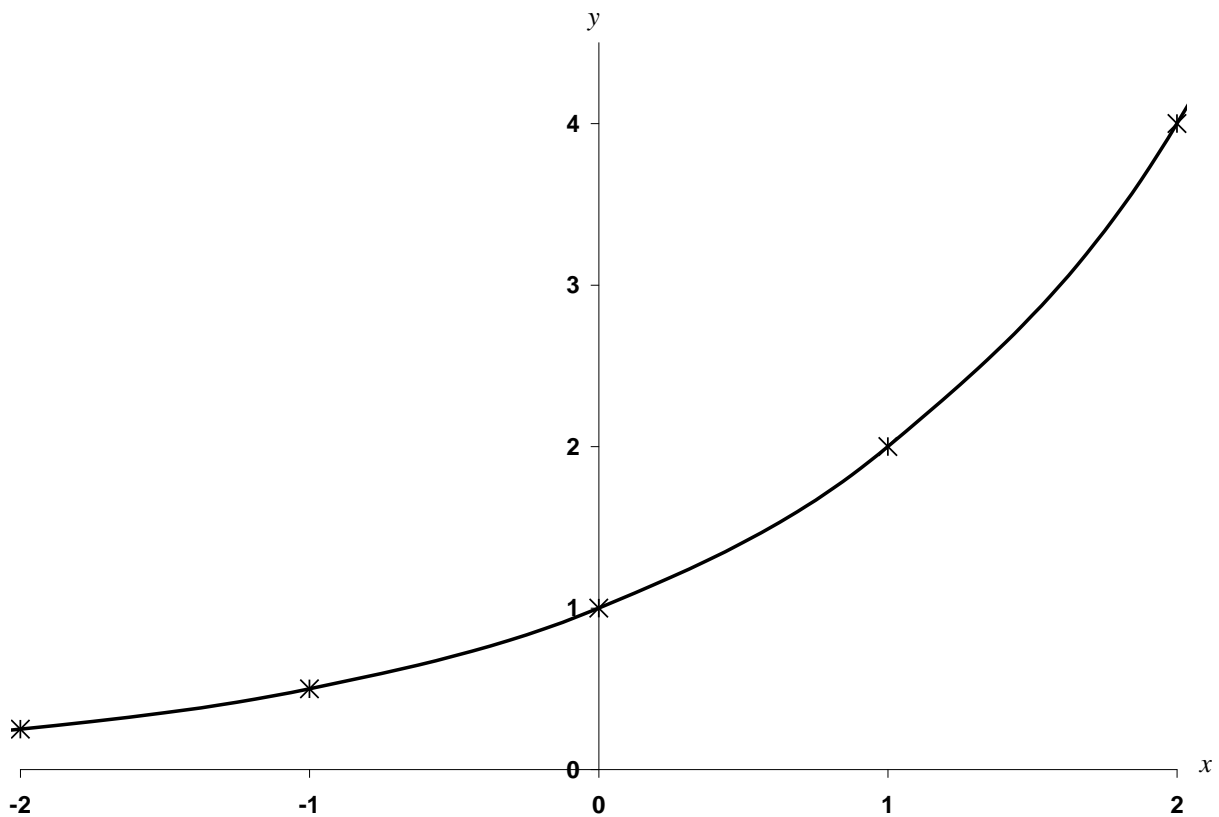Plot the graph of the following for $-2 \leq x \leq 2$

(a) $y = 2^x$

(b) $y = \left(\dfrac{1}{3}\right)^x$

(c) $y = e^x$

(d) $y = 3e^{\frac{x}{2}}$

Solutions

As usual, we substitute values of x within the range to obtain table of values

(a) $y = 2^x$

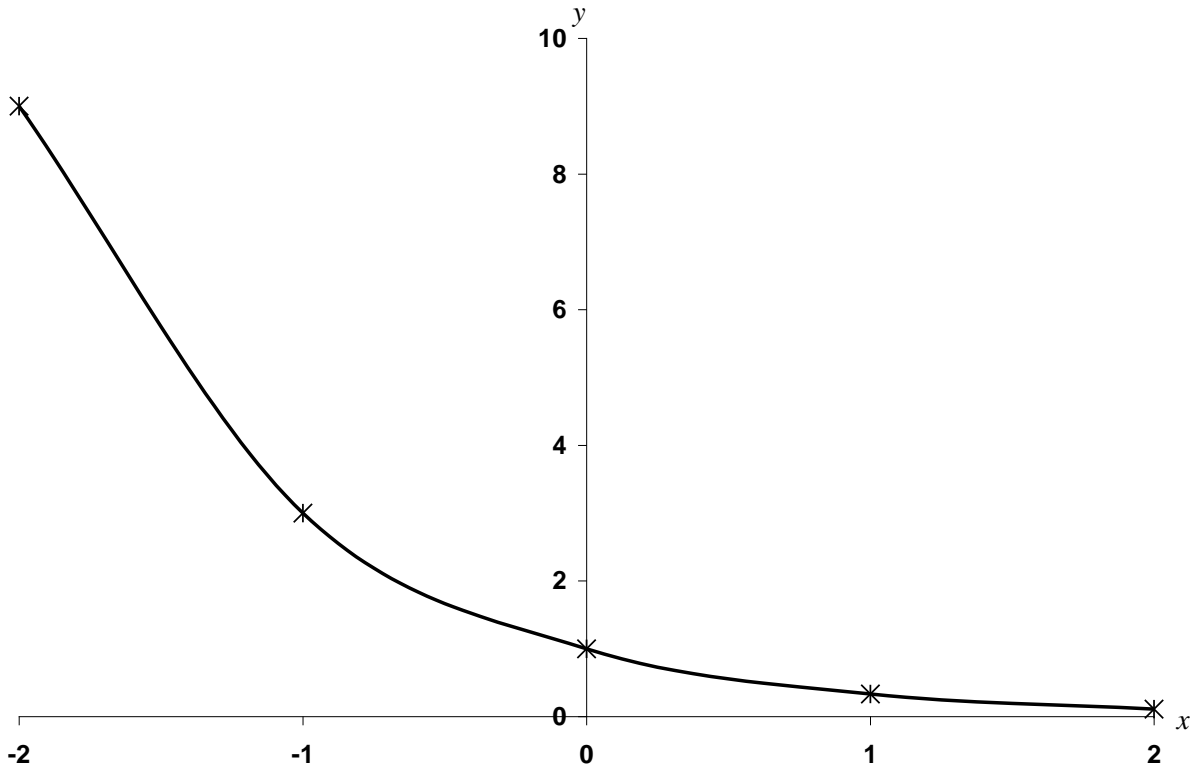| x	-2	-1	0	1	2
y	0.25	0.5	1	2	4

(b) $y = \left(\dfrac{1}{3}\right)^x$

x	-2	-1	0	1	2
y	9	3	1	0.33	0.11

(c) $y = e^x$

x	-2	-1	0	1	2
y	0.14	0.37	1	2.72	7.39
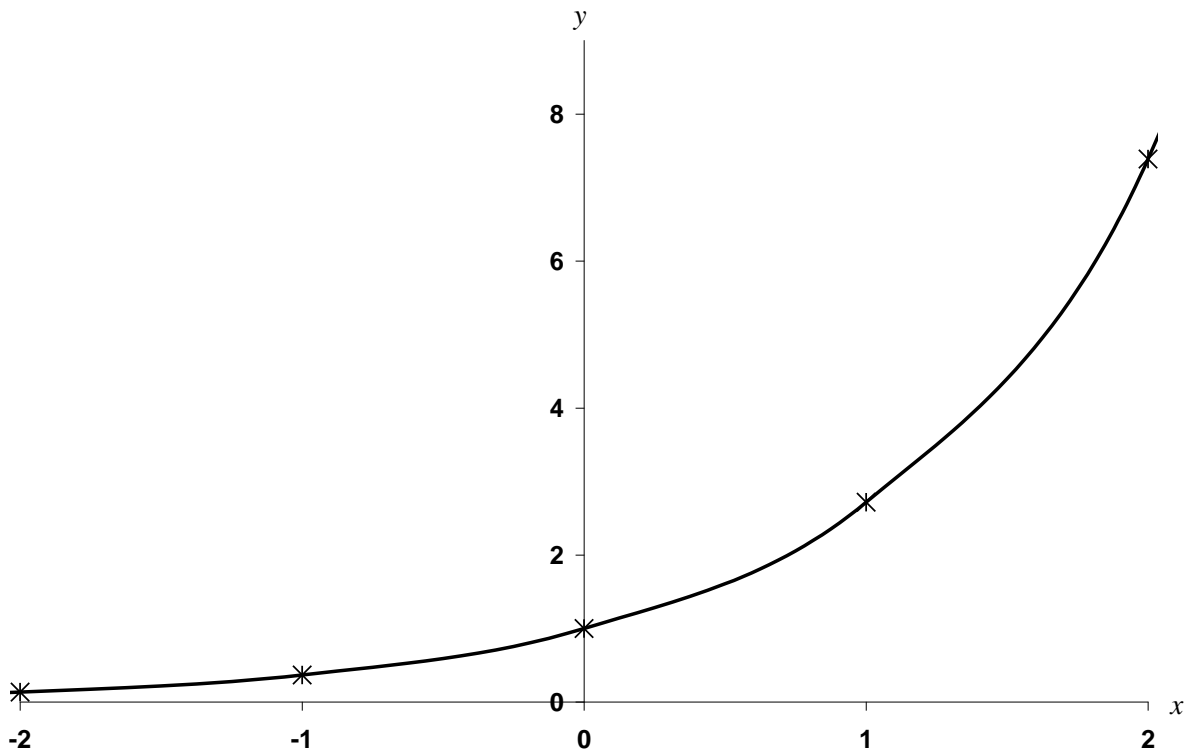
(d) $y = 3e^{\frac{x}{2}}$

x	-2	-1	0	1	2
y	1.1	1.82	3	4.95	8.15
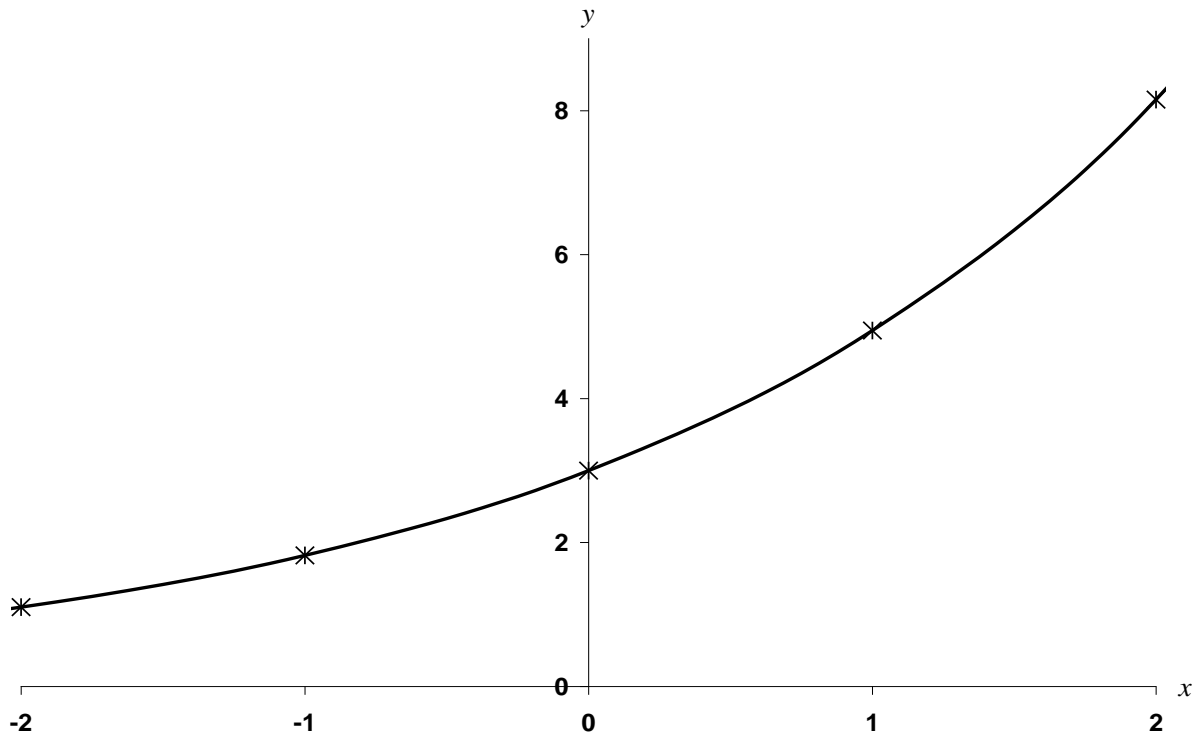
9.3   Equations

An equation is a mathematical expression of two equal quantities. It consists of variables which are referred to as unknowns and numbers which are called constants.

Examples

(a) $x + 5 = 2$

(b) $2x + 15 = x + 8$

(c) $x^2 + 6x + 9 = 0$

Rules for handling equations:

(i)   An equation is unchanged if

- the same number or expression is added to (or subtracted from) each side of an equation.

- each side of an equation is multiplied (or divided) by the same number or expression.

(ii)   The sign of an expression or a number changes when it crosses the equality sign.

9.3.1 Linear Equations

A linear equation in one variable is the simplest type of an equation and it contains one unknown, which is of the first degree. Examples (a) and (b) above are linear equations

9.3.1.1 Solutions of Linear Equations

A linear equation can be solved by applying the rules listed above

Example 9.5

Solve the following equations

(a) $2x + 3 = 5$

(b) $8a + 5 = 3a + 30$

(c) $\frac{1}{3}(y + 7) = 2(y - 1)$

(d) $\frac{2x}{5} - \frac{3}{14} = \frac{x}{10} + \frac{2x}{7}$

Solutions

(a) $2x + 3 = 5$

    Subtract 3 from both sides

    $2x = 2$

    Divide both sides by 2

    $x = 1$

(b) $8a + 5 = 3a + 30$

    $8a - 3a = 30 - 5$

    $5a = 25$

    $a = 5$

(c) $\frac{1}{3}(y + 7) = 2(y - 1)$

    Multiply each side by 3

    $y + 7 = 6(y - 1)$

    $y + 7 = 6y - 6$

    $y - 6y = -6 - 7$

    $-5y = -13$

$y \qquad = 2.6$

(d) $\qquad \dfrac{2x}{5} - \dfrac{3}{14} = \dfrac{x}{10} + \dfrac{2x}{7}$

Multiply each side by 70 (because the LCM of all the denominators is 70)

$28x - 15 \qquad = 7x + 20x$

$28x - 27x \qquad = 15$

$x \qquad\qquad = 15$

9.3.2 Linear Equations in two variables or unknowns (Linear Simultaneous Equations)

If the solutions to the equations containing two unknowns can be found at the same time, the equations are referred to as simultaneous equations.

In order for any set of simultaneous equations to have solutions, the number of equations and unknowns must be the same.

Any letter could be used to represent the unknowns but the letters x and y are commonly used.

Example

$2x + 3y = 13$

$3x + 2y = 32$

9.3.2.1 Solutions of Simultaneous Equations

The following three methods of solving simultaneous equations are discussed below

- Substitution method
- Elimination method
- Graphical method

The fourth method, which is the matrix method, is discussed in chapter 11

Example 9.6

Solve the following simultaneous equations

(a) $\qquad 2x + 3y = 13$

$\qquad 3x + 2y = 32$

by substitution method

(b) $5x - 4y = -6$

$6x + 2y = 20$

by   (i) elimination method

(ii) graphical method

Solutions

(a) $2x + 3y = 13$ ……………… (i)

$3x + 2y = 32$ ……………… (ii)

Express x (or y) in terms of y (or x), from equation (i)

$2x + 3y = 13$

$x = \dfrac{1}{2}(13 - 3y)$…………….. (iii)

Substitute the value of x from equation (iii) in equation (ii) to get

$\dfrac{3}{2}(13 - 3y) + 2y = 32$

$39 - 9y + 4y \quad = 64$

$-5y \qquad\qquad = 25$

$y \qquad\qquad\quad = -5$ ………………..……… (iv)

Use the result of (iv) in (iii) to obtain x, $i.e$

$x = \dfrac{1}{2}(13 - 3 \times -5)$

$x = \dfrac{1}{2}(13 + 15)$

$x = 14$

$\therefore$ The solutions are $x = 14$ and $y = -5$

Equation (ii) could also be used to express one variable in terms of the other. The common sense approach is to use the simpler of the two equations

(b) $5x - 4y = -6$……………… (i)

$6x + 2y = 20$……………… (ii)

(i)   By the elimination method

equation (ii) $\times$ 2 gives $12x + 4y \qquad = 40$ …………………….(iii)

$5x - 4y \qquad = -6$ ………………… (i)

clxxxii

equations (iii) + (i) gives $\quad 17x \qquad = 34$

$\qquad\qquad\qquad\qquad\qquad x \qquad\qquad = 2$ ………………..(iv)

Put the value of x from equation (iv) in equation (iii) to get

$$6(2) + 2y \qquad = 20$$
$$2y \qquad = 8$$
$$y \qquad = 4$$

$\therefore$ The solutions are $x = 2$ and $y = 4$

or putting the value of x from equation (iv) in equation (i),

$$5x - 4y \qquad = -6$$
$$5(2) - 4y \qquad = -6$$
$$-4y \qquad = -16$$
$$y \qquad = 4 \text{ as before.}$$

(ii)   Draw the graph of each of the two equations above on the same axes.

The point of intersection of the two lines is the solution

$$5x - 4y = -6 \dots\dots\dots\dots \text{ (i)}$$
$$6x + 2y = 20 \dots\dots\dots\dots \text{ (ii)}$$

From equation (i) $\qquad 5x - 4y = -6$

$$-4y \qquad = -5x - 6$$
$$y \qquad = \frac{5}{4}x + \frac{6}{4}$$
$$y \qquad = 1.25x + 1.5$$

When $x = 0$, $y = 1.5 \quad \square \ (0, 1.5)$

When $y = 0$, $x = 1.2 \quad \square \ (1.2, 0)$

From equation (ii) $\qquad 6x + 2y = 20$

$$2y \qquad = -6x + 20$$
$$y \qquad = -3x + 10$$

When $x = 0$, $y = 10 \quad \square \ (0, 10)$

When $y = 0$, $x = \dfrac{10}{3} \quad \square \ (3.3, 0)$
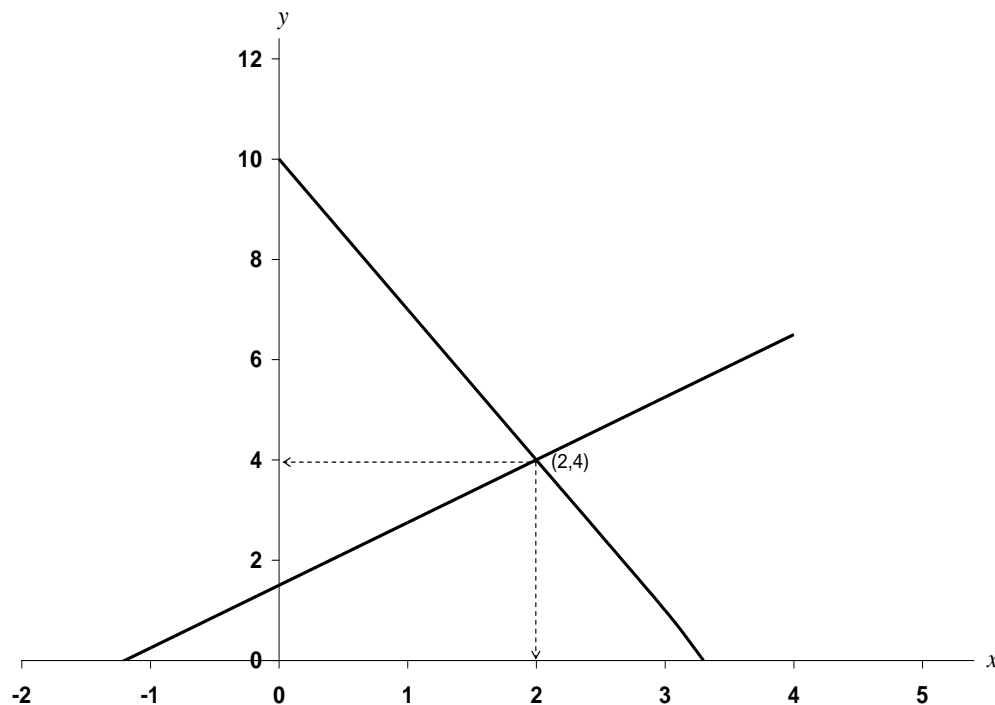
From the graph, the solution is at the point (2, 4) *i.e.* $x = 2$, $y = 4$ as before

Note: some of the time, graphical solutions may not be exactly the same as the algebraic solutions due to some approximations that might have crept in - they should not be too different though. In fact, graphical solutions are regarded as estimates.

9.3.3   Quadratic Equations

Equations in one variable and of the second degree are called quadratic equations. It is generally of the form

$$ax^2 + bx + c = 0, \text{ where } a\ (\neq 0),\ b \text{ and } c \text{ are constants}$$

$$e.g. \quad x^2 + x + 8 \quad = 0$$

$$3x^2 ++ 13x \quad = 0$$

$$4x^2 + 11 \quad = 0$$

9.3.3.1 Solutions of Quadratic Equations

The following three methods of solving quadratic equations are discussed below:

- Factorisation method
- Formula method
- Graphical method

Example 9.7

Solve the following quadratic equations

(a) $x^2 + 6x + 8 = 0$ by factorisation method

(b) $2x^2 + 13x - 16 = 0$ by formula method

(c) $x^2 + 3x - 7 = 0$ by graphical method

Solutions

(a) $x^2 + 6x + 8 = 0$

Find two numbers whose sum is 6 and product is 8 *i.e.* 2 and 4

So we have

$x^2 + 2x + 4x + 8 \qquad = 0$

$x(x + 2) + 4(x + 2) \qquad = 0$

i.e. $\qquad (x + 2)(x + 4) = 0$

$x + 2 = 0 \ \text{ or } x + 4 = 0$

$x = -2 \text{ or } -4$

This method is not applicable to equations that cannot be factorised.

(b) $2x^2 + 13x - 16 = 0$

The formula is

$$x = \frac{-b \pm \sqrt{b^2 - 4ac}}{2a}$$

In this case, $a = 2$, $b = 13$, $c = -16$

Substitute these in the formula to obtain

$$x = \frac{-13 \pm \sqrt{13^2 - 4(2)(-13)}}{2(2)}$$

$$x = \frac{-13 \pm \sqrt{169 + 128}}{4}$$

$$x = \frac{-13 \pm 17.23}{4}$$

i.e. $\qquad x = \dfrac{-13 + 17.23}{4} \ \text{ or } \ x = \dfrac{-13 - 17.23}{4}$
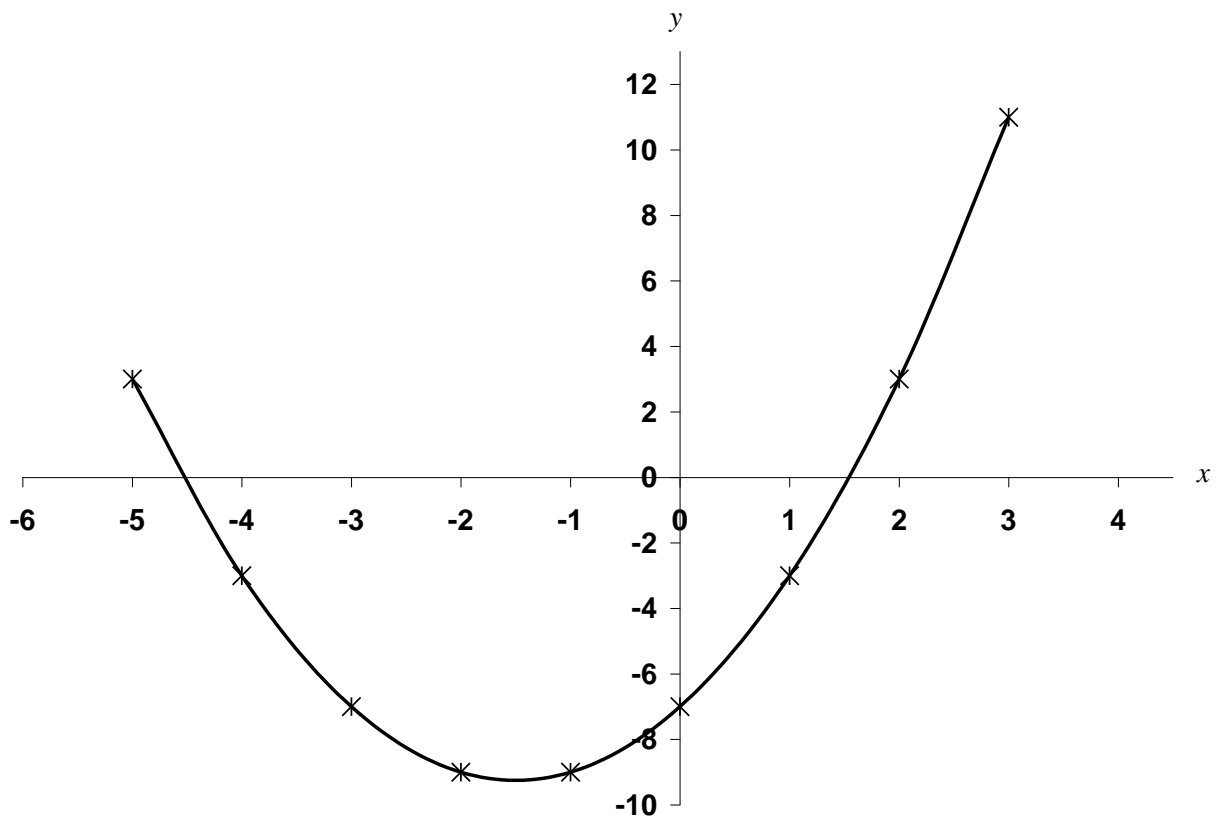
$x = -7.56 \text{ or } 1.06$

NOTE: This is a work in progress. All topics in the syllabus are covered but editing for necessary
corrections is in progress.
Thanks.

This method can be applied whether or not the equation can be factorised.

(c) $x^2 + 3x - 7$

Obtain the table of values as

x	-5	-4	-3	-2	-1	0	1	2	3
x^2	25	16	9	4	1	0	1	4	9
$3x$	-15	-12	-9	-6	-3	0	3	6	9
-7	-7	-7	-7	-7	-7	-7	-7	-7	-7
y	3	-3	-7	-9	-9	-7	-3	3	-11

The solutions are the points where the curve intersects the x-axis.

i.e. $x = -4.5$ or 1.5

This method can be applied to any quadratic equation.

9.4   Inequalities and their graphical solutions

When two quantities or expressions are not equal, then one has to be greater or less than the other.

They are known as inequalities.

The symbols used to express inequalities are

$<$	means "less than"
$\leq$	means "less than or equal to"
$>$	means "greater than"
$\geq$	means "greater than or equal to"

Example

(i) $19 < 30$

(ii) $2x + 5 > x - 4$

(iii) $3x + 2y \leq 6$

Rules for handling inequalities

(i)   The symbol of an inequality does not change if

- a number or an expression is added to (or subtracted from) both sides of the inequality.
- each side of the inequality is multiplied (or divided) by a *positive* number.

(ii)   the symbol of an inequality changes if both sides of the inequality is multiplied (or divided) by a *negative* number.

(iii) the sign of a number or an expression changes when it crosses the inequality symbol.

Generally, solutions to inequalities consist of range of values rather than point values, as is the case with equations.

Example 9.8

Indicate the region where each of the following inequalities is satisfied:
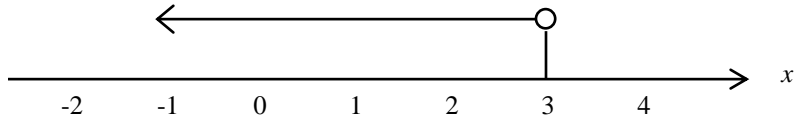
  a. $2x + 5 < x + 8$

  b. $4x + 7 \geq 2x + 15$

  c. $3x + 10 \leq 5x - 2$

  d. $x \square 0$

  e. $y \square 0$

  f. $x \square 0; y \square 0$

  g. $x \square 0; y \square 3$

  h. $x + 2y \square 6$

  i. $3x + 2y \square 12$

  j. $2x + 3y \square 9; 5x + 2y \square 10$

  k. $2x + y \square 18; 1.5x + 2y \geq 15; x \geq 0; y \geq 0$

Solutions

(a) $2x + 5 < x + 8$

$2x - x < 8 - 5$

$x < 3$

This is represented on the number line as:

O   indicates that 3 is not part of the solution

(b) $4x + 7 \geq 2x + 15$
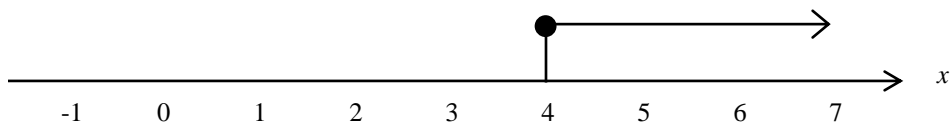
$4x - 2x \geq 15 - 7$

$2x \geq 8$

$x \geq 4$

This is represented on the number line as:

● indicates that 4 is part of the solution

(c) $3x + 10 \leq 5x - 2$

$3x - 5x \leq -2 - 10$

$-2x \geq -12$

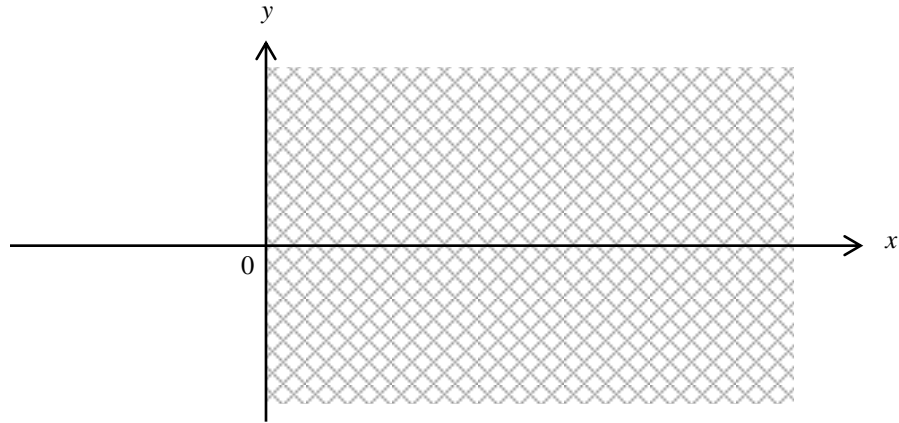$x \geq 6$ (the sign of the inequality changes since we have divided by –2)

This is represented on the number line as:
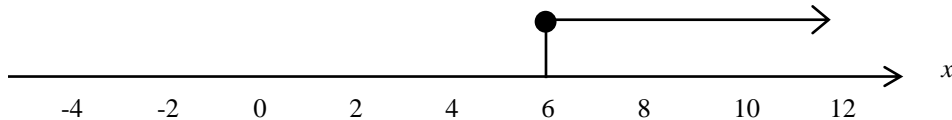
● indicates that 6 is part of the solution

(d) $x \geq 0$

(e) $y \geq 0$
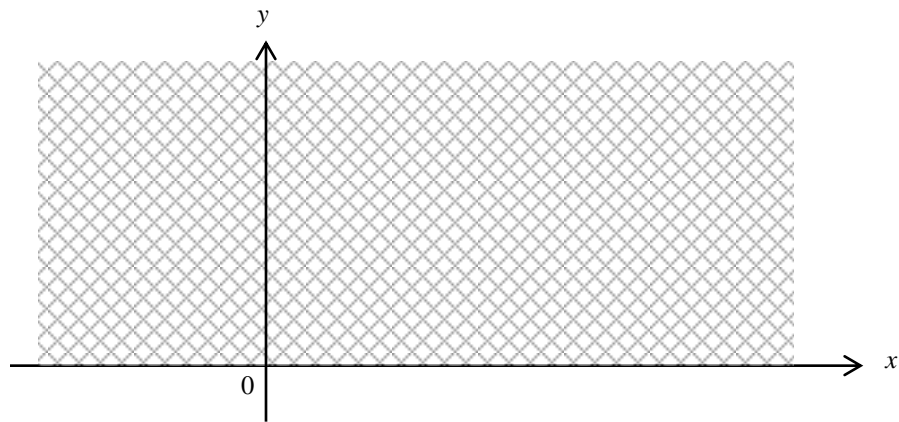
(f) $x \geq 0$ and $y \geq 0$

(g) $x \geq 0$ and $y \leq 3$
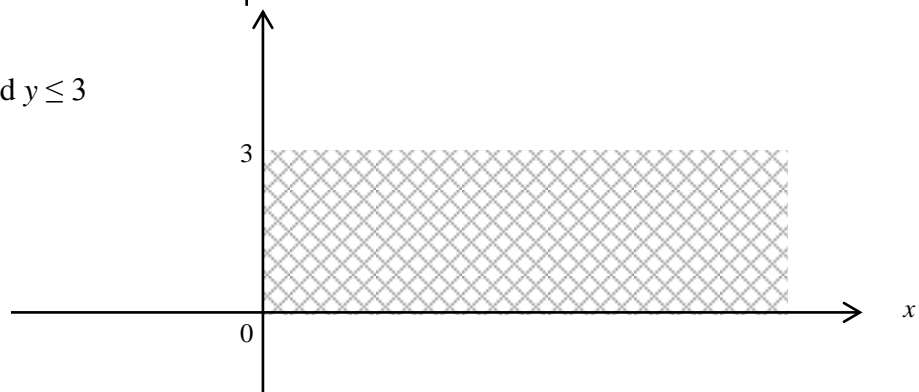
(h) $x + 2y < 6$

Draw the graph of $x + 2y = 6$

The line divides the whole region into two parts. The origin $(0, 0)$ is in one of the parts and is used to determine the part that satisfies the inequality.

$x + 2y = 6$

When $x = 0$, $y = 3$; = $(0, 3)$

When $y = 0$, $x = 6$; = $(6, 0)$

Now substitute the origin $(0, 0)$ in $x + 2y < 6 = 0 < 6$, which is true. Hence the part which satisfies the inequality contains the origin
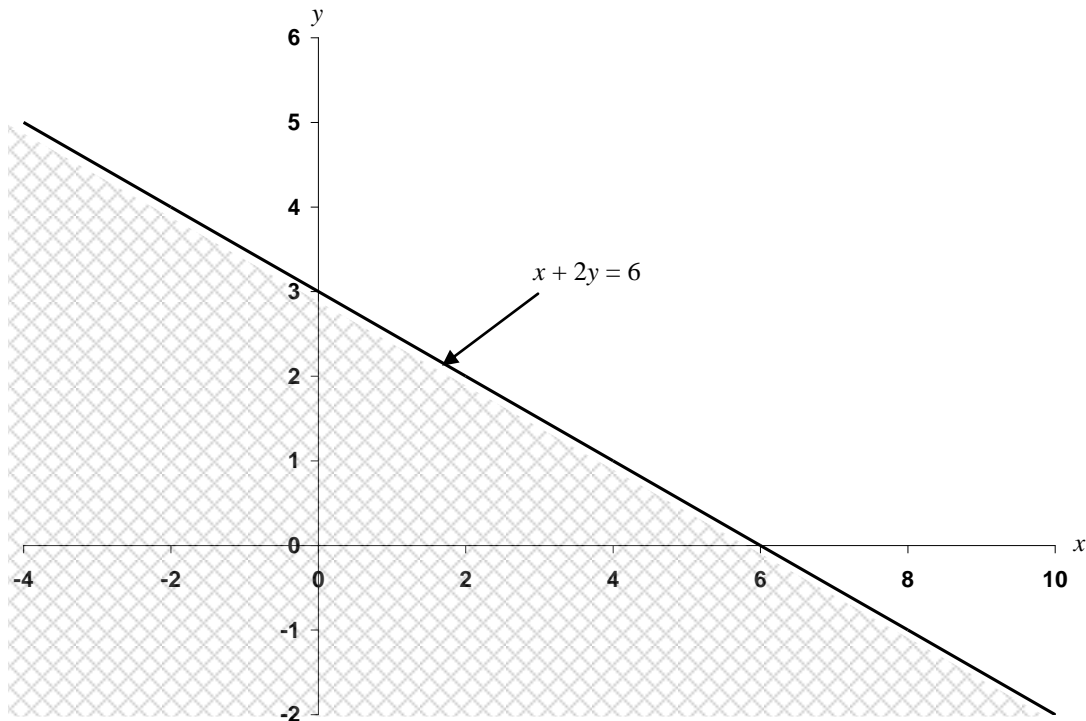
(i) $3x - 2y \geq 12$

Draw the line for $3x - 2y = 12$

When $x = 0$, $y = -6 = (0, -6)$

When $y = 0$, $x = 4$ $= (4, 0)$

Substituting the origin $(0, 0)$ in $3x - 2y \geq 12$ □ $0 \geq 12$, which is not true, hence the part in which the inequality is satisfied does not contain the origin.
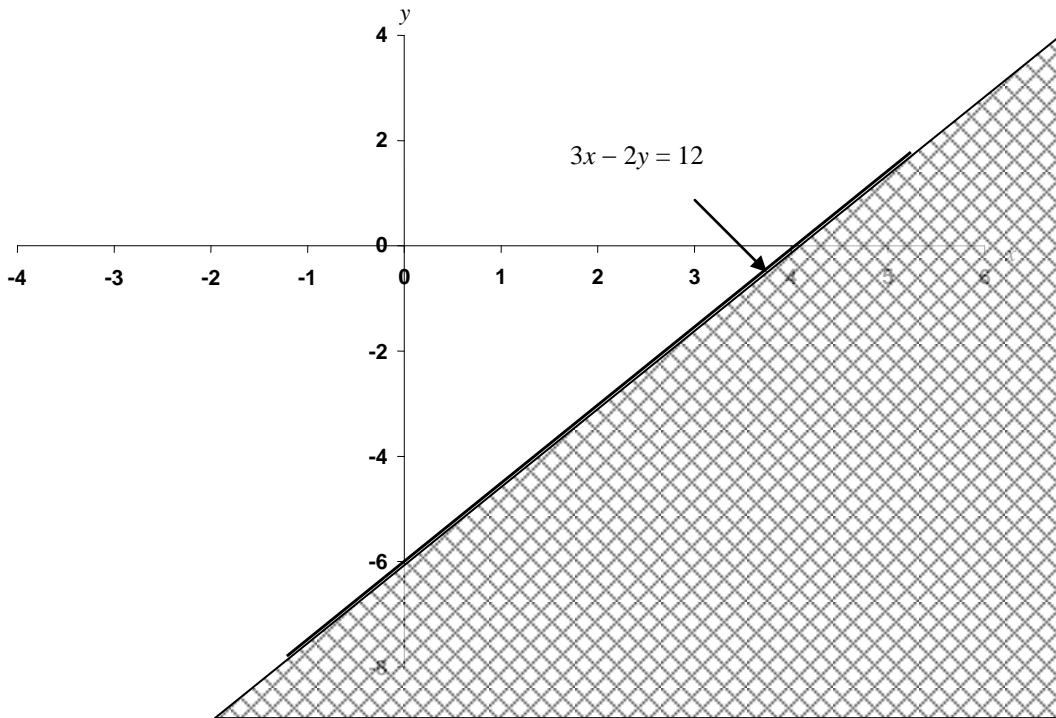
(j) $2x + 3y \leq 9$; $5x + 2y \leq 10$

Draw the lines for $2x + 3y = 9$ and $5x + 2y = 10$ on the same axes

For $2x + 3y = 9$,

When $x = 0$, $y = 3$ ⇨ (0, 3)

When $y = 0$, $x = 4.5$ ⇨ (4.5, 0)

For $5x + 2y = 10$,

When $x = 0$, $y = 5$ ⇨ (0, 5)

When $y = 0$, $x = 2$ ⇨ (2, 0)

The required region is where both inequalities are satisfied simultaneously.

(k) $2x + y \leq 18; 1.5x + 2y \geq 15; x \geq 0; y \geq 0$

For $2x + y = 18$,

When $x = 0$, $y = 18 \Rightarrow (0, 18)$

When $y = 0$, $x = 9 \Rightarrow (9, 0)$

For $1.5x + 2y = 15$,

When $x = 0$, $y = 7.5 \Rightarrow (0, 7.5)$

When $y = 0$, $x = 10 \Rightarrow (10, 0)$

The required region is where all the four inequalities are satisfied simultaneously.

NOTE: This is a work in progress. All topics in the syllabus are covered but editing for necessary corrections is in progress.
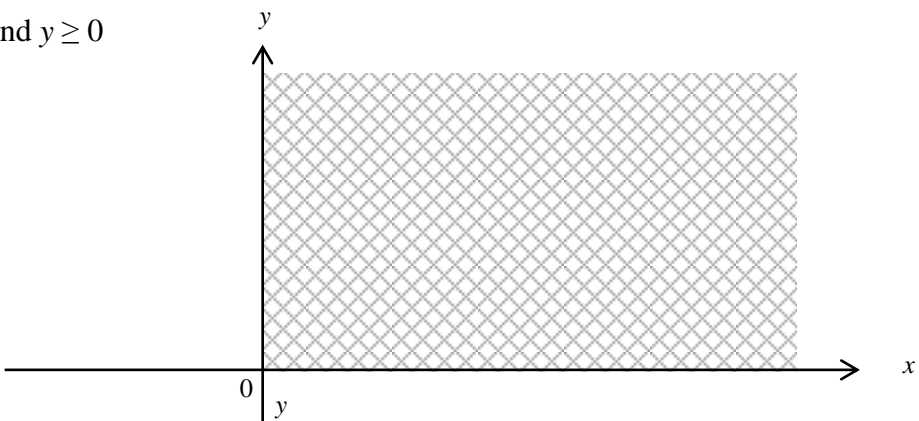
Thanks.
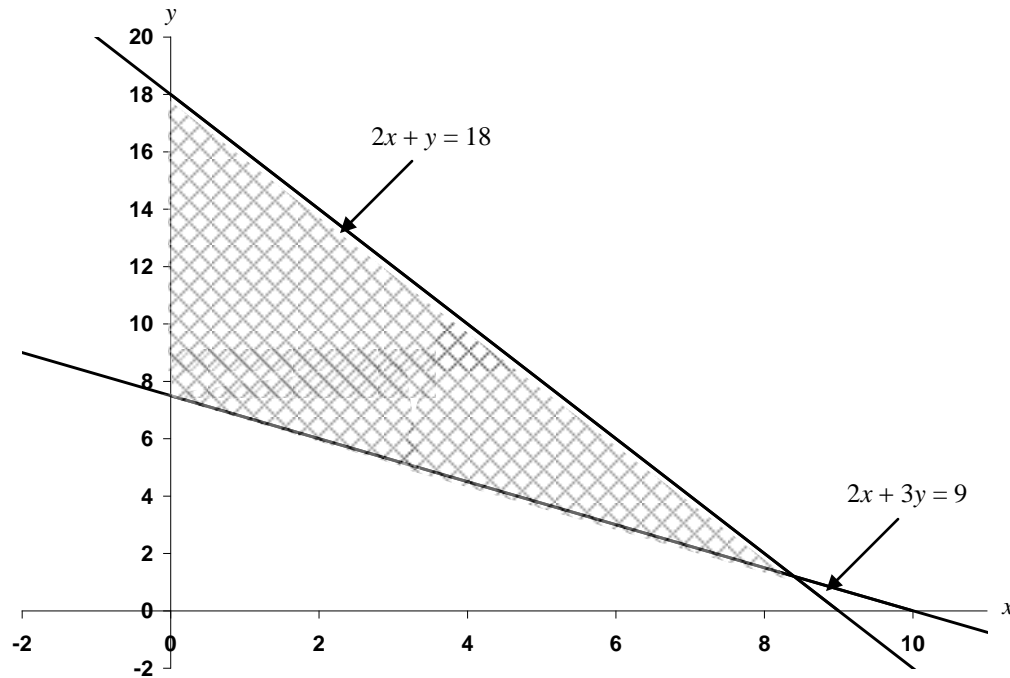
### 9.5 Applications to Business, Economic and Management Problems

The topics discussed in this chapter can be applied to various practical problems.

9.5.1 Cost, Revenue and Profit Functions

The **total cost** of a commodity, project, business etc consists of two parts, viz:

(i)   the **fixed cost** (cost of machinery, infrastructure etc), it can also be referred to as the **capital cost**

(ii)   the **variable cost** (cost of material, labour, utility etc)

The fixed cost is always constant. It represents the "set up cost" while the variable cost depends on the number of units of items produced or purchased.

i.e. **total cost = fixed cost + variable cost.**

Example

In a normal situation, the transport cost to a market is constant, and is independent of whatever commodities are purchased in the market. This is the fixed cost.

The variable cost is the cost of commodities purchased and it depends on the price and number of units of the commodities. If x represents the number of units of commodity purchased or items

produced, $C(x)$ called the **cost function,** can be expressed as follows

(a) $C(x) = a + bx$

This is a linear function, a represents the fixed cost, bx the variable cost and b is the price per unit of x.

(b) $C(x) = ax^2 + bx + c$

This is a quadratic function, a ($\neq 0$), b and c are constants; and c, which is independent of x, is the fixed cost.

Example 9.9

The cost of renting a shop is ₦120 000 per annum and an additional ₦25 000 is spent to renovate the shop. Determine the total cost, if the shop is to be equipped with 480 items at ₦75 per item.

Solution

The fixed cost $= ₦120\ 000 + ₦25\ 000$

$= ₦145\ 000$

∴ $C(x)$ $= ₦145000 + ₦75x$

If x $= 480$

$C(x)$ $= ₦145\ 000 + ₦75 \times 480$

$= ₦181, 000$

The **revenue function, R**(x) represents the income generated from the sales of x units of item.

R(x) totally depends on x and is always of the form:

R(x) = px where p is the sales price of an item.

There is nothing like fixed revenue.

Example 9.10

If 2,500 items are produced and sold at ₦120 per item, calculate the revenue that will accrue from the sales

Solution

$$\mathbf{R}(x) = px$$
$$= 120x$$
$$= ₦120 \text{ x } 2\,500$$
$$= ₦300,\,000$$

The **profit function** of any project or business is the difference between the revenue function and the cost function.

i.e. $\mathbf{P}(x) = \mathbf{R}(x) - \mathbf{C}(x)$

If $\mathbf{R}(x) > \mathbf{C}(x)$, then $\mathbf{P}(x) > 0$, *i.e.* positive (gain)

If $\mathbf{R}(x) < \mathbf{C}(x)$, then $\mathbf{P}(x) < 0$, *i.e.* negative (loss)

If $\mathbf{R}(x) = \mathbf{C}(x)$, then $\mathbf{P}(x) = 0$, *i.e.* no profit, no loss. This is the break-even case.

Example 9.11

A business man spends ₦1.5m to set up a workshop from where some items are produced. It costs ₦450 to produce an item and the sale price of an item is ₦1450. Find the minimum quantity of items to be produced and sold for the business man to make a profit of at least ₦800, 000.

Solution

It is always assumed that all the items produced are sold. Let x represent the number of items produced and sold.

Then

$$\mathbf{C}(x) = ₦1\,500\,000 + ₦450x$$
$$\mathbf{R}(x) = ₦1450x$$
$$\Box\mathbf{P}(x) = ₦1450x - ₦(1\,500\,000 + 450x)$$
$$= ₦1000x - ₦1\,500\,000$$

To make a profit of at least ₦800 000

$$\mathbf{P}(x) \geq 800\,000$$

i.e. $1000x - 1\,500\,000 \geq 800\,000$

NOTE: This is a work in progress. All topics in the syllabus are covered but editing for necessary corrections is in progress.

Thanks.

$$1000x \geq 2\ 300\ 000$$

$$x \geq 2300$$

i.e. at least 2300 items must be produced and sold to make a profit of at least ₦800, 000.

9.5.2 Break-even Analysis

A business or project breaks even when the revenue function is equal to the cost function.

i.e. $\mathbf{R}(x) = \mathbf{C}(x)$

or $\mathbf{R}(x) - \mathbf{C}(x) = 0$

but $\mathbf{R}(x) - \mathbf{C}(x) = \mathbf{P}(x)$

☐ $\mathbf{P}(x) = 0$

Hence the break-even point (b.e.p) is at the point where $\mathbf{P}(x) = 0$ *i.e.* when profit is zero

The general outlook of a break-even graph is as shown below.
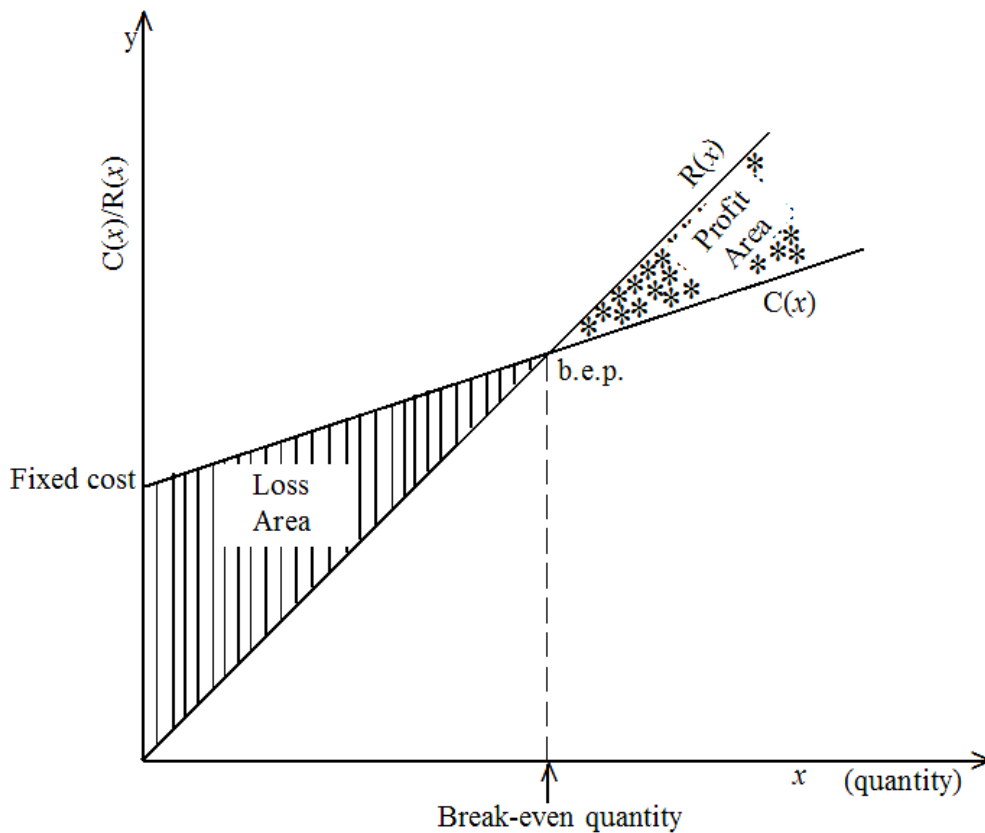
Example 9.12

Calculate the break-even quantity for example 9.11 above

Solution

$\mathbf{R}(x)$ $= 1450x$

$\mathbf{C}(x)$ $= 1\ 500\ 000 + 450x$

For the break-even situation

$\mathbf{R}(x)$ $= \mathbf{C}(x)$

i.e. $1450x = 1\ 500\ 000 + 450x$

$1000x = 1\ 500\ 000$

$x = 1500$

i.e. the break-even quantity is 1500

There are three possibilities

☐ $x < 1500$ ☐ negative profit*i.e.* loss

☐ $x = 1500$ ☐ break-even *i.e.* no profit, no loss

☐ $x > 1500$ ☐ profit

Example 9.13

The cost and revenue functions of a company are given by

$\mathbf{C}(x) = 400 + 4x$

$\mathbf{R}(x) = 24x$

Use the graphical method to determine the break-even quantity. Identify the profit and loss areas on your graph.

Solutions

The cost and revenue functions are drawn on the same graph. The quantity at the point of intersection of the two lines is the break-even quantity.

For $\mathbf{C}(x) = 400 + 4x$,

when $\mathbf{C}(x) = 0$, $x = -100$ ☐ $(-100, 0)$

when $x = 0$, $\mathbf{C}(x) = 400$ ☐ $(0, 400)$

For $\mathbf{R}(x) = 24x$

Since there is no fixed revenue, the graph of the revenue function always passes through the origin. So we need another point which is obtained by substituting any positive value of x in $\mathbf{R}(x)$.

When $x = 0$, $R(x) = 0$ $\Rightarrow$ $(0, 0)$

When $x = 25$, $R(x) = 600$ $\Rightarrow$ $(25, 600)$
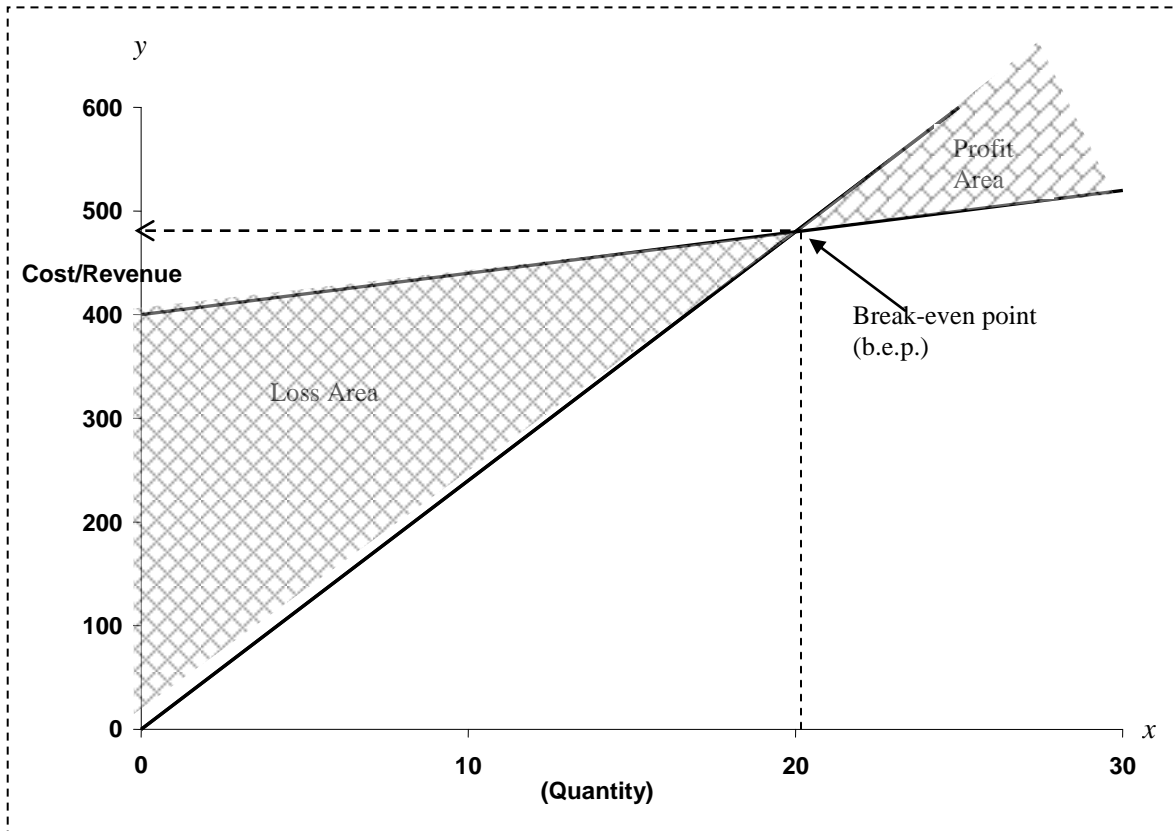
The break even point is as indicated on the graph.

i.e. the break-even quantity is 20.

The loss area is the region below the break-even point and within the lines while the profit area is the region above the break-even point and within the lines.

Alternatively, the loss area is the region where the revenue line is below the cost line and the profit area is the region where the revenue line is above the cost line.

Example 9.14

SEJAYEB (Nig Ltd) consumes 8 hours of labour and 10 units of material to produce an item thereby incurring a total cost of ₦2100. When 12 hours of labour and 6 units of material are consumed, the cost is ₦2700.

Find the cost per unit of labour and material.

Solution

Let x represent the cost per unit of labour and y the cost per unit of material.

Then we have the following simultaneous equations to solve:

$8x + 10y = 2100$ ………………….. (i)

$12x + 6y = 2700$………………………(ii)

equation (i) x 6 gives

$48x + 60y = 12600$………….(iii)

equation (ii) x 10 gives

$120x + 60y = 27000$………...(iv)

equation (iv) – equation (iii) gives $72x = 14400$

$x = 200$

Substitute x into (i)

i.e. $1600 + 10y = 2100$

$10y = 500$

$y=50$

i.e. the cost per unit of labour is ₦200 and the cost per unit of material is ₦50

Example 9.15

The Planning and Research department of a firm has estimated the sale function $S(x)$ to be $S(x) = 800x - 250$ and the cost function $C(x)$ as $50{,}000 + 200x^2 - 500x$, where x is the number of items produced and sold.

Determine the break-even quantity for the firm.

Solution

$\mathbf{C}(x)$ = $50{,}000 + 200x^2 - 500x$

$\mathbf{R}(x)$ = (number of items sold) (sale function)

= $x.\mathbf{S}(x)$

= $x(800x - 250)$

= $800x^2 - 250x$

For break-even quantity, $\mathbf{R}(x) = \mathbf{C}(x)$

i.e. $800x^2 - 250x = 50000 + 200x^2 - 600x$

$600x^2 + 350x - 50000 = 0$

$12x^2 + 7x - 1000 = 0$

Using the quadratic formula

$$x = \frac{-b \pm \sqrt{b^2 - 4ac}}{2a}$$

$a = 12, \, b = 7, \, c = -1000$

☐ $$x = \frac{-7 \pm \sqrt{7^2 - 4(12)(-1000)}}{2(12)}$$

$$x = \frac{-7 \pm \sqrt{7^2 - 4(12)(-1000)}}{2(12)}$$

$x = -9.43$ or 8.84

Since x cannot be negative because negative items cannot be produced, $x = 8.84 \,\square\, 9$

i.e. the break even quantity is 9.

Example 9.16

(a)   The total monthly revenue of DUPEOLU enterprise (in Naira) is given by the equation: $R = 200{,}000 \,(0.5)^{0.6x}$

Where x (₦'000) is the amount spent on overheads.

Calculate the

(i)   maximum revenue

(ii)   total revenue if ₦5000 is spent on overheads.

(b)   The production costs of GODSOWN Company are estimated to be
$C(x) = 200 - 80e^{-0.02x}$

Where x is the number of units of items produced.

Determine:

(i)   the fixed costs for the company

(ii)   the costs of producing 250 items.

(iii)   the percentage of the production costs in (ii) which are fixed.

Solutions

(a) (i) $R = 200{,}000(0.5)^{0.6x}$, maximum revenue occurs when nothing is spent on overheads ,

i.e. $x = 0$

$R = 200\,000(0.5)^{0}$

$R = ₦200,\, 000$

(ii) $x = \dfrac{5000}{1000} = 5$ since x is in thousands of naira

$R = 200{,}000(0.5)^{0.6(5)}$

$R = 200{,}000(0.5)^{3}$

$R = ₦25{,}000$

(b) (i) $C = 200 - 80e^{-0.02 x}$, fixed costs are the costs incurred when no items have been produced , *i.e.* $x = 0$

$C = 200 - 80e^0$

$C = 120$ *i.e.* ₦120 000

(ii) $C = 200 - 80e^{-0.02x}$

when $x = 250$,

$C = 200 - 80e^{-0.02(250)}$

$C = 199.461$ *i.e.* ₦199 461

(iii) %age required is $\dfrac{120000 \times 100}{199461} = 60.16\%$

9.5.3 Demand and Supply equations; Market Equilibrium.

Usually, demand and supply equations can be reasonably approximated by linear equations.

Demand Equation

Demand is inversely proportional to price i.e. quantity demanded decreases as price increases and it increases as price decreases.

For this reason, the slope of a demand curve is negative.

If the slope is zero, then the price is constant irrespective of demand and if the slope is undefined, it implies constant demand irrespective of price.

If x represents quantity demanded and y represents the price, both x and y must be positive.

The demand equation is of the form $y = a + bx$, where b is the slope.

The three situations are shown below
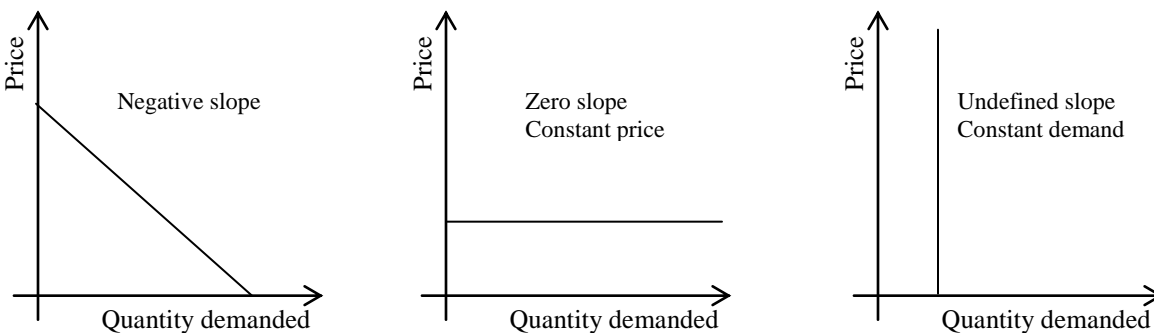
Example 9.17

500 watches are sold when the price is ₦2400 while 800 watches are sold when the price is ₦2000

(a)   obtain the demand equation.

(b)   how many watches will be sold if the price is ₦3000?

(c)   what is the highest price that could be paid for a watch?

Solutions

(a)   Let x be the quantity demanded and y the price, then the demand equation is

$y = a + bx$

When $x = 500$, $y = 2400$

i.e. $2400 = a + 500b$…….. (i)

when $x = 800$, $y = 2000$

i.e. $2000 = a + 800b$……...(ii)

(i) – (ii) gives

$400 = -300b$

$\square\ b = \dfrac{-4}{3}$

Substitute for b in equation (i)

i.e. $2400 = a - \dfrac{2300}{3}$

$a = \dfrac{9200}{3}$

Hence the demand equation is

$y = \dfrac{9200}{3} - \dfrac{4x}{3}$

i.e. $3y = 9200 - 4x$

(b)   if $y = 3000$, then

$9000 + 4x = 9200$

$4x = 200,\quad \square\ x = 50$

i.e. 50 watches will be sold when the price is ₦300

(c)   the highest price for a watch is attained when demand is zero

$3y + 4x = 9200$

When $x = 0$,

$3y = 9200$

$y = ₦3,066.67$

Supply Equation

Supply is directly proportional to price i.e quantity supplied increases as price increases and quantity supplied decreases as price decreases.

This implies that the slope of supply curve is positive. If the slope is zero, it means constant price irrespective of supply and if the slope is undefined, it implies constant supply irrespective of price.

If x represents the quantity supplied and y the price, both x and y must be positive.

The supply equation is also of the form

$y = a + bx$, where b is the slope.

The three situations are drawn below
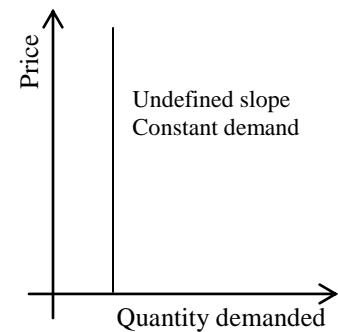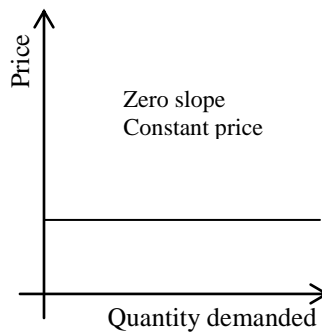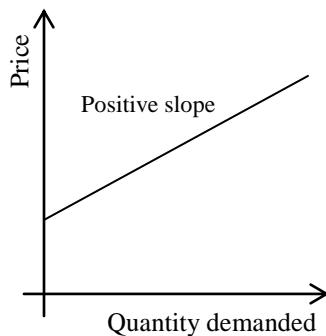
Example 9.18

2000 printers are available when the price is ₦15,000 while 3200 are available when the price is ₦21,000.

(a)   determine the supply equation

(b)   how many printers will be available if the price is ₦30,000?

(c)   if 8000 printers are available, what is the price?

(d)   what is the lowest price for a printer?

Solutions

(a)   Let y be the price and x the quantity supplied, and then the supply equation is

$y = a + bx$

when $y = 15000$, $x = 2000$

i.e. $15000 = a + 2000b$……..(i)

when $y = 21000$, $x = 3200$

i.e. $21000 = a + 3200b$ ……(ii)

equation (ii) $-$ equation (i) gives

$6000 = 1200b$

$\Rightarrow b = 5$

Substitute for b in equation (i)

i.e. $15000 = a + 10000$

$a = 5000$

hence the supply equation is

$y = 5000 + 5x$

i.e. or $y - 5x = 5000$

(b) if $y = 30000$, then

$30000 - 5x = 5000$

$-5x = -25000$

$\Rightarrow x = 5000$

i.e. 5000 printers will be available when the price is ₦30000

(c) if $x = 8000$, then

$y - 40000 = 5000$

$y = 45000$

(d) The lowest price is attained when supply is zero i.e. $y - 5x = 5000$ when $x = 0$,

$y = 5000$ hence the lowest price is ₦5000

Market Equilibrium

Market equilibrium occurs at the point where the quantity of a commodity demanded is equal to the quantity supplied.

The equilibrium price and quantity are obtained at that point. As mentioned earlier, neither *x* nor y can be negative hence equilibrium is only meaningful when the curves intersect in the first quadrant i.e. where both *x* and y are positive.

To find the equilibrium price or quantity, we solve the supply and demand equations simultaneously or graphically by drawing the supply and demand curves on the same graph. The point of intersection of the two curves is the equilibrium point.
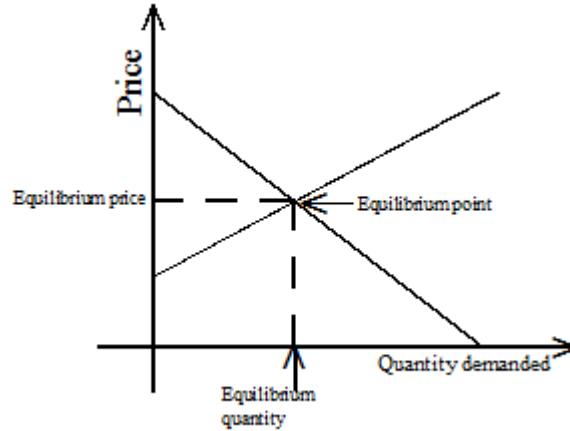
Example 9.19

The demand and supply equations for a commodity are respectively given by

$$20y + 3x = 335$$

$$15y - 2x = 60$$

Determine the point of equilibrium by solving the equations

(a) simultaneously

(b) using graphical method

Solutions

(a) $20y + 3x = 335$ ....(i)

      $15y - 2x = 60$.........(ii)

      (i)   × 2 gives $40y + 6x = 670$ ....(iii)

      (ii)   × 3 gives $45y - 6x = 180$ ....(iv)

      (iii) + (iv) gives $85y = 850$

                                     $y = 10$

      Substitute for y in equation (ii) to get

      $150 - 2x = 60$

         $-2x = -90$

           $x = 45$

      i.e. the equilibrium price is 10 and the equilibrium quantity is 45.

(b)   Draw the lines of the two equations on the same axes.

      For $20y + 3x = 335$,

      when $x = 0$, $y = 16.75$ $\Rightarrow (0, 16.75)$

      when $y = 0$, $x = 111.67$ $\Rightarrow (111.67, 0)$

      For $15y - 2x = 60$,

      when $x = 0$, $y = 4$ $\Rightarrow (0, 4)$

when $y = 6$, $x = 15$ $\Rightarrow (15, 6)$
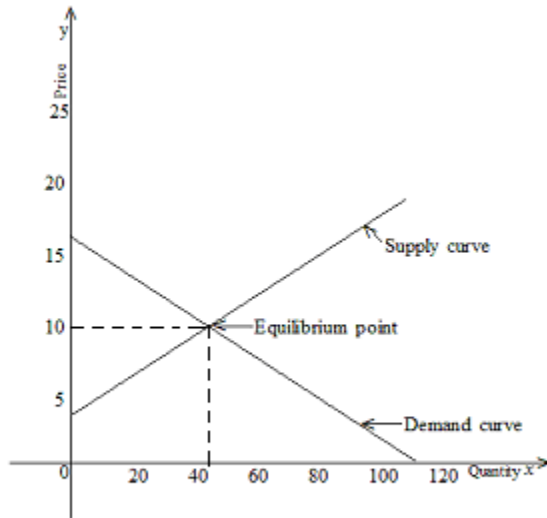
As could be seen from the graph, the equilibrium is at the point (44, 10) i.e. $x = 44$, $y = 10$

i.e. the equilibrium price is 10 while the equilibrium quantity is 44.

As mentioned earlier, graphical solutions are estimates and this accounts for the slight difference in the value of the quantity obtained.

9.6 Summary

A function has been described as a mathematical expression involving two or more variables. Linear, quadratic and exponential functions were discussed. Also discussed are different types of equations and the methods of solving them. Inequalities and their graphical solutions were discussed as well. Finally, comprehensive applications of these concepts to Business and Economics were treated.

9.7 Multiple-Choice and Short-Answer Questions

1. If $f(x) = 3x^3 - 4x^2 + 2x + 555$, then $f(-5)$ is
   A. 820
   B. 290
   C. 270
   D. 70
   E. 90

ccvi

2. The condition for the expression $ax^2 + bx + c$ to be a quadratic expression is
   A. $b \neq 0$
   B. $a \neq 0 \ \& \ b \neq 0$
   C. $b \neq 0 \ \& \ c \neq 0$
   D. $c \neq 0$
   E. $a \neq 0$

3. The cost function of a company is $C(x) = 500 + 2x^2 + 5x$ while the revenue function is $R(x) = 3x^2 - 10x$, where x is the number of items produced and sold. The profit that will accrue from 500 items is
   A. 24,200
   B. 242,000
   C. 124,000
   D. 224,000
   E. 257,000

4. If the demand and supply equations for a commodity are respectively given by
   $30y + 7x = 800$
   $14y - 5x = -40$
   the equilibrium point is then
   A. (15, 50)
   B. (50, 20)
   C. (50, 15)
   D. (20, 50)
   E. (25, 25)

5. On the graph of $y = a + bx$, a is the .... on the y-axis while b is the ....

6. An exponential function is a function which has a .... base and a .... exponent

7. The sign of a variable or a constant in an equation changes when it crosses the...................

8. When two expressions are not equal, then one has to be .... or ....than the other.

9. If for a business, $R(x) - C(x) < 0$, then the business is said to be running at a ....

10. The demand for a commodity is .... proportional while supply is .... proportional to the price of the commodity

Answers to Chapter Nine

1. $f(x) = 3x^3 - 4x^2 + 2x + 555$
    $f(-5) = 3(-5)^3 - 4(-5)^2 + 2(-5) + 555$
    $= -375 - 100 - 10 + 555$
    $= 70$ (D)

2. $a \neq 0$ (E)
    because if $a = 0$, the quadratic term will vanish.

3. Profit $= R(x) - C(x)$
    $= 3x^2 - 10x - 500 - 2x^2 - 5x$
    $= x^2 - 15x - 500$
    if $x = 500$, then profit $= (500)^2 - 15(500) - 500$
    $= 242,000$ (B)

4. $30y + 7x = 800$........... (i)
    $14y - 5x = -40$.............(ii)
    equation (i) x 5 gives
    $150y + 35x = 4000$.......(iii)
    equation (ii) $\times$ 7 gives
    $98y - 35x = -280$...... (iv)
    equation (iii) + (iv) gives
    $248y = 3720$
    $y = 15$
    Substitute for y in equation (i)
    $30(15) + 7x = 800$
    $7x = 350$
    $x = 50$
    So the equilibrium point (x, y) is (50, 15) (C)

5. Intercept, slope of the line.

6. Constant, variable

7. Equality sign

8. Less, greater

9. Loss

10. Inversely, directly.

CHAPTER 10

MATHEMATICS OF FINANCE

OBJECTIVE

At the end of this chapter, students and readers must be able to:

- define sequences and series

- identify Arithmetic Progression (AP)

- identify Geometric Progression (GP)

- find the nth term and sum of the first n terms of AP and G P

- understand simple interest and compound interest

- define and understand Annuities

- calculate Present Values (PV) and Net Present Values (NPV)

- apply all of the above to economic and business problems.

10.1 Sequences and series

A sequence is a set of numbers **that follow a definite pattern.** e.g

(a) 7,12,17,22 ... the pattern is that the succeeding term is the preceding time plus 5

(b) 256, 64,16,4: the pattern is that the succeeding term is the preceding term divided by 4.

If the terms in a sequence are connected by plus or minus signs, we have what is called a series.

e.g. the series corresponding to (a) above is $7 + 12 + 17 + 22 + ..,$.

10.1.1 Arithmetic progression (A.P)

A sequence in which each term increases or decreases by a constant number is said to be an A.P

The constant number is called the common difference.

The first term is usually represented by a while the common difference is represented by d

An A. P. is generally of the form

a, a+d, (a + d) + d, {(a + d) + d} + d, ...

i.e. a, a + d, a + 2d, a + 3d, ...

Note that

2nd term is a + d = a + (2 -1) d

3rd term is a+2d = a+(3 -1) d

4th term is a + 3d = a + (4-1) d and so on

following this pattern, the nth term will be a+(n-l)d

The sum of the first n terms of an A. P. is given by S_n = n {2a + (n - l)d} or n (a + *l*) d

$\qquad$ 2 $\qquad\qquad\qquad$ 2

Here *l* is the last term

Example 10.1

(a) $\qquad$ Find the 12th term of the A. P: 7,13,19,25, ...

(b) $\qquad$ Find the difference between the 8th and 52nd terms of the A. P: 210, 205, 200, ..

(c) $\qquad$ Find the sum of the first 10 terms of the A. Pin (a) and (b)

(d) $\qquad$ How many terms of the series 15 + 18 + 21 + ... will be needed to obtain a sum of 870?

Solutions

(a) $\qquad$ 7, 13, 19,25,

$\qquad$ d = 13 - 17 or 25 - 19 = 6

$\qquad$ a=7,n=12

$\qquad$ nth term $\qquad$ = $\qquad$ a + (n -1)d

$\qquad$ 12th $\qquad$ = $\qquad$ 7 + (12 - 1) 6

$\qquad\qquad\qquad$ = $\qquad$ 73

(b) $\qquad$ 210, 205, 200,

$\qquad$ a = 210, d =-5

$\qquad$ 8th term $\qquad$ = $\qquad$ 210 + (8 - 1)(-5)

$\qquad\qquad\qquad$ = $\qquad$ 175

$\qquad$ 52nd term $\qquad$ = $\qquad$ 210 + (52 - 1) (-5)

$\qquad\qquad\qquad$ = $\qquad$ **-45**

Difference $\qquad$ = $\qquad$ 175 - (- 45) = 220

(i) $\qquad$ 7, 13, 19,25, ...

$\qquad$ a = 7, d=6,n=1O

$\qquad$ S_n = n{2a+(n-l)d}

$$S_{10} = \frac{10}{2}\{2(7) + (10 - 1)6\}$$

$$= 340$$

(ii) 210, 205, 200, ...
a=210, d=-5, n=10
$$S_{10} = \frac{10}{2}\{2(210) + (10 - 1)-5\}$$
$$=1875$$

(d) 15 +18+21+…..
a = 15, d = 3, S_n = 870, n = ?
$$S_n = \frac{n}{2}\{2a + (n-1)\,d\}$$
$$870 = \frac{n}{2}\{30 + (n-1)3\}$$

i.e 1740 = n (27 + 3n)
i.e $3n^2 + 27n - 1740 = 0$
$n^2 + 9n - 580 = 0$
$n^2 + 29n - 20n - 580 = 0$
i.e n (n + 29) − 20 (n + 20) = 0

.: n = 20 or 29

but n cannot be negative since n is number of terms, .: n = 20

Example 10. 2

Mr. Emeka earns N240, 000 per annum with an annual constant increment of N25, 000.

a. How much will his annual salary be in the eight year

b. What will his monthly salary be during the 15th year?

c. If he retires after 30 years, and his gratuity is 250% of his terminal annual salary , how much will he be paid as gratuity?

Solution

(a) a = 240,000, d = 25,000, n = 8

n^{th} term = a + (n-1) d

8th term = 240,000 + (8-1) (25,000)

= N415,000

(b) $a = 240,000$, $d = 25,000$, $n = 15$

   15th term $= 240,000 + (15 - 1)(25,000)$

         $= ₦590,000$

   i.e his annual salary in the 15th year is N590,000

   .: his monthly salary in the 15th year is

   $\dfrac{N590,000}{12} = N49,166.67$

(c) $a = 240,000$, term $d = 25,000$, $n = 30$

   30th term $= 240,000 + (30 - 1)(25,000)$

         $= ₦965,000$

   i.e his annual salary $= ₦965,000$

   .: $\dfrac{250,000 \times 965,000}{100}$
       $= ₦2, 412, 500$

10.1.2 Geometric Progression (G.P.)

A sequence is said to be a G. P. if each of its terms increases or decreases by a constant ratio.

The constant ration is called the common ratio and is usually denoted by r while the first term is denoted by a.

The general form of a G. P.

a, ar, ar^2, ar^3, ar^4, ……..

The pattern is

2nd term $= ar^{2-1}$

3rd term $= ar^{3-1}$

4th term $= ar^{4-1}$

n^{th} term $= ar^{n-1}$

The sum of the first n terms of a GP

$S_n = a + ar + ar^2 + ar^3 + ….. + ar^{n-1}$

$S_n = \dfrac{a(1-r^n)}{1-r}$ $(r<1)$

If $r < 1$, r^n decreases as n increases and in fact tends to zero as n gets very big. In the extreme case where n is close to infinity (∞), r^n is approximately zero.

In such a case, we talk about the sum to infinity of a G.P. given by

$S_n = \underline{a\ (1\text{-}r^n)}$, if n $\rightarrow\infty$, then $r^n = 0$

$\therefore = S\infty = \dfrac{a}{1-r}$, s_n is said to converge to $\dfrac{a}{1-r}$

Example 10.3

    a. Find the 7th term of the G.P 8, 16, 32, …….

    b. Find the 6th term of the G.P 243, 81, 27, ……

    c. Find the sum of the first 15 terms of the G.P in (a) and (b)

    d. Find the sum to infinity for the G.P in (a) and (b)

Solutions

(a) 8, 16, 32, ………..
$r = \dfrac{16}{8}$ or $\dfrac{32}{16} = 2$
$a = 8$, $n = 7$
n^{th} term $= ar^{a-1}$
$\therefore 7^{th}$ **term** $= 8(2)^{7-1}$
    $=\ 512$

(b) 243, 81, 27
$r = \dfrac{81}{243} = \dfrac{1}{3}$
$a = 243$, $n = 6$
6th term $= 243\{\dfrac{1}{3}\}^{6-1}$
    $=\ 1$

(c.) (i) 8, 16, 32, ………..
    $a = 8$, $r = 2$, $n = 15$
    here r> 1, then
    $S_n = \dfrac{a(r^n - 1)}{r - 1}$

    $S_{15} = \dfrac{8(2^{15} - 1)}{2 - 1}$
    $S_{15} = 262,136$

(ii) **243,81, 27, ..………..**
    $a = 243$, $r = 1/3$, $n = 15$

$$S_{15} = \frac{243(1 - (1/3)^{15})}{1 - (1/3)}$$

$$S_{15} = 364.5$$

(d) (i) 8, 16, 32, ...

a = 8, r = 2

Sα is not responsible since r>1

(i) 243, 81, 27, ...

a = 243, r = 1/3

$$S\infty = \frac{a}{r-1}$$

$$= \frac{243}{1 - 1/3} = 364.5$$

S_{15}=Sα = 364.5 i.e. S_n has actually converged to 364.5 from n = 15

Example 10.4

a.　**A job is estimated to take 12 days. If the overhead costs were N5,000 for the first day, N7,500 for the second day, N11,250 for the third day and so on, how much will the total overhead costs be?**

b.　**The swamp in a village is highly infested by tsetse flies. After a lot of appeals by the head of the village, the state government took necessary actions and the population of the flies starts to decrease at a constant rate of 20% per annum. If at the initial stage, there were 104 million flies in the swamp, how long will it take to reduce the population of flies to 30m?**

Solutions

(a)　5000, 7500, 11250, ...

$$r = \frac{7500}{5000} = \frac{11250}{7500} = 1.5$$

:. the series is a G.P., then

a = 5000, n = 12 and common ratio = r = 1.5

$$S_n = \frac{a(r-1)}{r-1} \text{ since r>1}$$

$$= \frac{5000(1.5^{12} - 1)}{1.5 - 1}$$

$$= =1, 287,463.38$$

(b)   Decrease at the rate of 20% per annum means $r = 80\% = 0.8$

$a = 104m$, $ar^{n-1} = 30$, $n = ?$

$30 = ar^{n-1}$

$\quad = 104(0.8)^{n-1}$

$\dfrac{30}{104} = (0.8)^{n-1}$

i.e $(n-1)\log 0.8 = \log 0.28846$

$n - 1 = \dfrac{\log 0.28846}{\log 0.8}$

i.e $n = 6.57$ years

Example 10.5

An investment is estimated to grow at the rate of 15% per annum. If the worth of the investment now is N850,000

a.   What will its worth be in the 6th year?

b.   What is the percentage increase in its worth after 6 years?

c.   In what year will the investment worth N3.28m?

Solutions

(a)   growth of 15% per annum means

$r = I + 15/100$

$= 1.15$

$a = 850, 000$, $n = 6$

n^{th} term $= ar^{n-1}$

6th term $= 850,000(1.15)^5$

$= N1,709,653.61$

(b)   %age increase after 6 years is

$$\dfrac{1,709,653.61 - 850,000}{850,000} \times 100$$

$= 101.14\%$

(c) n^{th} term = ₦3.28m, a = 850000, r = 1.15, n = ?

$3,280,000 = 850,000 (1.15)^{n-1}$

(n-1) log 1.15 = log 3.8588

n-1 = log $\underline{3.8555}$

$\qquad\qquad$ log 1.15

$\qquad$ = 9.66

i.e. n = 10.66 years

Example 10.6

GODSLOVE Enterprises purchased a machine for N930,000 which is expected to last for 25 years.

If the machine has a scrap value of N65,000,

   a.   how much should be provided in each year if depreciation is on the straight line method?

   b.   what depreciation rate will be required if depreciation is calculated on the reducing balance method?

Solutions

Note: For this type of problem, number of terms is 1 more that the number of years because the cost is the value

at the beginning of the first year and the scrap value is at the end of the final year.

So, a = 930,000, n = 26, scrap value = 65000

   a.   This is an A. P.

$\qquad$ 65000 = a + (n- 1) d

$\qquad\qquad$ = 930000 + (26 – 1) d

$\qquad\qquad$ = 930000 + 25d

$\qquad$ ∴ d = $\underline{65000 – 930000}$

$\qquad\qquad\qquad$ 25

$\qquad\qquad$ = N34,600

   b.   65000 = ar^{n-1}

$\qquad\qquad$ = 930000 r^{25}

$\qquad$ i.e r^{25} = $\underline{65000}$

$\qquad\qquad\qquad$ 930000

$\qquad$ R= $25\sqrt{\dfrac{65000}{930000}}$

$\qquad$ = 0.90

NOTE: This is a work in progress. All topics in the syllabus are covered but editing for necessary corrections is in progress.

Thanks.

∴ Reducing balance depreciation rate is

$$1 - 0.90 = 0.1$$

i.e. 10%

10.2 SIMPLE INTEREST

Simple Interest is the interest that will accrue on the principal (original money invested/borrowed) for the period for which the money is invested/borrowed.

If P is the principal, r% is the rate of interest and n is the number of periods, then the simple interest is given

by

$$I = P. r. n$$

n can be in years, quarters, months etc.

the amount An at the end of the nth period is

$$A_n = P + I$$

$$= P + Prn$$

$$= P (1 + r.n)$$

Example 10.7

(a)

(i)   what is the interest that will accrue on N25000 at 12% simple interest at the end of 15 years?

(ii)   how much will it amount to at the end of this period?

(b)   How long will it take a money to triple itself at 9.5% simple interest?

Solutions

(a)   (i) $I = Pr.n$
P = *25000*, r = 0.12, n = 15
I = 25000 (0.12) *(15)*
= N45000

(ii)   Amount $= A_{15} = P + I = 25000 + 4500$
= N70,000

If the interest is not required, then

$A_{15} = P (1 + r. n)$

$\quad = 25000 \{1 + (0.12)\ (15)\}$

$\quad = N70,000$

a) If P is to triple itself, it becomes 3P, so we have

$\quad A n = P (1 + r.n)$

$\quad$ i.e. $\quad 3P = P(1 + r.n),$

$\qquad r = 0.095$

$\qquad 3P = P (1 + 0.095n)$

$\qquad 3 = 1 + 0.095n$

$\qquad n = \dfrac{2}{0.095}$

$\qquad = 21.05$ years

Note from $1 = Pr.n.$

$\qquad P = \dfrac{1}{r.n}$

$\qquad r = \dfrac{1}{P.n}$

$\qquad n = \dfrac{1}{P.\,r}$

The Present Value (PV) of any money is the current worth of its future amount based on prevailing conditions.

Example 10.8

Mohammed wants to purchase a house for ₦ 1.80m in 5 years' time Interest rate remains constant at 10% per annum computed by simple interest method. How much should he invest now?

Solution

$A n = P (1 + r.n)$

$A n = N1.80m, r = 0.10, n = 5, P = ?$

$1.80m = P \{1 + (0.10)\ (5)\}$

$\qquad = P(1.5)$

$.. \quad P = \dfrac{1.80m}{1.5}$

$\qquad ₦1.20m$

Note (a) ₦ 1.20m is the present value of ₦ 1.80m under the prevailing conditions.

(b) The question is the same as: What is the present value of ₦ 1.80m at 10% Simple Interest rate over 5 years?

Generally, the present value (PV) of a future amount (A_n) at r % simple interest for n periods is obtained as follows:

$A_n = P(1 + r.n)$

$$\therefore P = \frac{A_n}{1 + r.n}$$

for the last example, $A_n = 1.80m$, r =: 0.10, n = 5
so P $= \dfrac{1.80m}{1+(0.10)(5)}$

$= N1.20m$

10.3 Compound Interest

Compound interest can be referred to as multi-stage single period simple interest.

It is the type of interest commonly used in banks and financial institution. Simple interest on the initial principal for single period added to the principal itself (i.e, amount at the end of the first period) is the new principal for the second period. Amount at the end of the second period is the principal for the third period etc.

Example 10.9

How much will N200, 000 amount to at 8% per annum compound interest over 5 years?

Solution

Year	Principal	Interest I = P.r.n (n=1)	Amount
1	200000	16000	216000
2	216000	17280	233280
3	233280	18662.4	251942.4
4	251942.4	20155.4	272097.8
5	272097.8	21767.8	293865.6

.: required amount is ₦293,865.6

At simple interest

A_5 = 200000 (I + (0.08)5)

= ₦280,000

The accrued amount at simple interest is always less than that at compound interest. The compound interest formula is given by

$$A_n = P (1 + r)^n$$

where A_n **is the accrued amount after the nth period**

P is the Principal

r is the interest rate per period

n is the number of periods

Example 10.10

a) Use the compound interest formula to calculate the amount for example 10.9

b) What compound interest rate will be required to obtain ₦230,000 after 6 years with an initial principal of ₦120,000?

c) How long will it take a sum of money to triple itself at 9.5% compound interest.

d) How much will ₦250,000 amount to in 3 years if interest rate is 12% per Annum compounded quarterly?

Solution

a) $A_n = P (1 + r)^n$

   $P = 200000 , r = 0.08 , n = 5$

   $A_5 = 200000 (1 + 0.08)5$

   = 200000 (1.469328)

   = ₦293,865.60

b) $A_n = P (1 + r)^n$

   $A_n = 230,000, P = 120,000 , n = 6, r = ?$

   So, $230,000 = 120,000 (1 + r)^6$

$$(1 + r)^6 = 1.916667$$

$$1 + r = (1.916667)^{1/6}$$

$$= 1.1145$$

i.e $\quad r = 1.1145 - 1$

$$= 0.1145$$

i.e $\quad = 11.5\%$

c) If P is to triple itself, then we have 3P

$A_n = P(1 + r)^n$

$A_n = 3P, r = 0.095, n = ?$

$3P = P(1 + 0.095)^n$

$(1.095)^n = 3$

$n\log(1.095) = \text{Log } 3$

$n = \dfrac{\log 3}{\log(1.095)}$

$= 12.11$

Recall that the number of years under simple interest is 16.67 years (Ex. 6b). So, the number of years is less under compound interest (obviously?)

d) $A_n = P(1+r)^n$, P=250,000, $r = \dfrac{0.12}{4} = 0.03$, n = (4 X 3)=:12

$A_{12} = 250,000 (1 + 0.03)^{12}$

$= N356,440.22$

Example 10.11

Calculate the sum of money that will need to be invested now at 9% compound interest to yield N320,000 at the end of 8 years.

Solution

$A_n = P(1+r)^n$

$A_n = 320,000, r = 0.09, n = 8, P = ?$

$320,000 = P(1+0.09)^8$

$P = \dfrac{320,000}{(1.09)^8}$

$= N\,160,597.21$

Again, ₦ 160597.21 is the present value of ₦ 320,000 under the given situation.

Generally, from

$$A = p(1+r)^n$$

$$P = -\frac{A_n}{(1+r)}$$

This formula forms the basis for all discounting methods and it is particularly useful as the basis of discounted cash flow (DCF) techniques.

10.4   Annuities

An annuity is a sequence of constant cash flows received or paid.

Some examples are:

\(a) Weekly or monthly wages

(b) Hire-purchase payments

(c) Mortgage payments

10.4.1   Types of Annuity

- An **ordinary annuity** is an annuity paid at the end of the payment periods. This type of annuity is the one that is commonly used.
- A **due annuity** is an annuity paid at the beginning of the payment periods (i.e. in advance)
- A **certain annuity** is an annuity whose term begins and ends on fixed dates.
- A **perpetual annuity** is an annuity that goes on indefinitely.

10.4.2 Sum of an ordinary Annuity (Sinking fund)

The sum(s) of an ordinary annuity with interest rate of r% per annum compounded over n periods is given by

$$S = A + A(1+r) + A(1+r)^2 + A(1+r)^3 + \ldots\ldots\ldots + A(1+r)^{n-1}$$

This is a G.P with first term A and common ratio (1 +r) (see section 10.1 .2).

So, $S = \dfrac{A\{1-(1+r)^n\}}{1-(1+r)}$

$$= \frac{A\{1-(1+r)^n\}}{-r}$$

$$= \frac{A\{(1+r)^n-1\}}{r}$$

where A is the amount paid at the end of each period

Example 10.12

(a) Find the amount of an annuity ofN50,000 per year at 4% interest rate per annum for 7 years.

(b) Calculate the annual amount to be paid over 4 years for a sinking fund of N2,886,555 if the compound interest rate is 7.5% per annum

Solutions

(a) $S = \frac{A[(1 + r)^n - 1]}{r}$

A = 50,000, r = 0.04, n = 7

$S = \frac{50,000 [(1+0.04)^2-1]}{0.04}$

= N394,914.72

b) $S = \frac{A[(1+r)^{n-1}]}{r}$

$S = A[(1+r)^{n-1}]/r$

S = 2886555, r= 0.075, n = 4, A = ?

$2,886,555 = \frac{A[(1 + 0.075)^{7-1}]}{0.075}$

$\therefore A = \frac{(2886555) (0.075)}{(1.075)^{7-1}}$

= N645,339.91

10.4.3 Present Value of an Annuity

The present value of an annuity is the sum of the present values of all periodical payments. Thus, the present value of an annuity *(P)* with compound interest rate of r% per annum for n year is

$$P = \frac{A}{1+r} + \frac{A}{(1+r)^2} + \frac{A}{(1+r)^3} +.....+ \frac{A}{(1+r)^n}$$

This is a G.P with $\dfrac{A}{1+r}$ as the common ratio. Therefore,

$$P = \cfrac{\cfrac{A}{1+r}\left\{1-\left(\cfrac{1}{1+r}\right)^{n}\right\}}{1-\cfrac{1}{1+r}}$$

$$P = \cfrac{\cfrac{A}{1+r}\left\{1-(1+r)^{-n}\right\}}{1-\cfrac{1}{1+r}}$$

$$P = \cfrac{A\left\{1-(1+r)^{-n}\right\}}{r}$$

Example 10.13

Determine the present value of an annuity ofN45,000 for 9 years at 5.5% compounded annually

Solution

$$P = P = \frac{\left[1-(1+r)^{-n}\right]}{r}$$

A = 45000, r = 0.055n n =9
P = $\underline{45000 \; [1 - (1 +0.055)^{-9}]}$
       0.055
    = N312,848.79

10.4.4 Net Present Value (NPV)

The NPV is the sum of the present values (PV) of all future net cash flows of an investment. The net cash flow may be positive or negative.

The NPV can be used to determine the desirability of an investment. If the NPV is positive, the investment is desirable but if it is negative, the investment is not worth it.

$$NPV = A_0 \; \frac{A_1}{1+r} + \frac{A_2}{(1+r)^2} + \frac{A_3}{(1+r)^3} + ... + \frac{A_n}{(1+r)^n}$$

A_0 is the cost of the investment at year 0 and is always recorded as negative (cash outflow) in calculating the NPV

A_1 is the expected net cash flow for year 1

A_2 is the expected net cash flow for year 2

A_n is the expected net cash flow for year n

Generally for most business projects, it is usual to have to outlay (or invest) a sum at the start and then expect to receive income from the project (i.e revenue) at various times in the future. Supposing a company has the opportunity to buy a new machine now for ₦ 2.5m. It intends to use the machine to manufacture a large order for which it will receive ₦ 1.75m after 2 years and ₦ 1.25m after 3 years.

A project like this can be assessed by assuming a discount rate and then calculating the present values of all the flows of money, in and out. The total of the present values of money in (revenue0, less the total of the present values of the money out (costs) is the net present value (NPV) as discussed above.

Example 10.14

A project is presently estimated to cost ₦ 1.1 m. The net cash flows of the project for the first 4 years are estimated respectively as ₦ 225,000, ₦ 475,000, ₦ 655,000 and ₦ 300,000. If the discount rate is 12%.

(a) Calculate the NPV for the project. Is the project desirable?

(b) If ₦ 95,000 and ₦ 125,000 were spent on the project during the second year and fourth year respectively, will the project be desirable?

Solutions

(a) The usual set up is

$r = 0.12$

Year	Net Cash Flow (A) ₦	Discounting factor $\dfrac{1}{(1+r)^n}$	$PV= \dfrac{A}{(1+r)^n}$ (₦)
0	-1.1m	$\dfrac{1}{(1+0.12)^0} = 1$	-1.1m
1	225000	$\dfrac{1}{(1+0.12)} = 0.8929$	200,902.5
2	475000	$\dfrac{1}{(1+0.12)^2} = 0.7972$	378770
3	655000	$\dfrac{1}{(1+0.12)^3} = 0.7118$	466229

4	300000	$\dfrac{1}{(1+0.12)^4} = 0.6355$	190650
		NPV	₦136,551.50

The tables are available for the discounting factors.

NPV $=$N1,100,00+N200,902.5+N378,670+N466,229+190,650=N136,551.50

Since the NPV is positive, the project is desirable.

b)   The new net cash flow for 2nd year $=$ N475,000 – N95,000 $=$ N380,000

and for 4th year $=$ N300,000 – N125,000 $=$ N175,000

So, we now have

Year	Net Cash flow	Discounting factor	PV
0	-1.1m	1	-1.1m
1	225000	0.8929	200902.5
2	380000	0.7972	302936
3	655000	0.7118	466229
4	175000	0.6355	111212.5
		NPV	-18720

Since the NPV is negative, the project is not desirable.

Example 10.15

Supposing a company intends to buy a new machine now for N2.5m. It intends to use the machine to manufacture a large order for which it will receive N1.75m after 2 years and N1.25m after 3 years.

The table below calculates the net present value of the machine project assuming a discount rate of 5%.

End of year	₦	Present Value	₦
0	-2,500,000		-2,500,000
2	+1,750,000	$\dfrac{1,750,000}{1.05^2}$	1.587,301.6
3	+1,250,000	$\dfrac{1,250,000}{1.05^3}$	1,079,797
		.: Net Present Value	167,098.6

As the net present value (NPV) is positive , we then conclude that purchasing a new machine is

NOTE*:*   This is a work in progress. All topics in the syllabus are covered but editing for necessary corrections is in progress.

Thanks.

worthwhile

Example 10.16

An architectural outfit has an opportunity of buying a new office complex in Lekki selling for N75m. It will cost N25m to refurbish it which will be payable at the end of year 1. The company expects to be able to lease it out for N125m in 3 years time. Determine the net present value of this investment if a discount rate of 6% is allowed.

Solution

The corresponding table at a 6% discount rate is given as follows:

End of year	₦	Present Value	₦
0	-75m		-75,000,000
1	-25m	$\dfrac{-25,000,000}{1.06}$	-23,584,905.7
2	+125m	$\dfrac{125,000,000}{1.06^3}$	104,952,410.4
		Net Present Value	6,367,504.70

.: The investment is profitable because the NPV is positive.

10.4.5 The Concept of Internal Rate of Return

In example 10.15 above, a discount rate of 5% was used for the company that wanted to buy a new machine, we obtained a positive NPV of ₦167,098.60. If the discount rate were larger, the present value of the future revenue would reduce, which for this project would reduce the NPV.

Let's consider a discount rate of 7%, the calculations are tabulated as follows:

End of year	₦	Present Value	₦
0	-2.5m		-2,500,000
2	+1.75m	$\dfrac{1,750,000}{1.07^2}$	1,528,517.80
3	+1.25m	$\dfrac{1,250,000}{1.07^3}$	1,020,372.30
		Net Present Value	48,890.10

You can use that though the NPV is still positive, but it is far less that the case of 5% discount rate.

Further calculations show that when the discount rate is 8%, the NPV is obtained as follows:

End of year	₦	Present Value	₦
0	-2.5m		-2,500,000
2	+1.75m	$\dfrac{1,750,000}{1.08^2}$	1,500,342.90
3	+1.25m	$\dfrac{1,250,000}{1.08^3}$	992,290.30
		Net Present Value	-7,366.80

This reveals that the project breaks even, that is, the NPV is zero at a discount rate between 7% and 8%.

Therefore, the discount rate at which a project has a net value of zero is called Internal rate of Return (IRR). Suppose the discount rate is 7.8%, this means that the project is equivalent (as far as the investor is concerned), to an interest rate of 7.8% per year. So, if the investor can invest her money elsewhere and obtain a higher rate than 7.8%, or needs to pay more than 7.8% to borrow money to finance the project, then the project is not worthwhile and should not be invested on.

Example 10.17

If a financial group can make an investment of N85m now and receives N100m in 2 years' time, estimate the internal rate of return.

Solution

From the above analysis, the NPV of this project is

$$NPV = -85,000,000 + \frac{100,000,000}{(1+i)^2}$$

where i is the discount rate

since the IRR is the value of I which makes the NPV zero, then i can be calculated from

$$-85,000,000 + \frac{100,000,000}{(1+i)^2} = 0$$

$$\Rightarrow \quad \frac{100,000,000}{(1+i)^2} = \frac{100,000,000}{85,000,000}$$

$$\Rightarrow \quad (1+i)^2 = \frac{100,000,000}{85,000,000}$$

$$\Rightarrow \quad 1+i = \sqrt{\frac{100}{85}} \Rightarrow i = \sqrt{\frac{100}{85}} - 1 = 0.084652$$

giving the required IRR to be 0.084652

10.5   Summary

A sequence has been defined as a set of numbers that follow a definite pattern. Arithmetic Progression (A.P) and Geometric Progression (G.P) have been treated. Also treated are Simple and Compound Interests, Annuities and Net Present Value (NPV) when NPV is positive, it means the project is worthwhile otherwise it should not be invested on.

A special case is considered when the Net Present Value is zero. The discount rate that forces NPV to be zero is called the Internal Rate of Return (IRR). The relevance of all these concepts to Business and Economics was highlighted and applied.

10.6   Multiple-Choice and Short-Answer Questions

1)   If the 4th term of a geometric progress is $-$ 108 and the 6th term is -972, then the common ration is

A. +3

B. -3

C. 3

D. 9

E. -9

2)   Determine the present value of N450,000 in 3 years' time if the discount rate is 6% compounded annually.

A.   ₦377,826.68

B.   ₦386,772.68

C.   ₦268,386.68

D.   ₦338,277.68

    E.   ₦287,768.68

3)   The lifespan of a machine which costs N2.5m is 15years and has a scrap value of N150,000. If depreciation is on the straight line method, the amount of money to be provided for each year is

    A.   ₦ 157,666.67

    B.   ₦ 167,666.67

    C.   ₦ 156,777.67

    D.   ₦156,666.67

    E.   ₦167,777.67

4)   Mr. Nokoe earns ₦ 50,000 per month with an annual increment of 5%. What will his annual salary be in the 4th year?

    A.   ₦ 561,500

    B.   ₦ 651,500

    C.   ₦ 515,500

    D.   ₦ 615.500

    E.   ₦ 661,500

5)   If the discount rate of 6% in Question 2 is compounded six-monthly, the present value will be ……………

6)   A project is said to be desirable if the net present value is ……………………

7)   The present value of N2.8m at 15% simple interest rate over six years is ………… ………

8)   The time that a sum of money will take to triple itself Using simple interest is ………… than that of compound interest.

9)   An annuity is a sequence of …………………………cash flows …………or paid.

10)   The discount rate that occurs when the net present value is zero is known as…………………

Answers to Chapter Ten

1) 4th term is ar^3

     i.e. $ar^3 = -108$

     6th term is ar^5

i.e. $ar^5 = -972$

$\dfrac{ar^5}{ar^3} = \dfrac{-972}{-108}$

i.e. $r^2 = 9$

$r = +3$

but since $ar^3 = -108$, $r = -3$ (B)

2) $N\left(\dfrac{450,000}{1.06^2}\right) = 377,828.68$ (A)

3) $a = 2.5$, $n = 16$, scrap value $= 150,000$

i.e. $150,000 = a + (n - 1)d$

$\qquad\qquad = 2,500,000 + 15d$

$\qquad d = ₦156,666.67$ (D)

4) $N50,000$ per month $=> N600,000$ per annum

$a = 600,000$, $n = 4$, $r = 1.05$

$\therefore T4 = ar^3$

$\qquad = 600,000\ (1.05)^3$

$\qquad = N661,500$ (E)

5) PV $= \left(\dfrac{450,000}{1.03^6}\right) = N317,232.24$

6)   Positive

7) $A_n = P(1 + r.n)$

$2.8m = P\{1 + (0.15)(6)\}$

$= 1.9P$

$\therefore$ $P = \dfrac{2.8m}{1.9}$

$= N1.47m$

8)   Less

9)   Constant, received

10)   Internal Rate of Return (IRR)

NOTE: This is a work in progress. All topics in the syllabus are covered but editing for necessary corrections is in progress.
Thanks.

CHAPTER 11

DIFFERENTIAL AND INTEGRAL CALCULUS

At the end of this chapter, the students and readers should be able to

- find the first and second derivatives of functions

- find the maximum and minimum points for functions

- understand Marginal functions

- determine elasticity of demand

- find the indefinite integrals of functions

- evaluate the definite integrals for functions.

- find totals from marginals

- apply all the above to economic and business problems.

11.1   Differentiation

The gradient or slope of a line is constant. i.e. no matter what points are used, we obtain the same result. It is the increase in y divided by the increase in x (see section 9.2.1.1) The gradient of a curve at a point is the gradient of the tangent to the curve at that point i.e. different points will result in different gradients hence the gradient of a curve is not constant.

Let us consider the simplest quadratic curve,

$y = x^2$ shown below

X	−2	−1	0	1	2	3	4
y	4	1	0	1	4	9	16
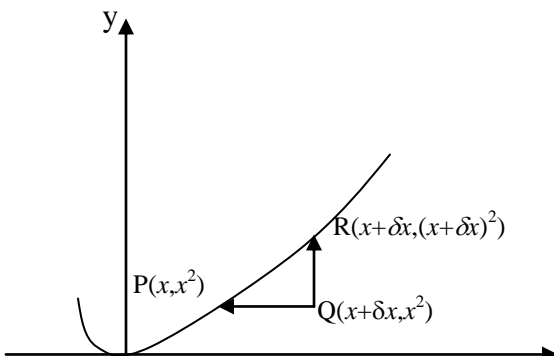
ccxxxiii

x

Let the tangent be drawn at point P (as shown). The gradient of the tangent is

$$\frac{RQ}{PQ} = \frac{\text{increase in } y}{\text{increase in } x}$$

$$= \frac{\delta y}{\delta x}$$

$$= \frac{(x + \delta x)^2 - x^2}{x + \delta x - x}$$

$$= \frac{x^2 + 2x\delta x + (\delta x)^2 - x^2}{\delta x}$$

$$= 2x + \delta x$$

In the limit as $\delta x \to 0$

$$\frac{\delta y}{\delta x} \to \frac{dy}{dx} \text{ , thus}$$

$$\frac{dy}{dx} = 2x$$

i.e. the gradient $= 2x$

The gradient is also referred to as the rate of change.

$\dfrac{dy}{dx}$ is in fact called the derivative of y with respect to x

$\dfrac{dy}{dx}$ can be written as y′ or $\dfrac{df(x)}{dx}$ or f′(x)

11.2 Basic Rule for Differentiation

Generally, if y= x^n, then $\dfrac{dy}{dx} = nx^{n-1}$

i.e. multiply x by its original index and raise x to one less than its original index.

Note:

Let c & k be constants

(i)   if $y = cx^n$, $\dfrac{dy}{dx} = nx^{n-1}$

(ii) $\dfrac{dc}{dx} = 0$, since $c = cx^0$, $\dfrac{dc}{dx} = 0.cx^{0-1} = 0$

(iii)   if $y = x^n + c$, $\dfrac{dy}{dx} = nx^{n-1}$

(iv)   if $y = cx^n + kx^m + x^r$
$\dfrac{dy}{dx} = ncx^{n-1} + mkx^{m-1} + rx^{r-1}$

i.e. the derivative of sum of functions is the sum of the derivative of each of the functions.

(v)   if $y = (cx^n)(kx^m)$

$\qquad = ck\,x^{n+m}$

$\dfrac{dy}{dx} = ck(n+m)x^{n+m-1}$

i.e. if y is the product of two functions, multiply out the functions and then differentiate or you can apply the product rule. This will be discussed later.

(vi)   if $y = (a+bx)^n$, $\dfrac{dy}{dx} = nb(a+bx)^{n-1}$

Note that this is an example of function of a function i.e.

Let $u = a + bx$, then $y = u^n$, $du/dx = b$, and $dy/du = nu^{n-1}$

$dy/dx = (dy/du).(du/dx) = (nu^{n-1})b = nbu^{n-1} = nb\,(a + bx)^{n-1}$

(vii)   if $y = e^{cx}$, $\dfrac{dy}{dx} = ce^{cx}$

Example 11.1

Find the derivative of each of the following functions

(a) $2x^2 + 5x$

(b) $4x^3 - 7x^2 + 6x - 14$

(c) $(5x+2)(4x-9)$

(d) $3e^{4x} - 8x^6$,

(e) $(5x + 13)^7$

(f) $\dfrac{1}{(6x - 20)^4}$

Solutions

(a) Let $y = 2x^2 + 5x$

$$\frac{dy}{dx} = (2)2x^{2-1} + (1)5x^{1-1}$$

$$= 4x + 5$$

(b) $y = 4x^3 - 7x^2 + 6x - 14$

$$\frac{dy}{dx} = (3)4x^{3-1} - (2)7x^{2-1} + (1)6x^{1-1} - (0)14x^{0-1}$$

$$= 12x^2 - 14x + 6$$

(c) let $y = (5x + 2)(4x - 9)$

$$= 20x^2 - 37x - 18$$

$$\frac{dy}{dx} = 40x - 37$$

(d) let $y = 3e^{4x} - 8x^6$

$$\frac{dy}{dx} = (4)3e^{4x} - (6)8x^{6-1}$$

$$= 12e^{4x} - 48x^5$$

(e) let $y = (5x + 13)^7$

$$\frac{dy}{dx} = (7)(5)(5x + 13)^{7-1} = 35(5x + 13)^6$$

or, let $u = 5x + 13$

$$\therefore \frac{du}{dx} = 5$$

$$\therefore y = u^7 \implies \frac{dy}{du} = 7u^6 = 7(5x + 13)^6$$

$$\therefore \frac{dy}{dx} = \frac{dy}{du} \cdot \frac{du}{dx} = 7(5x+13)^6 \cdot 5 = 35(5x+13)^6$$

(f) let $y = \dfrac{1}{(6x-20)^4} = (6x-20)^{-4}$

$$\frac{dy}{dx} = (-4)6(6x-20)^{-4-1}$$

$$= -24(6x-20)^{-5}$$

or let $u = 6x - 20 \Rightarrow \dfrac{du}{dx} = 6$

$\therefore y = u^{-4} \Rightarrow \dfrac{dy}{du} = -4u^{-5} = -4(6x - 20)^{-5}$

$\therefore \dfrac{dy}{dx} = \dfrac{dy}{du} \cdot \dfrac{du}{dx} = -4(6x - 20)^{-5}(6) = -24(6x - 20)^{-5}$

11.2.1 The Second Derivative

If $y = f(x)$, then $\dfrac{dy}{dx} = \dfrac{df(x)}{dx}$

If $\dfrac{dy}{dx}$ is differentiated again, we have the second derivative

i.e. $\dfrac{d}{dx}\left(\dfrac{dy}{dx}\right) = \dfrac{d^2y}{dx^2}$ or $\dfrac{d^2[f(x)]}{dx^2}$, (read as d two y by d x squared) or (the squared y d x squared). It can also be written as y'' or $f''(x)$.

Example 11.2

Find the second derivative of each of the following functions:

(a) $16x^2 - 5x$

(b) $7x^3 + 2x^2 + 3x - 21$

(c) $(x+4)(11x+1)(3x-2)$

Solutions

(a)   Let $y = 16x^2 - 5x$

$\dfrac{dy}{dx} = 32x - 5$

$\therefore \dfrac{d^2y}{dx^2} = 32$

(b)   Let $y = 7x^3 + 2x^2 + 3x - 21$

$\dfrac{dy}{dx} = 21x^2 + 4x + 3$

$\therefore \dfrac{d^2y}{dx^2} = 42x + 4$

(c) $y = (x + 4)(11x + 1)(3x+2)$

$= (11x^2 + 45x + 4)(3x-2)$

$= 33x^3 + 113x^2 - 78x - 8$

$\dfrac{dy}{dx} = 99x^2 + 226x - 78$

$$\frac{\mathrm{d}x}{}$$

$$\therefore \frac{\mathrm{d}^2y}{\mathrm{d}x^2} = 198x + 226$$

11.3 Maximum and Minimum points; Marginal Function, Elasticity .

12.3.1 Maximum and Minimum points

If the graph of $y = f(x)$ is drawn, the turning points on the graph could indicate maximum point, minimum point or point of inflexion.

For quadratic curves, we have the following curves;

Maximum point

Minimum point
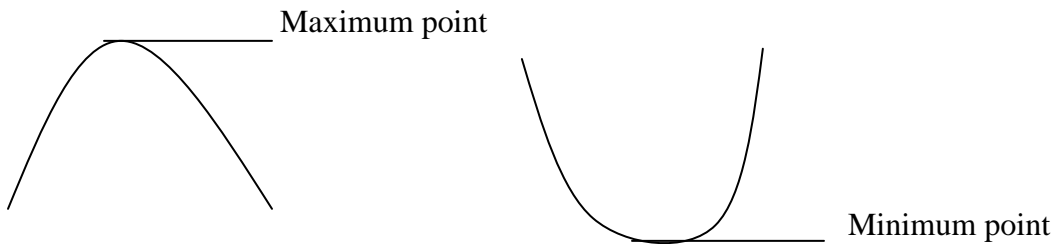
However, these can be obtained without drawing the graph as follows:

Step 1   Find $\frac{\mathrm{d}y}{\mathrm{d}x}$

Step 2   Equate $\frac{\mathrm{d}y}{\mathrm{d}x} = 0$ and solve for x. This gives the turning point(s.)

Step 3   Find $\frac{\mathrm{d}^2y}{\mathrm{d}x^2}$

Step 4 Substitute the value(s) of x obtained in (step 2) above into $\frac{\mathrm{d}^2y}{\mathrm{d}x^2}$ obtained in Step 3

Step 5 If the result in (step 4) is negative, the point is a maximum point but if the result is positive, the point is a minimum point. If the result is zero, it means it is neither a maximum nor a minimum point. It is taken as a point of inflexion.

i.e. * $\frac{\mathrm{d}^2y}{\mathrm{d}x^2} < 0 \Rightarrow$ minimum point

    * $\frac{\mathrm{d}^2y}{\mathrm{d}x^2} > 0 \Rightarrow$ minimum point

    * $\frac{\mathrm{d}^2y}{\mathrm{d}x} = 0 \Rightarrow$ point of inflexion

Example 11.3

Find the maximum or minimum points for each of the following functions

ccxxxviii

(a) $4x^2+9x-2$

(b) $12x - 3x^2 - 7$

(c) $5x-8.5x^2 - 4x + 18$

Solutions

(a)   Let $y = 4x^2 + 9x - 2$

$\dfrac{dy}{dx} = 8x + 9$

$\dfrac{dy}{dx} = 0 \Rightarrow 8x + 9 = 0, x = \dfrac{-9}{8}$

$\dfrac{d^2y}{dx^2} = 8 \Rightarrow x = \dfrac{-9}{8}$ is a minimum point since $\dfrac{d^2y}{dx^2}$ is positive (i.e. > 0)

(b)   Let $y = 12x - 3x^2 - 7$

$\dfrac{dy}{dx} = 12 - 6x$

$\dfrac{dy}{dx} = 0 \Rightarrow 12 - 6x = 0, x = 2$

$\dfrac{d^2y}{dx^2} = -6 \Rightarrow x = 2$ is a maximum point since $\dfrac{d^2y}{dx^2}$ is negative (i.e. < 0)

Example 11.4

The gross annual profit of DUOYEJABS Ventures has been estimated to be

$P(x) = 4x^3 - 10800x^2 + 5400000x$

where x is the number of products made and sold.

Calculate

(a)   The number of products to be made for maximum and or minimum profit.

(b)   The maximum and/or minimum profit.

Solutions

(a) $P(x) = 4x^3 - 10800x^2 + 5400000x$

$\dfrac{dP(x)}{dx} = 12x^2 - 21600x + 540000$

at the turning point, $\dfrac{dP(x)}{dx} = 0$

i.e. $12x^2 - 21600x + 5400000 = 0$

$$x^2 - 1800x + 450000 = 0$$

$$x^2 - 1500x - 300x + 450000 = 0$$

$$x(x - 1500) - 300(x - 1500) = 0$$

$$(x-1500)(x-300) = 0$$

i.e. $x = 1500$ or 300

$$\frac{d^2P(x)}{dx^2} = 24x - 21600$$

when $x = 1500$, $\frac{d^2P(x)}{dx^2} = 24(1500) - 21600$

$$= 14400 \Rightarrow \text{minimum profit.}$$

when $x = 300$, $\frac{d^2P(x)}{dx} = 24(300) - 21600$

$$= -13600 \Rightarrow \text{maximum profit.}$$

(b) $P(x) = 4x^3 - 10800x^2 + 5400000x$

when $x = 1500$

$$P(x) = 4(1500)^3 - 10800(1500)^2 + 5400000(1500)$$

$$= -918 \times 10^7, \quad \text{this is a loss}$$

when $x = 300$,

$$P(x) = 4(300)^3 - 10800(300)^2 + 5400000(300)$$

$$= 756 \times 10^6 \quad \text{is the maximum profit}$$

Though unusual, these results indicate that the profit decreases as the number of products made increases.

11.3.2 Marginal Functions

As before, let the cost function be $C(x)$, revenue function be $R(x)$ and profit function be $P(x)$, where x is the number of items produced and sold, then

$\frac{dC(x)}{dx} = C'(x)$ is the marginal cost function

$\frac{dR(x)}{dx} = R'(x)$ is the marginal revenue function

and $\frac{dP(x)}{dx} = P'(x)$ is the marginal profit function

Now, profit is maximized or minimized when

$\dfrac{dP(x)}{Dx} = P'(x) = 0$ i.e. when marginal profit is zero.

Also, revenue is maximized or minimized when marginal revenue is zero and cost is maximised or minimised when marginal cost is zero.

Recall, $P(x) = R(x) - C(x)$

$$\therefore P'(x) = R'(x) - C'(x)$$

$$P'(x) = 0 \Rightarrow R'(x) - C'(x) = 0 \Rightarrow R'(x) = C'(x)$$

i.e. profit is maximized or minimized when the marginal revenue equals the marginal cost.

Example 11.5

The demand and cost functions of a COAWOS (Nig Ltd) are given as follows:

$$p = 22500 - 3q^2$$
$$C = 5000 + 14400q$$

where p is the price per item

q is the quantity produced and sold

C is the total cost.

You are required to calculate

    (a)   the marginal revenue

    (b)   the quantity and price for maximum revenue

    (c)   the marginal profit function and hence the maximum profit

    (d)   the price for maximum profit

Solutions

(a) $p = 22500 - 3q^2$

    $\therefore$ Revenue $R = p.q$

        $= 22500q - 3q^3$

    Marginal Revenue $= \dfrac{dR}{dq}$, since R is now a function of q

        $= 22{,}500 - 9q^2$

(b)   Maximum Revenue is obtained when

        $\dfrac{dR}{} = 0$ i.e. $22500 - 9q^2 = 0$

$$\therefore 9q^2 = 22500$$

$$q^2 = 2500$$

$$q = \pm 50$$

since q is the quantity produced and sold, it cannot be negative, $\therefore q = 50$

Now $\dfrac{d^2R}{dq^2} = -18q$

When $q = 50$, $d^2R = -900$, this is less than zero which implies maximum revenue.

$\therefore$ Maximum Revenue $= 22500\,(50) - 3(50)^3$

$$= 75,000$$

Price $= \dfrac{\text{Revenue}}{\text{Quantity}}$

$= \dfrac{75000}{50}$

$= 15,000$

(c) $* R = 22500q - 3q^3$

C = 5000 + 14400q

$\therefore$ P = 22500q $- 3q^3 -$ 5000 $-$ 14400q

$= 8100q - 3q^3 - 5000$

$\therefore$ Marginal profit function is

$\dfrac{dP}{dq} = 8100 - 9q^2$

$*$ and $\dfrac{dP}{dq} = 0$ at the turning point

i.e. $8100 - 9q^2 = 0$

$$q^2 = 900$$

$$q = \pm 30$$

As explained above, q cannot be negative, therefore q = 30

$\dfrac{d^2P}{dq^2} = -18q = -540$ when $q = 30 \Rightarrow$ maximum profit

$\therefore$ maximum profit occurs at q = 30

$\therefore$ maximum profit $= P(30) = 8100 \times 30 - 3(30)^3 - 5000 = $ 157,000

(d)   Revenue when q = 30 is

$22500(30) - 3(30)^3 = 594000$

$\therefore$ price $= \dfrac{594000}{30} = 19{,}800$ when profit is maximum

ccxlii

It should be observed that the price for maximum revenue is not the same as the price for maximum profit. This is due to the effect of the cost function.

11.3.3 Elasticity

The point elasticity (or just elasticity) of the function y = f(x) at the point x is the ratio of the relative change in y (i.e. the dependent variable), to the relative change in x (i.e. the independent variable).

Elasticity of y with respect to x is given by

$$E_{y/x} = \frac{\text{relative change in y}}{\text{relative change in } x}$$

$$= \frac{dy/dx}{y/x}$$

$$= \frac{xdy}{ydx}$$

Elasticity is dimensionless i.e. it has no unit, and is represented by η.

11.3.3.1 Elasticity of Demand

Elasticities are most of the time used in measuring the responsiveness of demand or supply to changes in prices or demand.

In other words, a "business set up" will be faced with the problem of deciding whether to increase the price or not. The demand function will always show that an increase in price will cause a decrease in sales (revenue). So, to what extent can we go? Will a very small increase in price lead to increase or decrease in revenue?

Price elasticity of demand helps in answering these questions.

The price elasticity of demand is defined by

$$\eta = \frac{-p}{q} \bigg/ \frac{dp}{dq} \qquad (\text{ - sign is to make } \eta \text{ positive})$$

$$= \left(\frac{-p}{q}\right)\left(\frac{dq}{dp}\right)$$

Note:
(a)   A demand curve is

(i)   elastic if $|\eta| > 1$

(ii)   of unit elasticity if $|\eta| = 1$

(iii)   inelastic if $|\eta| < 1$

(b)   If (i) $\eta > 1$, an increase in price will cause a decrease in revenue.

(ii) $\eta < 1$, an increase in price will cause an increase in revenue.

Example 11.6

The demand function for a certain type of item is

$p = 40 - 0.03\sqrt{q}$

Investigate the effect of price increase when

(a)   3600

(b)   6400

items are demanded

Solutions

$$p = 40 - 0.03\sqrt{q}$$

$$\frac{dp}{dq} = -0.03 \frac{1}{2} q^{-1/2}$$

$$= \frac{-0.03}{2\sqrt{q}}$$

(a)   when q = 3600

$$\frac{dp}{dq} = \frac{-0.03}{2\sqrt{3600}}$$

$$= \frac{-0.03}{120}$$

$$= -0.00025$$

$$p = 40 - 0.03\sqrt{3600}$$

$$= 38.2$$

$$\therefore \eta = \left(\frac{-p}{q}\right)\left(\frac{dq}{dp}\right)$$

$$= \frac{-38.2}{3600(-0.00025)}$$

$$= \frac{38.2}{9}$$

$$= 4.24$$

Since $\eta > 1$, there will be a decrease in revenue if there is an increase in price when 3600 items are demanded.

(b)   when q = 6400

$$\frac{dp}{dq} = \frac{-0.03}{2\sqrt{q}}$$

$$= \frac{-0.03}{160}$$

$$= -0.0001875$$

$$p = 40 - 0.03\sqrt{6400}$$

$$= 37.6$$

$$\therefore \eta = \left(\frac{-p}{q}\right)\left(\frac{dq}{dp}\right)$$

$$= \frac{-37.6}{6400(-0.0001875)}$$

$$= \frac{-37.6}{-1.2}$$

$$= 31.33$$

Since $\eta > 1$, there will be a decrease in revenue if the price is increased when 6400 items are demanded.

11.4   Integration

When y = f(x), the derivative is $\frac{dy}{dx}$, the process of obtaining y back from $\frac{dy}{dx}$ will involve reversing what has been done to y i.e. **anti – differentiation.**

The process of reversing differentiation is called integration. i.e. **Integration is the reverse of differentiation.**

e.g. if $y = x^2$, $\frac{dy}{dx} = 2x$

For differentiation, we multiply by the old or the given index of x and decrease the index by 1.

Reversing this procedure means going from the end to the beginning and doing the opposite.

Increase the index by 1 and divide by the new index

Increasing the index by 1 gives $2x^{1+1} = 2x^2$ and dividing by the new index gives $\frac{2x^2}{2} = x^2$ which is the original function differentiated.

The integral sign is $\int$ (i.e. elongated S)

$$\therefore \int \frac{dy}{dx} \cdot dx = y$$

i.e. $\int 2x dx = x^2$

Check it out:

Differentiation:	decrease	old and	multiply
Integration:	increase	new and	divide

It is clear from the above that integration is the reverse of differentiation.

This integral is incomplete because x^2, $x^2 + 8$, $x^2 + 400$, $x^2 + c$ all have the same derivative $2x$ and the integral will give us just x^2 for each one of them. This is incomplete.

For this reason, there is the need to add a constant of integration i.e.

$\int 2x dx = x^2 + c$ where c is a constant that can take any value

11.5   Rules of Integration

Generally, if $y = x^n$, then
$$\int y dx = \int x^n dx$$
$$= \frac{x^{n+1}}{n+1} + c \quad \textbf{(provided } n \neq -1\textbf{)}$$

Note:

If a, b, c and k are constants, then

(i) $\int ax^n dx = a \int x^n dx = \frac{ax^{n+1}}{n+1} + c \quad (n \neq -1)$

(ii) $\int (ax^n + bx^m + kx^r)\, dx = \frac{ax^{n+1}}{n+1} + \frac{bx^{m+1}}{m+1} + \frac{kx^{r+1}}{r+1} + c$
$$(n \neq -1,\, m \neq -1,\, r \neq -1)$$

i.e. the integral of sum of functions is the sum of the integral of each of the functions.

(iii) $\int (x^m)(x^n)\, dx = \int x^{m+n}\, dx$
$$= \frac{x^{m+n+1}}{m+n+1} + c$$

(iv) $\int (ax + b)^n\, dx = \frac{(ax+b)^{n+1}}{a(n+1)} + c$

(v) $\int e^{kx}\, dx = \dfrac{e^{kx}}{k} + c$

(vi) $\int a\, dx = a x^0\, dx$

$$= \frac{a x^{0+1}}{0+1} + c$$

$$= ax + c$$

Example 11.7

Integrate each of the following functions with respect to x:

(a) $2x^3 - 7x^5 + 17$

(b) $(5x^2 + 6x)(4x - 15)$

(c) $(15x + 9)^4$

(d) $\dfrac{1}{(11x+10)^{12}}$

(e) $12e^{1.5x}$

Solutions

(a) $\int (2x^3 - 7x^5 + 17)\, dx$

$$= \int 2x^3\, dx - \int 7x^5\, dx + \int 17\, dx$$

$$= \frac{2x^{3+1}}{3+1} - \frac{7x^{5+1}}{5+1} + \frac{17x^{0+1}}{0+1} + c$$

$$= \frac{x^4}{2} - \frac{7x^6}{6} + 17x + c$$

(b) $\int (5x^2 + 6x)(4x - 15)\, dx$

$$= \int (20x^3 - 51x^2 - 90x)\, dx$$

$$= \frac{20x^4}{4} - \frac{51x^3}{3} - \frac{90x^2}{2} + c$$

$$= 5x^4 - 17x^3 - 45x^2 + c$$

(c) $(15x + 9)^4\, dx = \dfrac{(15x + 9)^{4+1}}{15\,(4+1)} + c$

$$= \frac{(15x + 9)^5}{75} + c$$

(d) $\dfrac{1}{(11x + 10)^{12}}\, dx = \int (11x + 10)^{-12}\, dx$

ccxlvii

$$= \frac{(11x + 10)^{-12+1}}{11(-12 + 1)} + c$$

$$= \frac{(11x + 10)^{-11}}{-132} + c$$

(e) $\int 8e^{1.5x}dx = \int 8e^{1.5x}dx = 8\int e^{1.5x}dx = \frac{8}{1.5} e^{1.5x} + c$

$$= \frac{16e^{1.5x}}{3} + c$$

11.5.1 Indefinite and Definite Integrals

- **Indefinite integral** is when the integration is done without any limits i.e. $\int f(x)\, dx$ is an indefinite integral.

  In fact, all the integrals we have discussed so far are indefinite integrals.

- **Definite integral** is when the integration is done within given limits

  e.g. $\int_{a}^{b} f(x)dx$, b > a, is a definite integral

  In this case, we talk about evaluating the integral and we obtain a number as our result, by subtracting the value of the integral for **a** (the lower limit) from the value of the integral for **b** (the upper limit)

Example 11.8

Evaluate each of the following integrals:

(a) $\int_{2}^{5} 15x^2 dx$

(b) $\int_{10}^{50} (3x^2 + 4x + 1)dx$

(c) $\int_{-\infty}^{0} 100e^{0.05x}\, dx$

Solutions

(a) $\int_{2}^{5} 15x^2 dx = \left[\frac{15x^{2+1}}{2+1} + c \right]_{2}^{5}$

$$= \left[5x^3 + c\right]_2^5$$

$$= [5(5)^3 + c] - [5(2)^3 + c]$$

$$= 625 + c - 40 - c$$

$$= 585$$

Observe that the constant of integration 'c' cancels out; hence it is not necessary to add the constant of integration when evaluating definite integrals.

(b) $\displaystyle\int_{10}^{50}(3x^2 + 4x + 1)dx = \left[x^3 + 2x^2 + x\right]_{10}^{50}$

$$= [(50)^3 + 2\,(50)^2 + 50] - [(10)^3 + 2\,(10)^2 + 10]$$

$$= 125000 + 5000 + 50 - 1000 - 200 - 10$$

$$= 128840$$

(c) $\displaystyle\int_{-\infty}^{0} 100e^{0.05x}dx = \left[\dfrac{100e^{0.05x}}{0.05}\right]_{-\infty}^{0}$

$$= \dfrac{100e^0}{0.05} - \dfrac{100^{-\infty}}{0.05}$$

$$= 200 \text{ (note that } e^0 = 1, \ e^{-\infty} = 0)$$

11.6   Totals from Marginal

The idea of integration as the reverse of differentiation is used to reverse marginal functions to obtain the total (original) functions.

Recall the marginal functions:

$\dfrac{dR(x)}{dx}$ is the marginal revenue

hence $\displaystyle\int_{a}^{b}\dfrac{dR(x)}{dx}dx = R(b) - R\,(a)$ is the total revenue

- Note that when no quantity is produced , nothing will be sold, hence no revenue. It means when x = 0, R (a) = 0

$\therefore$ $\displaystyle\int\dfrac{dR(x)}{dx}dx = R(x)$ always

also $\underline{dC(x)}$ is the marginal cost
     dx

$$\therefore \int_{m}^{n} \frac{dC(x)}{dx}dx = C(n) - C(m) \text{ is the total cost}$$

Note that $\int \frac{dC(x)}{dx}dx = C(x) + k$ where k is a constant

and $\underline{dP(x)}$ is the marginal profit
     dx

$$\therefore \int_{x_1}^{x_2} \frac{dP(x)}{dx}dx = P(x_2) - P(x_1) \text{ is the total profit}$$

Example 12.9

The marginal revenue function of a production company is given by

$$3x^2 - 5x - 50$$

while the marginal cost function is

$$6x^2 - 500x + 400$$

where x is the number of items produced and sold.

Calculate

(a)   the number of items that will yield maximum or minimum revenue.

(b)   the total revenue if 200 items are produced and sold

(c)   the total profit for the 200 items

Solutions

(a) $\underline{dR(x)} = 3x^2 - 5x - 50 = 0$
       dx

$\longrightarrow (3x + 10)(x - 5) = 0$

      $x = -10$ or 5

x can not be negative since it is the number of items produced and sold. Therefore, $x = 5$

$\dfrac{d^2R}{dx^2} = 6x - 5$

when $x = 5$, $\dfrac{d^2R}{dx^2} = 30 - 5 = 25 > 0 \Rightarrow$ minimum

i.e. minimum revenue is achievable when 5 items are produced.

(b)   total revenue when 200 items are produced is obtained as follows:

$$\int_0^{200} \frac{dR(x)}{dx} dx = \int_0^{200} (3x^2 - 5x - 50)dx$$

$$= \left[x^3 - \frac{5x^2}{2} - 50x \right]_0^{200}$$

- The lower limit MUST be ZERO, because when $x = 0$ (i.e nothing is produced), there will be no revenue

$$= (200)^3 - \frac{5(200)^2}{2} - 50\,(200) - 0$$

$$= 7,890,000$$

(c) $\dfrac{dP(x)}{dx} = \dfrac{dR(x)}{dx} - \dfrac{dC(x)}{dx}$

$= 3x^2 - 5x - 50 - (6x^2 - 500x + 400)$

$= -3x^2 + 495x - 450$

∴ total profit for 200 items is

$$\int_0^{200} \frac{dP(x)}{dx} dx = \int_0^{200} (-3x^2 + 495x - 450)dx$$

$$= \left[-x^3 + \frac{495x^2}{2} - 450x \right]_0^{200}$$

$$= (-200)^3 + \frac{495}{2}\,(200)^2 - 450(200) - 0$$

$$= 1,810,000$$

11.7 Consumers' surplus and producers' surplus

11.7.1 Consumers' surplus

Recall that a demand function represents the quantities of a commodity that will be purchased at various prices.

Some of the time, some consumers may be willing to pay more than the fixed price of a commodity. Such consumers gain when the commodity is purchased, since they payless than what they are willing to pay.

The total consumer gain is known as consumers' surplus.

If the price is denoted by y_0 and the demand by x_0, then

$$\text{Consumers' surplus} = \int_0^{x_0} f(x)dx - x_0 y_0$$

where $y = f(x)$ is the demand function

ccli

Also it could be defined as

$$\text{Consumers' surplus} = \int_{y_0}^{y_1} g(y)dy$$

where $x = g(y)$ is the demand function and y_1 is the value of y when $x = 0$

Either of these expressions could be used to determine the consumers' surplus.

Example 11.10

If the demand function for a commodity is $y = 128 + 5x - 2x^2$, find the consumers' surplus when

(a) $x_o = 5$

(b) $y_o = 40$

Solutions

(a) $y = 128 + 5x - 2x_2$

When $x_o = 5$, $y_o = 128 + 25 - 50 = 103$

$$\text{Consumers' surplus} = \int_0^5 (128 + 5x - 2x^2)dx - x_0 y_0$$

$$= \left[128x + \frac{5}{2}x^2 - \frac{2x^3}{3} \right]_0^5 - 5(103)$$

$$= 640 + \frac{125}{2} - \frac{250}{3} - 0 - 515$$

$$= 104.17$$

(b) $y = 128 + 5x - 2x^2$

when $y_o = 40$, $40 = 128 + 5x - 2x^2$

i.e $2x^2 - 5x - 88 = 0$

i.e $2x^2 - 16x + 11x - 88 = 0$

$\implies$ $2x(x-8) + 11(x - 8) = 0$

$(2x + 11)(x - 8) = 0$ i.e. $x = -11$ or 8

As explained earlier, x cannot be negative, $\therefore x_o = 8$

$$\text{Consumers' surplus} = \int_0^8 (128 + 5x - 2x^2)dx - x_0 y_0$$

$$= \left[128x + \frac{5}{2}x^2 - \frac{2x^3}{3} \right]_0^8 - 8(40)$$

$$= 1024 + 160 - 341.33 - 320$$

$$= 522.67$$

Example 11.11

The demand function for a commodity is $p = \sqrt{25 - q}$, where p represents price and q is the quantity demanded. If $q_o = 9$, show that

$$\int_0^{q_0} f(q)dq - q_0 p_0 = \int_{p_0}^{p_1} g(p)dp$$

where p_1 is the value of p when q=0

Solution

$$p = \sqrt{25 - q}, \quad \text{If } q_0 = 9$$

$$p_0 = \sqrt{25 - q} = \sqrt{25 - 9} = \sqrt{16} = 4$$

$$\therefore \quad \int_0^{q_0} f(q)dq = \int_0^9 (25 - q)^{1/2} dq - 9(4)$$

$$= \left[-\frac{(25 - q)^{3/2}}{3/2} \right]_0^9 - 36$$

$$= -\frac{2}{3}\left[(16)^{3/2} - (25)^{3/2} \right] - 36$$

$$= 4.67$$

Now $p = \sqrt{25 - q}$
$$p^2 = 25 - q$$
$$q = 25 - p^2$$

when q = 0, $p_1 = 5$, $p_0 = 4$ (as before)

$$\therefore \int_{p_0}^{p_1} g(p)dp = \int_4^5 (25 - p^2)dp$$

$$= \left[25p - \frac{p^3}{3} \right]_4^5$$

$$= \left(125 - \frac{125}{3} \right) - \left(100 - \frac{64}{3} \right)$$

$$= 4.67$$

i.e. $\int_0^{q_0} f(q)dq - q_0 p_0 = \int_{p_0}^{p_1} g(p)dp$

11.7.2 Producers' surplus

Recall that the supply function represents the quantities of a commodity that will be supplied at

ccliii

various prices.

At times, some producers may be willing to supply a commodity below its fixed price. When the commodity is supplied, such producers gain since they will supply at a price higher than the fixed price

The total producer gain is referred to as producers' surplus.

Let the fixed price be y_o and the supply be x_o then,

$$\text{Producers' surplus} = x_o y_o - \int_0^{x_0} f(x)dx$$

where the supply function is f(x).

It could also be expressed as

$$\text{Producers' surplus} = \int_{y_1}^{y_0} g(y)dy$$

where the supply function is g(y) and y_1 is the value of y when $x = 0$

Either expression could be used to determine the producers' surplus.

Example 11.12

If the supply function for a commodity is

$y = (x + 3)^2$, find the producers' surplus when

a. $y_o = 16$
b. $x_o = 5$

Solution

(a) $y = (x + 3)^2$

$x + 3 = \sqrt{y}$

$x = \sqrt{y} - 3$

When $x = 0$, $y = 9$

$$\text{Producers' surplus} = \int_{y_1}^{y_0} g(y)dy = \int_9^{16} (y^{1/2} - 3)dy$$

$$= \left[\frac{2}{3} y^{3/2} - 3y \right]_9^{16}$$

$$= \left(\frac{128}{3} - 48 \right) - \left(\frac{54}{3} - 27 \right)$$

$$= \frac{11}{3} \quad \text{or} \quad 3.67$$

Or

$y_o = 16$, $x_o = \sqrt{y} - 3 = \sqrt{16} - 3 = 4 - 3 = 1$

Producers' surplus $= x_o y_o - \int_0^{x_o} f(x)dx = 1(16) - \int_0^1 (x+3)^2 dx$

$$= 16 - \left[\frac{1}{3}(x+3)^3 \right]_0^1$$

$$= 3.67$$

(b)　$y = (x + 3)^2$
　　when $x_o = 5$, $y_o = (5 + 3)^2 = 64$

$\therefore$ producers' surplus $= x_o y_o - \int_0^{x_o} f(x)dx = 5(64) - \int_0^5 (x+3)^2 dx$

$$= 320 - \left[\frac{(x+3)^3}{3} \right]_0^5$$

$$= 149.33$$

Note:

If both the demand and supply functions for a commodity are given, the idea of equilibrium could be applied.

If the supply function S(q) and demand function D(q) are given, then the idea of equilibrium will be applied to obtain the equilibrium price p_0 and quantity q_0.

Then, we have

(a)　producers' surplus $= \int_0^{q_0} \{p_0 - S(q)\}dq$　and

(b)　consumers' surplus $= \int_0^{q_0} \{D(q) - p_0\}dq$.

Example 11.13

The supply function S(q) and the demand function D(q) of a commodity are given by

　　　　$S(q) = 400 + 15q$

　　　　$D(q) = 900 - 5q$

　　　　where q is the quantity

Calculate

(a)　the equilibrium quantity

(b)   the equilibrium price

(c)   the producers' surplus

(d)   the consumers' surplus

Solutions:

(a)   at equilibrium

$$S(q) = D(q)$$

i.e. $400 + 15q$ $=$ $900 - 5q$

$20q$ $=$ 500

q $=$ 25 which is the equilibrium quantity.

(b)   to obtain equilibrium price, we substitute $q = 25$ into either the supply or demand functions, thus:

p $=$ $400 + 15q$

$=$ $400 + 15(25)$

$=$ 775

or

p $=$ $900 - 5q$

$=$ $900 - 5(25)$

$=$ 775

(c)   producers' surplus $=$ $\displaystyle\int_{0}^{q_0} \{p_0 - S(q)\}dq$

$=$ $\displaystyle\int_{0}^{25} \{775 - (400 + 15q)\}dq$

$=$ $\displaystyle\int_{0}^{25} \{375 - 15q)\}dq$

$=$ $\left[375q - 15q^2/2\right]_{0}^{25}$

$=$ $375(25) - \dfrac{15(25)^2}{2}$

$=$ $9375 - \dfrac{9375}{2}$

$=$ 4687.5

(d)   Consumers' surplus $\quad = \quad \displaystyle\int_0^{q_0} \{D(q) - p_0\}\,dq$

$$= \quad \int_0^{25} \{900 - 5q - 775\}\,dq$$

$$= \quad \int_0^{25} \{125 - 5q\}\,dq$$

$$= \quad \left[125q - 5q^2/2\right]_0^{25}$$

$$= \quad 125(25) - 5(^{25}/_2)^2$$

$$= \quad 3125 - \frac{3125}{2}$$

$$= \quad 1562.5$$

11.8 Summary

Differentiation and Integration (which is the reverse of differentiation) were treated extensively. Marginal functions and elasticity of demand were also discussed. These concepts were thereafter applied to Business and Economic problems.

11.9 Multiple-Choice and Short-Answer Questions

1.   The second derivative of a function is used to determine its

      A.   nature of the turning point

      B.   maximum point only

      C.   minimum point only

      D.   value when it is squared

      E.   minimum or maximum point.

2.   The demand function for a certain item is, $p = 10 - 0.04\sqrt{q}$. Calculate the elasticity of demand when 1600 items are demanded.

      A.   8.4

      B.   −8.4

      C.   10.5

      D.   −10.5

      E.   16

3. If the marginal revenue function is $x^2 - 5x$, what is the total revenue when $x = 10$

   A. 50

   B. 833.33

   C. 83

   D. 83.33

   E. 500

4. Given that the profit function is $2x^2 - 8x$, the minimum profit is

   A. 8

   B. −2

   C. −8

   D. 4

   E. −16

5. If $y = (8 + 5x)^3$, then $\underline{dy} = \ldots\ldots$
   dx

6. The maximum cost is obtainable when the marginal cost function is ………..

7. A demand is elastic if the elasticity (η) is such that ……………

8. Consumers' surplus arises when consumers pay …………………….. than what they are ……………………. to pay

9. If the marginal profit function is $3x^2 - 2x$, then the total profit when $x = 5$ is …………

10. The equilibrium price is obtained at the point of intersection of …………… and ………… functions

Answers to Chapter Eleven

1.   its nature of the point   (A)

   recall that the first derivative is zero at the turning point and there are three possibilities: if the second derivative is less than zero, the turning point is a maximum, if it is more than zero, the turning point is maximum and if it is zero, the turning point is a point of inflection.

2. $p = 10 - 0.04\sqrt{q}$ $=$ $10 - 0.04q^{\frac{1}{2}}$

   $\dfrac{dp}{dq}$ $=$ $-0.04(\frac{1}{2})\,q^{-\frac{1}{2}}$

   $=$ $\dfrac{-0.04}{2\sqrt{q}}$

   when $q = 1600,$ $\dfrac{dp}{dq}$ $=$ $\dfrac{-0.04}{2\sqrt{1600}}$

   $=$ -0.0005

   p $=$ $10 - 0.04\sqrt{1600}$

   $=$ 8.4

   $\therefore$ η $=$ $(^{-p}/_q)\,(^{dq}/_{dp})$

   $=$ $\dfrac{-8.4}{1600(-0.0005)}$

   $=$ 10.5 (C)

3. $\dfrac{dR(x)}{dx}$ $=$ $x^2 - 5x$

   $R(x)$ $=$ $\displaystyle\int_{0}^{10}\left(x^2 - 5x\right)dx$

   $=$ $\left[\dfrac{x^3}{3} - \dfrac{5x^2}{2}\right]_{0}^{10}$

   $=$ 83.33

4. $P(x)$ $=$ $2x^2 - 8x$

   $\dfrac{dP(x)}{dx}$ $=$ $4x - 8$ $=$ 0 at turning points

   i.e. $4x - 8 = 0,$ $x = 2$

$$\frac{d^2P(x)}{dx^2} = 4 \quad \text{which is positive and hence minimum profit}$$

$$\therefore \quad \text{minimum profit} = 2(2^2) - 8(2) = -8 \quad \text{(C)}$$

5. $y = (8 + 5x)^3$

Let $u = 8 + 5x$, $\dfrac{du}{dx} = 5$

then $y = u^3$, $\dfrac{dy}{du} = 3u^2$

$$\frac{dy}{dx} = \left(\frac{dy}{du}\right)\left(\frac{du}{dx}\right)$$

$$= (3u^2)(5)$$

$$= 15(8 + 5x)^2$$

6.   Zero

7. $|\eta| > 1$

8.   Less, willing

9. $\dfrac{dP(x)}{dx} = 3x^2 - 2x,$

$$P(x) = \int_0^5 (3x^2 - 2x)dx = \left[x^3 - x^2\right]_0^5 = 100$$

10.   Demand, Supply (or vice-versa)

cclx

SECTION C

OPERATIONS RESEARCH

CHAPTER 12

INTRODUCTION TO OPERATIONS RESEARCH (OR)

At the end of this chapter, students or readers should be able to

- understand the concept of OR and how it can be applied to improve managerial decisions
- understand the principle behind modelling, different types of modelling and their nature.
- understand the major steps in OR
- know the various situations where OR can be applied.
- know the limitations of OR.

12.1   Meaning and Nature of OR

Decision-making is a day-to-day activity.

Individuals, societies, government, business organizations and so on do make decisions. In all these cases, decisions are made in order to benefit the decision makers and in most cases, those that the decisions will affect.

A problem would have arisen before a decision is made. In fact, decision-making is a response to an identified problem, which arises as a result of discrepancy between existing conditions and the organization's set objectives.

When a decision is to be made, a lot of factors have to be taken into consideration in order to ensure that the decision is not only the best under the existing conditions but also for the nearest future.

Simple as it may look, decision-making is not an easy task to perform. It is a complex issue. The manager of a business organization will have to decide a course of action to be taken when confronted with a problem. In deciding on this course of action, he has to take some risks since there is some uncertainty (however little it may be) about the consequences of such a decision, He will like to reduce such risks to the barest minimum.

12.1.1 Elements of Decision Making

The general set up for decision-making may be summarized as follows:

- The person, group of persons, organization and so on to make the decision i.e. a decision-making unit.

- Various possible actions that can be taken to solve the identified problem.

- States of nature that may occur

- Consequences associated with each possible action and state of nature that may occur.

- Comparison of the value of the decision and the consequences.

Steps in decision-making can be represented diagrammatically as follows:

```
┌─────────────────────────────────┐
│                                 │
│      Identify a problem         │
│                                 │
└─────────────────────────────────┘
                │
                ▼
┌─────────────────────────────────┐
│  Identify parameters of the     │
│  problem                        │
│       * Goals or objectives     │
│       *Controllable variables   │
│          involved               │
│       * Practical constraints   │
│                                 │
└─────────────────────────────────┘
                │
                ▼
┌─────────────────────────────────┐
│                                 │
│  Examine possible solutions     │
│  and decide on one              │
└─────────────────────────────────┘
                │
                ▼
┌─────────────────────────────────┐
│   Implement the solution        │
│                                 │
└─────────────────────────────────┘
```
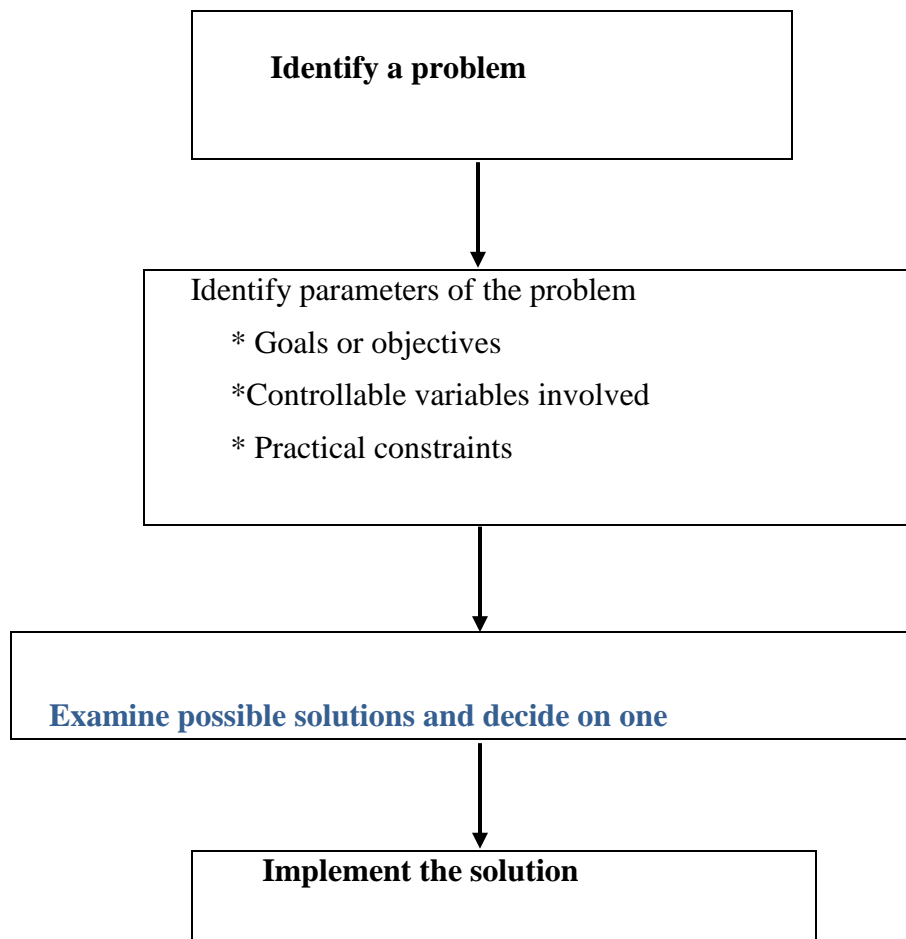
12.1.2 Need for operations research

From the foregoing, it could be deduced that decision-making is very important and should

therefore not be based on experience or mere intuition. Such decisions may not stand the test of time.

Thus, there is the need to find a method that will assist in making decisions that are objective and scientific. This method is called operations research.

The origin of OR dates back to the World War II when the need arose to allocate scarce resources, in an efficient manner, to the various military operations and the activities within each operation i.e. need to research into these operations hence operations research. Other terms used interchangeably with OR include system analysis, decision analysis, operation analysis, management science, decision science and cost- benefit analysis.

Operations Research can be simply defined as the scientific method of allocating scarce resources to competing activities in an optimum way.

More often than not, OR is a team work rather than a single person's work. It may involve Engineers, Scientists, Urban planners and Businessmen depending on the nature of the problem to be solved.

12.2 Techniques of OR

The main stages involved in OR are:

(i)   **identification of problems and objectives**.

The problems for which decisions are being sought must first of all be defined and the objectives clearly spelt out.

(ii)   **identification of variables**

It is very important to identify both the controllable or decision variables and the uncontrollable variables of the system.

The constraints on the variables and the system should be taken into recognition.

The 'bounds' of the system and the options open must also be established.

(iii)   **construction of a model**

This is the central aspect of an OR project.

A model has to be used because it would be impossible to experiment with the real life situations.

A suitable model to represent the system must first of all be established. Such a model should specify quantitative relationships for the objective and constraints of the problem in terms of controllable variables.

It must also be decided, based on the available information, whether the system is to be treated as a deterministic or a probabilistic one.

The model can be a mathematical one or a heuristic one.

Mathematical models are mostly used for OR

A major assumption is that all the relevant variables are quantifiable thus the model will be a mathematical function that describes the system under study.

Some mathematical models are:

(a)   **Allocation models:** these are concerned with sharing scarce resources among various competing activities. Linear programming, transportation and assignment are some examples of the allocation models.

(b)   **Inventory models:** these deal with policies of holding stocks of items of finished goods, ordering quantities and re- order level.

(c)   **Queuing models:** these are concerned with arrival at and departures from service points, and consequent development of queues of customers waiting for service.

(d)   **Replacement models:** these are concerned with determination of an optimal policy for

replacing "failed" items

(e)   **Simulation models**- simulation means to imitate or feign an original situation and so simulation models are based on probabilities of certain input values taking on (i.e. imitating) a particular value.

Random numbers are used most of the time for this type of model.

(a) and (b) are treated in details in subsequent chapters but (c)-(e) are beyond the scope of this study pack.

Heuristic models are essentially models that employ some intuitive rules to generate new strategies, which hopefully will yield improved solutions.

(iv)   **Solution of the model**

Once the model has been constructed, various mathematical methods can then be used to manipulate the model to obtain a solution.

If analytic solutions are possible we can then talk of optimal solutions but if simulation or heuristic models are used, we can only talk about "good" solution.

(v)   **Testing the model**

It is essential to validate the model and its solutions to determine if the model can reliably predict the actual system's performance. It must be ensured that the model reacts to change in the same way as the real system. The past data available for the system may be used to test the validity of the model.

The model for a system may be considered to be valid, if under similar input conditions, it can reproduce to a reasonable extent, the past performance of the system.

(vi)   **Implementation**

It is important to have those who will implement the result obtained on the team of the operations research study. If otherwise, the team should be on hand to give any necessary

advice in case any difficulties are encountered in the course of the implementation. A set of operating instructions manual may be necessary.

12.3   Uses and relevance of OR in business

OR has a very wide area of application in business, engineering, industry, government and science.

OR will always be relevant in any situation where resources do not merge the needs or requirements. Even where the resources are enough, there is still the need to allocate such resources in the best way (an optimal way).

In production planning, OR may be used to allocate various materials to production schedules in an optimal way. In transportation problems, OR may be used to decide on best routes (i.e. routes with minimum cost).

An accountant may apply OR to investment decisions where the fund available is not sufficient for all available projects- capital rationing.

Also, OR can be applied by an accountant in every situation for cost benefit analysis.

12.3 Limitation of OR

The use of OR (i.e. a formal approach) to decision making has its own limitations.

Some of these are:

- No matter how good or sophisticated the design of the system, it rarely provides all the information necessary for action.
- A formal approach involves a model which is more or less a simplified form of the reality but not the reality itself.
- Some of the time, lack of a suitable solution technique restricts the solution of an OR problem
- Optimal solution to an OR model is only optimal based on the assumptions used hence implementation has to consider behaviourial science along with the mathematical science used.

- The system, an OR model imitates, is constantly changing but the model itself is static.

12.5 Summary

The concept of Operations Research (OR) has been treated. Principle of modeling and different types of modeling were discussed. Major steps necessary in OR and real life application were also discussed.

12.6 Multiple-Choice and Short-Answer Questions

1. Decision-making is important in order to
  A. make profit for a business
  B. solve an identified problem
  C. please the customers
  D. perform a task
  E. please the management

2. An heuristic model is a
  A. mathematical model
  B. probabilistic model
  C. non – mathematical model
  D. constant model
  E. simulation model

3. Operations Research is relevant because
  A. resources do not always merge the needs
  B. resources have to be allocated
  C. all activities have to be taken care of
  D. in any operation, research is important
  E. resources have to be allocated in an optimal way

4. Operations Research team should consist of
  A. consultants
  B. managers and the chief executive officers
  C. those who will implement the result

     D.   experts in various disciplines

     E.   statisticians only

5.   Operations Research is a method which assists in making decisions that are ……. and ……………..

6.   Operations Research has a wide range of applications and hence, has no limitations.

          Yes   or   No

7.   In the financial circle, OR is used for ……………….....................…… rationing

8.   In OR, transportation problem can also be referred to as an ……………. problem

9.   OR assists to reduce the ……………….. involved in decision making

10.   The variables involved in any OR technique can be grouped into two: ……………. and …………………… variables

Answers

1.   B

2.   C

3.   E

4.   D

5.   Objective, Scientific

6.   No

7.   Capital

8.   Allocation

9.   Risks

10.   Controllable, Uncontrollable

CHAPTER 13

LINEAR PROGRAMMING (LP)

At the completion of this chapter, students and readers should be able to

- understand the basic concepts of linear programming.

- know the meaning of objective function and constraints in linear programming.

- understand the concept of optimal solution to a linear programming problem.

- know the assumptions underlying the linear programming.

- formulate linear programming problems

- solve linear programming problems by graphical method

- solve the dual problem of L.P.

13.1   Introduction

13.1.1 Meaning and Nature of Linear Programme

As mentioned in chapter 13, decision concerning allocation of scarce or limited resources to competing activities is a major one which managers of business organizations or executive directors of companies will have to take from time to time. These resources could be machines, materials, men and money (4M's) or a combination of two or more of them.

The main objective of a manager will be to use the limited resources to the best advantage. The linear programme model can assist in achieving this very important objective.

13.1.2 Linear Programming as a Resource Allocation Tool

Linear Programming can be defined as a mathematical technique which is concerned with the allocation of limited resources to competing activities. That is, it is a resource allocation tool.

It is mainly concerned with **optimization** of an **objective function** (e.g. to maximize profit or minimize cost) given some constraints (e.g. machine hours, labour hour, quantity of materials). Thus, the linear programming consists of an objective function and certain constraints based on the limited resources.

13.2 Concepts and Notations Of Linear Programming

13.2.1 Assumptions of Linear Programming (LP)

The two major assumptions are:

(1)   **linearity** of the objective function and the constraints.

This guarantees the **additivity** and **divisibility** of the functions involved.

(2)   **non – negativity** of the decision variables

i.e negative quantities of an activity are not possible.

Note that linear programming consists of two words: the linear part as stated above; and the 'programming' part, which is the solution method.

13.2.2 Formulation of Linear Programming Problems

The main steps involved in the formulation of linear programme problems are as follows:

i. define all variables and their units.

ii. determine the objective of the problem – either to maximize or minimize the objective function.

iii. express the objective function mathematically.

iv. express each constraint mathematically including the non-negativity constraints, which are the same for all linear programme problems.

The constraints are always expressed as inequalities.

13.3 Graphical Solution of Linear Programme Problems

There are two methods of solving linear programme problems viz:

a) graphical method

b) simplex method, which is beyond the scope of this study pack. However, computer packages are available for this method and it is particularly useful when the decision variables are more than two.

The graphical method is applicable only when problem involves two decision variables

The necessary steps are as follows:

- turn the inequalities into equalities

- draw the lines representing the equations on the same axes

- identify the region where each constraint is satisfied.

- identify the region where all the constraints are simultaneously satisfied. This is called the **feasible region.**

- find or read off the coordinates of the corner points of the boundary of the feasible region.

- calculate the value of the objective function for each of the points by substituting each of the coordinates in the objective function.

- determine the optimal solution, which is the point with the highest value of the objective function for maximization problems or the point with the lowest value of the objective function for minimization problems.

Inequalities and their graphs treated in chapter 9 (section 9.3) will be very relevant here.

Note that when there are more than two decision variables, the Simplex method (which is beyond the scope of this study pack) is used.

Example 13.1

        Maximize $800x + 600y$

        Subject to $2x + y < 100$

                          $x + 2y < 120$

                              $x \geq 0$

                              $y \geq 0$

Solution

        Maximize $800x + 600y$ - this is the objective function

        Subject to

$$\left.\begin{array}{c} 2x + y < 100 \\ 2x + 2y < 120 \\ x \geq 0 \\ y \geq 0 \end{array}\right\} \text{These are the constraints.}$$

Draw the graphs of

$$2x + y = 100$$

$$x + 2y = 120$$

on the same axes.

For line l_1: $2x + y = 100$, points are (0, 100) and (50, 0)

For line l_2: $x + 2y = 120$, points are (0, 60) and (120, 0)
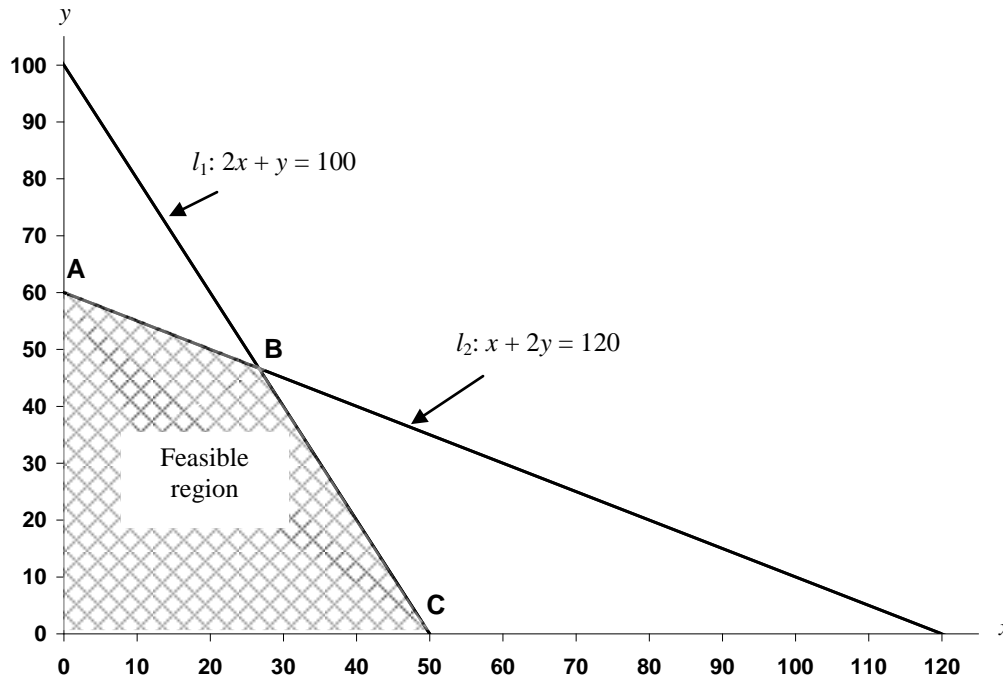
☐ From the graph, corner points of the boundary of the feasible region are A, B, C, (as indicated) with coordinates A(0,60), B(26, 47), C(50,0)

☐ Obtain the value of the objective function for each of the points identified above.

Coordinates	Value of $800x + 600y$ (₦)
A(0,60),	$800(0) + 600 (60) = (36,000)$
B(26, 47)	$800(26) + 600(47) = (49,000)$
C(50,0)	$800(50) + 600(0) = (40,000)$

Coordinates of B (26, 47) give the highest value of the objective function (i.e. N49,000). Consequently the best (optimal) combination is $x = 26$ and $y = 47$.

Note that (i) the corner point B is just the point of intersection of the two lines and usually gives the optimal solution for the two constraints.

(ii) if l_1 and l_2 are solved simultaneously, the coordinates of their point of intersection are

$$x = 26\tfrac{2}{3} \text{ and } y = 46\tfrac{2}{3}$$

Example 13.2

Maximization problem (Production Problem)

AWOSOYE furnitures produces two types of chairs – executive and ordinary by using two machines M_I and M_{II}. The machine hours available per week on M_I and M_{II} are 100 and 120 respectively.

An executive chair requires 4 hours on M_I and 3 hours on M_{II}, while the requirements for an ordinary chair are 1.25 hours on M_I and 2 hours on M_{II}. The company has a standing order to supply 15 ordinary chairs a week.

Profit (contribution) on (from) an executive chair is N2,500 while profit on an ordinary chair is N1,000.

You are required to:

a)   formulate the problem as a linear programming problem.

b)   use graphical method to solve the linear programming problem.

Solution

Let x units of the executive chairs and let y units of ordinary chairs be made.

Let P(N) be the total profit (contribution).

a)   The objective function is

$$P = 2{,}500x + 1000y \text{ (2500 from an executive chair and 1000 from ordinary chair)}$$

The constraints are:

$$4x + 1.25y < 100 \quad (M_I \text{ constraint})$$

$$3x + 2y \quad < 120 \quad (M_{II} \text{ constraint})$$

$$y \geq 15$$

$$x > 0 \quad (\text{Non-negativity constraint})$$

On the other hand, the information can be summarized first in tabular form, from where the objective function and the constraints will be obtained as follows:

| Type of chair | Machine hours per unit | | Profit (₦) |
	M_I	M_{II}	
X	4	3	2500
Y	1.25	2	1000
Available machine hours	100	120	

The problem is

Maximize $\quad$ P = 2500x + 1000y

Subject to $\quad$ 4x + 1.25y < 100 ----------(l_1)

$\qquad\qquad$ 3x + 2y $\quad$ < 120 ----------(l_2)

$\qquad\qquad\quad$ y > 15

$\qquad\qquad\quad$ x > 0

For line l_1: 4x + 1.25y = 100, points are (0, 80) and (25, 0)

For line l_2: 3x + 2y $\quad$ = 120, points are (0, 60) and (40, 0)
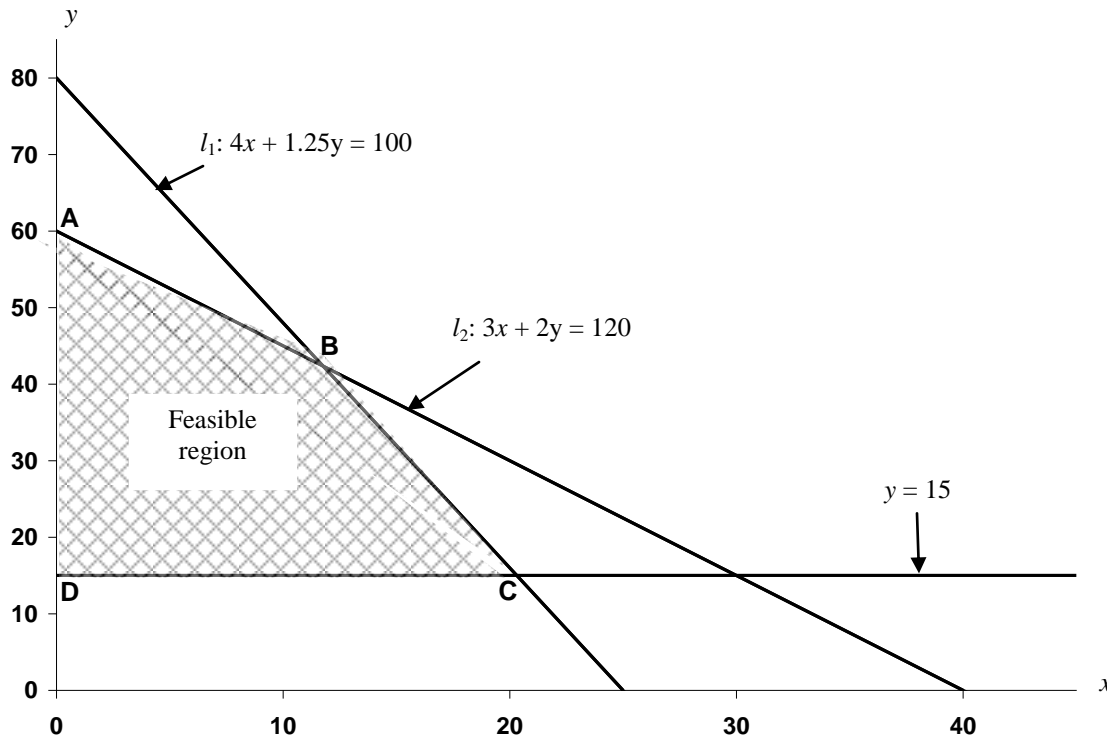
From the graph, the corner points of the boundary of the feasible region are A, B, C and D with the coordinates: A(0, 60); $\quad$ B(11.75, 42.5); C(20.25, 15); D(0, 15).

Coordinates	Value of the objective function $(2500x + 1000y)$	
A(0, 60)	$2500(0) + 1000(60)$	$= 60000$
B(11.75, 42.5)	$2500(11.75) + 1000(42.5)$	$= 71875$
C(20.25, 15)	$2500(20.25) + 1000(15)$	$= 65625$
D(0, 15)	$2500(0) + 1000(15)$	$= 15000$

Coordinates (11.75, 42.5) gives the highest profit of ₦71,875. Hence, the optimal combination to be produced is 12 executive chairs and 43 ordinary chairs.

Note that solving l_1 and l_2 simultaneously gives point B as (11.8, 42.4) and solving l_1 and $y = 15$ simultaneously gives point C as (20.3, 15).

Example 13.3

Minimization problem (Mix Problem)

A poultry farmer needs to feed his birds by mixing two types of ingredients (F_1 and F_2) which contain three types of nutrients, N_1, N_2 and N_3.

Each kilogram of F_1 costs ₦200 and 200 units of N_1, 400 units of N_2 and 100 units of N_3 while, each kilogram of F_2 costs ₦250 and contains 200 units of N_1, 250 units of N_2 and 200 units of N_3.

The minimum daily requirements to feed the birds are 14,000 units of N_1, 20,000 units of N_2 and 10,000 units of N_3.

a)   Formulate the appropriate linear programming problem.

b)   Solve the linear programming problem by graphical method.

Solution

Let x kg of F_1 and y kg of F_2 be used per day.

Let C (₦) be the total cost of daily feed mixture.

The problem is to meet the necessary nutrient requirement at minimum cost.

The objective function is

$C = 200x + 250y$

The constraints are

$200x + 200y > 14{,}000$ (N_1 constraint) ---------l_1

$400x + 250y > 20,000$ (N_2 constraint) ---------l_2

$100x + 200y > 10,000$ (N_3 constraint) ---------l_3

$x > 0$
$y > 0$ } Non-negativity constraints
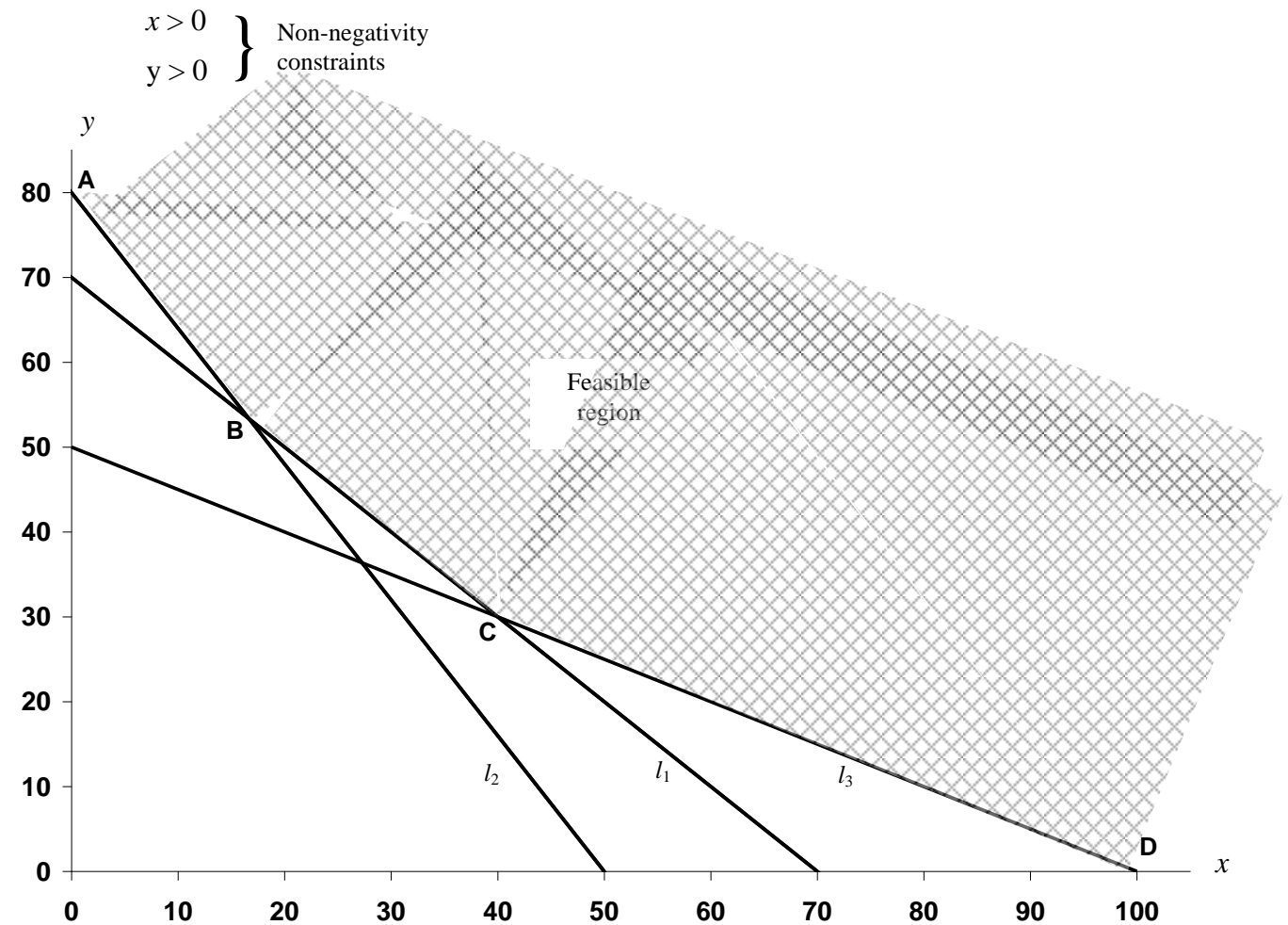
From the graph, the corner points of the boundary of the feasible region are A,B,C,D with the coordinates A(0, 80), B(16, 53), C(38, 31), D(100, 0)

Coordinates	Value of the objective function = $(200x + 250y)$
A(0, 80)	$200(0) + 250(80) = 20,000$
B(16, 53)	$400(16) + 250(53) = 16,450$
C(38, 31)	$200(38) + 250(31) = 15,350$
D(100, 0)	$200(100) + 250(0) = 20,000$

Coordinates (38, 31) give the lowest cost hence the optimal mix is 38kg of F_1 and 31kg of F_2.

Note that (i) solving l_1 and l_2 simultaneously gives point B as $(16\frac{2}{3}, 53\frac{1}{3})$; and

(ii) solving l_1 and l_3 simultaneously gives point C as (40, 30)

13.4 Duality problem

There is a **dual formulation** for every linear programming problem. The original linear programming is known as **primal problem**.

If the original linear programming problem is a minimizing one then the dual formulation is a maximizing one and vice versa.

13.4.1 Shadow Costs

Only the binding constraints have shadow costs. They are also known as shadow prices or dual prices.

The binding constraints are the two constraints which intersect at the optimal solution point.

The shadow cost of a binding constraint is the amount by which the objective function decreases (or increases) as a result of availability of one unit less or more of the scarce resource.

The solutions to the dual problem of the primal problem give the shadow costs (prices) hence the alternative term dual costs (prices).

The shadow costs help management of a business organization to carry out sensitivity analysis on the availability of scarce resources.

Solving the dual problem is the same as carrying out sensitivity analysis.

Example 13.4

ROYEJAS Sewing Institute produces two types of dresses, shirt and blouse. A shirt has a contribution of ₦800 while a blouse has a contribution of ₦600.

A shirt requires 2 units of materials and 1 hour of labour while a blouse requires 1 unit of materials and 2 hours of labour.

If 100 units of materials and 120 labour hours are available per week;
a) formulate and solve the linear programming problem

b)   find the shadow cost of a

     i)   unit of materials

     ii)   labour hour

c)   advise the Sewing Institute accordingly.

Solutions

a)   Let x units of shirts and y units of blouse be produced

Then the Linear Programming is

Maximize: $800x + 600y$

Subject to: $2x + y < 100$ (materials constraint)

     $x + 2y < 120$ (labour hours constraints)

$$\left. \begin{array}{l} x > 0 \\ y > 0 \end{array} \right\} \begin{array}{l} \text{Non-negativity} \\ \text{constraints} \end{array}$$

This problem is the same as Example 14.1 with solution (26, 47), which gives the highest revenue of ₦49 000. i.e. the optimal combination (mix) to be produced per week is 26 shirts and 47 blouses.

b)   The binding constraints are

     $2x + y < 100$ (materials constraint)

     $x + 2y < 120$ (labour constraints)

i)   Materials

- Increase the units of materials by 1 while the labour hours remain unchanged
- Solve the resulting simultaneous equation to obtain new values for x and y
- Calculate the resulting difference in contribution. This is the shadow cost.

The binding constraints now became

     $2x + y = 101$

     $x + 2y = 120$

Solving these simultaneous equation (see chapter 9 section 9.3.1.2), we obtain

     $x = 27.33$ and $y = 46.34$

Now, substitute these values into the objective function ($800x + 600y$) to obtain,

800 (27.33) + 600 (46.34)	=	₦49668
Original contribution	=	₦49000
Difference	=	₦668

i.e. 1 extra unit of material has resulted in an increase of ₦668 in the contribution.

Thus, the shadow cost per unit of materials is ₦668

ii)   **Labour hours**

Here, we increase the labour hours by 1hour while the units of materials remain unchanged.

The new binding constraints will now be:

$2x + y = 100$

$x + 2y = 121$

which when solved will give

$x = 26.33$ and $y = 47.33$

and the new contribution is

800(26.33) + 600(47.33)	=	₦49,466.67
original contribution	=	₦ 49,000
Difference	=	₦ 466.67

i.e. 1 extra labour hour has resulted in an increase of ₦ 466.67 in contribution i.e. the shadow cost per hour of labour is ₦ 466.67

c)   The sewing institute can consequently be advised to increase the units of materials rather than to increase the hours of labour because the increase in the contribution brought about by a unit increase in materials is bigger than that of the labour.

13.4.2 Limitations of Linear Programming

Limitations of OR methods have been enumerated in chapter 13 (section 13.4) so, linear programming being one of the OR methods has its own limitations.

Some of these are:

- Linear programming cannot be applied if the objective function and/or the constraint are not linear.

- Linear programming involves a model which is just a simplified form of the reality but not the reality itself.
- Solution of the linear programming model is only valid based on the assumptions used.
- Linear programming is static and therefore not robust enough to accommodate changes.
- In particular, solution of linear programming by the graphical method can not be applied if the decision variables are more than two although other methods such as Simplex can be adopted. Note that the Simplex method is complex.

13.5 Summary

Linear programming has been described as a resource allocation tool. Methods of formulating and solving graphically Linear programming problems from a given set of data were discussed. Practical applications of Linear programming were also treated.

13.6 Multiple-Choice and Short-Answer Questions

1. The objective function for a minimisation LP problem is $500x + 700y$ and the coordinates of the corner points of the feasible region are (0, 40), (30, 50) and (60, 0). The optimal solution is

   A. 28 000
   B. 30 000
   C. (60, 0)
   D. (0, 40)
   E. (30, 50)

2. Linear programming can be applied only if the objective function and/or constant are:

   A. linear
   B. non-linear
   C. complex
   D. independent functions
   E. dependent functions

3. Find the value of z in the following linear programming problem:

   Minimize $z = 50x + 60y$

   Subject to $\quad 2x + 4y \geq 80$

   $\qquad\qquad x + y \geq 30$

   $\qquad\qquad x, y \geq 0$

A. 2000

B. 1600

C. 1800

D. 1400

E. 1200

4. If the original linear programming problem is a minimising one then the _____ is a maximising one and vice versa

A. shadow occurrence

B. anal formulation

C. simplex method

D. objective function

E. inequality constraint

5. The region where all the constraints are simultaneously satisfied is called the _____

A. equilibrium point

B. point of maximum profit

C. feasible region

D. point of maximum profit

E. break-even point

6. Linear programming is concerned with the allocation of ……………… to ……………………..

7. The solution to a linear programming problem is always at …………………. point of the boundary of the ……………….. region.

8. The linear programming cannot be applied when there are more than two decision variables.

Yes or No?

9. If the primal problem of a linear programming is a minimizing one, the dual problem will be a ……………

10. The two major assumptions in linear programming are ……………… of the objective function and …………. of the decision variables.

Answers to Chapter Twelve

1 **Coordinates**   **Value of the objective function**

   (0, 40) $500(0) + 700(40) = 28000$

   (20, 50) $500(30) + 700(50) = 50000$

   (60, 0) $500(60) + 700(0) = 30000$

   28000 is the least, hence (0, 40) is the optimal solution (D).

2. Linear (A)

3. Recall that since only two constraints are involved, the solution is at the point of intersection of the constraints.

   So, solving the equations simultaneously

   $$2x + 4y = 80 \ldots\ldots\ldots\ldots\ldots(i)$$

   $$x + y = 30 \ldots\ldots\ldots\ldots\ldots(ii)$$

   equation (ii) × 2 gives

   $$2x + 2y = 60 \ldots\ldots\ldots\ldots...(iii)$$

   equation (i) – equation (iii) gives

   $$2y = 20, y = 10$$

   substitute for y in equation (ii)

   $$x + 10 = 30, x = 20$$

   value of $z = 50x + 60y$

   $$= 50(20) + 60(10)$$

   $$= 1600 \text{ (B)}$$

4   dual formulation (B)

5   feasible region (C)

6   scarce resources, competing activities

7   corner, feasible region

8   No

9   maximizing one

10   linearity, non-negativity

CHAPTER 14

INVENTORY PLANNING AND PRODUCT CONTROL

At the end of this chapter, students and readers should be able to

- explain the meaning and functions of inventory

- understand the basic concepts in inventory control

- distinguish between deterministic and stochastic inventory models

- calculate the Economic Order Quantity (EOQ) under various situations

14.1 Meaning and Functions of Inventory

Inventory could mean a list of items in a shop or a house or a company. It could also mean the stock of items available in an organisation. The items could be raw materials, partly finished products or finished products. Inventory taking also means stock taking.

There are three major motives for holding stocks, these are

a. Transactions motive

This is to meet the demand at any given time. The quantity demanded is known with certainty and when stock-out occurs, replenishment of stocks is immediate.

b. Precautionary motive

This is to avoid loss of sales due to some uncertainties. Buffer or safety stocks are held so as not to run out of supply.

c. Speculative motive

This is in anticipation of shortage from the supplier or price increase by the supplier. Current stock may be increased.

14.1.1 Reasons for holding Stock.

The three major objectives discussed above may be broken down into reasons why a business needs to hold stock. Some of the reasons are to

i. act as a buffer for variations in demand and usage

ii. take advantage of quantity discount by buying in bulk

iii. take advantage of seasonal and price fluctuations

iv. keep to the barest minimum the delay in production process which may be caused by lack of raw materials

v. take advantage of inflation or possible shortages

vi.   ensure no stock-outs.

The main objective of inventory control is to maintain stock levels so as to minimise the total inventory costs. Two main factors are to be established – when to order and what quantity to order.

14.2   Inventory Costs

Inventory costs are of four types

  (a) **Ordering or procuring costs** – These are all the costs relating to the placement of orders for the stocks. They could be internal or external.

   i.   administrative costs associated with the departments involved in placing and receiving orders

   ii. transport costs

   iii. production set-up costs where goods are manufactured internally – cost associated with production planning, preparing the necessary machinery and the work force for each production run.

   (i)   and (ii) are external while (iii) is internal.

  (b) **Stock costs** – these are the suppliers' price or the direct costs of production. These costs need to be considered especially when

   i.   bulk purchase discounts are available

   ii. savings in the direct costs of production are possible with longer "batch runs"

  (c) **Holding costs** – these are also known as **carrying costs** and include the following:

   i.   cost of capital tied up including interest on such capital

   ii. handling and storage costs

   iii. insurance and security costs

   iv. loss on deterioration and /or obsolescence

   v.   stock taking, auditing and perpetual inventory costs

   vi. loss due to pilferage and vermin damage

  (d) **Shortage or stock-out costs** – as a result of running out of stock, a company will normally incur some loss. Stock-out costs include:

   i.   loss of customers

   ii. loss of sale and contribution earned from the sale

iii. loss on production stoppages

iv. loss on emergency purchase of stock at a higher price.

14.3 Definition of Terminology

- **Lead time or Procurement time** is the time expressed in days, weeks or months, which elapses between ordering and eventual delivery.

  A supply lead time of one week means that it will take one week from the time an order is placed until the time it is supplied.

- **Physical stock** is the number of items physically in stock at the time of inventory.

- **Free stock** is the physical stock added to awaiting orders less unfulfilled demands.

- **Maximum stock** is the selected stock level to indicate when stocks have risen too high

- **Stock-outs** refer to a situation where there is a demand for an item of stock but the warehouse is out of stock.

  Four stock-outs means that there is a demand for 12 times but only eight items are available

- **Buffer stock or safety stock or minimum stock** is the level to indicate when stock has gone too low and is usually held to safeguard against stock-outs.

- **Economic Order Quantity (EOQ) or Economic Batch Quantity(EBQ)** is the ordering quantity of an item of stock which minimises the costs involved
  (see section 16.4)

- **Re-order quantity** is the number of units of item in one order.

- **Re-order level** is the level of stock at which a new order for more units of items should be placed.

- **Demand** is the number of units of stock required within a particular period of time.

Illustration of Simple Stock
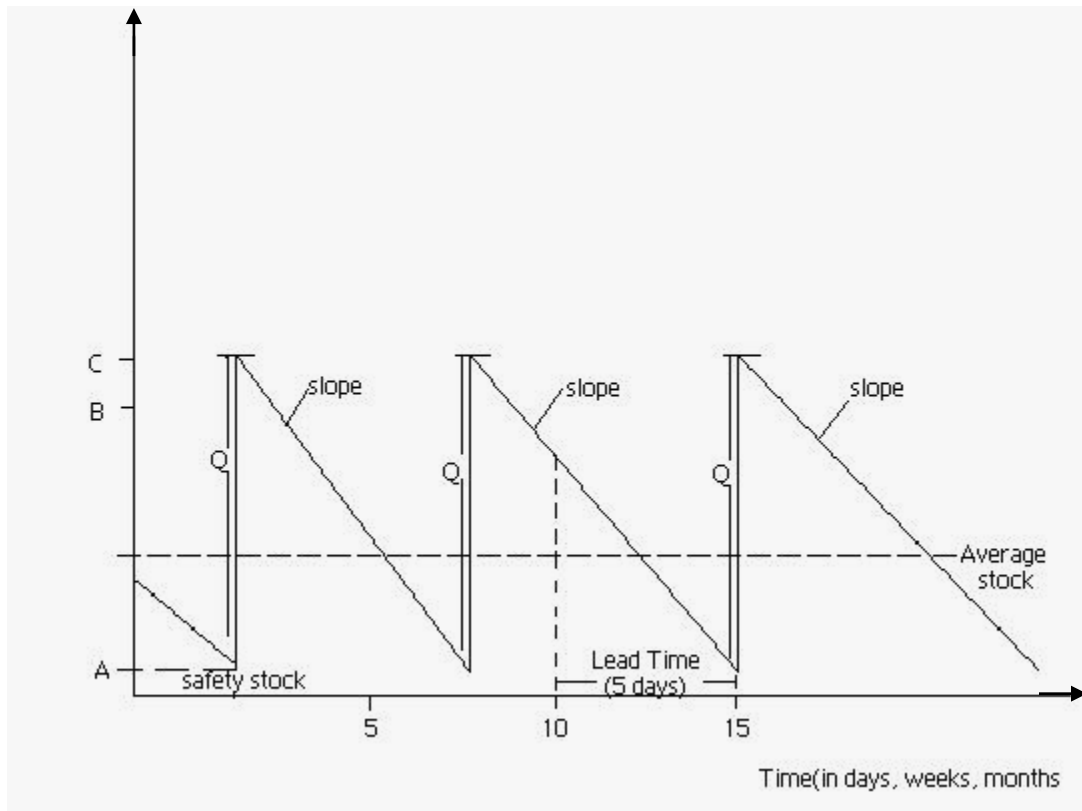
- Average Stock = $\dfrac{Q}{2}$ i.e. $\dfrac{C-A}{2}$

- Re – order level = B

- Re – order quantity (Q) = C – A

- The slopes show anticipated rates of demand

- Note that safety stock is allowed here.

Simple Calculations

Example 14.1

Data on a given stock item are as follows:

Normal usage	2400 per week
Minimum usage	1,600 per week
Maximum usage	3,500 per week
Lead time	20 -23 weeks
EOQ	50,000

Calculate the various control levels.

Solutions

The control levels are

Reorder level

Minimum level

Maximum level

Reordered level = maximum usage x maximum lead time

$$= 3, 500 \text{ x } 23$$

$$= 80, 500$$

Maximum level = re-order level + EOQ – (minimum usage x minimum lead time)

$$= 80, 500 + 50, 000 – (1600 \text{ x } 20)$$

$$= 98, 500$$

Minimum level = re – order level – (normal usage x average lead time)

$$= 80,500 – (2400 \text{ x } 21.5)$$

$$= 28, 900$$

14.4 General Inventory Models.

Inventory models help to decide how to plan and control stock in order to minimise costs. An inventory model may be a deterministic or stochastic one.

a) **Deterministic model**

It is one in which all factors like demand, ordering cost, holding cost etc are known with certainty. No element of risk is involved. Deterministic models may or may not allow for stock-outs.

b) **Stochastic Models**

A stochastic model is one in which one or all of the factors is/are not known with certainty. It is also known as probabilistic model. Because of the uncertainty in this type of model, there may be necessity for buffer stock which is not necessary in deterministic model.

Stochastic models may be analysed as

1. A **Periodic Review System (PRS)** in which stock levels for all parts are reviewed at fixed time intervals and varying quantities ordered at each interval as necessary i.e. varying quantities ordered at fixed intervals.

Advantages

   i   Since stock items are reviewed periodically, obsolete items will be eliminated from time to time

   ii It allows for the spreading of the "load" of purchasing department more evenly

   iii Production planning will be more efficient since orders will be in the same sequence

   iv Range of stock items will be ordered at the same time thus allowing for larger quantity discount and reduction in ordering costs.

Disadvantages

   i   Larger stocks are required.

   ii EOQ not taken into consideration hence reorder quantities are not economical

   iii Changes in consumption not taken care of

   iv May be difficult to set up especially if demands are not consistent

2. **Re-order Level System (RLS)** in which a fixed or predetermined quantity, usually EOQ, is ordered at irregular time intervals. This may involve what is referred to as a "two-bin system" where the stock is segregated into two places (bins) and usage is from the first bin so that an order is placed when it becomes empty while usage continues form the second bin.

Advantages

   i   Lower stocks are required

   ii EOQ used hence re-order quantities are economical

   iii It is more responsive to changes in consumption

   iv It takes care of differing types of inventory

Disadvantages

   i   Overloading of purchasing department when many items reach re-order level at the same time.

   ii No particular sequence as items reach re-order level randomly

iii   EOQ may not be appropriate some of the time.

14.5   Basic Economic Order Quantity (EOQ) Model

The EOQ has earlier been defined as the ordering quantity which minimises all the costs involved.

14.5.1   Assumptions of the Model

There are some assumptions which the model is based on. These are:

i.   rates of demand are known

ii.   stock holding cost is known and constant

iii.   price per unit is known and constant

iv.   no stock-outs are allowed

v.   ordering cost is known and constant

vi.   no part-delivery; ordered batch is delivered at once.

The symbols used are:

d: the annual demand

Q: the re-order quantity

c: the ordering cost for single order

h: the cost of holding a unit of stock for one year..

Note:

- Number of orders in a year is $\dfrac{d}{Q}$

- Ordering cost in a year is $c.\dfrac{d}{Q}$

- Average stock is $\dfrac{Q}{2}$

- Holding cost per annum is $h.\dfrac{Q}{2}$

- Length of inventory cycle is 52/*No. of orders i.e* $\dfrac{52Q}{d}$ weeks or $\dfrac{12Q}{d}$ months

- Total cost per annum = ordering cost p.a+ holding cost p.a   i.e. $\dfrac{cd}{Q}+\dfrac{Qh}{2}$

14,5.2   The EOQ formula

The total cost per annum is

$$T = \frac{cd}{Q} + \frac{Qh}{2}$$

by the definition of EOQ, we want to minimise T

$$T = cdQ^{-1} + \frac{Qh}{2}$$

$$\frac{dT}{dQ} = -cdQ^{-2} + \frac{h}{2}$$

but $\dfrac{dT}{dQ} = 0$ at the turning point

this implies that

$$-cdQ^{-2} + \frac{h}{2} = 0$$

$$\frac{h}{2} = \frac{cd}{Q^2}$$

$$Q^2 h = 2cd$$

$$Q^2 = \frac{2cd}{h}$$

$$\therefore Q = \sqrt{2cd/h}$$

$$\frac{d^2T}{dQ^2} = 2cdQ^{-3}$$

$$\therefore when \; Q = \sqrt{2cd/h} \, , then \; \frac{d^2T}{dQ^2} = \frac{2cd}{\left(\dfrac{2cd}{h}\right)^{3/2}} = \frac{h^{3/2}}{\sqrt{2cd}} > 0 \; for \; all \; values$$

of h, c and d $\sin ce$ *none of h, c & d can be negative,*

then $Q = \sqrt{2cd/h}$ *gives the* $\min imum$ *total* $\cos t.$

Example 14.2

(a) The demand for an item is 60, 000 per annum. The cost of an order is ₦25 and holding cost per item is ₦2·00 per annum. You are required to find

(b) the number of orders per year and the associated ordering cost

(c) length of inventory cycle

(d) total cost per annum

Solutions

C = 25, d = 60, 000, h = 2

(a)

$$Q = \sqrt{\frac{2cd}{h}}$$

$$Q = \sqrt{\frac{2 \times 25 \times 60000}{2}}$$

$$= 1225 \text{ items}$$

(b) No. of orders per year $= \dfrac{d}{Q}$

$$= \frac{60000}{1225}$$

$$= 49 \text{ orders}$$

associated cost $= 49 \times 25 = ₦1225$

(c) Length of inventory cycle $= 52/49 = 1.06$ weeks

(d) Total cost per annum $= 25 \times 49 + (60000 \times 2)/2$

$$= ₦61\ 225$$

EOQ with Gradual Replenishment (Stock holder Manufacturer)

In this situation, the manufacturer serves as the stock holder. i.e. the manufacturer holds stocks of finished good item and the production is not continuous.

If *r* represents the production rate per annum, then

$$Q = \sqrt{\frac{2cd}{h(1 - \dfrac{d}{r})}}$$

Example 14.3

A firm has decided to manufacture its own items for stock at the rate of 16000 per month. Demand for the item is at the rate of 8000 in one month. Each production run costs ₦40 and holding cost per item is 80k per month.

Determine:

(a) the Economic Production Run

(b) the length of the inventory cycle

(c) number of production runs per month

(d) total cost of production per month.

Solutions

(a)

$$Q = \sqrt{\frac{2cd}{h(1 - \frac{d}{r})}}$$

note that the values have been given on monthly basis. If the periods are different, all values must be converted to the same period before proceeding.

So c = 40, d = 8000, h = 0·80, r = 16,000

$$Q = \sqrt{\frac{2 \times 40 \times 8000}{0 \cdot 8\left(1 - \frac{8000}{16000}\right)}}$$

$$= 1265$$

(b) Length of inventory circle is $\frac{Q}{d}$ months $= \frac{1265}{8000} = 0 \cdot 16$ months

(c) Number of production runs is $\frac{d}{Q} = \frac{8000}{1265} = 6 \cdot 32$

(d) Production cost = ordering cost + holding cost

$$= 6 \cdot 32 \times 40 \ + \ Q \frac{\left(1 - \frac{d}{r}\right)h}{2}$$

$$= 252{\cdot}8 + 1265 \left[\frac{\left(1 - \dfrac{8000}{16000}\right)}{2} \right] (0.8)$$

$$= 252{\cdot}8 + 253$$

$$₦\,505.8$$

14.6   EOQ with Stock-out

When stock-out is allowed, the stock-out costs must be known. If c_s represents the stock-out per

item, then $Q = \left(\sqrt{\dfrac{2cd}{h}}\right)\left(\sqrt{\dfrac{h + c_s}{c_s}}\right)$

Example 14.4

The demand for an item is 1000 units per week.it costs ₦30 to place an order and holding cost

per item is ₦1.50 per annum. Cost associated with stock-out is 90k per item.

Calculate the EOQ

Solution

$$Q = \left(\sqrt{\frac{2cd}{h}}\right)\left(\sqrt{\frac{h + c_s}{c_s}}\right)$$

c =   30,d = 1000×52 = 52,000(per annum), h = 1·50, c_s = ₦0.90

$$Q = \left(\sqrt{\frac{2 \times 30 \times 52000}{1 \cdot 50}}\right)\left(\sqrt{\frac{1 \cdot 5 + 0 \cdot 9}{0 \cdot 9}}\right)$$

$$= 2{,}356.$$

14·7   EOQ with bulk discounts

Some of the time, the assumption of constant price per item may not be realistic because usually, some form of discount can be obtained by ordering large quantities. In order to determine whether the discount offered is worth it or not, comparisons of costs at EOQ (i.e. no discount) and at different levels of discount are made. The best quantity to order is the one with the least costs. Note that discount may not necessarily imply reduction in total cost at all times.

Example 15.5

The demand for an item is 45,000 per annum, the ordering cost is ₦50 per order, holding cost per item is ₦2.50 per annum and the price per unit of item is ₦10.

The supplier offers a discount rate of 3% for orders between 10,000 and 31,999 and 8% discount rate for orders of 32,000 and above.

Determine the best quantity to order.

$$Q = \sqrt{\frac{2cd}{h}}$$

c = 50, d = 45,000, h = ₦ 2.50

$$\therefore Q = \sqrt{\frac{2 \times 50 \times 45000}{2.50}}$$

$$= 1,341.640787$$

$$\text{numbers of orders} = \frac{45000}{1,341.640787}$$

$$= 33.54101966$$

No discount (E0Q = 1342 items)

Cost of purchase	= 45000×10	= ₦450, 000
Holding Cost	= 1,341.640787×2.5	= 3, 354.10
Ordering cost	= 23.54×50	= <u>1677</u>
	Total Cost	= ₦455,031.10

3% (10,000 items)

Cost of purchase = 450,000×0.97 = ₦436, 500

Holding Cost = 10,000×2.5×0.97 = 24, 250

$$\text{Ordering cost} = \frac{45000}{10000} \times 50 = \underline{225}$$

Total Cost = ₦460,975

8% (32,000 items)

Cost of purchase = 450,000×0.92 = ₦414, 000

Holding Cost $= 32{,}000 \times 2.5 \times 0.92 =$ 73, 600

Ordering cost $= \dfrac{45000}{32000} \times 50$ $=$ $\underline{75}$

Total Cost $= \underline{\underline{48{,}767.00}}$

So, advantage of no discount rate should be taken.

Note that the discount does not affect the ordering cost since this does not concern the supplier (see section 14.2(a))

14.8   Summary

The meaning and function of inventory have been discussed. Deterministic and stochastic inventory models were also treated. Calculation of Economics Order Quantity (EOQ) and its usefulness in taking decisions under Bulk Purchase discounts were treated.

14.9   Multiple-Choice and Short-Answer Questions

1.   Which of these is not a reason for holding stock?

A.   to take advantage

B.   to act as a buffer for variations in demand and usage

C.   to ensure that the store is filled up at all times

D.   to take advantage of inflation

E.   to ensure no stock – outs

2.   Physical stock is the number of items in stock at

A.   the end of each week

B.   the end of each month

C.   the time after customers have been supplied

D.   the time of inventory

E.   the end of the year.

3.   Stock-out refers to a situation when

A.   an item is not in the store

B.   no item is in the store

C.   the store runs out of stock

    D.   there is no demand for an item

    E.   there is demand but the item is not in store.

4.   The demand for an item is 3600 units per annum, the cost of an order is ₦16 and holding cost per unit of an item is ₦2 per annum. The number of orders per year is:

    A.   240

    B.   15

    C.   225

    D.   220

    E.   25

5.   The main objective of inventory control is to maintain stock levels to …………………… the total ………………….. costs

6.   The lead time is the time that elapses between ……………….. and …………………

7.   The reorder level is the product of ……………. usage and maximum ……………….

8.   If discount rates are allowed for bulk purchase, the highest discount rate is always the best.
           Yes   or   No?

9.   Economic Order Quantity is the ordering quantity which ………….. all the costs involved.

10.   Given that the annual demand is 50,000, the re-order quantity is 2000, ordering cost is ₦20 per order and holding cost per item is ₦2 per annum, then total cost per annum is ………………………

Answers to Chapter Fourteen

1. C

2. D

3. E

4. $Q = \sqrt{\dfrac{2cd}{h}}$

   $= \sqrt{\dfrac{(2)(16)(3600)}{2}}$

   $= 240$

   No of orders per year is

   $\dfrac{\text{Demand}}{Q}$

   $= \dfrac{3600}{240}$

   $= 15$ (B)

5. Minimise, Inventory

6. Ordering, Delivery

7. Maximum, Lead Time

8. No

9. Minimises

10. Total Cost per annum is

    ordering cost per annum + holding cost per annum

    i.e. $\dfrac{cd}{Q} + \dfrac{Qh}{2}$

    $= \dfrac{(20)(50000)}{2000} + \dfrac{(2000)(2)}{2}$

    $= ₦2,500$

CHAPTER 15

NETWORK ANALYSIS

At the end of this chapter, the students and readers should be able to

- explain the concept of Network Analysis
- define Activity, Event and Dummy Activity
- draw a network diagram
- identify the paths in a Network diagram and calculate their durations
- identify the critical path and the critical activities
- calculate the shortest time for the completion of a project
- calculate the Earliest Start Time (EST), Latest Start Time (LST), Earliest Finish time (EFT)

  and Latest Finish Ti me (LFT) for an activity.

  calculate floats and interprete their values

15.1 Introduction

Network analysis is another Operations Research (OR) method of approach to management problems. It involves management of large projects in an optimal way.

It is a technique for planning, scheduling and controlling of projects.

Some examples of such projects include:

- setting up of a new business
- construction of projects
- maintenance of buildings and machines
- personnel training etc.

Its primary objective is to complete the project within the minimum time

The project manager has to decide on the "flow diagram" for the project by identifying

- the tasks that must be done first before others can start i.e the tasks which precede other tasks
- the tasks that can be done simultaneously
- the tasks which are "crucial" to the project.

ccc

Some of other terms used for Network analysis are:

- Critical Path Method (CPM)

- Critical Path Analysis (CPA)

- Project Evaluation and Review Technique (PERT)

15.1.1 Definition of Terms

- The tasks are called activities

  An activity is represented by an arrowed line ⟶ (from left to right and not drawn to scale): it runs between two events.

  It consumes time and resources e.g prepare a set of accounts

- **An event** is the start and/or completion of an activity and is represented by a circle called a node. It is usually numbered.

Tai/Start   Head/End
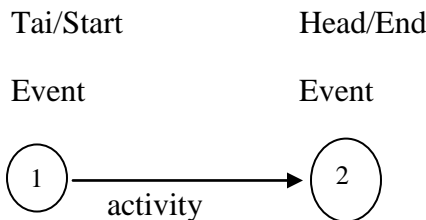
Event   Event

The activity is between event 1 and event 2

- **A dummy activity** is an activity of "circumstance". It consumes neither time nor resources and is only used to ensure non-violation of the rules for drawing a Network diagram.

  It is usually represented by a dotted line ⇢ The use of a dummy activity will ensure that two different activities do not have the same starting and finishing nodes.
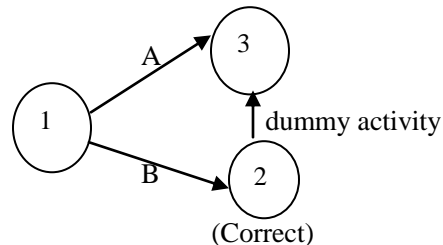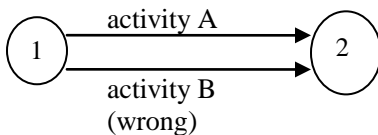
- **A path** of a Network is a sequence of activities which will take one from the start to the end of the Network.

- The **critical path** is the path of a Network with the longest duration and gives the shortest time within which the whole project can be completed. It is possible to have more than one critical path in a Network.

- **Critical activities** are the activities on the critical path.

   They are very crucial to the completion of a project. There must not be any delay in starting and finishing these activities otherwise the duration of the whole project will be extended.

15.2 Construction/Drawing of a Network diagram

A Network diagram is a combination of activities and events in a logical sequence for the completion of a project.

In order to draw a Network diagram, the following must be noted:

- all activities with their durations must be known or estimated

- the logical sequence of the activities must be put in place i.e which activities must be done one after the other (preceding activities) and which ones can be done simultaneously

- all activities must contribute to the progression of the project otherwise they should be discarded.

- a network diagram must have one starting event and one finishing event.

Example 15.1

The activities involved in a project are as shown below:

Activity	*Preceding Activity*
A	-
B	-
C	B
D	C
E	A
F	E

Draw the Network diagram for the project.

Solution

When a Network diagram is to be drawn, do not look at the entire Network to start with but rather take one step after the other. It is also advisable to use a pencil (so that any mistake can easily be erased) and then trace out.

Start off with the activity (ies) that have no preceding activity(ies), then follow through the given information on the activities.

(try to draw your own Network diagram before looking at the solution)
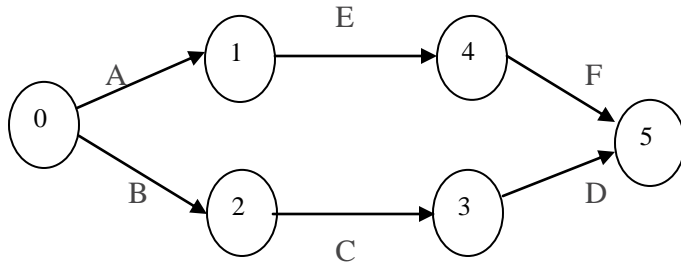
Note: The paths of the Network are: B,C, D and A,E,F

The numbers inside the nodes are the event numbers: The arrows must be directed (>) otherwise the whole diagram is meaningless.

Example 15.2

In the last example, if C is preceded by B and F is preceded by E; draw the Network diagram.

Solution

To draw the Network diagram, the use of dummy activities will be necessary.
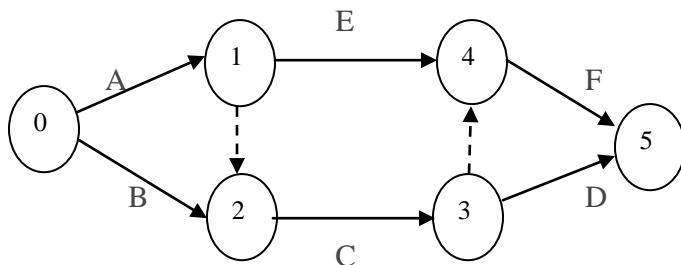
The path s of the Network are now.

A, E, F

A, Dummy, C, D

B, C, D

B, C, Dummy, F

15.2.1 Duration of Paths/Critical Path

The following example will be used to explain both the duration along paths of a Network and the critical path.

Example 15.3

The following activities with their durations are necessary to complete a project.

Activity	Preceding Activity	Duration (Weeks)
A	-	4
B	A	6
C	A	3
D	B, C	5
E	A	7
F	D	1
G	D	4
H	E, F	6

a) Draw the Network diagram for the project

b) Identify all the paths and calculate the duration for each path

c) Identify the critical path and the critical activities

Solutions

a)
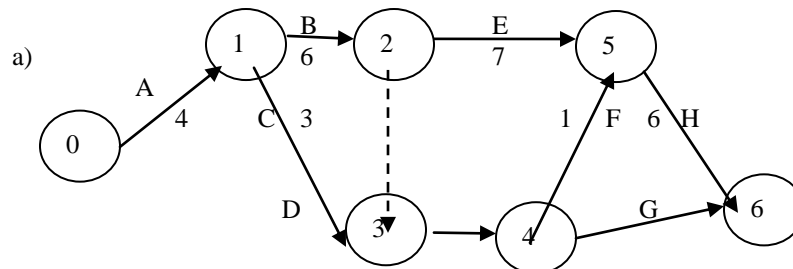
The activities and their durations are as shown e.g. for activity B we have

$$\xrightarrow[\text{6 (duration)}]{\text{B (activity)}}$$
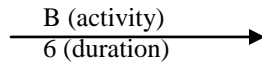
b)

Path	Duration
A, B, E, H	$4 + 6 + 7 + 6 = 23$ weeks
A, C, D, G	$4 + 3 + 5 + 4 = 16$ weeks
A, C, D, F, H	$4 + 3 + 5 + 1 + 6 = 19$ weeks
A, B, Dummy, D, G	$4 + 6 + 0 + 5 + 4 = 19$ weeks
A, B, Dummy, D, F, H	$4 + 6 + 0 + 5 + 1 + 6 = 22$ weeks

c)   The critical path is A, B, E, H with duration of 23 weeks.

The critical activities are A, B, E H

d)   The duration of the project is 23 weeks

15.3   Earliest Start Time (EST), Latest Start Time (LST) and Floats

The **Earliest Start Time** is the earliest possible time that a succeeding activity can start while the Latest Start Time is the latest possible time that a preceding activity must be completed in order not to increase the project duration.

The **float** of an activity is the amount of spare time associated with the activity. Only non-critical activities have floats. These activities can start late and/or take longer time than specified without affecting the duration of the project.

The three types of float are:

- **Total float** is the amount of time by which the duration of an activity could be extended without affecting the project duration.

- **Free float** is the amount of time by which the duration of an activity can be extended without affecting the commencement of subsequent activities .

- **Independent float** is the amount of time by which the duration of an activity can be extended without affecting the time available for succeeding activities or preceding activities.

15.3.1 Calculation of ESTs, LSTs and Floats

The following example will be used for all the calculations

Example 16.4

Use the diagram in example 3 to calculate

(a) (i) The ESTs

(ii) LSTs for all activities

(b) floats for all activities

Solution

Key: Figures in the nodes are Identified as:
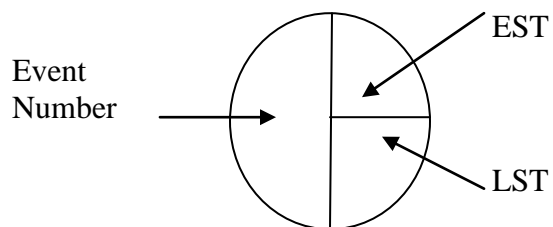
(a) *Calculation of ESTs (Forward pass)*

- Start with the event zero and work forward (hence the term "forward pass") through the Network.
- The EST of a head event is the sum of the EST of the tail event and the duration of the linking
  activity
- Where two or more activities have the same head event, the largest time is chosen
- The EST in the finish event gives the project duration.

Thus, the ESTs are as shown
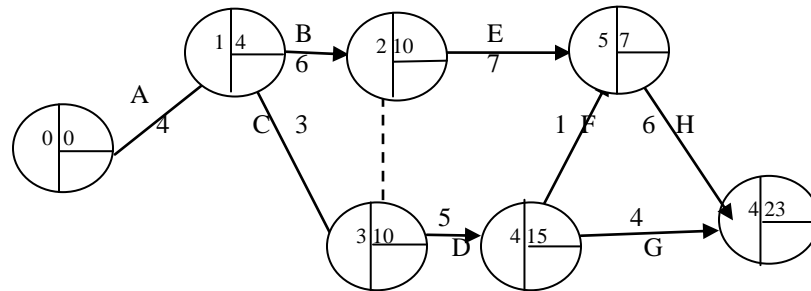
Explanation

Event Nos	EST
1	$0 + 4 = 4$
2	$4 + 6 = 10$
3	either 4+3= 7 or 10+0 (Dummy) = 10, pick 10 because it is the larger of the two i.e. 10> 7
4	10+5 =15
5	either $10 + 7 = 17$ or $15 + 1 = 16$, pick 17 because it is the larger of the two i.e. $17 > 16$
6	either $15 + 4 = 19$ or $17 + 6 = 23$, pick 23 because it is the larger of the two i.e. $23 > 19$

Calculation of LSTs (Backward pass)

- Start at the finish event using its EST and work backwards (hence the term "backward pass")
  through the network.
- Deduct each activity duration from the previous LST
- Where there are two or more LSTs, select the shortest time

Thus, the LSTs are as shown in the positions as in the key.
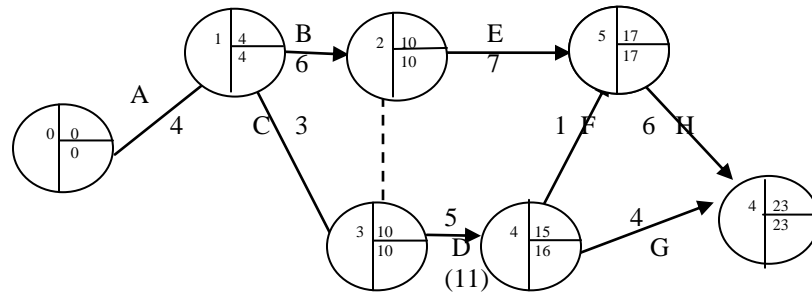
Note: The bracketed numbers along some activities showed the activities with two or more LSTs

e.g.

$10 = 10 - 0;$ $19 = 23 - 4;$ $16 = 17 - 1;$ $11 = 16 - 5$

Explanation

Event Nos	LST
6	23 (EST)
5	$23 - 6 = 17$
4	either $23 - 4 = 19$ or $17 - 1 = 16$ (the smaller)
2	$10 - 6 = 4$
3	either $16 - 5 = 11$ or $10 - 0 = 10$, pick 10 (the smaller)
1	$10 - 6 = 4$
0	$4 - 4 = 0$

Note:

The ESTs and LSTs along the critical path are equal.

This is another way of identifying the critical path. In the last example, events 0, 1, 2, 5 & 6 are on the critical path since the relevant ESTs and LSTs are equal.

b) **Calculation of floats**

To calculate floats, we need to know two other values apart from the ESTs and the LSTs.

These are the Earliest Finish Time (EFT) and the Latest Finish Time (LFT) which are read directly from the head node of an activity.

These values are shown below;

Note that the EFT is just smaller of the two values while the LFT is the larger.

Consequently, the EST of an activity is the EFT of the preceding activity and the LST of an activity is the LFT of the preceding activity.

Activity	EFT	LFT
A	4	4
B	10	10
C	10	10
D	15	16
E	17	17
F	17	17
G	23	23
H	23	23

All the values are now combined to calculate the float.

Explanation: Both the EST and LST for an activity are read directly from the head node of the activity. EFT is the smaller of the two numbers.

- For activities A, B, C, E, G and H, the ESTs and LSTs are equal e.g. head node of activity B is
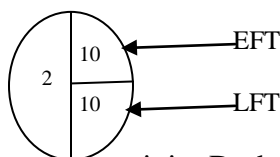
•For activity D, the head node is
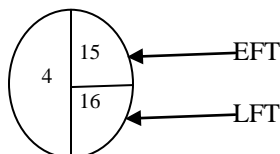
Activity	EFT	LFT	EST	LST	Duration D	Total Float LFT-EST-D	Free Float EFT-EST-D	Independent Float EFT-LST-D
A	4	4	0	0	4	0	0	0
B	10	10	4	4	6	0	0	0
C	10	10	4	4	3	3	3	3
D	15	16	10	10	5	1	0	0
E	17	17	10	10	7	0	0	0
F	17	17	15	16	1	1	1	0
G	23	23	15	16	4	4	4	3
H	23	23	17	17	6	0	0	0

As could be seen, all floats of the critical activities are zeros confirming the fact that only non-critical activities have floats.

An alternative method to <u>activity-on-arrow</u> diagram is <u>activity-on node</u> diagram which is also known as precedence diagram.

15.4 Summary

The concept of Network Analysis has been discussed in details. Steps to follow in drawing the Network diagram for a project were enumerated and explained. Identification of critical path and calculation of floats were treated as well. Practical applications of Network work analysis were also treated.

15.5 Multiple-Choice and Short-Answer Questions

1. In a network diagram, two different activities must not have
   A. the same duration
   B. the same starting nodes
   C. the same finishing nodes
   D. the same starting and finishing nodes
   E. the same preceding activity
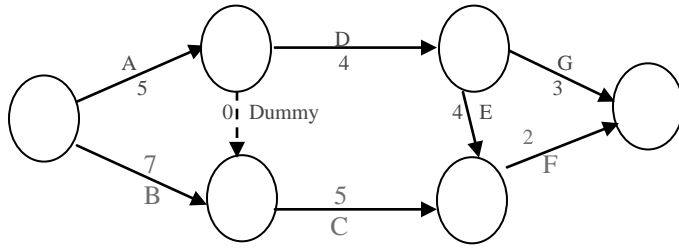
The diagram above shows the network diagram to complete a project. The durations are in weeks.

Use this to answer questions (2) and (3).

2. The shortest time within which the project can be completed is:
    A.   12 weeks
    B.   14 weeks
    C.   10 weeks
    D.   11 weeks
    E.   15 weeks

3. The number of paths of the network is:
    A.   2
    B.   4
    C.   3
    D.   5
    E.   6

4. In a network diagram, the critical activities have:
    A.   free float only
    B.   independent float and free float
    C.   independent float only
    D.   total float
    E.   no float.

5.   The primary objective of a Network Analysis is to ……………… a project within the ……………… time

6.   A dummy activity neither consumes nor …………………………. :   .

7.   The critical path of a network is the path with the ……………….. duration.

8.   A Network diagram can be drawn using events on the arrows method only
            Yes   or No?

9.   The calculation of the latest start time (LST) is also referred to as …………   pass

10.   The float of an activity is the amount of ………….. time associated with the activity.

Answers to Chapter Fifteen

1. D

2. The paths with their durations are:

A, D, G	12 weeks
A, Dummy, C, F	12 weeks
A, D, E, F	15 weeks
B, C, F	14 weeks (E)

3. See 2,   4 paths   (B)

4. E

5. Complete, minimum

6. Time, resources

7. Longest

8. No

9. Backward

10. Spare

CHAPTER 16

SOLVING TRANSPORTATION PROBLEMS

16.0 LEARNING OBJECTIVES

At the end of the chapter, students and readers should be able to

- Know that transportation problem is a special class of linear programming problem;
- Understand the concept of balanced and unbalanced transportation model;
- Obtain the initial basic feasible solution using the following methods: northwest corner rule, least cost, Vogel's approximation

16.1 INTRODUCTION: Nature of Transportation Model

Transportation model is a special class of linear programming problem in which the objective is to transport or distribute a single commodity or goods from various sources/origins to different destinations at a minimum total cost.

An example of a transportation problem is given below:

Balanced Transportation Problem

Destination					
	1	2	3	4	Supply
Origin 1	22	19	26	22	85
2	30	27	22	28	60
3	25	23	37	24	48
Demand	51	72	36	34	193

From the above mathematical model, we could see that a transportation problem is a special class of a linear programming problem.

Also, a transportation problem is said to be balanced if the total quantity demanded by the destination = total quantity available at the origin, otherwise the transportation problem is said to be unbalanced.

When the problem is not balanced, there is need to create a dummy origin (row) or destination (column) for the difference between total supply and demand with zero cost in order to create the balance.

A typical example of balancing an unbalanced problem is given below:

Destination					
Origin	1	2	3	4	Supply
1	18	16	20	21	40
2	16	14	16	20	60
3	20	23	21	22	40
Dummy	0	0	0	0	30
Demand	40	80	30	20	170

In this table, the given total supply is 140 while the total demand is 170. Therefore, a dummy row is created for the difference of 30 (as supply) with zero costs in order to create the balance.

Solving a transportation problem involves choosing a strategy of shipping programme that will satisfy the destination and supply requirements at the minimum total cost. Hence, there is the need to apply the Transportation methods of solving the model.

16.2   METHODS OF OBTAINING INITIAL BASIC FEASIBLE SOLUTION FOR A TRANSPORTATION PROBLEM

The solution to Transportation problem involves two phases. The first phase is to obtain the initial basic feasible solution: while obtaining the optimal solution is the second phase. In this chapter, we shall concentrate on this first phase, there are various methods of obtaining the initial basic feasible solution. These methods are

a) North-West corner Rule,

b) Least cost Method, and

c) Vogel's Approximation (or penalty) Method.

Note that any of the three methods must satisfy the following conditions in order to obtain the initial solution:

a) The problem must be balanced as discussed above

b) The number of cell's allocation must be equal to $m + n - 1$, where m and n are numbers of rows and columns respectively.

Any solution satisfying the two conditions is termed "Non-degenerate Basic Feasible solution" otherwise, it is called "Degenerate solution".

(a)   North-West Corner Rule (NCWR)

The method is the simplest but most inefficient as it has the highest total transportation cost in comparison to all other methods. The main reason that can be attributed to this is that the method does not take into account the cost of transportation for all the possible alternative routes.

These steps needed to solve a transportation problem by NCWR are:

Step 1: Begin by allocating to the North-West cell (i.e variable X_{11}) of transportation matrix the allowable minimum of the supply and demand capacities of that cell. i.e. Min (a_1, b_1)

Step 2: Check if allocation made in the first step is equal to the supply (demand) available at the first row (column), then cross-out the exhausted row (column) so that no further assignment can be made to the said row (column). Move vertically (horizontally) to the next cell and apply Step 1.

Step 3: Step 2 should be continued until exactly one row or column is left uncrossed in the transportation matrix. Then make allowable allocation to that row or column and stop. Otherwise, return to Step 1.

(b)   Least Cost Method (LCM)

In this method the cheapest route is always the focus for allocation. It is a better method compared to NCWR because costs are considered for allocation. The algorithm is stated thus:

Step 1: Assign as much as possible to the smallest unit cost (ties are broken arbitrarily). Also bear in mind the idea of allowable minimum of supply and demand capacities as done in NWCR.

Step 2: Cross-out the exhausted row or column and adjust the supply and demand accordingly. If both row and column are exhausted simultaneously, only one is crossed-out (in order to avoid degenerating case).

Step 3: Look for the smallest cost in the uncrossed row or column and assign the allowable quantity. Repeat this process until left with exactly one uncrossed row or column.

(c)   **Vogel's Approximation Method (VAM)**

VAM, which is also called penalty method, is an improvement on the LCM method that generates a better initial solution.

It makes use of opportunity cost (penalty) principles in order to make allocation to various cells by minimizing the penalty cost. The steps in this method are:

Step 1: Compute for each row (column) the penalty by subtracting the smallest unit from the next smallest unit cost in the same row (column).

Step 2: Select the row or column with the highest penalty and then allocate as much as possible to the variable with least cost in the selected row or column. Any ties should be handled arbitrarily.

Step 3: Adjust the supply and demand and cross-out the exhausted row or column. If a

row and a column are simultaneously exhausted, only one of the two is crossed-out and the other assigned zero supply (demand).

Step 4: Compute the next penalties by considering uncrossed rows (columns) and go to step 2.

Step 5: Exactly when one row (column) remains uncrossed, then allocate the leftover and stop.

Example 16. 1

SAO company has 3 plants or locations (A, B, C) where its goods can be produced with production capacity of 50, 60, 50 per month respectively for a particular product. These units are to be distributed to 4 points (X, Y, W, Z) of consumption with the demand of 50, 70, 30 and 10 per month respectively.

The following table gives the transportation cost (in Naira) from various plants to the various points of consumption:

	Destination			
	X	Y	W	Z
Source/Plant A	21	18	27	22
B	19	18	24	20
C	24	25	28	25

Obtain the Initial Basic Feasible solution by

(a) NWCR   (b) LCM   (c) VAM

SUGGESTED SOLUTION 16.1

The NWCR algorithm applied to this problem gives the following table:

Table:

Destination					
Plant	X	Y	W	Z	Supply
A	(50) 21	18 / 0	27	22	50 0
B	19	(60) 18	24	20	60 0
C	24	25 / (10)	28 / (30)	25 / (10)	50 40 10 0
Demand	50 / 0	70 / 10 / 0	30 / 0	10 / 0	160

Total Cost = 50 (21) + 60 (18) + 10 (25) + 30 (28) + 10 (25)
= ₦3470

Note that in all transportation tables, figures in small boxes are the transportation costs from destinations to destinations while the circled figures represent the allocation.

Explanation to the above allocation

Beginning from North-West (i.e Cell X_{11}) corner

- Allocate 50 to cell X_{11} in order to satisfy the minimum of demand and supply capacities. Zero balance is left for both demand and supply. Therefore, row 1 and column 1 are crossed out.

- Move to cell X_{22} and allocate 60. The supply balance is zero while the demand balance is 10. Therefore, row 2 is crossed out.

- Next move to cell X_{32} and allocate 10 giving the balance of zero for demand and 40 for the supply. Column 2 is therefore crossed out.

- Then move to cell X and allocate 30 to exhaust the demand and having 10 balance for supply. Therefore, column 3 is crossed out.

- Finally, 10 is allocated to cell X_{34}

(b)   Use the LCM to determine the Initial Basic Feasible solution

Using the LCM algorithm, gives the following table:

Plant	X		Y		W		Z		Supply
					Destination				
A		21	50	18		27		22	~~50~~ 0
B	40	19	20	18		24		20	~~60~~ ~~40~~ 0
C	10	24		25	30	28	10	25	~~50~~ ~~40~~ ~~30~~ 0
Demand	~~50~~ ~~10~~ 0		~~70~~ ~~20~~ 0		~~30~~ 0		~~10~~ 0		160

Total Cost = 50 (18) + 40 (19) +20 (18)+ 10 (24) + 30 (28) + 10 (25)

= ₦3350

Explanation to the above allocation

- The lowest cost cell is X_{12} with 18. Then 50 is allocated to that cell in order to satisfy minimum of the demand and supply. Zero balance is left for supply while that of demand is 20. Therefore row one is crossed out.

- For the remaining un-crossed cells, cell X_{22} has lowest cost of 18. Then assign 20 to this cell in order to give zero balance for demand and 40 balance for supply. Therefore, column two is crossed out.

- The next cell with lowest cost, for the uncrossed cells, is X_{21} with 19. Allocate 40 to this cell in order to give zero balance to supply and 10 to demand. Therefore cross out row two.

- Move to cell X_{31} because it has lowest cost. (24) among the uncrossed cells and then allocate 10. The balance of zero is obtained for demand while that of supply is 40. Column one is exhausted and it is crossed out.

- Lastly, allocate 30 and 10 respectively to the remaining two cells X_{33} and X_{34}

(c)   Use the VAM algorithm to obtain the initial basic feasible solution

Applying the steps of VAM, gives the following tables:

Plant	X		Y		W		Z		Supply
						Destination			
A		21	50	18		27		22	50 0
B	40	19	20	18		24		20	60 40 0
C	10	24		25	30	28	10	25	50 40 30 0 / 50 40 10 0
Demand	50 10 2		70 20 0		30 0		10 0		160

Penalty table

Iteration	Rows			Columns				Allocation
1	3 1 1			2 0 3 2				$X_{12} = 50$

(2)

cccxx

NOTE: This is a work in progress. All topics in the syllabus are covered but editing for necessary corrections is in progress.
Thanks.

2	*	1	1	5		4	5	$X_{22} = 20$
3	*	1	1	⑤	*	4	5	$X_{21} = 40$
4	*	*	1	-	*	-	-	$X_{31} = 10$
								$X_{33} = 30$
								$X_{34} = 10$

The principle of obtaining the difference between the two least costs along row (column) is applied throughout the penalty table in order to make allocation as stated in step 2

In the Penalty Table, the circled figures represent highest penalty in each iteration; * represents the crossed-out row or column, and (-) stands for penalty not possible.

Note that in the penalty table, 3 is obtained under row A (in iteration 1) by the difference to other rows (B and C) to obtain 1 and 1 respectively.

For the column also, 2 is obtained under column x (in iteration 1) by the difference of costs 21 and 19.

Total cost = 50 (18) + 40 (19) + 20 (18) + 10 (24) + 30 (28) + 10 (25)

$$= ₦3350$$

Example 12.2

A company with three factories (X, Y, Z) and five warehouse (A, B, C, D, E) in different locations has the transportation costs (in Naira) from factories to warehouses. Factory capacities and warehouse requirements are stated below:

Factories	Warehouses					Factory Capacities
	A	B	C	D	E	
X	5	8	6	4	3	800
Y	4	7	8	6	5	600
Z	8	4	7	5	6	1100
Warehouse Requirements	350	425	500	650	575	2500

Determine the initial basic feasible solution by

a) North West Corner Rule (NWCR)

b) Least Cost Method (LCM)

c) Vogel's Approximation Method (VAM)

SUGGESTED SOLUTION 16. 2

(a) **North West Corner Rule**

cccxxi

Factories	Warehouses					Factory Capacities
	A	B	C	D	E	
X	5 ⭕350	8 ⭕425	6 ⭕25	4	3	~~800~~ ~~450~~ ~~25~~ 0
Y	4	7	8 ⭕475	6 ⭕125	5	~~600~~ ~~125~~ 0
Z	8	4	7	5 ⭕525	6 ⭕575	~~1100~~ ~~575~~ 0
Warehouse Requirements	~~350~~ 0	~~425~~ 0	~~500~~ ~~475~~ 0	~~650~~ ~~525~~ 0	~~575~~ 0	2500

Total Cost = (350 x 5) + (425 x 8) + (25 x 6) + (475 x 8) + (125 x 6) + (525 x 5) + (575 x 6)
= ₦15,925

Explanation to the above allocation

Beginning from North-West (i.e cell X_{11}) corner

- Allocate 350 to cell X_{11} in order to satisfy the minimum of factory capacity and warehouse requirement. Zero balance is left for warehouse requirement while that factory capacity along X_{11} is 450. Therefore, column one is crossed out.

- Move to cell X_{12} and allocate 425. The warehouse balance along cell X_{12} is zero while we have the balance of 25 for the factory capacity. Therefore, column two is crossed out.

- Next, move to cell X_{13} and allocate 25 giving the balance of zero for factory capacity and 475 for the warehouse along that cell (i.e X_{13}). Row one is therefore crossed out.

- Move to cell X_{23} to allocate 475. It then gives a balance of zero for the warehouse while that of factory capacity is 125. Column three is crossed out.

- We also move to cell X_{24} and allocate 125 to that cell (X_{24}) giving the balance of zero for the factory capacity and 525 for the warehouse. Then cross-out row two.

- Finally, allocate 525 and 575 respectively to the remaining two cells X_{34} and X_{35}.

(b) Least Cost Method

Factories	Warehouses					Factory Capacities
	A	B	C	D	E	
X	5	8	6	4 225	3 575	~~800~~ 225 0
Y	4 475	7	8 125	6	5	~~600~~ 125 0
Z	8	4 425	7 250	5 425	6	~~1100~~ 425 ~~250~~ 0
Warehouse Requirements	~~350~~ 0	~~425~~ 0	~~500~~ ~~250~~ 0	~~650~~ ~~525~~ 0	~~575~~ 0	2500

Total Cost   = (225 x 4) + (575 x 3) + (350 x 4) + (250 x 8) + (425 x 4) + (250 x 7) + (425x5)

= ₦11,600

Explanation to the above allocation

The cell that has lowest cost is X_{15}. Allocate 575 to that cell giving the balance of zero for warehouse requirement and 225 for factory capacity. Column five is crossed out.

- For the remaining uncrossed cells, cells X_{14}, X_{21} and X_{32} have the lowest cost of 4. Arbitrarily, cell X_{32} is allocated with 425 giving the balance of zero for warehouse requirement and 675 for the factory capacity. Column two is crossed-out.

- Next, cell X_{21} is allocated 350 giving the balance of zero for warehouse requirement and 250 for the factory capacity. Column 1 is also crossed-out. We allocate 225 to cell X_{14} to give the balance of zero for factory capacity and 425 for the warehouse requirement. Row one is crossed out.

- The left-over cells are X_{23}, X_{24}, X_{33} and X_{34}. We allocate 425 to cell X_{34} to give the balance of zero for warehouse requirement and 250 for the factory capacity. Therefore, column four is crossed out.

- Finally, the remaining two cells X_{23} and X_{33} are respectively allocated 250 each.

Vogel's Approximation Method

Factories	Warehouses					Factory Capacities
	A	B	C	D	E	
X	5	8	6	4 / 225	3 / 575	~~800~~ ~~225~~ 0
Y	4 / 350	7	8	6 / 250	5	~~600~~ ~~250~~ 0
Z	8	4 / 425	7 / 500	5 / 175	6	~~1100~~ ~~675~~ ~~250~~ 0
Warehouse Requirements	~~350~~ 0	~~425~~ 0	~~500~~ ~~250~~ 0	~~650~~ ~~425~~ ~~175~~ 0	~~575~~ 0	2500

Penalty table

	Rows			Columns					
Iteration	X	Y	Z	A	B	C	D	E	Allocation
1	1	1	1	1	(2)	1	1	2	$X_{32} = 425$
								(3)	$X_{15} = 575$
2	1	1	1	1	*	1	1		$X_{21} = 350$
3	1	(2)	2	1	-	1	1	*	$X_{14} = 225$
									$X_{24} = 250$
	(2)								$X_{33} = 500$
4		2	2	*	-	1	1	-	$X_{34} = 175$
	*	(2)	2	-	-	1	1	-	
5									

The principle of obtaining the difference between the two least costs along row (column) is applied throughout the penalty table in order to make allocation as stated in step 2.

In the penalty table, the circled figures represent highest penalty in each iteration.

(*) represents the crossed out row or column, and (-) stands for penalty not possible. Also, note in the penalty table that 1 is obtained under row X (in iteration 1) by the difference between 4 and 3. Similarly for row Y, 1 is obtained from the difference between 5 and 4. The difference computations continued in this manner for both rows and columns in each iteration.

Total cost = (225 x 4) + (575 x 3) + (350 x4) + (250 x 6) + (425 x 4) + (500 x 7) + (175 x 5)

= ₦11,600

16.3 SUMMARY AND CONCLUSIONS

This is a special class of linear programming problem in which the objective is to transport or distribute a single commodity from various sources to different destinations at a minimum cost. The transportation problem can only be solved if it is balanced. The condition under which a transportation problem is balanced is that total quantity demanded by the destination must be equal to the total quantity available at the origin. The initial basic feasible solution of a transport problem can be obtained by the following methods:

(a)   The North-West Corner Rule (NWCR);

(b)   Least Cost Method (LCM); and

(c)   Vogel's approximation Method (VAM).

16.4 MULTIPLE-CHOICE AND SHORT-ANSWER QUESTIONS

Pick the appropriate and correct answers to the following:

1.   The main objective of a transportation problem is to

A.   Transport goods to the supply points

B.   Allocate goods to people

C.   Distribute a single good from various sources to different destinations at minimum total cost

D.   Sell goods to consumers

E.   To bring goods nearer to the people

2. Which of the following is not a method of obtaining initial feasible solution of a transportation problem?

    A. NCWR

    B. Least Cost Method

    C. Vogel's approximation Method

    D. Stepping Stone Method

    E. Simplex Method

3. Least Cost Method is better than NCWR because it is

    A. Faster in computation and allocation

    B. Concerned with handling of computation and allocation

    C. Straight forward for allocation and computation

    D. Needed for cost for allocation and computation

    E. Having fewer iterations

4. The rim condition in transportation problem

    A. Is the addition of dummy to destination of source

    B. Means the total demand equals total supply

    C. Means that shipment to a dummy source represents surplus

    D. Means all of the above

    E. Is the absence of dummy

5. With the supply capacities: $a_1 = 120$, $a_2 = 70$, $a_3 = 60$, $a_4 = 80$, and demand capacities: $b_1 = 120$, $b_2 = 70$, $b_3 = 90$, $b_4 = 110$; determine the amount of dummy capacity to add to the source or destination.

    A. 60 (dummy) for supply (source)

    B. 60 (dummy) for demand (destination)

    C. 70 (dummy) for supply (source)

    D. 70 (dummy) for demand (destination)

    E. 50 (dummy) for supply (source)

6. The major difference between LCM and VAM is that VAM uses _____ costs

7. If m and n are numbers of origins and destinations respectively, the number of cells allocation that does not conform to m+n-1 will lead to _____ solution.

8. The necessary condition for transportation problem to be solvable is that it must be

    _____

9. Solution to a typical transportation problem involves _____ phases.

10. By Least Cost method, determine initial feasible solution for the following cost table (in naira)

16.4.2 **Multi-Choice and Short-Answer Questions: Solutions**

1. C

2. D

3. D

4. B

5. A

6. Penalty or opportunity

7. Degenerate

8. Balanced

9. Two

10. The table is not balanced. Hence, we make it balanced by creating a dummy destination for demand.

	Destination				
Origin	A	B	C	Dummy	Supply
I	4	30 5	20 6	20 0	~~70 50 20~~ 0
II	20 3	8 20	7	0	~~40 20~~ 0
III	5	6 60	8	0	~~60~~ 0

Demand	~~20~~ 0	~~30~~ 0	~~100~~ ~~80~~ ~~60~~ 0	~~20~~ 0	

Total Cost = 30(5) + 20(6)+20(0)+20(3)+20(7)+60(8)

= 150+120+0+60+140+480

= ₦950

CHAPTER 17

REPLACEMENT ANALYSIS

17.0 LEARNING OBJECTIVES

At the end of this chapter, students and readers should be able to know the

- Meaning and purpose of replacement theory;
- Concept and technique of replacement;
- Replacement policy of those items that wear out gradually;
- Replacement policy of the items that fail suddenly; and
- Difference between individual and group replacement policies

17.1 INTRODUCTION

In a real life, all equipments used in industries, military, and even at homes have a limited life span. By passage of time, these equipments can fail suddenly or wear out gradually. As these machines or equipments are wearing out, the efficiency of their functions continues to decrease with time and in effect it affects the production rates or economic/social benefits of the system. When the equipments become old or when they get to the wearing-out stage, they will definitely require higher operating costs and more maintenance costs due to repairing and replacement of some parts. And at that point, replacing them completely may be a better option.

Examples of the equipment's are transportation vehicles (such as cars, lorries or aircraft), machines used in industries, tires. These wear out gradually whereas highway tube lights, electric bulbs, and contact set, used by vehicles wear out suddenly.

The two essential reasons for the study of replacement analysis are to

a) ensure efficient functioning of the equipments
b) know when and how best the equipments can be replaced in order to minimize the total costs of maintaining them.

The two major replacing policies are

a) Replacement of equipments/items that deteriorate or wear-out gradually; and
b) Replacement of equipments/items that fail suddenly.

17.2   REPLACEMENT OF EQUIPMENTS/ITEMS THAT DETERIORATE OR WEAR-OUT GRADUALLY

The efficiency of equipments/items that deteriorate with time will be getting low. As the items deteriorate, gradual failure sets in. The gradual failure is progressive in nature and thus affects the efficiency and it results in

a) a decrease in the equipments production capacity;

b) increasing maintenance and operating costs; and

c) decrease in the value of the re-sale (or salvage) of the item

Due to these effects, it is reasonable and economical to replace a deteriorating equipment/item with a new one. As the repair and maintenance costs are the determining factors in replacing policy, the following two policies are to be considered.

a) **Replacement of items that Deteriorate and whose maintenance and repair costs increase with time, ignoring changes in the value of money during the period.**

This is a simple case of minimizing the average annual cost of an equipment when the maintenance cost is an increasing function time but the time value of money remains constant.

In order to determine the optional replacement age of a deteriorating equipment/item under the above conditions, the following mathematical formulae shall be used:

Let C   = Capital or purchase cost of the equipment/item

S   = Scrap (or salvage) value of the equipment/item at the end of t years

TC(t) = Total cost incurred during t years

ATC   = Average annual total cost of the equipment/item

n   = Replacement age of the equipment/item: i.e. number of equipment years the equipment/item at time t.

cccxxx

Case 1:   When time "t" is a continuous variable)

Here, we need to find the value of <u>n</u> that minimizes the total cost incurred during the first n years.

TC(t) = Capital cost – scrap value + <u>maintenance cost</u>

$$= C\text{-} S + \int f(t)\, dt \qquad \text{--------------------}17.2.1$$

But for average annual total cost, we have

$$\text{ATC}(n) = [1/n\ C - S + \int f(t)\, dt] \ldots\ldots\ldots\ldots\ldots\ldots..17.2.2$$

To obtain optimal value for n for which ATC(n) is minimum, we differentiate ATC (n) with respect to n and equate the result to zero so that.

$$f(n) = 1/n\ [\ C - S + \int f(t)\, dt], n \neq 0$$

$$= \text{ATC}(n) \ \text{--}17.2.3$$

Hence, the equipment/item is to be replaced when the average annual cost equals the current maintenance cost.

Case 2 (When time "t" is a discrete variable).

In a similar manner, the total cost incurred during n years is

$$TC(n) = C - S + \sum_{t=0}^{n} f(t) \ldots\ldots\ldots\ldots\ldots\ldots\ldots\ldots\ldots\ldots\ldots\ldots\ldots.17.2.4$$

while the average annual cost on item is given as

$$TC(n) = 1/n\ [C - S + \sum_{t=0}^{n}\ ] f(t)\ldots\ldots\ldots\ldots\ldots\ldots\ldots\ldots\ldots\ldots.17.2.5$$

$$\therefore \frac{d[TC(n)]}{dn} = \frac{-C}{n^2} + \frac{S}{n^2} - \frac{f(t)}{n^2} \ldots\ldots\ldots\ldots\ldots..17.2.6$$

NOTE: This is a work in progress. All topics in the syllabus are covered but editing for necessary corrections is in progress.

Thanks.

For this case, a tabular method for obtaining the value of n is usually used

Example 1

An owner of a grinding machine estimates from his past records that cost per year of operating his machine is given below:

Year	1	2	3	4	5	6	7	8
Operating Cost ($\mathbb{N}$)	250	550	850	1250	1850	2550	3250	4050

If the cost price is $\mathbb{N}$12,300 and the scrap value is $\mathbb{N}$250, when should the machine be replaced?

Solution

C = $\mathbb{N}$12,300 and S = $\mathbb{N}$250

Year of Service n	Running Cost ($\mathbb{N}$) f(n)	Cumulative Running Cost ($\mathbb{N}$) $\sum$f(n)	Depression Cost Price ($\mathbb{N}$) = C - S	Total Cost ($\mathbb{N}$) TC	Average Cost ($\mathbb{N}$) ATC(n)
Col 1	Col2	Col 3	Col 4	Col 3 + Col 4 = Col 5	Col 6 = Col 5/n
1	250	250	12050	12300	12300
2	550	800	12050	12850	6425
3	850	1650	12050	13700	4566.67
4	1250	2900	12050	14950	3737.50
5	1850	4750	12050	16800	3360
6	2550	7300	12050	19350	**3225 ***
7	3250	10550	12050	22600	3228.57
8	4050	14600	12050	26650	3331.25

It is observed from the table that the average annual cost ATC (n) is minimum in the sixth year.

Hence, the machine should be replaced at the end of sixth year of usage.

cccxxxii

Example 2

A company is considering replacement of a machine which cost ₦70000. The maintenance cost and resale values per year of the said machine are given below:

Year	1	2	3	4	5	6	7	8
Maintenance Cost (₦)	9000	12000	16000	21000	28000	37000	47000	59000
Resale Cost (₦)	40000	20000	12000	6000	5000	4000	4000	4000

When should the machine be replaced?

Solution

Year of Service n	Resale Value (₦) (S)	Purchase Price resale value (₦) C - S	Annual Maintenance Cost f(t)	Cumulation Maintenance Cost (₦) $\sum_{0}^{n} f(t)$	Total Cost (₦) $TC\left[C - S + \sum_{0}^{n} f(t)\right]$	Average Cost (₦)
Col 1	Col2	Col 3 =70000 – Col 2	Col 4	Col 5	Col 6 = Col 3 + Col 5	Col 7 = Col 6/n
1	40000	30000	9000	9000	39000	39000
2	20000	50000	12000	21000	71000	35500
3	12000	58000	16000	37000	95000	31666.67
4	6000	64000	21000	58000	122000	30500
5	5000	65000	28000	86000	151000	**30200 ***
6	4000	66000	37000	123000	189000	31500
7	4000	66000	47000	170000	236000	33714
8	4000	66000	59000	229000	295000	36875

From the table, it is observed that the average cost ATC(n) is minimum in the fifth year. Hence, the machine should be replaced by the end of 5th year.

b) **Replacement of items that deteriorate and whose maintenance cost increase with time with the value of money also changing with time**

This replacement policy can be seen as a value of money criterion. In this case, the replacement decision is normally based on the equivalent annual cost whenever we have time value of money effect.

Whenever the value of money decreases at constant rate, the issue of depreciation factor or ratio comes in as in the computation of present value (or worth). For example, if the interest rate on ₦100 is r percent per year, the present value (or worth) of ₦ 100 to be spent after n years will be:

$$D = \frac{100}{100 + r^n} \quad \dots\dots\dots\dots\dots\dots\dots\dots\dots\dots\dots\dots 17.2.6$$

where D is the discount rate or description value. With the principle of the discount rate, the critical age at which an item should be replaced can be determined.

Example 3

The yearly costs of two machines A and B when money value is neglected are is given below:

Year	1	2	3
Machine A (₦)	1400	800	1000
Machine B (₦)	24000	300	1100

If the money value is 12% per year, find the cost patterns of the two machines and find out which of the machines is more economical.

Solution

The discount rate per year = (d) = $\dfrac{1}{1 + 0.12}$ = 0.89

The discounted cost patterns for machines A and B are shown below:

For A: $\dfrac{1}{(1 + 0.12)^1}$

For B: $\dfrac{1}{(1 + 0.12)^2}$

Year	1	2		3		Total Cost (₦)
Machine A (Discounted Cost in ₦)	1400	800 × 0.89 712	=	1000 × $(0.89)^2$ 792.1	=	2,904.1
Machine B (Discounted Cost in ₦)	24000	300 × 0.89 267	=	1100 × $(0.89)^2$ 871.31	=	25,138.3

Decision: Machine A is more economical because its total cost is lower.

17.3   REPLACEMENT OF EQUIPMENTS/ITEMS THAT FAIL SUDDENLY

In a real life situation, we have some items that do not deteriorate gradually but fail suddenly. Goofdexamples of these items are electric bulbs, contact set, plugs, resistor in radio, television, computer, etc. Majority of items in this category are not usually expensive, but a quick attention or preventive replacement should be given in order not to have a complete breakdown of the system.

The items that experience sudden failure normally give desired service at variant periods. Service periods follow some frequency distributions which may be random, progressive and retrogressive.

There are two types of policies in the sudden failure category. These are

a)   The individual replacement policy; and

b)   The group replacement policy

a) **The individual replacement policy**

This is a policy in which an item is replaced immediately it falls. The life span of this item is uncertain and it is assumed that failure occurs only at the end of its life span (say period t). Therefore, an establishment of probability distribution for the failure (through the past experience) is required for the problem. This will enable us to find the period that minimizes the total cost involved in the replacement.

Example 4

The computer sets that are used in a company have resistors with a life span of five months. The failure rates (in percentages) of these resistors are given in the following table:

Months	1	2	3	4	5
Percentage Failures	10	30	35	20	05

Given that 798 resistors are fixed for use at a time, each resistor costs ₦14, if group replacement and ₦ 60 if replacements is done individually, determine the cost of individual monthly replacement.

The following steps are to be taken:

a) Compute the cost the cost of monthly individual replacement which is represented by C i.e. C = RK, where R is the average number of monthly replacements, K is the cost per item.

b) If it is not directly possible to have R, it can be computed or obtained from the following relationship:

R = $\dfrac{\text{Total number of items used}}{\text{Average life span of the item}}$

i.e. R = $\underline{N}$, where N and t represent total number of items used and average life span of item respectively t

we are to calculate t, the average life span of the resistor, as follows:

Monthly (X_i)	Percentage failure	Probabilities (P_i)	P_iX_i
1	10	0.10	0.10
2	30	0.30	0.60
3	35	0.35	1.05
4	20	0.20	0.80
5	05	0.05	0.25
	100	1.00	2.80

From the above table

$$t = \frac{\sum P_i X_i}{\sum P_i} = \frac{2.8}{1} = 2.80 \text{ months}$$

Average number of monthly replacements is equal to

$$R = \frac{N}{T} = \frac{798}{2.8} = 285$$

.: The individual cost replacement is equal to

$$C = RK$$
$$= 285 \times 60$$
$$= ₦17,100$$

(b) **The Group Replacement Policy**

This is a policy where an en-masse replacement of items is made. It is cheaper and safer to apply the policy when there is a large number of identical items which are more likely to fail within a particular time period.

The usual practice in the group replacement policy is to fix a time interval during which replacement can be made, that is, we make a replacement of all items at a fixed interval of time period t, whether the items failed or not, and at the same time to replace the individual failed items during fixed interval.

Next, we ensure that group replacement is made at the end of tth period is more than the average cost per unit time through the end of t period.

Example 5

Use the data in example 4 to determine the

a) Best interval period between group replacement

b) Cost of group replacement

Solution

Given that N = 798 and the computed P_i(I = 1, 2, …….5) is shown as

P_1	P_2	P_3	P_4	P_5
0.10	0.30	0.35	0.20	0.05

Let N_i represent the number of items replaced at the end of the ith month and $N_0 = 798$

Therefore, we have the following replacements in the subsequent months:

$N_0 = N_0 = 798$ (initial)

At the end of 1st Month, $N_1 = N_0P_1 = 798 \times 0.10 = 79.8 \approx 80$

At the end of 2nd Month, $N_2 = N_0P_2 + N_1P_1 = 798 \times 0.3 = 80 \times 0.1 = 247$

At the end of 3rd Month, $N_3 = N_0P_3 + N_1P_2 + N_2P_1 = 798 \times 0.35 = 80 \times 0.30 + 247 \times 0.10 = 328$

At the end of 4th Month, $N_4 = N_0P_4 + N_1P_3 + N_2P_2 + N_3P_1$

$$= 798(0.20) + 80(0.35) + 247(0.30) + 328(0.10)$$

$$= 295$$

At the end of 5th Month, $N_5 = N_0P_5 + N_1P_4 + N_2P_3 + N_3P_2 + N_4P_2$

$$= 798(0.05) + 80(0.20) + 247(0.35) + 328(0.30) + 295(0.1)$$

$$= 270$$

At the end of 6th Month, $N_6 = N_0P_6 + N_1P_5 + N_2P_4 + N_3P_3 + N_4P_2 + N_5P_1$

$$= 798(0) + 80(0.05) + 247(0.20) + 328(0.35) + 295(0.30) + 270(0.10)$$

$$= 284$$

From the computed values of N_i, we could see that the number of resistors failing each month increases till third month, then decreases and again increases from sixth month. Hence, N_i will oscillate continuously until the system reaches a steady state.

Now let's obtain the Total cost of group replacement (TC) by:

TC = Number of group replacement x Group Cost + Number of individual replacement x individual Cost.

End of Month	Total Cost of group replacement	Average Cost per Month
1	798 x 14 + 80 x 60 = 15972	15972
2	798 x 14 + 60(80 + 247) = 30792	**15396***
3	798 x 14 + 60(80 + 247 + 328) = 50472	16824
4	798 x 14 + 60(80 + 247 + 328 + 295) = 68172	17043
5	798 x 14 + 60(80 + 247 + 328 + 295 + 270) = 84372	16874

a) From the table, it is the second month that we have the average minimum cost. Hence, it is optimal to have group replacement after every two months.

b) The cost of the group replacement is ₦15,396

17.4   SUMMARY AND CONCLUSIONS

The essential reasons for the study of replacement analysis are to

- ensure efficient functioning of the equipments
- know when and how best the equipments can be replaced in order to minimize the total costs of maintaining them.

The following replacing policies are considered:

- replacement of equipments that deteriorate or wear-out gradually. Examples of these items include vehicles (such as cars, lorries), machines (used in industries) and tires. The maintenance and repair costs increase with time.
- Replacement of equipments/items that fail suddenly. Examples of these items include electric bulbs, contact set, plugs, resistors in radio and television, etc.

Two types of replacement policies under sudden failure are used.

- The individual replacement policy where an item is replaced immediately it fails.
- Group replacement policy where an en-masse replacement of items is made.

17.5   MULTIPLE COICE AND SHORT-ANSWER QUESTIONS

1. Which of the following is the reason for the study of replacement theory?
   A. To ensure efficient functioning of the equipments
   B. To know when and how best the equipments can be replaced
   C. To minimize the costs of maintenance
   D. (a) and (b) only
   E. (a), (b) and (c)

2. Which of the following is a policy in the replacement of equipments or items that fail suddenly?
   A. Gradual replacement policy
   B. Individual replacement policy
   C. Group replacement policy
   D. (b) and (c) only
   E. (a) and (b) only

3. Which of the following is not a resulting effect of gradual failure or deterioration of items?
   A. The output of the equipment
   B. Its production capacity
   C. The maintenance and operating costs
   D. The value of the re-sale price of the item
   E. The efficiency of the equipment

Use the following information to answer the next 3 questions
Given the following failure rates of certain item:

Month	1	2	3	4	5	6	7	8
Probability of failure	0.05	0.08	0.12	0.18	0.25	0.20	0.08	0.04
Cumulative probability	0.05	0.13	0.25	0.43	0.68	0.88	0.96	1.00

cccxl

If the total number of items is 1000 and the individual and group costs of replacement are ₦2.25 and 60 kobo per item respectively, assuming a replacement of all items simultaneously at fixed intervals and the individual items replace as they fail. Then

4. The average number of failures per month is approximately _____

5. The average cost of individual replacement is _____

6. The best interval between group replacement is _____

7. At the old age of an operating machine, state the reason why it will definitely require higher operating costs and more maintenance costs

8. The two common replacing policies are replacement of equipment items that:

   i.   _____

   ii.   _____

9. State any two consequences of equipments/items that deteriorate with time having its efficiency getting low.

10. State two characterizing features of the policy that governs an equipments that is replaced immediately it fails.

MULTIPLE-CHOICE and Short-Answer Questions - Solutions

1. E

2. D

3. B

4. Expected value $= (1 \times 0.05) + (2 \times 0.08) + (3 \times 0.12) + \ldots \ldots (8 \times 0.04)$

   $= 0.05 + 0.16 + 0.36 + 0.72 + 1.25 + 1.2 + 0.56 + 0.32$

   $= 4.62$

   .: Average number of failures per month $= 1000/4.62 \approx 216$

5. Average cost of individual replacement $= ₦(216 \times 2.25)$

$$= ₦486$$

6. Minimum cost of group replacement per month $= ₦208.30$ in the 3rd Month

7. The reason is due to repairing and replacement of some parts

8. i)   Deteriorate or wear-out gradually and

   ii)   fail suddenly

9. i)   A decrease in its production capacity

   ii)   The increasing maintenance and operating costs

   iii) Decrease in the value of the re-sale price (for salvage) of item

10. i)   item's life span is uncertain

    ii) it is as summed that failure occurs only at the end of its life span

EXERCISES

Question 1.

Marks scored by 50 students of the department of Accountancy in Mathematics are given below:

36	54	12	23	53	54	14	45	50	38
24	41	28	36	21	12	19	49	30	15
32	62	17	61	09	20	42	51	42	53
34	43	24	32	37	55	27	16	20	21
45	06	19	47	27	45	50	32	51	42

a.   Use The intervals 01 – 10, 11 – 20, 21 – 30, ……., etc to construct the frequency distribution of the data above.

b.   Construct a table of cumulative frequency

c.   Construct the ogive and histogram of the frequency data

d.   Calculate the mean, median and mode of the frequency data

e.   Compute the coefficient of variation of the frequency data

Question 2:

The following table gives the distribution of students by age grouping:

Age (in years)	0 - 10	10 - 20	20 - 30	30 - 40	40 - 50	50 - 60	60 - 70	70 – 80
No of Students	8	7	10	17	9	8	5	6

From the table calculate the

    a. Mean deviation

    b. Variance

    c. Semi-interquantile range

    d. Coefficient of variation

Question 3

Suppose a bag contains 7 white and 9 black balls. A ball is drawn at random. What is the probability that the ball is black?

    i. Find the probability of having black ball?

    ii. Find the probability of having white ball?

Question 4

From past experience, a stock broker finds that 70% of the telephone calls he receives during the business hours, are orders and the remaining are for other business. What is the probability that out of first 8 telephone calls during a day

    i. exactly 5 calls are order calls?

    ii. at least 6 are order calls?

Question 5:

The marks obtained (over 10) by ten students in Mathematics (x) and Accounts (y) are given below

Candidates	A	B	C	D	E	F	G	H	I	J
Accounts (x)	7	4	10	3	6	8	5	9	1	2
Accounts (y)	6	7	9	1	4	10	3	8	5	2

    a. Find the product moment correlation coefficient

    b. Obtain the rank correlation coefficient

Question 6

Fit a least squares line to the data in the table below assuming (a) x as independent variable and (b) x as dependent variable, (c) find y if x = 17, 18 and x if y = 13, 14

X	3	5	6	8	10	11	13	16
Y	2	3	5	5	6	8	9	10

Question 7:

The table below shows the daily expenses (₦'000) of a company for four weeks with 5 working days

Days of the week

Week	Mon	Tue	Wed.	Thur	Fri
1	13	17	23	34	42
2	32	32	47	51	54
3	46	50	57	64	71
4	63	68	71	75	83

a. Draw the time - plot of the data and comment on time series components displayed.

b. Determine five - day moving averages.

c. Compute the seasonal variation.

Question 8:

The following table shows the unit prices and quantities of an item required by a family in Ibadan during the period of years 2002 and 2006.

Commodity	2002		2006	
	Unit Price (N'00)	Quantity	Unit Price (N'00)	Quantity
Milk (tin)	0.60	24	0.80	30
Egg (create)	0.50	2	0.70	4
Butter (tin)	0.25	4	0.40	6
Bread (loaf}	1.10	80	1.30	100
Meat (kg)	1.12	100	1.50	120

From the table using 2002 as the base year, determine

a. Laspeyre index   b. Paasche index   c. Fisher index, and

d. Marshall Edgeworth index.

Question 9:

The following table shows the profit of 10 randomly selected small scale enterprises in Ibadan, Nigeria.

Company	J	K	L	M	N	O	P	Q	R	S
Profit (₦m)	16	6	14	18	10	12	20	22	24	18

a.   Estimate the mean profit of all the enterprises and its standard error.

b.   Test the hypothesis that the sample mean is significant or not to the population mean of 15 at 5% significance level.

Question 10:

a.

Class Interval	Frequency
5 – 9	4
10 – 14	4
15 – 19	10
20 – 24	8
25 – 29	17
30 – 34	20
35 – 39	16
40 – 44	15
45 – 49	12
50 – 54	<u>14</u>
	<u>120</u>

Using the information above, calculate

D_2, D_3, P_{45}, P_{55}

b.   If all possible samples of size two are drawn with replacement from population consisting of numbers 12, 13, 17, 18, determine the

    I.   Mean of the population (u).

    II.   Standard deviation of the population;

    III.   The mean of the sampling distribution of mean ($\bar{x}$).

    IV.   The standard deviation of the sampling distribution or mean.

Question 11:

(a)   ALLAHDEY company spent ₦1.6m to set up its production machinery. Each item costs ₦5,000 to produce while it is sold for ₦30,000

    You are required to calculate

    i.   the cost of producing 500 items

    ii.   the profit on the 500 items.

    iii.   the break-even quantity

    iv.   the number of items to be produced if a profit of W21.52m is targeted.

(b)   The project development department of DUJASEYEB ventures has estimated the cost function of the firm to be

$$C(x) = 25,500 - 30X2 + 600x$$

Where x is the number of products made.

If each product can be sold for ₦600, use the graphical method to

    i.   show the loss and profit areas

    ii.   determine the break-even quantity.

(c)   The demand and supply curves for a commodity are respectively given by

$$5y + x = 135$$

$$37y-4x=315$$

where x represents quantity demanded and y the price.

You are required to determine

    i.   the lowest and highest prices for the commodity.

    ii.   the equilibrium price and quantity by

       •   algebraic method

- graphical method

Question 12:

(a)   The rental agreement between JARUS and sons, an estate management firm, and his tenants provides for a constant annual increment of N15,000. If a tenant paid N95,000 in the first year, calculate the rent for the tenant in the 5th year.

(b)   Chief Akanbi, a business man, has two options of investing his money:

Option 1 -Involves simple interest of 19% per annum while

Option II - Is a compound interest of 15.5% per annum.

If he has N3,550,000 to invest for 4 years, which option should he go for?

(c)   Alhaji Owonikoko wants to buy a property worth N2.8m in the nearest future and therefore sets up a savings scheme which increases by 18% every year. If he saves N520,000 each year, how long will it take him to buy the property?

Question 13:

(a)   The research and planning unit of AWOS (Nig Ltd) has analysed the company's operating conditions concerning price and costs and developed the sale function S(q) in

N̶

and cost function C(q) in N̶ as follows:

$$S(q) = 900 - 6q$$
$$C(q) = 2q2 + 260q + 500$$

where q is the number of items produced and sold.

You are required to find

(i)   the marginal revenue

(ii)   the marginal cost

(iii)   the maximum profit, using the marginals obtained in (i) and (ii) above

(iv)   the sale price of an item for the maximum profit

(b)   The demand function for a particular commodity is $p = 770 - 3q$ and the cost function is $C(q) = q2 + 290q$

If a tax of ₦10 per unit of the commodity is imposed, determine the maximum profit and the corresponding tax.

(c). The demand and supply functions for a commodity are respectively

$P = 14 - q^2$ and $p = 2q^2 + 2$

Find the

(i) consumers' surplus

(ii) producers' surplus

Question 14:

a. i) Discuss the importance of decision making in an organization

ii) State five elements of decision-making

b. i) List four mathematical models and their purpose

ii) State three limitations of OR

c. i) Explain how OR can be used by an Accountant.

ii) Why is it necessary to test validity of solutions in OR?

Question 15:

The table below shows details of availability at four depots: W, X, Y, and Z and requirements from four destinations: A, B, C and D. Respective transportation costs (₦) are as shown at the top left-hand corner of each cell.

Destination Deport	A	B	C	D	Available supplies
W	6	5	4	3	5000
X	4	3	5	7	3500
Y	3	6	7	2	4500
Z	5	7	4	4	4000
Demands	3000	2500	4500	4000	

a) Use the Least Cost method to solve the transportation problem and calculate the total transportation cost.

b) Use the Vogel's approximation method to solve the transportation problem and calculate the total transportation cost

Question 16:

a)   RONKUS manufacturing company has set up machinery to produce its own commodity for stock at the rate of 5,000 units per week. Demand for the commodity is ta the rate of 3,000 units per week. Each production run costs N100 while holding cost per commodity is 50k per week.

You are required to find:

i.   the Economic production run

ii. the number of production runs per month

iii. length of the inventory cycle in weeks

iv. total costs

b) The demand for an item is 20,000 per quarter. The ordering cost is N60 per order, holding cost is N7.50 per item per annum and the basic per item is N20.

c) The supplier offers discount rate of 5% for items between 20,000 and 39,999 while 40,000 items and above attract 7.5% discount rate. Do you think the discount rates are worth considering?

(b)   The demand for an item is 20, 000 per quarter. The ordering cost is N60 per order, holding cost is N7.50 per item per annum and the basic price per item is N20.

The supplier offers discount rate of 5% for items between 20,000 and 39,999 while 40,000 items and above attract 7.5% discount rate. Do you think the discount rates are worth considering?

Question 17:

Activity	Immediate Duration (Months)	Preceding activity
A	-	3
B	A	2
C	-	4
D	A, C	1
E	B	2
F	D, E	5
G	B	3
H	G	2
I	F, H	3
J	G	3

a. Draw a network for the project

b. List all the paths and calculate their durations

c. Identify the critical path and give the shortest time for the completion of the business center.

d. Calculate

    I. All the ESTs and LSTs

    II. Float for all activities

e. Interpret the floats obtained for activities A, C, and J

Solutions

1.

(a)

Interval	Tally	Frequency
01-10	ll	2
11- 20	ﷻ ﷻ	10
21 - 30	ﷻ llll	9
31-40	ﷻ lll	8
41- 50	ﷻ ﷻ ll	12
51 - 60	ﷻ ll	7
61-70	ll	2

cccl

(b).

Interval	Frequency	Less **than** class boundary	Cumulative Frequency
01 - 10	2	10.5	2
11 - 20	10	20.5	12
2] - 30	9	30.5	21
31-.40	8	40.5	29
41- 50	12	50.5	41
51- 60	7	60.5	48
61 -70	2	70.5	50

(c   0give
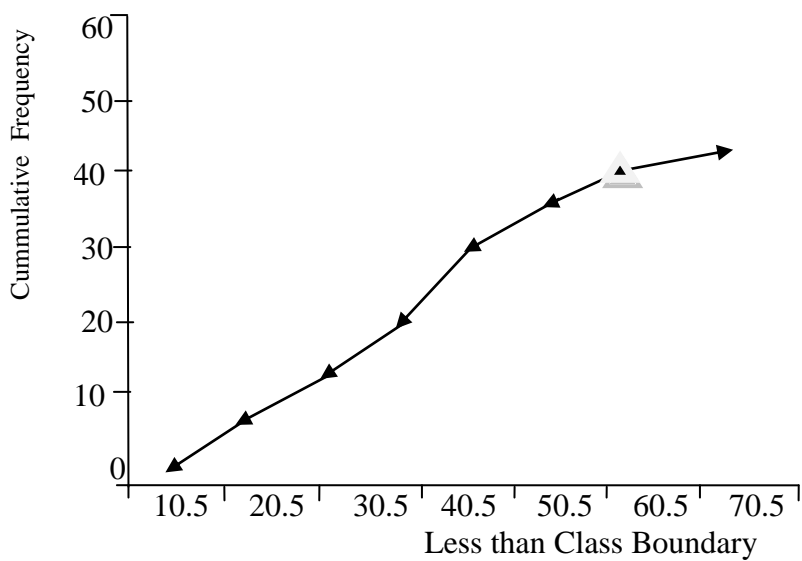
Histogram
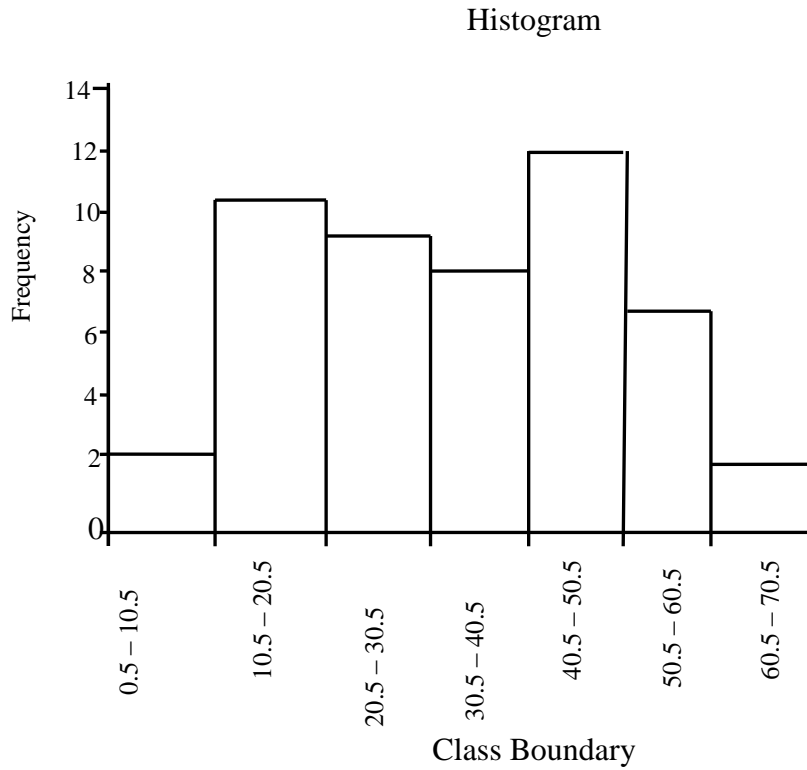
d) To calculate the mean, median and mode

Interval	Frequency	Cumulative Frequency	X	fx	X^2	fx^2
01 – 10	2	2	5.5	11.0	30.25	60.50
11 – 20	10	12	15.5	155.0	240.25	2402.50
21 – 30	9	21	25.5	229.5	650.25	5852.25
31 – 40	8	29 →Medial class	35.5	284.0	1260.25	10082
41 – 50	12	41 →Modal class	45.5	546.0	2070.25	24843
51 – 60	7	48	55.5	388.5	3080.25	21561.75
61 – 70	2	50	65.5	131.0	4290.25	8580.50
	50			1745		73382.50

ccclii

Mean $= \dfrac{\sum fx}{\sum f} = \dfrac{1745}{50} = 34.9$

Median $= L_I + \left[\dfrac{\dfrac{N}{2} - \sum f_i}{F_{median}} \right] C$

If $\dfrac{N}{2} = \dfrac{\sum f}{2} = \dfrac{50}{2} = 25$

$F_{median} = 8$, $\sum f_i = 21$, $C = 10$, $L_i = 30.5$

∴ Median $= 30.5 + \left[\dfrac{25 - 21}{8} \right] 10$

$= 30.5 + \left[\dfrac{4}{8} \right] 10$

$= 30.5 + (0.5)\,10$

$= 30.5 + 5$

$= 35.5$

Mode $= Li + \left[\dfrac{\Delta_1}{\Delta_1 + \Delta_2} \right] C$

$Li = 40.5,$ $\Delta_1 = 12 - 8 = 4$, $\Delta_2 = 12 - 7 = 5$, $C = 10$

Mode $= 40.5 + \left[\dfrac{4}{4 + 5} \right] 10$

$= 40.5 + \left[\dfrac{4}{9} \right] 10$

$= 40.5 + (0.4)10$ $= 40.5$

$= 44.5$

(e)   Variance $= \dfrac{\sum fx2}{\sum f} - \left(\dfrac{\sum fx2}{\sum f} \right)$

$= \dfrac{73382.5}{50} = \left(\dfrac{1745}{50} \right)^2$

$= 1467.65 - 1218.01$

$= 249.64$

∴ $\delta = 15.8$

∴ Coefficient of variation $= \dfrac{\delta}{\bar{x}} \times \dfrac{100}{1}$

$= \dfrac{15.8}{34.9} \times \dfrac{100}{1} \% = \underline{45.27}$

2.

a.

Interval	Mid value (x)	Frequency (f)	Fx	[x - 37.3] $\bar{x} = 37.3$	f[x - 37.3]
0 – 10	5	8	40	32.3	258.4
10 – 20	15	7	105	22.3	156.1
20 – 30	25	10	250	12.3	123
30 – 40	35	17	595	2.3	39.1
40 – 50	45	9	405	7.7	69.3
50 – 60	55	8	440	17.7	141.6
60 – 70	65	5	325	27.7	38.5
70 - 80	75	6	450	37.7	226.5
		70	2610		1152.5

$$\bar{x} = \sum \frac{fx}{f} = \frac{2610}{70} = 37.2857 \approx 37.3$$

$$\therefore \text{Mean deviation} = \frac{\sum f[x - \bar{x}]}{\sum f} = \frac{\sum f[x - 37.3]}{\sum f}$$

$$= \frac{1152.5}{70} = 16.46443 \approx 16.46$$

b.Interval	Mid value (x)	Frequency (f)	Fx	X^2	fx^2	Cf
0 – 10	5	8	40	25	200	8
10 – 20	15	7	105	225	1575	15
20 – 30	25	10	250	625	6250	25→Q_1
30 – 40	35	17	595	1225	20825	42
40 – 50	45	9	405	2025	18225	51
50 – 60	55	8	440	3025	24200	59→Q_3
60 – 70	65	5	325	4225	21125	64
70 - 80	75	6	450	5625	33750	70
		70	2610		1152.5	

cccliv

Variance $\sigma = \dfrac{\sum fx^2}{\sum f} - \left(\dfrac{\sum fx}{\sum f}\right)^2$

$= \dfrac{126150}{70} - \left(\dfrac{2610}{70}\right)^2$

$= 1802.1429 - \dfrac{6812100}{4900}$

$= 1802.1429 - 1390.2245$

$= 411.9184$

$\sigma \quad = \sqrt{411.9184}$

$= 20.2958 = 20.296$

c.   Semi-Interquantile range $\approx$ ½ $(Q_3 - Q_1)$

$Q_3 = L_3 + \left[\dfrac{\dfrac{3N}{4} - \sum f_3}{F_3}\right]$

$\dfrac{3N}{4} = \dfrac{3 \times 70}{4} = \dfrac{210}{4} = 52.5$

$L_1 = 50$, $\sum f_3 = 51$, $F_3 = 8$, $C = 11$

$Q_3 \quad = 50 + \left[\dfrac{52.5 - 51}{8}\right]11 = 50 + \left[\dfrac{1.5}{8}\right]11$

$= 50 + [0.1875]11 \qquad = \qquad 50 + 2.0625$

$Q_3 \quad = 52.0625$

$Q_1 = L_1 + \left[\dfrac{\dfrac{N}{4} - \sum f_1}{F_1}\right]C$

$\dfrac{N}{4} = \dfrac{70}{4} = \dfrac{70}{4} = 17.5$

$L_1 = 20$, $\sum f_1 = 15$, $F_1 = 10$, $C = 11$

$Q_1 \quad = 20 + \left[\dfrac{17 - 15}{10}\right]11 = 20 + \left[\dfrac{2}{10}\right]11$

$= 20 + (0.2)11$

$= 20 + 2.2$

$Q_1 \quad = 22.2$

. :   Semi − Interquantile Range = ½ $(Q_3 - Q_1)$

$= $ ½ $(52.0625 - 22.2)$

$= $ ½ (29.8625)

$= 14.93125$

d.   Coefficient of variation $= \dfrac{\sigma}{\bar{x}} \times \dfrac{100}{}$

Recall since $\sigma = 20.296$

$= 37.3$

$\therefore CV = \dfrac{20.296}{37.3} \times \dfrac{100}{}$

$= 0.5441 \times 100$

$= 54.4129$

$= 54.41$

3.   The total no of balls in the bag $= 7+9 = 16$

  i.   Probability of having black ball $= 9/16 = 0.5625$

 ii.   Probability of having white ball $= 7/16 = 0.4375$

        $\therefore$ P(of having white or black balls)$= 0.4375 + 0.5625$

                              $= 1$

4.   By using Binomial distribution

    P(X=x) $= \dbinom{n}{x} p^{x} q^{n-x}$ for $x = 1, \text{----}, n$

        $p = 70\% = 0.7$, $n = 8$

        $q = 1 - 0.7 = 0.3$

  i.   P(exactly 5 calls are order calls) $= P(x=5)$

        P(X=5) $= \dbinom{8}{5}(0.7)^{5}(0.3)^{8-5}$

                $= \dfrac{8!}{(8-5)!5!} \times 0.1681 X (0.3)^{3}$

                $= \dfrac{8!}{3!5!} X\ 0.1681$ $\times\ 0.27$

                $= \dfrac{8 \times 7 \times 6 \times 5! \times 0.1681 \times 0.027}{3! \times 5!}$

                $= 56 \times 0.1681 \times 0.027$

                $= 0.2541$

 ii.   P(at least 6 are order calls) $= P(6) + P(7) + P(8)$

ccclvi

$$P(x=6) = \binom{8}{6}(0.7)^6 (0.3)^{8-6}$$

$$= \frac{8!}{(8-6)!6!} \ X \ 0.7^6 \ X \ (0.3)^2$$

$$= \frac{8 X 7}{2} \ X \ (0.7)^6 (0.3)^2$$

$$= 0.2965$$

$$\rightarrow \quad P(X=7) \quad = \binom{8}{7}(0.7)^7 (0.3)^{8-7}$$

$$= \frac{8!}{(8-7)!7!} \ X \ 0.7^7 \ X \ (0.3)^1$$

$$= \frac{8!}{1!7!} \ X \ 0.0.0824 \ X \ 0.3$$

$$= \frac{8 \times \cancel{7 \times 6 \times 5 \times 4 \times 3 \times 2 \times 1}}{\cancel{7 \times 6 \times 5 \times 4 \times 3 \times 2 \times 1}} \ x \ 0.0824 \ x \ 0.3$$

$$-= 8 \ x \ 0.0824 \ x \ 0.3$$

$$= 0.1978$$

$$P(X=8) \quad = \binom{8}{8}(0.7)^8 (0.3)^{8-8}$$

$$= \frac{8!}{(8-8)!8!} \ X \ 0.7^8 \ X \ (0.3)^0$$

$$= \frac{8!}{0!8!} \ X \ 0.0.0576 \ X \ 1$$

$$= \frac{\cancel{8 \times 7 \times 6 \times 5 \times 4 \times 3 \times 2 \times 1}}{\cancel{8 \times 7 \times 6 \times 5 \times 4 \times 3 \times 2 \times 1}} \ x \ 0.0576 \ x \ 1$$

$$= 1 \ x \ 0.0576 \ x \ 1$$

$$= 0.0576$$

P(at least 6 are order calls) $= P(6) + P(7) + P(8)$

$$= 0.2965 + 0.1978 + 0.0576$$

$$= 0.2965 + 0.2554$$

$$= 0.5519$$

5.a.

X	y	xy	X^2	Y^2
7	6	42	49	36
4	7	28	16	49
10	9	90	100	81
3	1	3	9	1
6	4	24	36	16
8	10	80	64	100
5	3	15	25	9
9	8	72	81	64
1	5	5	1	25
2	2	4	4	4
55	55	363	385	385

$$r = \frac{n\sum xy - \sum x \sum y}{\sqrt{[n\sum x^2 - (\sum x)^2][n\sum y^2 - (\sum y)^2]}}$$

$$= \frac{10 x 363 - 55 x 55}{\sqrt{[10 x 385 - (55)^2][10 x 385 - (55)^2]}}$$

$$= \frac{3630 - 3025}{\sqrt{[3850 - 3025][3850 - 3025]}} = \frac{605}{\sqrt{825 X 825}}$$

$$= \frac{605}{\sqrt{680625}} = \frac{605}{825} = 0.7333$$

∴ r = 0.733

b.

Candidate	X	y	Rank of x	Rank of y	d_i = rank of x − Rank of y	d_i^2
A	7	6	7	6	1	1
B	4	7	4	7	-3	9

C	10	9	10	9	1	1
D	3	1	3	1	2	4
E	6	4	6	4	2	4
F	8	10	8	10	-2	4
G	5	3	5	3	2	4
H	9	8	9	8	1	1
I	1	5	1	5	-4	16
J	2	2	2	2	0	0
						44

$$R = 1 - \frac{6\sum d_i^2}{n(n^2-1)}$$

$$= 1 - \frac{6x44}{10(10^2-1)} = 1 - \frac{264}{10(100-1)} = 1 - \frac{264}{10(99)}$$

$$= 1 - \frac{264}{990} = 1 - 0.2667$$

$$= 0.7333$$

6.

a. since x is an independent variable, the equation becomes

$$y_i = a + bx_i; \quad \text{where } a = \sum y/n - b\sum x/n$$

$$b = \frac{n\sum xy - \sum x\sum y}{n\sum x^2 - (\sum x)^2}$$

x	Y	xy	x^2	y^2
3	2	6	9	4
5	3	15	25	9
6	5	30	36	25
8	5	40	64	25
10	6	60	100	36
11	8	88	121	64
13	9	117	169	81
16	10	160	256	100
72	48	516	780	344

∴ $b = \dfrac{8x516 - 72x48}{8x780 - (72)^2}$

$= \dfrac{4128 - 3456}{6240 - 5184} = \dfrac{672}{1056} = 0.6364$

$a = \sum \dfrac{y}{n} - b\sum \dfrac{x}{n}$

$= \dfrac{48}{8} - b(\dfrac{72}{8})$ $= \dfrac{48}{8} - 0.6364(\dfrac{72}{8})$

$= 6 - 0.6364(9) =$ $6 - 5.7276$

$= 0.2724$

∴ The equation is $y_i = 0.2724 + 0.6363x_i$

b. since x is a dependent variable, the equation becomes

$x_i = a + by_i$ where $a = \sum \dfrac{y}{n} - b\sum \dfrac{x}{n}$

$b = \dfrac{n\sum xy - \sum x \sum y}{n\sum y^2 - (\sum y)^2}$

$b = \dfrac{8 \times 516 - 72 \times 48}{8 \times 344 - (48)^2}$

$$= \frac{4128 - 3456}{8 x 344 - 2304} = \frac{4128 - 3456}{2752 - 2304}$$

$$= \frac{672}{448} = 1.5$$

$$\mathbf{a} = \sum \frac{y}{n} - b\sum \frac{x}{n}$$

$$= \frac{48}{8} - 1.5\left(\frac{72}{8}\right) \qquad = \qquad 6 - 1.5(9)$$

$$= \mathbf{6 - 13.5} \qquad\qquad = \qquad \mathbf{-7.5}$$

:. The equation is $x_i = -7.5 + 1.5y_i$

c. to find y if x = 17 and 18

From $y_i = 0.2724 + 0.6364x_i$

If x = 17

$y = 0.2724 + 0.6364(17)$

$y = 0.2724 + 9.546$

$y = 9.8184$

If x = 18

$y = 0.2724 + 0.6364(18)$

$y = 0.2724 + \mathbf{11.4552}$

$Y = 11.7276$

To find x if y = 13 and 14

If y = 13

From $x_i = -7.5 + 1.5y_i$

$x = -7.5 + 1.5(13)$

$x = -7.5 + 19.5$

$x = 12$

If y = 14

$x = -7.5 + 1.5 (14)$

$x = -7.5 + 21$

$x = 13.5$

7.

a.
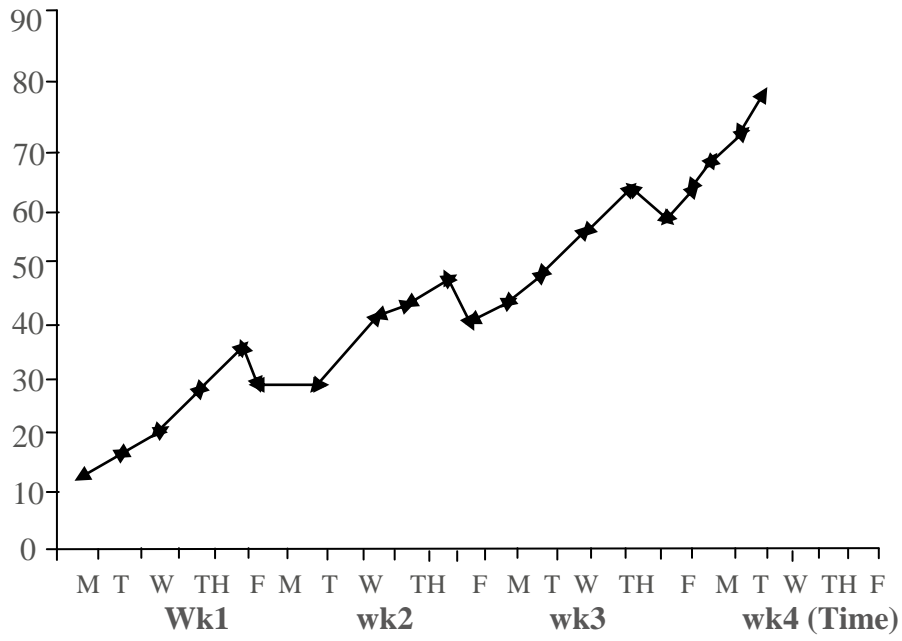
b&c.

Day (x)	Observation	5 - day moving totals	5 - day moving averages	Sv = Y-T Seasonal variation
Mon	13			
Tue	17			
Wed	23	129	25.8	-2.8
Thu	34	148	29.6	4.4
Fri	42	163	32.6	9.4
Mon	32	187	37.4	-5.4
Tue	32	166	33.2	-1.2
Wed	47	216	43.2	3.8
Thu	51	230	46	5
Fri	54	248	49.6	4.4
Mon	46	258	51.6	5.6
Tue	50	271	54.2	-4.2
Wed	57	288	57.6	-0.6
Thu	64	305	61	3
Fri	71	323	64.6	6.4
Mon	63	337	67.4	-4.4
Tue	68	348	69.6	-1.6
Wed	71	360	72	..1
Thu	75			
Fri	83			

8.

Commodity	2002		2006		P_0q_0	P_0q_1	P_1q_0	P_1q_1
	Unit price (P_0)	Quantity (q_0)	Unit price (P_1)	Quantity (q_1)				
Milk (tin)	0.60	24	0.80	30	14.4	18	19.2	24
Egg (crate)	0.50	2	0.70	4	1	2	1.4	2.8
Butter (tin)	0.25	4	0.40	6	1	1.5	1.6	2.4
Bread (loaf)	1.10	80	1.30	100	88	110	104	130
Meat (kg)	1.12	100	1.50	120	120	114	150	180
					224.4	245.5	276.2	339.2

a. Laspeyre's index $= \dfrac{\sum p_1q_0}{\sum p_0q_0} x100$

$= \dfrac{276.2}{224.4} x100$

$= 123.838$ ≈ 123.08

b. Paache's index $= \dfrac{\sum p_1q_1}{\sum p_0q_1} x100$

$= \dfrac{339.2}{245.5} x100$

$= 138.1670$ ≈ 138.17

c. Fisher's index $= \sqrt{\dfrac{\sum p_1q_0 \ X \ \sum p_1q_1 x100}{\sum p_0q_0 \ X \ \sum p_0q_1}}$

$= \sqrt{123.08 x 138.17}$

$= \sqrt{17005.9636}$

$= 130.4069$

≈ 130.41

d. Marshall Edgeworth's method using 2002 as base year

$$\frac{\sum p_1(q_0 + q_1)}{\sum p_0(q_0 + q_1)} \ X \ 100$$

P_0	q_0	P_1	q_1	q_0+q_1	$P_1(q_0+q_1)$	$P_0(q_0+q_1)$
0.60	24	0.80	30	54	43.2	32.4
0.50	2	0.70	4	6	4.2	3
0.25	4	0.40	6	10	4	2.5
1.10	80	1.30	100	180	234	198
1.20	100	1.50	120	220	330	264
					615.4	499.9

$$\therefore \ \text{Marshal Edgesworth's method} = \frac{\sum p_1(q_0 + q_1)}{\sum p_0(q_0 + q_1)} \ x100$$

$$= \frac{615.4}{499.9} \ x100$$

$$= \ 123.1046 \quad = \ 123.10$$

9.a

Profit (x)	$x-\bar{x}$	$(x - \bar{x})^2$
16	0	0
6	-10	100
14	-2	4
18	2	4
10	-6	36
12	-4	16
20	4	16
22	6	36
24	8	64
18	2	4
160	0	

a.   Estimate of the mean profit after tax is

$$\bar{x} = \frac{\sum x}{n} = \frac{160}{10} = 16$$

and the variance of the sample mean is

$$\sigma_{\bar{x}}^2 = \frac{S^2}{n}, \text{ where } S^2 = \frac{\sum(x-\bar{x})^2}{n-1}$$

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

$$= \frac{280}{10-1} = \frac{280}{9}$$

$$\therefore \sigma_{\bar{x}} = \frac{S^2}{n}$$

$$= \frac{31.1111}{10} = 3.1111$$

Which gives standard error of the mean as $\sqrt{3.111}$

$$= \quad 5.5777$$

$$\approx \quad 5.58$$

b. $t_{cal} = \dfrac{\bar{x}-\mu_0}{S/\sqrt{n}}$ (since n < 30)

$H_0: \mu = \bar{x}$ i.e. $H_0 : 15 = 16$

$H_1: \mu \neq = \bar{x}$ i.e. $H_1 : 15 \neq 16$

$\alpha = 0.05x$

$S^2 = 3.1111 \times 10 \Rightarrow S = \sqrt{3.1111x10} = 5.5777$

$\bar{x} = 16, \quad n = 10$

$$t_{cal} = \frac{16-15}{5.5777/\sqrt{10}} = \frac{1}{5.5777/3.1623}$$

$$= \frac{1}{0.5578} = 0.56695$$

Table value of t at 5% significance level (for two – tailed test) = 2.26

Decision: The problem is of two – tail test (see the alternative hypothesis), the $t_{cal} < t_{table}$ i.e 0.57 <2.26, we accept H_0 and conclude that no significant difference between the sample mean the population mean.

10.a.

Class Interval	Frequency	CF
5 – 9	4	4
10 – 14	4	8
15 – 19	10	18
20 – 24	8	26 $\rightarrow D_2$
25 – 29	17	43 $\rightarrow D_3$
30 – 34	20	63 $\rightarrow P_{45}$
35 – 39	16	79 $\rightarrow P_{55}$
40 – 44	15	94
45 – 49	12	106
50 – 54	<u>14</u>	120
	<u>120</u>	

a.i. $\quad D_2 \quad = \quad \dfrac{2}{10}$ x 120th item

$\qquad\qquad = \quad 24^{th}$ item

$D_3 \quad = \quad \dfrac{3}{10}$ x 120th item

$\qquad\qquad = \quad 36^{th}$ item

$$D_2 \quad = \quad L_2 + \left[\frac{\frac{2}{10} X (120)^{th} - F}{f_i} \right] C \qquad\qquad L_2 = 19.5, C = 5 , F = 18 , f_i = 8$$

$$= \quad 19.5 + \left[\frac{24-18}{8_i}\right]5 \qquad = \quad 19.5 + (6/8)5$$

$$= \quad 19.5 + (0.75)5 \qquad = \quad 19.5 + 3.75 \qquad = \quad 23.25$$

$$D_3 \quad = \quad L_3 + \left[\frac{\frac{3}{10} X (120)^{th} - F}{f_i}\right]C \qquad L_3 = 24.5, C = 5, F = 26, f_i = 17$$

$$= \quad 24.5 + \left[\frac{36-26}{17}\right]5 \qquad = \quad 24.5 + (10/17)5$$

$$= \quad 24.5 + (0.5882)5 \qquad = \quad 24.5 + 2.9412 \qquad = \quad 27.4412$$

ii. $P_{45} \quad = \quad \dfrac{45}{100} X N \quad = \quad \dfrac{45 \times 120}{100} = \quad 54^{th}$ item

$$P_{55} \quad = \quad \frac{55}{100} X N \quad = \quad \frac{55 \times 120}{100} = \quad 66^{th} \text{ item}$$

$$P_{45} \quad = \quad 29.5 + \left(\frac{54-63}{20}\right)5 \qquad = \quad 29.5 + 2.75$$

$$= \quad 32.25$$

$$P_{55} \quad = \quad 34.5 + \left(\frac{66-63}{16}\right)5 \qquad = \quad 34.5 + 0.9375$$

$$= \quad 35.4375$$

b.

i.   Population mean (μ) $\quad = \quad \dfrac{12+13+17+18}{4} = \quad 15$

ii.   Standard deviation of population (σ)

$$\sigma^2 \quad = \quad \frac{(12-15)^2 + (13-15)^2 + (17-15)^2 + (18-15)^2}{4}$$

$$= \quad \frac{(-3)^2 + (-2)^2 + (2)^2 + (3)^2}{4} = \frac{9+4+4+9}{4} = \frac{26}{4}$$

$$\sigma^2 \quad = \quad 6.5$$

$$\sigma \quad = \quad \sqrt{6.5} \quad = \quad 2.549$$

$$P_{45} = 29.5 + \left(\frac{54 - 63}{20}\right)5 = 29.5 + 2.75$$

$$= 32.25$$

$$P_{55} = 34.5 + \left(\frac{66 - 63}{16}\right)5 = 34.5 + 0.9375$$

$$= 35.4375$$

iii. Selection of 2 samples with replacement

(12,12)   (12,13)   (12, 17)   (12,18)

(13,12)   (13,13)   (13, 17)   (13,18)

(17,12)   (17,13)   (17, 17)   (17,18)

(18,12)   (18,13)   (18, 17)   (18,18)

The corresponding sample means are

12.0   12.5   14.5   15.0

12.5   13.0   15.0   15.5

14.5   15.0   17.0   17.5

15.0   15.5   17.5   18.0

∴ The sampling distribution of means can be grouped as:

Number (x)	12.0	12.5	13.0	14.5	15.0	15.5	17.0	17.5	18.0
Frequency (f)	1	2	1	2	4	2	1	2	1

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

$$= \frac{12.0 \text{ x } 1 + 12.5 \text{ x } 2 + 13.0 \text{ x } 1 + 14.5 \text{ x } 2 + 15.0 \text{ x } 4 + 15.5 \text{ x } 2 + 17.0 \text{ x } 1 + 17.5 \text{ x } 2 + 18.0 \text{ x } 1}{16}$$

$$= \frac{240}{16} = 15$$

iv. $$S^2 = \frac{(12 - 15)^2 + (12.4 - 15)^2 + - - - + (18.0 - 15)^2}{16}$$

$$= \frac{39}{16} = 2.4375$$

$$S^2 = 2.4375$$

S $= \sqrt{2.4375} = 1.5612495$

11. (a) Let x represent the number of items produced and sold, then
$C(x) = 1,600,000 + 5,000x$
$R(x) = 30,000x$

i. When x = 500, then
$C(x) = 1,600,000 + 5,000\,(500)$
$= ₦\,4,100,000 = ₦\,4.1m$

ii. When x = 50
$R(x) = 30,000\,(500)$
$= 15,000,000\;(₦\,15m)$
:. Profit $= R(x) - C(x)$
i.e. $15,000,000 - 4,100,000$
$= ₦\,10,900,000 = ₦10.9m$

iii. Break-even quantity is the quantity when profit is zero
i.e. $R(x) - C(x) = 0$
$R(x) = C(x)$
$30000x = 1\,600\,000 + 5000x$
$25000x = 1\,600\,000$
$x = 64$

iv. Profit of $₦\,21.52m$ implies
$R(x) - C(x) = 21.52$
i.e. $30\,000x - 1\,600\,000 - 5000x = 21\,520\,000$
$25\,000x = 23\,120\,000$
$x = 924.8m$
$\cong 92.5$

(b)i. The revenue function $R(x) = 600x$
When $x = 0$, $R(x) = 0 \Rightarrow (0,0)$
$X = 25$, $R(x) = 15\,000 \Rightarrow (25,15000)$
The table of values for
$C(x) = 25500 - 30x^2 + 600x$ is obtained as

x	0	10	20	30	40
y('000)		25.5	28.5	25.5	16.5	1.5
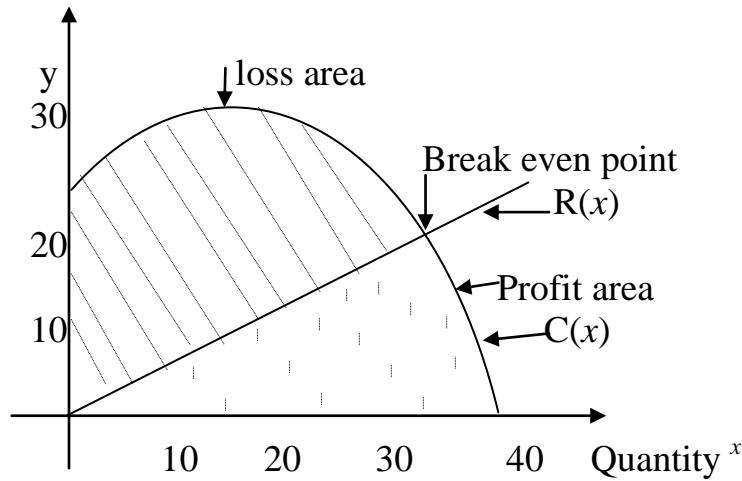
iii.   The break-even quantity is 29.5 as indicated on the graph\

(c)i   The lowest price is obtained when the supply is zero

i.e.   put x = 0 in 37y – 4x   = 315

$$37y = 315$$
$$y = ₦8.51$$

The highest price is obtained when the demand is zero

i.e. put   x = 0 in 5y + 3x   = 135

$$5y = 135$$
$$y = ₦27$$

ii.   algebraic method

5y + x = 135 …………………..(i)

37y - 4x = 315 ………………..(ii)

Equation (i) x 4 gives

20y + 4x = 540 ………………..(iii)

Equation (iii) + equation (ii) gives

57y = 855

Y = 15 i.e. the equilibrium price is ₦15

Substitute for y in equation (i)

5(15) + x = 135

    X = 60 i.e. equilibrium quantity is 60

-   graphical method – draw the graphs of the two equations on the same axes thus:
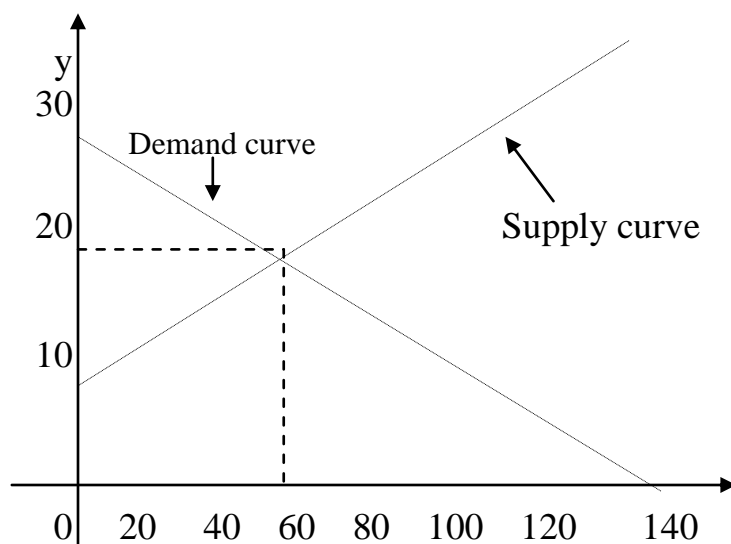
From the above graph, the equilibrium price is ₦16.5 while the equilibrium quantity is 60.

12. (a)   This is an A.P. with
$A = 95,000$, $d = 15,000$, $n = 5$
$T_5 = a + (5\text{-}1)d$
$= 95,000 + 4 \times 15,000$
$= ₦155,000$

(b) $P = ₦3,550,000$, $n = 4$, $r = 19$
S. I. will yield $\dfrac{PRT}{100} = \dfrac{3,550,000 \ X \ 19 \ X \ 4}{100}$
$= ₦2,698,000$

With C. I., $P = N3,550,000$, $n = 4$, $r = 0.155$
$A = P \ (1 + r)^n$
$= 3,550,000 \ (1.155)^4$
$= ₦6,317,660.59$

∴ interest $= ₦ \ 6,317,660.59 - ₦ \ 3,550,000$
$= ₦ \ 2,767,660.59$

Since the interest for C. I. is greater than that of S. I., he should go for option II

(c) $S = \dfrac{A\{1 + r)^n - 1\}}{r}$

i.e. 2,800,000 $= \dfrac{520,000\{(1.18)^n - 1\}}{0.18}$

$$\frac{2,800,000 x 0.18}{520,000} = \{(1.18)^n - 1\}$$

$$\therefore \{(1.18)^n\} \quad = 1.9692$$

$$\therefore \quad n \quad = \log 1.9692/\log 1.18$$

$$= 4.09 \text{years}$$

13. (a)(i) $S(q) = 900 - 6q$

$R(q) = q.S(q) = q(900 - 6q) = 900q - 6q^2$

$\therefore$ Marginal Revenue $= \dfrac{dR(q)}{dq} = 900 - 12q$

(ii) $C(q) = 2q^2 + 260q + 500$

$\dfrac{dC(q)}{dq} = 4q + 260$

(iii) Using the marginals, Profit is maximum when $\dfrac{dR(q)}{dq} = \dfrac{dC(q)}{dq}$

i.e. $900 - 12q = 4q + 260$

$\Rightarrow \quad 16q = 640$

$\Rightarrow \quad q = 40$

Testing for minimum or maximum, we have,

$\dfrac{dP(q)}{dq} = \dfrac{dR(q)}{dq} - \dfrac{dC(q)}{dq} = (900 - 12q) - (4q + 260) = 640 - 16q$

$\dfrac{d^2 P(q)}{dq} = -16$ which is negative and hence $q = 40$ gives maximum profit

Maximum profit, $\pi = R - C$ at $q^* = 40$

$\Rightarrow \pi = 900q - 6q^2 - 2q^2 - 260q - 500$

$= 640q - 8q^2 - 500$

$= 640(40) - 8(40)^2 - 500$

i.e. maximum profit $= ₦12,300$

(iv) at q $= 40$

$R(q) = 900q - 6q^2$

$= 900(40) - 6(40)^2 = ₦26,400$

$\therefore$ Sales Price $= \dfrac{26,400}{40} = ₦660$

$S(q) = 900 - 6q$

i.e. $S(40) = 900 - 6(40)$

$= 900 - 240 = ₦660$

(b) $P = 770 - 3q$

$R(q) = q(770 - 3q) = 770q - 3q^2$

$C(q) = q^2 + 290q$

$\therefore P(q) = 770q - 3q^2 - (q^2 + 290q)$

$= 480q - 4q^2$

$$\frac{dP(q)}{dq} = 480 - 8q = 0 \text{ at the turning point}$$

i.e $480 - 8q = 0$

$\qquad 8q = 480, \text{ i.e. } q = 60$

$$\frac{d^2 P(q)}{d(q)} = -8 \Rightarrow \text{maximum profit}$$

At maximum profit, tax $= 10 \times 60 = 600$

$\qquad P(q) = 480q - 4q^2$

When $q = 60$, $P(60) = 480(60) - 4(60)^2 = 14,400$

Hence, maximum profit $= 14400 - 600$

$\qquad\qquad = ₦13,800$

(c) $P = 14 - q^2$ (demand function)

$P = 2q^2 + 2$ (supply function)

At equilibrium, demand function = supply function

i.e. $14 - q^2 = 2q^2 + 2$

$\qquad -3q^2 = -12$

$\qquad q^2 = 4; \ q = 2$

when $q = 2; \ p = 14 - 2^2 = 10$

(i) Consumers' Surplus $= \displaystyle\int_0^{qo} \{D(q) - P_0\} dq$

$\qquad\qquad\qquad = \displaystyle\int_0^2 \{14 - q^2 - 10\} dq$

$\qquad\qquad\qquad = \displaystyle\int_0^2 \{4 - q^2\} dq$

$\qquad\qquad\qquad = \left[4q - \dfrac{q^3}{3} \right]_0^2$

$\qquad\qquad\qquad = 8 - \dfrac{8}{3} = \dfrac{16}{3} \text{ Or } 5.33$

(ii) Producers' Surplus $= \displaystyle\int_0^{qo} \{P_0 - S(q)\} dq$

$\qquad\qquad\qquad = \displaystyle\int_0^2 \{10 - 2q^2 - 2\} dq$

$\qquad\qquad\qquad = \left[8q - \dfrac{2q^3}{3} \right]_0^2$

$\qquad\qquad\qquad = 16 - \dfrac{16}{3}$

$$= \quad \frac{32}{3} \quad \text{Or } 10.67$$

14. (a)

(i) Decision-making is a day-to-day activity in an organization. An organization needs to make decisions to tackle problems which may arise as a result of discrepancies between existing conditions and the organisation's set objectives.

The Manager of an organization will have to decide on a course of action when confronted with a problem. Such decisions are normally taken to benefit the organisation and enhance the actualization of the set objectives.

(ii)

- decision-making unit e.g. an organization
- various possible actions that can be taken to solve the identified problem.
- states of nature that may occur
- consequences associated with each possible action
- comparison of the value of the decision and the consequences

b.(i) Mathematical Models

- allocation models – to share scarce resources among competing activities
- inventory models – to hold stock of items and re-order when necessary at minimum cost
- queuing models – to monitor arrival at and departure from service points
- replacement models – to determine an optimal policy for replacing failed items

(ii) Limitations of OR

- formation of a model is a simplified form of the reality but not the reality itself.
- the system which an OR model imitates is constantly changing but the model itself is static
- optimal solution to an OR model is only optimal based on the assumptions made

(c)

i. An Accountant can apply OR to investment decisions where the fund available is not

sufficient for all available projects i.e. capital rationing.

Also an Accountant can apply OR in every situation for cost-benefit analysis.

ii.   It is necessary to validate solutions of OR in order to determine if the model used can reliably predict the actual system's performance. Also to ensure that the model reacts to changes, in the same way as the real system.

15. a) As could be seen, the total available is 17,000 and the total demand is 14,000 hence we have 3000 extra. Therefore, there is the need to create a dummy destination or depot with transportation cost of zero in each of the cells.

The final tableau is as shown below. The allocations are explained thereafter.

Destination Depot	A		B		C		D		Dummy		Available supplies
W	6		5		4	4500	3		0	500	~~5000~~ ~~500~~ 0
X	4	1000	3	2500	5		7		0		~~3500~~ ~~1000~~ 0
Y	3	500	6		7		2	4000	0		~~4500~~ ~~500~~ 0
Z	5	1500	7		4		4		0	2500	~~4000~~ ~~2500~~ 0
Demands	~~3000~~ 0		~~2500~~ 0		~~4500~~ 0		~~4000~~ 0		~~3000~~ 0		~~17000~~ 0

Explanation

Note that the dummy column or row is not normally considered. It is just there fill up the balance after all the allocations have been computed.

Begin from cell YD with the least cost of 2 i. e. allocate 4000 to D from 4500 of Y. D is satisfied but Y has 500 left. Continue the allocation as explained earlier by considering the next least cost and so on. If there is a tie, pick any one with deficit in the row or column. At the end , all demands are satisfied but we are left with 500 from depot W and 2500 from depot Z. These are to be allocated to the dummy column at the appropriate rows i. e. row W and row Z.

Total transportation cost is

4 x 4500 + 0 x 500 + 4 x 1000 + 3 x 2500 + 3 x 500 +2 x 4000 + 5 x 1500 + 0 x 2500 = 46,500

c) The table is reproduced with the calculated penalties.

Destination Depot	A		B		C		D		Dummy		Available supplies	Penalties
W	6		5		4	4500	3		0	500	5000	1
X	4	1000	3	2500	5		7		0		~~3500~~ 1000	1
Y	3	500	6		7		2	4000	0		4500	1
Z	5	1500	7		4		4		0	2500	4000	1
Demands	3000		~~2500~~ 0		4500		4000		3000			
Penalties	1		2		1		1					

Ignore the dummy column (since it is just to fill up the balances at the end) and calculate the penalties as shown to obtain the penalty table. Column B has the highest penalty and is therefore picked. Allocate 2500 to row cell XB because it has the cheapest route in that column.

Continue as before and when all the penalties are equal, pick the row or column with the least route.

1st iteration

Destination Depot	A		B		C		D		Dummy		Available supplies	Penalties
W	6		5		4		3		0		5000	1
X	4		3	2500	5		7		0		~~3500~~ 1000	1
Y	3		6		7		2		0		4500	1
Z	5		7		4		4		0		4000	1
Demands	3000		~~2500~~ 0		4500		4000		3000			
Penalties	1		2		1		1					

2nd iteration

Destination Depot	A		B		C		D		Dummy		Available supplies	Penalties
W	6		5		4		3		0		5000	1
X	4		3	2500	5		7		0		~~3500~~ 1000	1
Y	3		6		7		2		0		4500	1
Z	5		7		4		4		0		4000	1
Demands	3000		~~2500~~ 0		4500		~~4000~~ 0		3000			
Penalties	1		-		1		1					

3rd iteration

Destination Depot	A		B		C		D		Dummy		Available supplies	Penalties
W	6		5		4		3		0		5000	2
X	4		3		5		7		0		~~3500~~ 1000	1
Y	3	500	6		7		2		0		~~4500~~ ~~500~~ 0	4
Z	5		7		4		4		0		4000	1
Demands	~~3000~~ 2500		~~2500~~ 0		4500		~~4000~~ 0		3000		17000	
Penalties	1		-		1		-					

4th iteration

Destination Depot	A		B		C		D		Dummy		Available supplies	Penalties
W	6		5		4	4500	3		0		5000 500	2
X	4		3		5		7		0		~~3500~~ 1000	1
Y	3		6		7		2		0		~~4500~~ ~~500~~ 0	-
Z	5		7		4		4		0		4000	1
Demands	~~3000~~ 2500		~~2500~~ 0		4500		~~4000~~ 0		3000		17000	
Penalties	1		-		1		-					

After the 4th iteration, columns B, C and D and row Y have been exhausted. Essentially, it only remains unit cells 6, 4 and 5 in column A. Minimum cell 4 will take 1000 to exhaust row X; next minimum cell 5 will take 1500 to exhaust column A; row W and dummy will take 500 and row Z and dummy will take 2500, producing the final tableau shown as

Destination Depot	A		B		C		D		Dummy		Available supplies	Penalties
W	6		5		4	4500	3		0	500	~~5000~~ ~~500~~ 0	2
X	4	1000	3	2500	5		7		0		~~3500~~ ~~1000~~ 0	1
Y	3	500	6		7		2		0		4500 ~~500~~ 0	4
Z	5	1500	7		4		4	4000	0	2500	~~4000~~ ~~2500~~ 0	1
Demands	3000 2500 1500 0		~~2500~~ 0		~~4500~~ 0		~~4000~~ 0		3000		17000	
Penalties	1		-		1		-					

The total cost of transportation is

4 x 4500 + 4 x 1000 + 3 x 2500 + 3 x 500 + 2 x 4000 + 5 x 1500 = ₦46,500

16.   (a)i. $c = 100, \ h = 0.5, \ d = 3000, \ r = 5000$

$\therefore$ Economic production run $Q = \sqrt{\dfrac{2cd}{h(1-\dfrac{d}{r})}}$

$$= \sqrt{\dfrac{2 \times 100 \times 3000}{0.5(1-\dfrac{3000}{5000})}}$$

$$= \quad 1{,}732.05$$

$$= \quad 1{,}733 \text{ per week}$$

ii   Number of production runs per month is $\dfrac{d}{Q} x4$ (since there are 4 weeks in a month)

$$= \quad \dfrac{3000}{1733} x4$$

☐   ☐ ☐ ☐ ☐

$\cong$ ☐

(iii)   Length of the inventory cycles is $\dfrac{Q}{d} = \dfrac{1733}{3000}$ ☐ since all values are given in weeks)

$$= \quad 0.58 \text{ weeks}$$

(iv)   Total costs = production cost + holding cost

$$= \quad 7 \times 100 + \frac{1733}{2}\left(1 - \frac{3000}{5000}\right)0.5x4$$

$$= \quad ₦1,393.20$$

b.   Demand per quarter is 20000. Therefore, demand per annum, d = 20000 x 4 = 80000 and with

c = 60 , h = 7.5, we have

$$Q = \sqrt{\frac{2cd}{h}}$$

$$Q = \sqrt{\frac{2x60x80000}{7.5}}$$

$$= \quad 1132$$

Number of orders $= \dfrac{80000}{1132} = 71$ orders

Now, we calculate different costs

No discount (EOQ = 1,132 items)

Cost of purchase	= 80,000 x 20	= 1,600,000
Holding Cost	$= \dfrac{1132}{2}x7.5$	= 4,245
Ordering Cost	= 71 x 60	= 4,260
	Total	**=₦1,608,505**

(C)   **5% (20,000 items)**

Cost of purchase	= 80,000 x 2 x 0.95	= 1,520,000
Holding Cost	$= \dfrac{20000}{2}x7.5x0.95 =$	71,250
Ordering Cost	$= \dfrac{80000}{20000}x60$	= 240
	Total Cost	**=₦ 1,591,490**

Since the 5% gives the smallest total cost, therefore, the 5% discount rate is worth considering.

17   (a)
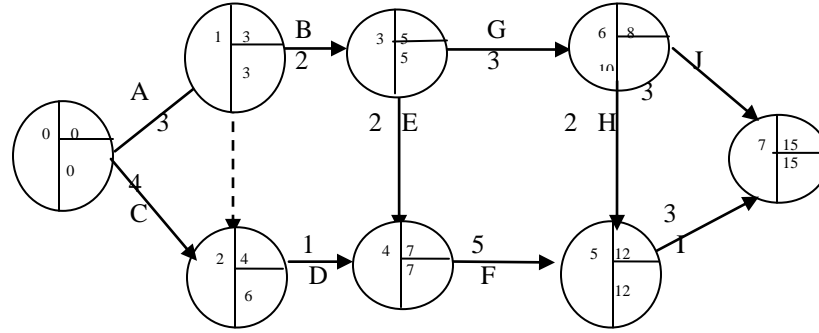
b)   Path   Duration

A, B, G, J   11

A, dummy, D, F, I   12

C, D, F, I   13

A, B, E, F, I   15

A, B, G, H, I   13

(c)   The critical path is along A, B, E, F, I and the shortest time for the completion of the project is 15 months.

d (i)
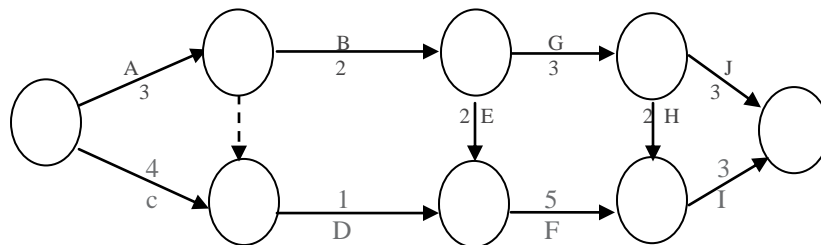
ii.

Activ	EFT	LFT	EST	LST	Duration	Total Float	Free Float	Independent Float
A	3	3	0	0	3	0	0	0
B	5	5	3	3	2	0	0	0
C	4	6	0	0	4	2	0	0
D	7	7	4	6	1	2	2	0
E	7	7	5	5	2	0	0	0
F	12	12	7	7	5	0	0	0
G	8	10	5	5	3	2	0	0
H	12	12	8	10	2	2	2	0
I	15	15	12	12	3	3	0	0
J	15	15	8	10	3	4	4	2

e) A: Since A is a critical activity all its floats are zeros.

   C: Total float is 2 months i.e the duration of activity C can be extended by 2 months without affecting the duration of the project. Other floats are zeros.

   J: Total float is 4 months i.e activity J can be extended by 4 months without affecting the project duration.

Free float is 4 months i.e activity J can be extended by 4 months without affecting the commencement of subsequent activities. In this case, no subsequent activities.

Independent float is 2 months i.e activity J can be extended without affecting the time available for activity G.

WRITTEN & REVIEWED BY:

Professor C. O. A. Awosope

Covenant University

Canaanland, Ota.

Ogun State, Nigeria

NOTE: This is a work in progress. All topics in the syllabus are covered but editing for necessary corrections is in progress.
Thanks.

IBLIOGRAPHY

Adamu, S. O. and Johnson, T. L. (1983): STATISTICS FOR BEGINNERS, Evans, Nigeria.

Ahuja, R., Magnati, T., and Orlin, J. (1993): NETWORK FLOWS-THEORY, ALGORITHMS, AND APPLICATIONS, Upper Saddle River: Prentice-Hall Inc.

Anderson M. Q et al (1986): QUANTITATIVE MANAGEMENT: AN INTRODUCTION, Kent Publishing Company, Boston.

Antin, H et al (1988): MATHEMATICS WITH APPLICATIONS FOR THE MANAGEMENT LIFE AND SOCIAL SCIENCES, Harcourt Brace Jovanorich Publishers, Florida.

Awosoga, A. O. (2000): STATISTICS FOR SCIENCES, JAS Publishers, Lagos.

Barry, R. and Star, R. M. (1997): QUANTITATIVE ANALYSIS FOR MANAGEMENT, New Jersey: Prentice-Hall Inc.

Bittinger, M. L. et al (1980): BUSINESS MATHEMATICS, Addison – Wesley Publishing Company, Reading.

Chartered Association of Certified Accountants (1987): STUDY TEXT, QUANTITATIVE ANALYSIS, BPP Publishing Ltd., London.

Esan, E. O. (1994): INTRODUCTION TO SAMPLING THEORY AND PRACTICE, JAS Publishers, Lagos

Esan, E. O. and Okafor, R. O. (1995): BASIC STATISTICAL METHODS, JAS Publishers, Lagos

Francis A. (2000): BUSINESS MATHEMATICS AND STATISTICS, Continuum, London.

Funk J. (1980): BUSINESS MATHEMATICS, Allyn and Bacon Inc., Boston

Gaughan, E. D. et al (1986): INTRODUCTION TO BUSINESS MATHEMATICS, Kent Publishing Company, Boston.

Hazarika P. (2005): A CLASS TEXTBOOK OF BUSINESS STATISTICS, Chaud & Company Publisher, New Delhi

Ignizio, J. P. and Cavalier, T. M. (1994): LINEAR PROGRAMMING, Upper Saddle River, N.J; Prentice-Hall Inc.

Inanga, E. L. and Osanyimwese, I.: MATHEMATICS FOR BUSINESS, Onibonoje Press.

Lapin, L. (1994): QUANTITATIVE METHODS FOR BUSINESS DECISIONS (6TH Edition), New York, The Dryden Press.

Levin, R. I. et al (1994): QUANTITATIVE APPROACH TO MANAGEMENT, Mc-Graw-Hill Lipschutz,, S probability: Schaum's outlines series Mc Graw-Hill.

Lucey, T (2002): QUANTITATIVE TECHNIQUES, ELST/Continuum, London.

Marssland, M.W. : QUANTITATIVE TECHNIQUES FOR BUSINESS, Polytechnic.

Nering, E. and Tucker A. (1992): LINEAR PROGRAMMING AND RELATED PROBLEMS, Boston Academy Press.

Parsons, J. A. (1977) :MODERN BUSINESS : THE USE OF MATHEMATICS IN BUSINESS, Alexander Hamilton Institute Incorporated, USA.

Sasiem, M. et al: OPERATIONAL RESEARCH – METHODS AND PROBLEMS, New York

Wiley.Sayibo, A. and Adekanye, F.: LINEAR PROGRAMMING FOR BUSINESS AND FINANCE, f & a. Publishers Ltd.

Winston, W. I. (1994): OPERATIONS RESEARCH – APPLICATIONS AND ALGORITHMS, California Duxubury Press.

Zameeruddin, Qazi, Khanna, V. K. and Bhambric, S. K. (2004): BUSINESS MATHEMATICS, Vikas Publishing House, New Delhi.

APPENDIX

NORMAL DISTRIBUTION

The table gives the area under the normal curve between the mean and a point z standard deviations above the mean.

Z	.00	.01	.02	.03	.04	.05	.06	.07	.08	0.09
0.0	.0000	.0040	.0080	.0120	.0160	.0199	.0239	.0279	.0319	.0359
0.1	.0398	.0438	.4878	.0517	.0557	.0596	.0636	.0675	.0714	.0753
0.2	.0793	.0832	.0871	.0910	.0948	.0987	.1026	.1064	.1103	.1141
0.3	.1179	.1217	.1255	.1293	.1331	.1368	.1406	.1443	.1480	.1517
0.4	.1554	.1591	.1628	.1664	.1700	.1736	.1772	.1808	.1844	.1879
0.5	.1915	.1950	.1985	.2010	.2054	.2088	.2123	.2157	.2190	.2224
0.6	.2257	.2291	.2324	.2357	.2389	.2422	.2454	.2486	.2517	.2549
0.7	.2580	.2611	.2642	.2673	.2704	.2734	.2764	.2794	.2823	.2852
0.8	.2881	.2910	.2939.	.2967	.2995	.3023	.3051	.3078	.3106	.3133
0.9	.3159	.3186	.3212	.3238	.3264	.3289	.3315	.3340	.3365	.3389
1.0	.3413	.3438	.3461	.3485	.3508	.3531	.3554	.3577	.3599	.3621
1.1	.3643	.3665	.3686	.3708	.3729	.3749	.3770	.3790	.3810	.3830
1.2	.3849	.3869	.3888	.3907	.3925	.3944	.3962	.3980	.3997	.4015
1.3	.4032	.4049	.4066	.4082	.4099	.4115	.4131	.4147	.4162	.4177
1.4	.4192	.4207	.4222	.4236	.4251	.4265	.4279	.4292	.4306	.4319
1.5	.4332	.4345	.4357	.4370	.4382	.4394	.4406	.4418	.4429	.4441
1.6	.4452	.4463	.4474	.4484	.4495	.4505	.4515	.4525	.4535	.4545
1.7	.4554	.4564	.4573	.4582	.4591	.4599	.4608	.4616	.4625	.4633
1.8	.4641	.4649	.4656	.4664	.4671	.4678	.4686	.4963	.4699	.4706
1.9	.4713	.4719	.4726	.4732	.4738	.4744	.4750	.4756	.4751	.4767
2.0	.4772	.4778	.4783	.4788	.4793	.4798	.4803	.4808	.4812	.4817
2.1	.4821	.4826	.4830	.4834	.4838	.4842	.4846	.4850	.4854	.4857
2.2	.4861	.4864	.4868	.4871	.4875	.4878	.4881	.4884	.4887	.4890
2.3	.4893	.4896	.4898	.4901	.4904	.4906	.4909	.4911	.4913	.4916
2.4	.4918	.4920	.4922	.4925	.4927	.4929	.4931	.4932	.4934	.4936
2.5	.4938	.4940	.4941	.4943	.4945	.4946	.4948	.4949	.4951	.4952
2.6	.4953	.4955	.4956	.4957	.4959	.4960	.4961	.4962	.4963	.4964
2.7	.4965	.4966	.4967	.4968	.4969	.4970	.4971	.4972	.4973	.4974
2.8	.4974	.4975	.4976	.4977	.4977	.4978	.4979	.4979	.4980	.4981
2.9	.4981	.4982	.4982	.4983	.4984	.4984	.4985	.4985	.4986	.4986
3.0	.4987	.4987	4987	.4988	.4988	.4989	.4989	.4989	.4990	.4990

STUDENT – t DISTRIBUTION

The values of t_0 given in the table has a probability of being exceeded.

d. f	$\tan_{.100}$	$\tan_{.050}$	$\tan_{.025}$	$\tan_{.010}$	$\tan_{.005}$	d.f
1	3.078	6.314	12.706	31.821	63.657	1
2	1.886	2.920	4.303	6.965	9.925	2
3	1.638	2.353	3.182	4.541	5.841	3
4	1.533	2.132	2.776	3.747	4.032	4
5	1.476	2.015	2.571	3.365	3.707	5
6	1.440	1.943	2.447	3.143	3.499	6
7	1.415	1.895	2.365	2.998	3.335	7
8	1.397	1.860	2.306	2.896	3.250	8
9	1.383	1.833	2.262	2.821	3.169	9
10	1.372	1.812	2.228	2.764	3.106	10
11	1.363	1.796	2.201	2.718	3.055	11
12	1.356	1.782	2.179	2.681	3.012	12
13	1.350	1.771	2.160	2.650	2.977	13
14	1.345	1.761	2.145	2.624	2.947	14
15	1.341	1.753	2.131	2.602	2.921	15
16	1.337	1.746	2.120	2.583	2.898	16
17	1.333	1.740	2.110	2.567	2.878	17
18	1.330	1.734	2.101	2.552	2.861	18
19	1.328	1.729	2.093	2.539	2.845	19
20	1.325	1.725	2.086	2.528	2.831	20
21	1.323	1.721	2.080	2.518	2.819	21
22	1.321	1.717	2.074	2.508	2.807	22
23	1.319	1.714	2.069	2.500	2.797	23
24	1.318	1.711	2.064	2.492	2.787	
25	1.316	1.708	2.060	2.485	2.779	24
26	1.315	1.706	2.056	2.479	2.771	25
27	1.314	1.703	2.052	2.473	2.763	26
28	1.313	1.701	2.048	2.467	2.756	27
29	1.311	1.699	2.045	2.462	2.57	28
inf	1.282	1.645	1.960	2.326	2.326	29
						inf