Chalimbana University

Integrity, Service, Excellence

SCHOOL OF BUSINESS AND ENTREPRENEURSHIP

MTH1100 – BUSINESS MATHEMATICS AND STATISTICS

2018

AUTHOR: **DR. W. PHIRI**

PhD (Yuseong), MA (Unza), MBA (U.K/CU), BA ed (Unza), BSc (Yuseong), Dip-Mat (Unza), Dip-GCP (Unza)

PUBLISHER: Chalimbana University

RATIONALE OF THE MODULE

This module is meant to provide students with a comprehensive understanding of Business Mathematics and Statistics. The core purpose is to allow graduates of Business to apply mathematics and statistical concepts to various emerging business situations. Students are encouraged to go through all the ten units in the order they are arranged. Note that each unit has some activities which are meant to provide an opportunity to students for enough practice. Our esteemed students in the School of Business and Entrepreneurship are further encouraged to work out the examples given in each unit before attempting the activity questions. Finally a list of recommended books for further reading has been given in order to allow students interact and supplement course content covered in this module. Furthermore, students are advised to consult and read latest published scholarly articles or works as business market environments are always dynamic.

The module contains ten units of which unit three is the longest. That is unit three has four sub units: We wish you our esteem students the very best and hope you will enjoy this module.

Welcome to Business Mathematics and Statistics –MTH 1100

TABLE OF CONTENT

UNIT ONE

BASIC ARITHMETICS AND CURRENCY CONVERSION

Introduction

This unit will introduce you to arithmetic techniques.

Unit Contents

This unit contains;

- Operations on rational and real numbers
- Proportions
- Foreign currency conversion
- Weights and measures

Unit Learning Objectives

By the end of this unit you should be able to:

- Be able to add, subtract, multiply, divide numbers with like and unlike signs

Operations on rational & Real Numbers

To add, subtract, multiply and divide integers.
In your earlier studies you might have covered some of the concepts portrayed in this unit. However, the emphasis is on the rules which are important to follow in the manipulation of algebraic equations.
The rules are as follows;

Rule 1: Addition of numbers with the same sign

To add numbers with the same sign, add the absolute values of the numbers. Then attach the sign of the addends. For example:

(i) Add: -65 + (-48)

-65 + (-48)	The signs are the same, add the absolute values of the numbers, then attach the sign of the addends
|-65| + |-48|	
65 + 48 = 113	
-113	

(ii) Add: +5 + (+20)

|5| + |20| = 25

+25

Rule 2: Addition of numbers with different signs

To add numbers with different signs, find the absolute value of each number, Subtract the smaller of the two numbers from the larger number. Then attach the sign of the number with the larger numerical value. For example:

(i) Add: 27 + (-53) **|27| = 27 and |-53| = 53**	The signs are different. Find the absolute value
53 – 27 = 26 **27 + (-53) = -26**	Subtract the smaller number from the integer. Because the (-53) >(-27), attach the sign 0f -53
(ii) Add: 16 + (-23) **|16| = 16 and |-23| = 23** **23 – 16 = 7** **16 + (-23) = -7**	

*Subtraction is defined as addition to the additive inverse.

<u>Rule 3: Rule for subtraction of Real Numbers</u>

This deals with how to subtract a number, and it says to subtract a number, you simply change the sign and add. If a and b are real numbers, then $a - b = a + (-b)$.
For example:

(i) Subtract: 48 – (-22) **48 – (-22)** **48 + 22 = 70**	Change the subtraction minus sign to the addition plus sign. The negative of a 22 is then a positive 22.
(ii) Subtract: - 31 – 18 **- 31 – 18** **-31 + (-18) = -49**	Change the subtraction minus sign to an addition plus sign. The negative of a positive 18 is a negative 18.

Note that **++=+, – – = +, + – = –, – + = –**

<u>Rule 4: Multiplication of Real Numbers</u>

The **product** of two numbers with the same sign is positive.

The **product** of two numbers with different signs is negative.

Example:

(i) Multiply: - 4 · (-9) **- 4 · (-9) = 36**	The product of two numbers with the same sign is positive.
(ii) Multiply: 84 · (-4) **84 · (-4) = - 336**	The product of two numbers with different signs is negative.

The **multiplicative inverse** of a nonzero real number *a* is $\frac{1}{a}$. This number is also called the reciprocal of *a*. For instance, the reciprocal of 2 is $\frac{1}{2}$ and the reciprocal of $-\frac{3}{4}$ is $-\frac{4}{3}$

Division of a real number is defined in terms of multiplication of the multiplicative inverse.

Rule 5: Division of Real Numbers:

If *a* and *b* are real numbers and $b \neq 0$, then $a \div b = a \cdot \frac{1}{b}$ .Because division is defined in terms of multiplication, the sign rules for dividing real numbers are the same as for multiplying. For Example:

(i) Divide: $\frac{-59}{9}$ $\frac{-59}{9}$ **= - 6**	The quotient of two numbers with different signs is negative.
(ii) **Divide**: **(-21) ÷ (-7)** (-21) ÷ (-7) = (-7)	The quotient of two numbers with same signs is positive

<u>Addition, Subtraction, multiplication and division of rational numbers</u>

A rational numbers are written in the fraction form $\frac{x}{y}$, where x and y are integers and $y \neq 0$. Examples of rational numbers are $-\frac{5}{9}$ and $\frac{12}{5}$. The number $\frac{9}{\sqrt{7}}$ is not a rational number because $\sqrt{7}$ is not an integer. Integers, Terminating and repeating decimals are rational numbers.

Add: -12.34 + 9.059	
12.340 – 9.059 = 3.281	The signs are different. Subtract the absolute values of the numbers
-12.34 + 9.059 = -3.281	Attach the sign of the number with the larger absolute value.

Multiply: (-0.23) · (0.04)	
(0.23) · (0.04) = -0.00092	The signs are different. The product is negative
Divide: -(4.0764) ÷ (-1.72)	
(-4.0764) ÷ (-1.72) = 2.37	The signs are the same. The quotient is positive.

Proportions and Foreign Currency Conversion

An exchange rate is how much it costs to exchange one currency for another. Exchange rates fluctuate constantly throughout the week as currencies are actively traded. This pushes the price up and down, similar to other assets such as gold or stocks. The market price of a currency. For instance, how many U.S. dollars it takes to buy a Zambian Kwacha for example, it is the different than the rate you will receive from your bank when you exchange currency. Here's how exchange rates work, and how to figure out if you are getting a good deal.

Market Exchange Rates

Traders and institutions buy and sell currencies 24 hours a day during the week. For a trade to occur, one currency must be exchanged for another. To buy British Pounds (GBP), another currency must be used to buy it. Whatever currency is used will create a currency pair. If U.S. dollars (USD) are used to buy ZMK, the exchange rate is for the ZMK/USD pair.

Reading an Exchange Rate

If the USD/ZMK exchange rate is 11.80 (as at 10th November 2018, Bank of Zambia Exchange Rate), that means it costs 11.80 Zambian Kwacha for 1 U.S. dollar. The first currency listed (USD) always stands for one unit of that currency; the exchange rate shows how much of the second currency (ZMK) is needed to purchase that one unit of the first (USD).

This rate tells you how much it costs to buy one U.S. dollar using Zambian Kwacha. To find out how much it costs to buy one Zambian Kwacha using U.S. dollars use the following formula: 1/exchange rate.

In this case, 1 / 11.80 = 0.08474576. It costs 0.08474576 U.S. dollars to buy one Zambian Kwacha. This price would be reflected by the ZMK/USD pair; notice the position of the currencies has switched.

Conversion Spreads

When you go to the bank to convert currencies, you most likely won't get the market price that traders get. The bank or currency exchange house will markup the price so they make a profit, as will credit cards and payment services providers such as PayPal, when a currency conversion occurs.

If the USD/ZMK price is 11.80, the market is saying it costs 11.80 Zambian Kwacha to buy 1 U.S. dollar. At the bank though, it may cost 11.98 or 12.16 Zambian Kwacha. The difference between the market exchange rate and the exchange rate they charge is their profit. To calculate the percentage discrepancy, take the difference between the two exchange rates, and divide it by the market exchange rate: 12.16 – 11.80 = 0.36/11.80 = 0.03050847. Multiply by 100 to get the percentage markup: 0.03050847 x 100 = 3.05%.

A markup will also be present if converting U.S. dollars to Zambian Kwacha. Banks and currency exchanges compensate themselves for this service. The bank gives you cash, whereas traders in the market do not deal in cash. In order to get cash, wire fees and processing or withdrawal fees would be applied to a forex account in case the investor needs the money physically. For most people looking for currency conversion, getting cash instantly and without fees, but paying a markup, is a worthwhile compromise.

Shop around for an exchange rate that is closer to the market exchange rate; it can save you money. Some banks have ATM network alliances worldwide, offering customers a more favorable exchange rate when they withdraw funds from allied banks.

Exchange rates always apply to the cost of one currency relative to another. The order in which the pair are listed (USD/CAD versus CAD/USD) matters. Remember the first currency is always equal to one unit and the second currency is how much of that second currency it takes to buy one unit of the first currency. From there you can calculate your conversion requirements. Banks will markup the price of currencies to compensate themselves for the service. Shopping around may save you some money as some companies will have a smaller markup, relative to the market exchange rate, than others.

Activity

1. Visit the Bank of Zambia Website and extract current exchange rates between US. Dollar, Euro and Great British Pound and the Zambian Kwacha
2. Conduct a simple survey of 10 Bureau of Exchange within your city and compare their rate of exchange to those of the BOZ
3. Compute the differences of those in (2) from those in (1), hence calculate in percentage form the markup each Bureau is charging

UNIT TWO

ALGEBRA AND EQUATIONS

INTRODUCTION

This unit will introduce you to the concept of algebra.

Unit Contents

This unit contains:

- Numbers and exponents
- Powers and Roots
- Indices and Logarithms
- Equations
    - Simple equations
    - Simultaneous equations
    - Quadratic equations

Unit learning Objectives

By the end of this unit you should have acquired knowledge and skills by being able to;

- Solve basic algebra

Numbers and Exponents

Exponential notation is used to show repeated multiplication. When $x \cdot x \cdot x \cdot x \cdot x$ is written as x^5 , x is called the base and 5 is called the exponent. An exponent is then a positive integer, written to the right and slightly above the base, which indicates the number of times the base is to appear as a factor. For example:

(a) $\boldsymbol{a^3 = a \cdot a \cdot a}$
(b) $\mathbf{16 = 2 \cdot 2 \cdot 2 \cdot 2 = 2^4}$
(c) $\mathbf{216 = 2^3 \cdot 3^2}$

You will now turn your attention to the laws of exponents as follows;

Let a and b be real numbers, variables, or algebraic expressions and let ***m*** and ***n*** be integers and $a \neq 0$, we have

Property	*Example*	
1. $a^m \cdot a^n = a^{m+n}$	$3^2 \cdot 3^4 = 3^{2+4} = 3^6$	Product Rule for Exponents - When multiplying with same bases, add the powers
2. $\frac{a^m}{a^n} = a^{m-n}$	$\frac{x^7}{x^4} = x^{7-4} = x^3$	The Quotient Rule for Exponents - When dividing with the same bases, subtract the powers
3. $(ab)^m = a^m b^m$	$(3x)^3 = 3^3 x^3 = 27x^3$	Power of a Product
4. $\left(\frac{a}{b}\right)^m = \frac{a^m}{b^m}$	$\left(\frac{4}{x}\right)^3 = \frac{4^3}{x^3} = \frac{64}{x^3}$	Power of a Quotient
5. $(a^m)^n = a^{mn}$	$(x^3)^4 = x^{3(4)} = x^{12}$	Power of a Power
6. $a^{-n} = , \frac{1}{a^n}$	$x^{-2} =, \frac{1}{x^2}$	Definition of negative exponent
7. $a^0 = 1, a \neq 0$	$(x^3 + 1)^0 = 1$	Definition of zero exponent
8. $a^{\frac{1}{n}} = \sqrt[n]{a, n}$	$(16)^{\frac{1}{2}} = \sqrt{16} = 4$	

You have to take note that the above rules apply to products and quotients only. The rules are not for simplifying sums or differences of numbers. These rules cannot be used to simplify or reduce $3^x + 3^y$ for example $3^4 + 3^5 \neq 3^9$

Powers and Roots

One of the methods if simplifying expressions is using powers. An equation such as **5+5+5+5+5+5+5 = 35** could be simplified as $5 \cdot 7 = 35$, where as $5 \cdot 5 \cdot 5 \cdot 5 \cdot 5 \cdot 5 \cdot 5$ could be simplified as 5^7**.** One way of describing powers is to know how many times the base number is multiplied by itself.

- $7 \cdot 7 \cdot 7 \cdot 7$ is the same is 7^4 which is also **2,401**.
- From 7^4 the **7** is called the base, The **4** is called the exponent, the index or the power.

One of the common ways to describe this expression (7^4) is "Seven to the power four." Special names are used when the index is 2 or 3, which are called "squared" and "cubed" respectively. Thus 5^2 is called "**Five squared**" and "9^3" is called "**nine cubed**." When no index is shown, the power is 1, example 3 means 3^1**.**

Roots

We have looked at powers: $\mathbf{7^2}$ **= 49.** Roots are used to find an unknown base: $\sqrt{\mathbf{49}}$ = **7**. The symbolic form $\sqrt{\mathbf{49}}$ is expressed as **"What is the square root of 49?" or "what number multiplied by itself equals 49?"** The value of a square root is the value of the base which when multiplied by itself gives the number. If **3 x 3 = 9**, then $\sqrt{\mathbf{9}}$ **= 3**. However, **(- 3) x (- 3) = 9**, so $\sqrt{\mathbf{9}} = \mathbf{-3.}$ Since there are always two answers when finding the square root of a number, a + and a – sign are placed in front of the answer to the square root problem. Therefore $\sqrt{\mathbf{9}}$ = $\pm\mathbf{3}$, $\sqrt{\mathbf{25}}$ = $\pm\mathbf{5}$.

Like exponents, the two common roots ($\sqrt{\ }$ and $\sqrt[3]{\ }$) are called square root and cube root.

$\sqrt[3]{\mathbf{27}}$ is expressed as the cube of root 27. We know that **3** cubed is 27 $(\mathbf{3 \cdot 3 \cdot 3})$ so **3** is the cube root of **27;**

If a ≥ 0 and b ≥ 0: we have $\sqrt{\boldsymbol{a} \cdot \boldsymbol{b}}$ = $\sqrt{\boldsymbol{a}} \cdot \sqrt{\boldsymbol{b}}$. However this only works for multiplication. For example $\sqrt{\mathbf{5}} \cdot \sqrt{\mathbf{4}} = \sqrt{\mathbf{5} \cdot \mathbf{4}} = \sqrt{\mathbf{20}}$.

Therefore $\sqrt{\boldsymbol{a}+\boldsymbol{b}} \neq \sqrt{\boldsymbol{a}} + \sqrt{\boldsymbol{b}}$.

A real number $\boldsymbol{r}$ is called an $\boldsymbol{n}^{th}$ root of **b if** $\boldsymbol{r}^{\boldsymbol{n}}$ = **b.**

Let's assume $\boldsymbol{n}$ is a positive number and $\boldsymbol{n}$ a positive integer: $\boldsymbol{b}^{\frac{1}{n}}$ or $\sqrt[n]{\boldsymbol{b}}$ is defined to be the positive $\boldsymbol{n}^{th}$ root of$\boldsymbol{b}$, thus $\left(\boldsymbol{b}^{\frac{1}{n}}\right)^{n}$ = $\boldsymbol{b}$.

Indices and Logarithms

The rules regarding Logarithms are: the logarithm, base b, of a positive number N (written $\log_b N$) is the exponent $\boldsymbol{x}$ such that $\boldsymbol{b}^{\boldsymbol{x}} \neq \boldsymbol{N}$**.** See the following examples to understand:

Rule 1: For logarithm of the product of two or more positive numbers is the sum of the logarithms of the numbers represented by $\mathbf{log}_{\boldsymbol{b}} \boldsymbol{M} + \mathbf{log}_{\boldsymbol{b}} \boldsymbol{N} \Leftrightarrow \mathbf{log}_{\boldsymbol{b}} \boldsymbol{MN}$

For example given log 2 = 0.3010 and log 3 = 0.477121; then

$\log 12 = \log(3x4) = \log(3x2^2)$

$= \log 3 + \log 2^2$

$= \log 3 + 2\log 2$

= 0.477121 + 2(0.3010)

= 1.079121

Rule 2: The Logarithm of the quotient of two positive numbers is the logarithm of the numerator minus the logarithm of the denominator represented **by** $\log_b M - \log_b N$ $\Leftrightarrow \log_b(\frac{M}{N})$

For example:

$\log 0.12 = \log 3\,(\frac{12}{100}) = \log 12 - \log 100$

= 1.079181 – 2

= -0.920819

Rule 3: The logarithm of a power of a positive number is the exponent of the power times the logarithm of the power represented by $\log_b(M^r) = r\log_b(M)$

For example

$\log \sqrt[3]{0.81} = \log(0.81)^{\frac{1}{3}}$

$= \frac{1}{3}\log 0.81$

$= \frac{1}{3}(\log 3^4 - \log 10^2)$

$= \frac{1}{3}(4\log 3 - 2\log 10) = \frac{1}{3}(4(0.477121)-2)$

= - 0.03051

Equations

What is an equation? An equation is simply a statement that two quantities are equal or simply a statement of equality between two expressions. So the statement $\mathbf{a + b = c + d}$

means $\mathbf{a + b}$ and $\mathbf{c + d}$ represent the same number. For example (6 + 9) and (7 + 8) both represent the number 15 which can be written as (6 + 9 = 7 + 8). The equations of interest in algebra are those that involve at least one variable or unknown on either (or both) sides of the equal sign.

Properties of Equality

Let a, b and c be real numbers, variables, or algebraic expressions

(i) Reflexive: a = a

(ii) Symmetric: if a = b, b = a

(iii) Transitive: if a = b and b = c, then a = c.

Simple equations (or equations of the first degree**)** are those in which an unknown quantity is raised only to the power 1. To 'solve an equation' means to find the value of the 'unknown'**.** Any arithmetic operation may be applied to an equation **as long as the equality of the equation is maintained.**

<u>Examples</u>

(i) Solve the equation $6x = 24$

$\frac{6x}{6} = \frac{24}{6}$

$x = 4$

The Left-hand side (LHS) and the Right-Hand side (RHS) equation has been maintained.

(ii) Solve: $\frac{2x}{5} = 6$

$5(\frac{2x}{5}) = 5(6)$

The fraction can be removed my multiplying both sides of the equation by 5.

2x = 30

Then canceling

$\frac{2x}{2} = \frac{30}{2}$

= 15

Then divide both sides of the equation by 2 giving 15

(iii) Solve x – 8 = 3

x – 8 + 8 = 3 + 8

x = 11

The result of the procedure is to move the '-8' from the LHS of the original equation, across the equals sign to the RHS but the sign changes to +

(iv) Solve 2x - 1 = 5x + 11

2x – 5x = 11 + 1

-3x = 12

$\frac{-3x}{-3} = \frac{12}{-3}$

x = -4

For such equations terms containing x are grouped on one side and the remaining terms are grouped on the other side of the equation. As in the previous example, changing from one side of the equation to the other side means the signs must be changed.

Simultaneous Equations; these are equations that are solved together to find unique values of unknown quantities, for which are true for each of the equations. When finding the value of single unknown quantities one equation is necessary. However when there are two unknown quantities in an equation, there is an infinite number of solutions. Similarly, for three unknown quantities it is necessary to have three equations in order to solve for a particular value of each unknown quantities and so on.

There are two methods of solving simultaneous equations which are; (i) by **Substitution,** and (ii) by **Elimination.**

Simultaneous Equations in two Unknowns

Solve the following equations for x and y,

(a) By substitution, and (b) by elimination **(1) x + 2y = -1, (2) 4x + 3y = 18**

(a) **By substitution**

(1) x = -1 – 2y Substituting this expression for x into equation (2) gives

4(-1 -2y) – 3 = 18

This is now a simple equation in y. Removing the brackets gives:

-4 -8y -3y = 18

-11y = 18 + 4 = 22

y = $\frac{22}{-11}$ = -2

Substituting y = -2 into equation (1) gives:

x + 2(-2) = -1

x – 4 = -1

x = -1 + 4 = 3

Therefore x = 3 and y = -2 are solutions to the equations

(b) By Elimination

Multiplying equation (1) by 4, the coefficient of x will be the same as in equation (2) giving:

(3) **4x + 8y = -4**

Subtract equation (3) from equation (2) giving us;

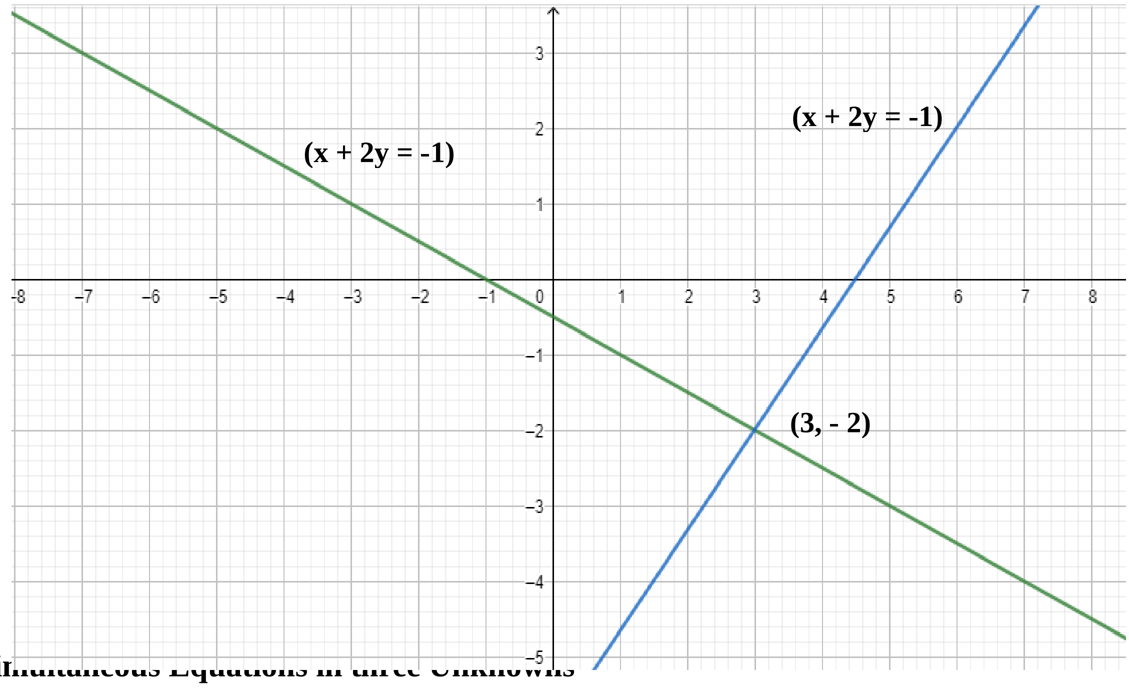
4x – 3y = 18
4x + 8y = -4
0 – 11y = 22

***In the above subtraction 18 – (-4) = 18 +4 = 22**

Hence **y =** $\frac{22}{-11}$ **= -2**

Substituting y = -2 into either equation (1) or (2) will give x = 3. The solution x = 3, and y = -2 are the values that satisfy the equations.

Since the point (3,-2) satisfies equations (1) and (2), then this point is at the point of intersection of the lines represented by the equations the two equations. Where the two lines intersect plotted in the graph below, the point of intersection is the solution.



Simultaneous Equations in three Unknowns

The methods used previously are used for solving two unknowns which can be extended to three unknowns, four unknowns, etc. When solving the equations we can use elimination or substitution method to make a system of two equations in two variables. For example:

$(1)\ \boldsymbol{x + 2y - z = 4}$

$(2)\ \boldsymbol{2x + y + z = -2}$

$(3)\ \boldsymbol{x + 2y + z = 2}$

First let’s add equation (1) and (2) to make an equation that has two variables. Then we subtract the third equation from the second so that another equation is gotten with two variables.

$\boldsymbol{x + 2y - z = 4}$ **(1)**
$\underline{\boldsymbol{2x + y + z = -2}}$
(2)
$\underline{\boldsymbol{3x + 3y + 0 = \ 2}}$
(4)

Now we have two equations with two variable

$\boldsymbol{2x + y + z = -2}$ **(2)**
$\underline{\boldsymbol{x + 2y + z = 2}}$ **(3)**

x – y – 0 = - 4 (5)

$$3x + 3y = 2 \quad (4)$$
$$x - y = -4 \quad (5)$$

$$3x + 3y = 2 \quad (4)$$
$$\underline{3x - 3y = -12} \quad (5)$$

Then multiply the second equation with 3 on both sides.

$$6x = -10$$
$$\frac{6x}{6} = \frac{-10}{6}$$

$$x = \frac{-10}{6}$$

We plug this value into the **3x + 3y = 2** equation to find our y-value

$$3 \cdot \frac{-10}{6} + 3y = 2$$
$$-5y + 3y = 2$$
$$3y = 7$$
$$y = \frac{7}{3}$$

Last we place the x value and the y value into one of the equations in order to find the z value.

$$x + 2y - z = 4$$
$$\frac{-10}{6} + 2 \cdot \frac{7}{3} + z = 2$$
$$3 + z = 2$$
$$z = -1$$

$$\therefore x = \frac{-10}{6}, \quad y =, \frac{7}{3}, \quad z = -1$$

Quadratic Equations; one in which the highest power of the unknown quantity is 2. For example $ax^2 + bx + c = 0$ is a quadratic equation where a ≠ 0. There are several methods for solving quadratic equations, namely;

- **(i)** By factorization
- **(ii)** By 'completing the square'
- **(iii)** By using 'the quadratic formula'
- **(iv)** Or graphically

Factorization

A useful method of solving quadratic equations is based on factoring. Multiplying out $(2x + 1)(x - 3)$ gives out $2x^2 - 6x + x - 3 = 2x^2 - 5x - 3$. The reverse process of moving from $2x^2 - 5x - 3$ to $(2x + 1)(x - 3)$ is called factorizing. If the quadratic equation can be factorized, this provides the simplest method of solving it. This is demonstrated in the following example:

$$x^2 + x - 12 = 0$$

The left side factors easily

$$(x - 3)(x + 4) = 0$$

Think of this as two quantities $x - 3$ and $x + 4$ whose product is zero. Whenever the product of two or more quantities is zero, at least one of the quantities must be zero. This means that either

$$x - 3 = 0 \quad or \quad x + 4 = 0$$

Solving these gives $x = 3$ and $x = -4$ respectively. Thus the roots of the equation are 3 an -4 and the solution set is {-4, 3}.

This technique of factorizing is often one of 'trial and error'.

Completing the square

An expression such as x^2 or $(x + 2)^2$ or $(x - 3)^2$ is called a perfect square

If $x^2 = 3$ then $x = \pm\sqrt{3}$

If the quadratic equation can be rearranged so that one side of the equation is a perfect square and the other side of the equation is a number, then the solution of the equation is readily obtained by taking the square roots of each side. The process is by rearranging one side of the quadratic equation into a perfect square before solving.

Example: Solve $2x^2 + 5x = 3$

- Rearrange the equation so that all terms are on the same side of the equals sign

  $$2x^2 + 5x - 2 = 0$$

- Make the coefficient of the x^2 term unity. This is achieved by diving throughout by 2

  $\frac{2x^2}{2} + \frac{5x}{2} - \frac{3}{2} = 0 \qquad x^2 + \frac{5x}{2} - \frac{3}{2} = 0$

- Rearrange the equations so that x^2 and x terms are on one side of the equals sign and the constant is on the other side.

  $$x^2 + \frac{5x}{2} = \frac{3}{2}$$

- Add to both sides of the equation (half the coefficient of x^2). In this case the coefficient of x is $\frac{5}{2}$. Half the coefficient squared is $\left(\frac{5}{4}\right)^2$
- This $x^2 + \frac{5}{2}x + \left(\frac{5}{4}\right)^2 = \frac{3}{2} + \left(\frac{5}{4}\right)^2$
- The LHS is now a perfect square

$$\left(x + \frac{5}{4}\right)^2 = \frac{3}{2} + \left(\frac{5}{4}\right)^2$$

- Evaluate the RHS $\left(x + \frac{5}{4}\right)^2 = \frac{3}{2} + \frac{25}{16} = \frac{24+25}{16} = \frac{49}{16}$
- Taking the square root of both sides of the equation remembering the square root of a number gives a $\pm$ answer.

$$\sqrt{x + \frac{5}{4}^2} = \sqrt{\frac{49}{16}}$$

$$x + \frac{5}{4} = \pm\frac{7}{4}$$

- Solve the simple equation. Thus

$$x = -\frac{5}{4} \pm \frac{7}{4}$$

$$x = -\frac{5}{4} + \frac{7}{4} = \frac{2}{4} = \frac{1}{2}$$

$$x = -\frac{5}{4} - \frac{7}{4} = -\frac{12}{4} = -3$$

Hence

$$x = \frac{1}{2} \ or \ x = -3$$

Quadratic Equation

Let the general form of a quadratic eqation be given by:

$$ax^2 + bx + \mathrm{c} = 0$$

Where a, b and c are constants, Dividing $ax^2 + bx + \mathrm{c} = 0$ by a gives:

$$x^2 + \frac{b}{a}x + \frac{c}{a} = 0$$

Rearranging gives

$$x^2 + \frac{b}{a}x = -\frac{c}{a}$$

Adding to each side of the equation the square of half the coefficient of the term x to make the LHS a perfect square gives

$$x^2 + \frac{b}{a}x + \left(\frac{b}{2a}\right)^2 = \left(\frac{b}{2a}\right)^2 - \frac{c}{a}$$

Rearranging gives: $\left(x + \frac{b}{a}\right)^2 = \frac{b^2}{4b^2} - \frac{c}{a} = \frac{b^{2-}4ac}{4a^2}$

Taking the square root of both sides gives:

$$x + \frac{b}{2a} = \sqrt{\frac{b^{2-}4ac}{4a^2}} = \frac{\pm\sqrt{b^2 - 4ac}}{2a}$$

Hence $x = -\frac{b}{2a} \pm \frac{\sqrt{b^2-4ac}}{2a}$

$$x = \frac{-b \pm \sqrt{b^2 - 4ac}}{2a}$$

Example: Solve $x^2 + 2x - 8 = 0$ using quadratic formula.

Comparing $x^2 + 2x - 8 = 0$ $with$ $ax^2 + bx + \mathrm{c} = 0$ gives $a = 1,\ b = 2, c = -8$

Substitute those values in the quadratic formula

$$x = \frac{-2 \pm \sqrt{2^2 - 4(1)(-8)}}{2(1)}$$

$$= \frac{-2 \pm \sqrt{4 + 32}}{2} = \frac{-2 \pm \sqrt{36}}{2}$$

$$= \frac{-2 \pm 6}{2} = \frac{-2 + 6}{2} = \quad or \quad \frac{-2 - 6}{2}$$

Hence $x = \frac{4}{2} = \mathbf{2}$ or $x = \frac{-8}{2} = \mathbf{-4}$

Activity

1. Evaluate $\log_e(2)$
2. Express as a single logarithm: 2log(a) -2log (2a)
3. Solve the equations:

a. $2^{2/x} = 32$

b. $3^{2(1+x)} - 28 \times 3^x + 3 = 0$

4. Solve the equations simultaneously:

   $P - q + r = 2$

   $P + 2q - 2r = 1$

   $-p + 2q + 2r = 9$

5. Solve for x in the equation $2^x = 64$

6. The roots of the equation $2x^2 - 7x + 4 = 0$ are α and β. Find an equation with integral coefficients whose roots are α/β, β/α

UNIT THREE

STATISTICS ONE

DATA COLLECTION TECHNIQUES

Introduction

Statistics is a scientific discipline which involves applying a set of rules and procedures to reduce large masses of data to manageable proportions thereby making it possible to draw reasonable conclusions from those data. In this unit, we will look at different techniques which are related to data collection. These techniques include the different sampling methods, the design of data collection instruments such as questionnaires and the various methods of collecting data.

Objectives

By the end of the unit, you should be able to:

- describe the different sampling methods
- describe methods of collecting data
- design good instruments of data collection
- conduct a survey for a particular enquiry

Note that most of the work in this section was adopted from Munzara (2017)

Sampling

Munzara (2017) referred sampling to the process of selecting a smaller set of study units (sample) from the entire collection of units of interest (population). This is usually necessitated by limited financial and time resources to carry out a population study. A census, which is the term used for a study of the whole population, is usually costly and requires more time and labour compared to a sample survey.

The data that is collected from a sample is meant to provide an insight into the population from which it was drawn. It is, therefore, necessary that the sample ought to be representative of the underlying population. A representative sample is one which reflects the characteristics of the underlying population as closely as possible. It has to be as large as resources would permit because the larger the sample the more accurate the results would be. To ensure the use of a representative sample in a survey, the sample elements are selected randomly using a suitable random sampling method.

The methods of selecting samples are classified into two categories namely random sampling and non- random sampling. In random sampling, also known as probability sampling, every element of the population has equal chance of being included or excluded from the selected sample, that is, selection of study units is free from personal bias. This is not the case in non-random sampling where the researcher exercises human judgement in selecting study units resulting is samples which are biased.

Random Sampling Methods

In this section we are going to consider the following random sampling methods: Simple random sampling, Stratified random sampling, Systematic random sampling, and Cluster sampling.

Simple random sampling

To select a sample using simple random sampling one has to come up with a numbered list of all the elements in the population of interest. This list is called a sampling frame. The sampling frame makes it possible for us to draw elements from the population by randomly generating the numbers of the elements to be included in the sample using a scientific calculator, a computer or random number tables. Random digit tables are provided in the appendices.

In some studies, obtaining a frame of all the units in the population is impossible. In these studies, a random sample can still be obtained by randomising some aspect of the study such as the location and the time of collecting the observations.

Example

Using random number tables, obtain a simple random sample of 30 accounts from a population of 500 bank accounts.

Solution

You start by numbering the accounts from 001 to 500. In the table of random digits, you randomly choose a row or column to start from. Since 500 is a three digit number, you draw random numbers with three digits, ignoring any numbers greater than 500. You should also ignore any number already obtained.

Suppose you select row six as the starting point. The first account to be included in the sample is the account with identification number 428. The next three digit numbers 664 and 627 are ignored since they are greater than 500. Continuing systematically in the same row you get the following random numbers: 343, 621 (ignore), 936 (ignore), 362, 358, 259, 351, 298, 285, 300, 606 (ignore), 004. You will continue this process until you have the desired sample of 30 accounts.

Stratified random sampling

In stratified random sampling, the population is divided into two or more subpopulations which are called strata, and then simple random samples are taken in each stratum and then combined to give the desired sample.

This method is suitable for selecting a sample from a population which is made up of distinct subpopulations. Each subpopulation consists of study units with similar characteristics and has to be adequately represented in the final sample. In order to ensure that this is the case, the sample sizes in each stratum are made to be proportional to their frequency in the population.

Example

A researcher would like to hear the opinion of students regarding a proposed business plan. On a particular day, the sample comprised of 250 students from CHAU University, 200 students from UNZA and 50 students from CBU. How would the researcher select a random sample of 40 students for use in the study using stratified random sampling?

Solution

CHAU: $\frac{250}{500} \times 40 = 20$ members

UNZA: $\frac{200}{500} \times 40 = 16$ members

CBU: $\frac{50}{500} \times 40 = 4$ members

A sampling frame of each subpopulation is prepared, and then simple random sampling is used to fill its allocation.

Systematic random sampling

Systematic random sampling involves selecting study units at periodic intervals, for example, selecting every 10^{th} item from a production run for inspection. The first study unit is the only one selected at random while subsequent units are selected following a specific order. For this reason, systematic random sampling is viewed as a quasi-random sampling method.

Suppose a population has N elements and a random sample of size n is to be selected using systematic random sampling. The steps to be followed are:

- number the elements in the population from 1 to N,
- calculate the sampling ratio(£) by dividing the population size N by the desired sample size n,
- randomly select a number between 1 and *k* inclusive to get the starting point, and
- Select additional elements at evenly spaced intervals (every A:"[1] element thereafter) until the desired sample is obtained.

Example

A bank has 2 000 account holders. A random sample of 40 accounts is to be selected using systematic random sampling. List the accounts that would make up the sample if the first account randomly selected is 12.

Solution

You begin by numbering the accounts from 1 to 2 000. Then you calculate the sampling ratio *k* as follows:

N 2000

$$K = \frac{N}{N} = \frac{200}{40} = 50$$

Now, starting from 12, every 50^{th} account thereafter is included.

The first account is 12, the second account is 12 + 50 = 62, the third account is 62 + 50 = 112 and you continue adding 50 like this until a sample of 40 bank accounts is obtained.

Cluster sampling

Where the population consists of homogeneous subpopulations the population can be divided into these subpopulations called clusters. Unlike in stratified random sampling where there is homogeneity in terms of the variable of interest, elements within clusters exhibit different characteristics. The sampling frame is a complete list of all the clusters that make up the population. A random sample of clusters is selected and all the study units in the sampled clusters are selected for the survey.

The advantages of cluster sampling are that the method is cost less costly and less time consuming. The amount of travelling by interviewers to reach the respondents is reduced and as a result time wastages are minimal. The disadvantages are that the selected sample may not reflect the diversity of the population elements because some sections of the population may be under-represented or over-represented.

Non- Random Sampling Methods

Non- random sampling methods are non-probabilistic in nature. In this section, we will consider the following non-random sampling methods: Judgemental sampling, Convenience sampling, and Quota sampling.

Judgemental sampling

Judgemental sampling is used where special skills are the criteria used to form a sample. Usually the sample is very small. For example, only renowned economists may be asked to comment on the economic status of a country. These economists are not randomly selected but deliberately selected due to their expertise. The method is subjective in that it is purely based on the researcher's personal discretion to choose the sample.

Convenience sampling

The sample is drawn to suit the convenience of the researcher. The researcher selects those elements which are easily accessible to him, making the sampling process easy and cheaper. For example, suppose a researcher would like to find out the level of customer service satisfaction of shoppers in a supermarket. The researcher will simply interview shoppers coming out of the supermarket rather that following up on customers who once bought goods from the supermarket. However, the sample chosen is usually not representative of the population.

Quota sampling

In quota sampling, the sample is drawn on the basis of more specific guidelines about which study units should be drawn. To start with, you decide on the proportion of the population to be included in the sample. Then you identify distinct groups within the population. You go on to determine the quota of respondents needed from each group. You then fill each quota by finding enough elements from each group of the population.

Quota sampling is almost similar to stratified random sampling in that you start by identifying strata in the population and dividing the population accordingly. It differs from stratified random sampling in that it does not require a sampling frame. It is therefore much quicker, cheaper and easier compared to stratify random sampling. Its major drawback is that it is susceptible to bias.

Data Collection Methods

Data collection can be defined as the process of counting or enumerating or measuring together with the recording of results. One important factor that is considered when selecting a suitable method of data collection is the source of the data. Sources of data can be classified as primary source and secondary source giving rise to primary data and secondary data respectively.

Primary data are those data which are originally collected by a researcher from the field for the first time and for a specific purpose. The data are first hand information and hence are up-to-date, more reliable and accurate. However, primary data is usually expensive to collect and requires more time and labour to gather.

Secondary data are those data which are already collected, processed and used by someone else for their own purpose which may not necessarily be the same as for the current researcher. Secondary data is obtained from journals, newspapers, company newsletters and from publications made by government agencies. Although it is much faster and cheaper to obtain secondary data, the data might be outdated and therefore, not suitable for the current study.

In this section, we will look at four principal ways of collecting primary data.

Direct observation

The researcher records some observations on the study units. If the study units are people, they should not realize that they are being observed. In some cases, the researcher becomes a member of the group being studied.

Advantages of direct observation

The method is easy to implement.

It is faster and cheaper.

Disadvantages of direct observation

The researcher would not be able to give reasons for certain observed behaviour by merely observing without interviewing study units.

If study units come to realise that they are being observed they may change their usual behaviour.

1.5.2 Personal interview

Large scale surveys are conducted using this method. The method involves face-to-face interviews between field investigators and the subjects of the study. The interviewer will ask questions from a prepared interview schedule and will record the responses of the interviewee on spaces provided on the schedule. Interviewers must undergo training so that there is some degree of uniformity in the manner in which questions are asked and the mode of filling the schedules.

Advantages of personal interviews

- A high response rate
- There is room for probing and clarification of misunderstood questions.
- Non-verbal responses can be noted.
- Questionnaires are completed in a standard way by the interviewer.
- The presence of interviewer discourages collective responses.

Disadvantages of personal interviews

- The exercise is costly due to training expenses and the amount of travelling done by enumerators.
- It is time consuming.
- Respondent anonymity is lost.
- Interviewer bias may arise if the interviewer suggests answers to the respondent through use of suggestive gestures or asking leading questions.
- The interviewer may also encounter hostile respondents who refuse to be interviewed.

Questionnaire Design

The questionnaire is the major instrument used to collect primary data (from human beings). It is important to ensure that the questionnaire is well structured in order to solicit for the required information.

Well designed questionnaires should:

- have clear instructions to guide the respondent when completing it,
- not be unnecessarily long,
- have clear and unambiguous questions,
- have questions arranged in logical sequence,
- have short questions rather than long questions,
- avoid questions which require calculations, and
- not contain double-barrelled questions.

The questionnaire should be tested in a pilot survey before the actual survey. This is necessary for the following reasons:

- to help fine-tune the data organisation procedures
- to test the questionnaire for clarity of instructions and questions
- to test the data analysis tools beforehand
- to determine the appropriate sample size for the actual study
- to train enumerators in questioning skills

Summary

In this unit we looked at the various methods of selecting random samples from underlying populations. Random sampling methods are those in which population elements have the same probability of becoming part of the sample. We then looked at the three principal ways of obtaining primary data which are observation, personal interview and postal questionnaires. The advantages and disadvantages of these methods were discussed. We ended the unit by describing the attributes of well-designed questionnaires.

Activity

1. (a) Distinguish between primary data and secondary data.

   (b) What are the advantages of primary data over secondary data?

2. Explain the advantages and the disadvantages of administering a survey questionnaire online.

3. Suggest ways of improving the response rate in postal questionnaire interviews.

4. Design a questionnaire that you would use to collect data concerning student’s perception towards Business Mathematics and Statistics course in Chalimbana University.

5. Use random digits tables to select a random sample of 50 elements from a population of 500 elements.

6. A popular government primary school has 1 000 students of whom 600 are in boarding and the rest are day scholars. An opinion poll is to be conducted concerning the time to begin lessons during winter. How would you select a random sample of 200 students to take part in the poll using stratified random sampling?

STATISTICS 2

MEASURES OF CENTRAL TENDENCY AND DISPERSION

Introduction

This unit aims to help students grasp the principles of statistics and the statistical methods of statistics in a way that emphasizes working with data and mastering statistical reasoning.

Unit Contents

- Introduction, Collection and Classical Information
- Sources of Statistics
- Frequency Distributions
  - Simple
  - Cumulative
- Results presentation
- Measures of Central Tendency
- Moving and weighted averages
- Measures of dispersion

Statistics defined and described

- Scientific study of handling quantitative information (processed data)
- Methodology of collection, classification, description and interpretation of data obtained through the conduct of surveys and experiments.
- Main purpose is to describe and draw inferences about the numerical properties of populations
- **Descriptive Statistics**
  - Procedures for organizing, summarising and describing quantitative data about samples or about the populations where complete population data are available.
  - It does not involve the drawing of an inference from a sample to its population
  - E.g. Measures of central tendency (mean, mode, median) and measures of variability (range, variance, standard deviation)

- **Inferential statistics**
    - Making inferences (conclusions) from small (sample) to large (population) groups of subjects or events on the basis of probability theory
    - Note that inferences about population drawn from sample measure may involve some error or discrepancy, the magnitude of which can be estimated on the basis of the probability theory

Population and Sample are basic concepts of statistics. **Population** is characterized as the set of individual persons, items or objects under investigation or considered in a statistic study. **Sample** is a part of the population from which information is collected. Population always represents a target and we can learn about the population by sampling. In some cases the population could be physically listed (Finite Population) e.g. Books in a library, Members in a community. In other cases the population is abstract(Hypothetical Population) e.g. a mine producing copper, if the mine keeps using the same equipment, raw materials and methods of production also in the future then the copper that will be produced in the mine constitute to hypothetical population.

Descriptive and Inferential Statistics

There are three branches of statistics;

- Descriptive statistics – This involves the methodologies for summarizing and describing of data. It involves the use of graphs, charts and tables and the calculation of various descriptive measures such as measures of variation, percentiles and averages.
- Inferential Statistics – This is concerned with the use of sample data to make an inference. Methods such as point estimation, interval estimation and hypothesis testing which are all based on probability testing are used.

Example

Consider an event of tossing dice. The dice is rolled 100 times and the results are forming the sample data. Descriptive statistics is used to group the sample data to the following table

Outcome of the roll	Frequencies in the sample data
1	10
2	20
3	18
4	16
5	11
6	25

Inferential statistics can be used to verify whether the dice is fair or not.

Collection and Classification

Characteristics that vary from one thing or one person to another is called a variable. A variable is any characteristic that varies from on individual of the population to another. Examples of variables for humans are height, weight, number of siblings, sex, marital status and eye color. The first three yield numerical information and are examples of **Quantitative Variables**, the last three yield non-numerical information and are examples of **Qualitative Variables**.

Quantitative variables are classified as either **Discrete** or **Continuous**.

A discrete variable is one that results from a counting process and holds values that are distinct and certain. For example, the number of male students in a class is a discrete value, e.g. 5 or 10 male students. The variable usually does not assume an in-between value, e.g. 6.5 male students.

A continuous random variable on the other hand usually results from a measuring process. For example, height, length and weight are measurements that can assume values within a range of values (continuum).

Scales

There are four levels of measurement that we can use – nominal, ordinal, interval and ratio. It is important to know what the scale of our data is, as the techniques used to analyse the data depends on the scale of the data.

Nominal scale

A scale in which individuals or items are assigned to categories that have no numerical properties. This scale is used for categorical data. Numerical values can be assigned to the levels of a nominal variable. For example, gender we can code the data, that is, Male as 1 and Female as 2. There is no scale involved with this measurement. This enables computation using a statistical package. The values 1 and 2 do not have any scalar importance. That is, 2 is not bigger than 1.

Ordinal scale

A Scale in which individuals or items are categorized and the categories are ranked. For example when you are asked whether a services offered are – very bad, bad, satisfactory, good, very good – the categories are ranked to reflect an increasing scale of satisfaction. Thus if you were to label very bad as '1', bad as '2' and so on, the value 2 holds a higher ranking than the value 1.

Interval scale

A scale in which the units of measurement (intervals) between the numbers on the scale all equal in size. An example of an interval scale is temperature. The difference between 28 degrees Celsius and 30 degrees Celsius is a constant unit of 2 degrees Celsius. The interval scale however uses an arbitrary zero point. That is, when the temperature in a room is zero, it does not mean that the room is void of heat.

Ratio scale

A scale in which, addition to order and equal units of measurement, there is an absolute zero that indicates an absence of the variable being measured. Observations such as time, height and weight are measured in a ratio scale.

Sampling and Data Collection

Where do data sets come from? There are various sources of data varying according to cost, convenience, how well they will satisfy the investigators needs and how strong the conclusions can be when using them. When you control the design of the data-collection plan, you obtain **Primary Data.** When you use data previously collected by others for their own purposes, you obtain **secondary data.**

Primary Data

Primary data is collected and used by the investigator for himself for the purpose of a specific study or inquiry. You can design the data gathering process by designing the types of questions and measurements. Sources of primary data are either **Censuses** or **Samples**.

Frequency Distributions

We describe how numeric data can be organized into frequency distributions of various types, their geographical presentation and their interpretation and use.

Before data is obtained from a statistical survey or investigation has been worked on, it is known as **Raw Data.** Raw data has little information as they are especially if there are hundreds or thousands of data items. . One of the ways of extracting some information is by arranging the data into size order know as **Data Array.** In this way some information can be known such as the lowest or highest of the data collected. The standard form in which the data is organized is known as **frequency distribution.**

Simple Frequency Distribution.

A simple frequency distribution contains a list of data values, each showing the number of items having that value (called frequency). This distribution is normally applied to discrete raw data since the data values are repeated many times. This structure is not suitable for continuous data since the values repeated could be small. The values obtained can be recorded and presented with the aid of a tally chart.

Example: The frequency distribution resulting from examining the number of children in the families of 47 workers in a company.

Raw data:

1 1 3   2 0 2   0 1 2   2 1 3   5 2 4   0 0 3   4 1 1   2 2 0
3 0 0   2 1 3   6 0 2   1 0 3   2 2 2   1 0 0   1 1 3   1 4

Tally Chart

Data Value	Tally Marks	Total
0													11
1														12
2															13
3								6
4					3
5			1
6			1

Frequency Distribution Table

Number of Children in family	Number of workers
0	11
1	12
2	13
3	6
4	3
5	1
6	1

The tally chart has been constructed examining each value and recording its occurrence with a stroke from the frequency distribution table.

Cumulative Frequency

There are times when we are interested in finding the relative position of a given value in a distribution. For example we may be interested in finding how many or what percentage of our sample was older than 40 or younger than 60. The distributions can be presented in a cumulative way. Therefore, A Cumulative Frequency distribution shows the frequencies at or below each category of a variable. It describes the number of items that have values either below or above a particular level. Cumulative Frequencies are appropriate only for variables that are measured at an ordinal level or higher. The data is presented graphically either as a pie chart or bar graph (Horizontal or Vertical) for discrete data, Histogram for continuous data. One disadvantage of graphs maybe that the values may not read accurately.

Example: A doctor's office staff studied the waiting times for patients who arrive at the office with a request for emergency service. The following data with waiting times in minutes were collected over a one month period.

2	**5**	**10**	**12**	**4**	**4**	**5**	**17**	**11**	**8**
9	**8**	**12**	**21**	**6**	**8**	**7**	**13**	**18**	**3**

Using classes of 0 – 4, 5 – 9 and so on we construct a frequency distribution table. To construct a frequency distribution start with the lowest class interval or lowest score if the data is ungrouped and add it to the frequencies in the next highest class interval. The cumulative frequency in the last class interval will be equal to the total number of cases.

No. of minutes	**Frequency(*f*)**	**(*cf*)**
0 - 4	4	4
5 - 9	8	12
10 - 14	5	17
15 - 19	2	19
20 - 24	1	20
Total	**20**	

Results Presentation

Data are often collected in raw form. These are then not useable unless summarized. There are certain guidelines for data summarization such as summarization

- should be as useful as possible,
- should represent data fairly, and
- should be easy to interpret.

After collection of data (primary or secondary), it is necessary to summarize them suitably and present in such forms as can facilitate subsequent analysis and interpretation. There are two major tools/techniques for presentation of data as follows:

– Presentation in tabular form
– Presentation in graphical form.

Tabular Presentation

Data may be presented in the form of statistical tables. In one table only simple frequencies can be shown. Also, in the same table cumulative frequencies, relative frequencies, and cumulative relative frequencies can be shown. Relative frequencies and cumulative frequencies are defined as follows:

- Relative frequency: It means the ratio of the frequency in the category of concern to the total frequency in the reference set

$$Relative\ frequency\ of\ x_1 = \frac{actual\ frequency\ of\ x_1}{sum\ of\ all\ frequencies} = proportion$$

- Cumulative frequency of x_1 = sum of all frequencies of all values up to and including x_1.

In the same table the simple frequencies combined with one or more but not all, of the cumulative frequencies, relative frequencies, and cumulative relative frequencies may be shown.

Graphical Presentation

Data presented in the form of tables give good information in concise form. Tables provide all relevant information of the data. Apart from tabular presentation, graphical presentation of data has also become quite popular. It gives visual information in addition to magnitudes. Furthermore, comparisons and changes in the data can be well visualized when presented in graphical form. A very useful part of graphical presentation is the interpretation of the graphs. In every graph we should try to interpret the data.

With the help of computer software packages such as Harvard Graphics, Lotus 123, Energraphics, etc., graphical presentation of data can be made in a variety of ways. But these may broadly be categorized into the following:

(a) Bar chart - Bar charts are used for categorical data or metric data that are transformed into categorical data. Categories are shown on the horizontal axis. Frequency, percentage, or proportion is shown on the vertical axis. Bars are separated from each other to emphasize the distinctness of the categories. The bars must be of the same width. The length of each bar is proportional to the frequency, percentage, or proportion in the category. Levels ought to be provided on both axes.

(b) Pie chart - Like bar charts, pie charts are also used for categorical data. A circle is divided into segments, the areas of which are proportional to the values in the question. But the areas are proportional to the angles the corresponding segments make at the center of the circle. Thus, segments of the circle are cut in such a way that their values are proportional to the angles. In one pie chart only values of one variable can be shown. However, two or more pie charts may be constructed side by side for comparison or to study the change over time

(c) Histogram - are used for metric data but converted to categories. These are somewhat similar to bar charts. However, there are some important features in histograms. The blocks in histograms are placed together one after another. These are not separated. Classes are ordered on the horizontal axis, with scores increasing from left to right. Areas of the blocks are proportional to the frequencies. If the class intervals are of equal width, the heights of the blocks/rectangles are proportional to the frequencies. If the class intervals are of unequal width, the blocks/rectangles are drawn in such a way that the areas of the blocks/rectangles are proportional to the frequencies. However, it is easier to interpret the histograms, if the class intervals are of equal width.

Measures of Central Tendency

So far we have learnt that frequency distributions and graphical techniques are useful tools when presenting information. The advantage of using frequency distributions graphs is that information can easily be understood, but when we need to describe a large data set that involves many variables, they may not be the most efficient tools. Another way of describing the distribution is by selecting a single number that summarizes the distribution concisely.

Measures of tendency are numbers that describe what is average or typical of the distribution. There are three measures of tendency: **The mode, the median** and **the mean.** Each of them describe what is typical, central or representative of a distribution.

The Mean – This the widely known and most used average. It is the arithmetic average obtained by adding up all the values and dividing by the total number of values. If the *n* observations are $x_1, x_2, x_3, \dots, x_n$, their mean is

That is $arithmetic\ mean = \frac{\text{the sum of all the values}}{\text{the number of values}}$

Here's an example where a salesman completes 4, 5, 12, 18, 12 sales in consecutive weeks. Find the mean.

$$mean = \frac{4 + 5 + 12 + 8 + 2}{5}$$

$$= \frac{31}{5} = 15.5$$

Usually the symbol $\bar{x}$ (pronounced as 'x bar') to represent the mean sample. The general formula for the mean sample of n items is

$$\bar{x} = \frac{x_1, + x_2 + \ x_3 + \ \dots, x_n}{n}$$

Or in a compact notation

$$\bar{x} = \frac{1}{2} \sum x_i$$

The Greek symbol Σ in the formula for the mean is short for 'add them all up' or 'Summation.'

The Mode – This describes the category or value with the highest frequency (percentage) in a distribution/data set. To obtain the mode(s) of a variable, we first construct a frequency distribution for the data, then the mode can be determined easily from the frequency distribution. For example we have collect test scores for students arranged from lowest to highest: 2 , 3 ,3, 4, 4 ,4, 5, 5, 6. The score of 4 is the mode because it occurs more frequently than the other scores.

The Median – This is the value of a variable in a data set that divides the set of observed values in half, so that the observed values in one half are less than or equal to the median value and the observed values in the other half are greater or equal to the median value. Here's a rule for finding the median:

1. Arrange all observations in order of size, from smallest to largest
2. If the number of observations n is odd, the median is the center observation in the ordered list. Find the location of the median by counting (n+1)/2 observations up from the bottom of the list
3. If the number of observations n is even, the median M is the mean f the two center observations in the ordered list. The location of the median is again (n+1)/2 from the bottom list

The formula (n+1)/2 does not give the median, just the location of the median in the ordered list.

Example: Participants in a bike race had the following finishing times in minutes: 28, 22, 26, 29, 21, 23, 24. What is the median?

Placing out the data in increasing order – 21, 22, 23, 24, 26, 28, 29

$$\frac{7+1}{2} = 4$$

The position of the median is 4 hence the median is 24.

Measures of Dispersion

Previously we were concerned with the measures of location or position. However, having obtained a measure of location or position of a distribution, we need to know how the data is spread about that point. Two data sets of the same variable may exhibit similar positions of center but maybe remarkably different.

Range – The sample range of the variable is the difference between its maximum and minimum values in a data set. The sample range is obtained by computing the difference between the largest observed value of the variable in the data set and the smallest one.

Range = Max – Min

Example: 8 participants in a horse race had the following finishing times in minutes: 28, 22, 26, 29, 21, 23, 24, 50. What is the Range? 50 – 21 = 29

Interquartile Range

Measures that split the data into four quarters. The first quartile measures where 25% of the data is and the third quartile measures where 75% of the data is. In order to calculate the quartiles we must first arrange the numbers in ascending order. Let us use the numbers below.

10 12 14 15 15 15 16 17 18 20

To see where the first quartile is, we use the following equation:

$$Q_1 = \frac{(n+1)}{4}$$

Since there are 10 numbers, n is equal to 10. Q_1 works out to be 11/4 = 2.75. This number should be rounded up to 3. The value 3 is a positional value. That is, the first quartile is the third number from the left. Therefore, the first quartile is 14.

The third quartile is calculated as:

$$Q_3 = \frac{3(n+1)}{4}$$

For the above data, Q_3 works out to be 33/4=8.25. This is rounded to 8. Therefore, the third quartile is 8 numbers from the left which is 17.

IQR is the difference between the third quartile and the first quartile. A general rule is necessary to set out for locating the first, second, and third quartiles in a set of data. If there are n observations arranged in ascending order, then the location of the first quartile is (n + 1)/4, the location of the second quartile, i.e., the median is (n + 1)/2 and the location of the third quartile is 3(n + 1)/4. If (n + 1) is not an integer multiple of 4, then the quartiles are to be found out by interpolation.

- Standard Deviation
  - Tells us how much observations in our sample differ from the mean value within our sample
- Standard Error
  - Tells us how much the sample mean represents the population mean.
  - It is the standard deviation of the sampling distribution of a statistic

Standard Deviation

The standard deviation is the most widely used measure of dispersion, since it is directly related to the mean. If you chose the mean as the most appropriate measure of central location, then the standard deviation would be the natural choice for a measure of dispersion. The standard deviation measures the differences from the mean; a larger value indicates large variation. The standard deviation is in the same units as the actual observations. For example if the observations are in cm, even the standard deviation will be in cm. To calculate the standard deviation, we follow the following steps;

i. compute the mean $\bar{x}$

ii. Calculate the differences from the mean $(x - \bar{x})$

iii. Square the differences $(x - \bar{x})^2$

iv. Sum the squared difference i.e. $\sum\ (x - \bar{x})^2$

v. Take the average of the sum of the squared differences in (iv) to find the variance i.e.

$s^2 = \frac{\sum\ (x-\bar{x})^2}{n-1}$ For a sample and $\sigma^2 = \frac{\sum\ (x-\bar{x})^2}{N}$ for a population

vi. Square root of the variance gives the standard deviation

$\boldsymbol{x} = \sqrt{\frac{\sum\ (x-\bar{x})^2}{n-1}}$ For a sample and $\sigma = \sqrt{\frac{\sum\ (x-\bar{x})^2}{n-1}}$ for a population

Variance

Variance is defined as the square of standard deviation i.e $variance = SD^2$. If $SD = 4$, then variance is 16. It is an important Concept for statistical test called Analysis of Variance.

The variance of a population data is the mean of the squares of the deviations of the observations of all the values from the mean. In other words, the variance is the square of the standard deviation. It takes every observation into account and it is based on an average deviation from the central value.

Strengths of Variance

- Variance makes the measure independent of the length of the list of numbers.
- As long as the mean and standard deviation are provided, it is easy to calculate.
- It is not affected by changes in absolute values or units of measurements.
- Although it cannot be applied to other statistical analysis, it is however part of other calculations such as standard deviation and ANOVA.

Weaknesses of Variance

- It is measured in square units as compared with the number in the original list.
- It is not used in further statistical analysis apart from a few like standard deviation and ANOVA.
- It does not make sense for a single data set.

STATISTICS THREE

ESTIMATION AND SAMPLING DISTRIBUTIONS

Interval Estimate

An interval estimate is defined by two numbers, between which a population parameter is aid to lie. For example, $a < x < b$ is an interval estimate of the population mean μ. It indicates that the population mean is greater than a but less than b.

Confidence Interval

- The drawback with point estimators is that:
  - In practice the point estimator will not be exactly equal to the population parameter it is meant to measure
  - We need some estimation method or procedure which gives us an appreciation of the size of error we make when we adopt a certain estimator.

This is where confidence intervals come in

- A confidence interval is an interval in which the probability that a population parameter lies in it is known.
- It is constructed from sample observations
- Usually we talk of a 100(1-α)% confidence intervals.

Confidence interval for the difference of two population means

Introduction

From each of two populations an independent random sample is drawn. Sample means x_1 and x_2 are calculated. The difference is $x_1 - x_2$, which is an unbiased estimator of the difference between the two population means $\mu_1 - \mu_2$. The variance of the estimator is $(\frac{\sigma^2{}_1}{n_1} + \frac{\sigma^2{}_2}{n_2})$.

Conditions for use

Assuming the populations are normally distributed, there are three situations where we would determine the 100 (1-α) percent confidence interval for $(\mu_1 - \mu_2)$.

- where the population variances are known (use z)
- where the population variances are unknown but equal (use t)
- Where the population variances are unknown but unequal (use t'). There is some acceptance of t' for cases such as these.

The concept of t' is noted here so that readers are aware of its existence but it will not be treated further in this section.

When the population variances are known, the 100 (1-α) percent confidence interval for $(\mu_1 - \mu_2)$ is given by $(x_1 - x_2) \pm\ z_{(1-\frac{\alpha}{2})\sqrt{\frac{\sigma^2{}_1}{n_1}+\frac{\sigma^2{}_2}{n_2}}}$

Example

A research team is interested in the difference between serum uric acid levels in patients with and without Down's syndrome. In a large hospital for the treatment of the mentally retarded, a sample of 12 individuals with Down's syndrome yielded a mean of = 4.5 mg/100 ml. In a general hospital a sample of 15 normal individuals of the same age and sex were found to have a mean value of = 3.4 mg/100 ml. Construct the Confidence Interval for the Difference of Two Population Means assuming that the two populations of values are normally distributed with variances equal to 1 and 1.5, at 95% confidence level.

1. **Data**: $n_1 = 12, x_1 = 4.5, \sigma^2{}_1$ and $n_2 = 15, x_2 = 3.4, \sigma^2{}_2 = 1.5$
2. **Calculations**
   - The point estimate for $\mu_1 - \mu_2$ is $x_1 - x_1 = 4.5 - 3.4$
     $= 1.1$
   - The standard error is $\sqrt{\frac{\sigma^2{}_1}{n_1} + \frac{\sigma^2{}_2}{n_2}} = \sqrt{\frac{1}{12} + \frac{1.5}{15}}$
     $= 0.4282$
   - The 95% confidence interval is $(x_1 - x_2) \pm z_{\left(1-\frac{\alpha}{2}\right)}\sqrt{\frac{\sigma^2{}_1}{n_1} + \frac{\sigma^2{}_2}{n_2}}$
     = $1.1 \pm 1.96(0.4282)$
     = $(0.26,1.94)$
3. **Discussion**: As this is a z-interval, we know that the correct value of z to use is 1.96. We interpret this interval that the difference between the two population means is 1.1 and we are 95% confident that the true mean lies between 0.26 and 1.94.

Population variances are unknown but can be assumed to be equal (t is used)

If it can be assumed that the population variances are equal then each sample variance is actually a point estimate of the same quantity. Therefore, we can combine the sample variances to form a pooled estimate.

The pooled estimated of the common variance is made using weighted averages. This means that each sample variance is weighted by its degrees of freedom. The pooled estimate of the variance comes from the formula: $S^2{}_p = \frac{(n_1-1)S^2{}_1+(n_2-1)S^2{}_2}{n_1+n_2-2}$

Activity

Calculate the standard variance and deviation for the following data set:

a) **32 37 29 40 35 26 45 37 34 29 30 34 56 74 40 48 45 43 32 35**

b)

Class Interval	Frequency
2-9	2
10-17	6
18-25	12
26-33	5
34-41	3
42-49	2

STATISTICS FOUR

HYPOTHESIS TESTING

Introduction

Hypothesis testing is that branch of statistical inference that is used to verify claims made concerning population parameters. In this unit, you will be introduced to the terminology used in hypothesis testing. You will learn how to conduct hypothesis tests concerning the mean and proportion of a single population. You will also learn about significance tests for the difference between two population means for both independent and correlated (paired) samples and for difference of two population proportions.

Objectives

By the end of the unit, you should be able to:

- define a statistical hypothesis
- formulate the null hypothesis and the alternative hypothesis for a given situation
- distinguish between Type I and Type II errors
- distinguish between one-tailed and two-tailed tests
- outline the steps followed in the procedure of hypothesis testing
- conduct hypothesis tests concerning the population mean and the population proportion
- distinguish between matched and independent samples
- conduct hypothesis tests for difference of two population means and for difference of two population proportions
- carry out a test for the mean difference of paired samples

Statistical Hypothesis

A statistical hypothesis is a claim or guess or an assumption or a statement, which may or may not be true, made concerning a population parameter.

Hypothesis testing involves gathering evidence from a random sample drawn from the population of interest in order to decide whether the null hypothesis is likely to be true or false. The hypothesis is rejected if evidence from the sample is not consistent with the stated hypothesis, otherwise it is accepted. However, the acceptance of the stated hypothesis does not necessarily imply that it is true, rather it is a result of insufficient evidence to reject it.

Types of hypothesis

There are two types of hypothesis which are called the:

- null hypothesis, and
- alternative hypothesis

A null hypothesis is an assertion about the value of a population parameter. It is a formal statement of the claim being made concerning a population parameter. The null hypothesis is denoted by H_0.

The alternative hypothesis, denoted by H_1, is the negation of the null hypothesis. For example, a null hypothesis might assert that the population mean is equal to a specified value μ_0. We write this as $H_0: \mu = \mu_0$. The alternative hypothesis oppose this assertion and it is written as $H_1: \mu \neq \mu_0$.

In this case, the alternative hypothesis suggests that the mean takes values that are below μ_0 or above it. Therefore, to investigate H_0 we conduct a non-directional test which is known as a two-tailed test.

A null hypothesis might assert that the population mean is at least equal to a certain specified value μ_0. We write $H_0: \mu \geq \mu_0$. In this case, the alternative hypothesis would consist of value below μ_0, that is, $H_1: \mu < \mu_0$. Similarly, if the null hypothesis assert that the population mean is less than or equal to a specified value μ_0, that is, $H_0: \mu \leq \mu_0$, the alternative will be $H_1: \mu > \mu_0$. In both these cases, since the alternative hypotheses consist of values either below or above the specified value μ_0, we conduct a one-sided test or a one-tailed test.

Deciding on the null hypothesis

Determining what the null hypothesis should be in a given situation may prove to be difficult. However, if the null hypothesis is wrongly formulated then the test will be pretty meaningless. The following notes will be handy in deciding what the null hypothesis should be:

- The null hypothesis is always has an element of equality: either an equal to (=) sign or a greater or equal to (≥) sign or a less or equal (≤) sign is used in the expression of H_0.
- The null hypothesis is usually an expression of a claim made by someone. However, if the claim does not include an equal to (=) sign it becomes the alternative hypothesis.
- The null hypothesis is the hypothesis that we formulate with the hope of rejecting.
- If the null hypothesis is true, then no corrective action would be necessary, whereas if the alternative hypothesis is true, some corrective action would be necessary.

Example 1

A statistician claims that the average age of a student at Chalimbana University is 21. However, the Dean of students doubts this claim. Set up the null hypothesis if the DOSA wishes to show that it is not 21.

Solution

$H_0: \mu = 21$

$H_1: \mu \neq 21$

Example 2

A G & G bakery claims that the average cost of producing a standard loaf of bread is K8 kwacha. If you suspect that the claim is exaggerates the cost, how would you set up the null and alternative hypothesis?

Solution

$H_0 : \mu \geq 8$

$H_1 : \mu < 8$

Activity

1. A ZRA officer at Nakonde border post claims that it takes, on average, at most 2 days for a truck driver to clear his consignment. You suspect that the average is greater than 2 days and you want to test the claim. State the null and alternative hypothesis for this test.
2. An ice cream vending machine is set to dispense 100 grams per cup. You suspect that the machine is under-filling the cups. Set up the null and alternative hypothesis to investigate this case.

6.3 Type I and Type II Errors

In deciding to reject or accept a null hypothesis, there will be chances for erroneously rejecting or accepting it. Such errors may be due to faulty sampling procedures

A type I error is committed when a true null hypothesis is rejected. The probability of committing a type I error is called the level of significance and it is denoted by α. It is common to use 1%, 5% and 10% level of significance in calculations. If $\alpha = 0.05$ then there are only 5 in 100 chances of committing this error.

A type II error is committed if we accept the null hypothesis when it is false. The probability that the test will be able to detect a false null hypothesis is called the power of a test. In other words, the power of the test is the probability of rejecting H_0 when indeed H_0 is false.

6.4 Steps Followed in Hypothesis Testing

The following steps should be followed when conducting a hypothesis test:

Step 1: State the null and alternative hypothesis

The null and alternative hypotheses are specified at the initial stage before gathering any evidence. It would be unethical and rather manipulative to formulate the H_0 and H_1 at one’s convenience after gathering evidence; a practice that we refer to as data snooping.

Step 2: Identify the distribution

For problems in this unit, you have to choose between two distributions namely the z-distribution and the t-distribution.

When testing for the distribution mean μ we use:

a) The z-distribution when the:

- population standard deviation σ is known
- population standard deviation is unknown and n is large $(n \geq 30)$

b) The t-distribution when the population standard deviation σ is unknown and the sample size n is small $(n \geq 30)$.

When testing for a population proportion and sample size is large we use the z-distribution.

Rejection Region

- Rejection Region: is a set of values of a statistic for which a null hypothesis is rejected.
- Critical Values: are values of the boundaries of the rejection region on which we base our decision to reject the Null hypothesis or not.
- Significance Level: is the probability of rejecting H_0 when it is true. It must be specified prior to performing the test and it is denoted by α. The most common value of alpha (α) is 0.05 or 5%.

Hypothesis Testing of the Difference between Two Population Means

B) Hypothesis testing of the difference between two population means

This is a two sample z test which is used to determine if two population means are equal or unequal. There are three possibilities for formulating hypotheses.

I. $H_0 : \mu_1 = \mu_2$: H_A: $\mu_1 \neq \mu_2$

2. $H_0 : \mu_1 \geq \mu_2$; H_A: $\mu_1 < \mu_2$

3. $H_0 : \mu_1 \leq \mu_2$; H_A : $\mu_1 > \mu_2$

Procedure

The same procedure is used in three different situations

• Sampling is from normally distributed populations with known variances

$$Z = \frac{(x_1 - x_2) - (\mu_1 - \mu_2)0}{\sqrt{\frac{\sigma^2{}_1}{n_1} + \frac{\sigma^2{}_2}{n_2}}}$$

• Sampling from normally distributed populations where population variances are unknown

◦ Population variances equal: $t = \frac{(x_{1-x_2}) - (\mu_1 - \mu_2)0}{\sqrt{\frac{S^2{}_p}{n_1} + \frac{S^2{}_p}{n_2}}}$

This is with t distributed as Student's t distribution with $(n_1 + n_2$ -2) degrees of freedom and pooled variance.

When population variances are unequal, a distribution of t is used in a manner similar to calculations of confidence intervals in similar circumstances.

• Sampling from populations that are not normally distributed

If both sample sizes are 30 or larger the central limit theorem is in effect. The test statistic is

Hypothesis Testing of the Difference Between Two Population Means

If the population variances are unknown, the sample variances are used.

Sampling from normally distributed populations with population variances known

Example

Given that $x_1 = 4.5, n_1 = 12, \sigma^2 = 1, x_2 = 3.4, n_2 = 15, {\sigma^2}_2 = 1.5\ and\ \alpha = 0.05,$ Is there a difference between the means between individuals with Down's syndrome and normal individuals?

Solution

(1) Data

x_1= 4.5, n_1 = 12, ${\sigma^2}_1$ = 1, x_2= 3.4, n_2 = 15, ${\sigma^2}_2$ = 1.5 and α= .05

(2) Assumptions

• two independent random samples

• each drawn from a normally distributed population

(3) Hypotheses

H_0: $\mu_1 = \mu_2$

H_A: $\mu_1 \neq \mu_2$

(4) Test statistic

This is a two sample z test.

(a) Distribution of test statistic, H_0 is true.

If the assumptions are correct and is true, the test statistic is distributed as the normal distribution.

(b) Decision rule

With α = .05, the critical values of z are -1.96 and +1.96. We reject if z < -1.96 or z > +1.96.

(5) Calculation of test statistic

$$Z = \frac{(x_1 - x_2) - (\mu_1 - \mu_2\ )0}{\sqrt{\frac{{\sigma^2}_1}{n_1} + \frac{{\sigma^2}_2}{n_2}}}$$

$$= \frac{(4.5 - 3.4) - 0}{\sqrt{\frac{1}{12} + \frac{1.5}{15}}}$$

$$= \frac{1.1}{0.4282}$$

= 2.57

(6) Statistical decision

Reject H_0 because 2.57 > 1.96.

(7) Conclusion

From these data, it can be concluded that the population means are not equal. A 95% confidence interval would give the same conclusion.

p = .0102.

Sampling from normally distributed populations with unknown variances

With equal population variances, we can obtain a pooled value from the sample variances.

Example

Lung destructive index

We wish to know if we may conclude, at the 95% confidence level, that smokers, in general, have greater lung damage than do non-smokers given the following set of data:

Smokers: $x_1 = 17.5, n_1 = 16, {S^2}_1 = 4.4752.$

Non-smokers: $x_2 = 12, n_2 = 9, {S^2}_2 = 4.8492$

(1) Data

Smokers: x_1 = 17.5, n_1 = 16, ${S^2}_1$ = 4.4752

Non-Smokers: = 12.4 = 9 = 4.8492
α = 0.05

Calculation of Pooled Variance: $S^2{}_p = \frac{(n_1-1)S^2{}_1+(n_2-1)S^2{}_2}{n_1+n_2-2}$

$= \frac{15(4.4711)+8(4.8492)}{16+9-2}$

$= \frac{299.86+188.12}{23}$

$= 21.2165$

(2) Assumptions
• Independent random samples
• Normal distribution of the populations
• Population variances are equal

(3) Hypotheses
H_0: $\mu_1 \leq \mu_2$
H_A: $\mu_1 > \mu_2$

(4) Test statistic: t = $\frac{(x_1-x_2)-(\mu_1-\mu_2)0}{\sqrt{\frac{S^2p}{n_1}+\frac{S^2p}{n_2}}}$

(a) Distribution of test statistic
If the assumptions are met and H_0 is true, the test statistic is distributed as Student's t distribution with 23 degrees of freedom.

(b) Decision rule
With α = 0.05 and df = 23, the critical value of t is 1.7139. We reject if t> 1.7139.

(5) Calculation of test statistic: t = $\frac{(17.5-12.4)-0}{\sqrt{\frac{21.2165}{16}+\frac{21.2165}{9}}}$

$= \frac{5.1}{1.92}$

$= 2.6563$

(6) Statistical decision
Reject H_0 because 2.6563 > 1.7139.

(7) Conclusion
On the basis of the data, we conclude that $\mu_1 > \mu_2$.

Paired comparisons

Sometimes data comes from non-independent samples. An example might be testing "before and after" of cosmetics or consumer products. We could use a single random sample and do "before and after" tests on each person. A hypothesis test based on these data would be called a paired comparisons test. Since the observations come in pairs, we can study the difference, d, between the samples. The difference between each pair of measurements is called di.

Test statistic

With a population of n pairs of measurements, forming a simple random sample from a normally distributed population, the mean of the difference, is tested using the following implementation of t.

Paired comparisons

Example

Very-low-calorie diet (VLCD) Treatment

Table gives B (before) and A (after) treatment data for obese female patients in a weight-loss program.

We calculate di = A-B for each pair of data resulting in negative values meaning that the participants lost
weight.
We wish to know if we may conclude, at the 95% confidence level, that the treatment is effective in causing weight reduction in these people.
n = 9
= .05
Activity
These data were obtained in a study comparing persons with disabilities with persons without disabilities. A scale known as the Barriers to Health Promotion Activities for Disabled Persons (BHADP) Scale gave the data. We wish to know if we may conclude, at the 99% confidence level, that persons with disabilities score higher than persons without disabilities based on the following set of data:
Disabled: x_1 = 31.83, n_1 = 132 and S_1 = 7.93
Nondisabled: x_2 = 25.07, n_2 = 137 and S_2 = 4.80, given that α= .01

UNIT FOUR

PROBABILITY

The **probability** of something happening is the likelihood or chance of it happening. Values of probability lie between 0 and 1, where 0 represents an absolute impossibility and 1 represents an absolute certainty. The probability of an event happening usually lies somewhere between these two extreme values and is expressed either as a proper or decimal fraction.

If p is the probability of an event happening and q is the probability of the same event not happening, then the total probability is p C q and is equal to unity, since it is an absolute certainty that the event either does or does not occur, i.e. ***p* Y *q* = 1.**

In probability, we define an experiment, an outcome, the sample space and an event as follows

Experiment - An experiment is an action or process that can be repeated. We do not know the result of an experiment before it is carried out, but we can identify the results that occur. Example, Flipping two coins ad record the sides which are facing upwards

Outcome – an outcome is simply the result of an experiment example, Heads or tails when flipping a coin

Sample Space - The sample space is the set of all possible outcomes of an experiment. We use s to denote the sample space of an experiment. Example S= {HH, HT, TH, TT} if a coin is flipped

Event – an event is a collection of one or more outcomes of an experiment. It is always the subset of the sample space. A simple event consists of a single element in the sample space. An event which consists of one or more than one element in the sample space is called a compound event. Example the sides facing upwards when a coin is flipped

Probability is expressed as a **fraction**: the denominator is the total number of ways things can occur and the numerator is the number of things that you are hoping will occur:

$$probability = \frac{number\ of\ ways\ the\ event\ can\ occure}{number\ of\ possible\ outcomes}$$

Experiment: Roll a die and record the number facing upwards
Event: ‘the number facing upwards less than 5’
Sample space: S= {1, 2, 3, 4, 5, 6}

$$P(number\ on\ die\ less\ than\ 5) = \frac{number\ of\ element\ in\ the\ sample\ space\ that\ are\ less\ than\ 5}{number\ of\ elements\ in\ the\ sample\ space}$$

$$p = \frac{4}{6}$$

The **expectation**, E, of an event happening is defined in general terms as the product of the probability p of an event happening and the number of attempts made, n, i.e. ***E* = *pn*.**

Thus, since the probability of obtaining a 3 upwards when rolling a fair dice is 1 6, the expectation of getting a 3 upwards on four throws of the dice is $\frac{1}{6}x4 = \frac{2}{3}$1 6 ð 4. Thus expectation is the average occurrence of an event.

Dependent event

A **dependent event** is one in which the probability of an event happening affects the probability of another ever happening

Independent event

An independent event is one in which the probability of an event happening does not affect the probability of another event happening.

Conditional probability

Conditional probability is concerned with the probability of say event B occurring, given that event A has already taken place. If A and B are independent events, then the fact that event A has already occurred will not affect the probability of event B. If A and B are dependent events, then event A having occurred will affect the probability of event B.

The addition law of probability

The multiplication law of probability

Normal Approximation to the binomial. The standardized normal random variable Z has a mean of zero and a standard deviation of 1.

Binomial Distribution

If P is the probability of success in an individual trial, then n independent trails the probability of x successes is given by the Binomial formula:

P(x) = ${}^{n}C_{x}P^{X}(1-\boldsymbol{P})^{\boldsymbol{n-x}}$ where x = 0, 1, 2, 3,…., n. The number of successes has a mean of np and standard deviation = $\sqrt{np(1-P)}$The numbers of trials are fixed. We donate this by the number n as in the formula above

- The n trails are independent and repeated under identical conditions
- Each trail has only two outcomes: success, denoted by S and failure denoted by F.
- For each individual trial, the probability of success is the same. We donate the probability of success by p and that of failure by q. Since each trail results in either success or failure $p + q = 1$ and $q = 1 - p$.
- The main problem of binomial experiment is to find the probability of x success out of the n trials.

Alternatively, when we multiply out the power of the binomial brackets such as $(a + \boldsymbol{b})^{\boldsymbol{n}}$ we say that we are expanding the term and the result is called a binomial expansion.

$$\boldsymbol{(a+b)^1 = a + b}$$
$$\boldsymbol{(a+b)^2 = a^2 + 2ab + b^2}$$
$$\boldsymbol{(a+b)^3 = a^3 + 3a^2b + 3ab^2 + b^2}$$
$$\boldsymbol{(a+b)^4 = a^4 + 4a^3b + 6a^2b^2 + 4ab^3 + b^4}$$
$$\boldsymbol{(a+b)^5 = a^5 + 5a^4b + 10a^3b^2 + 10a^2b^3 + 5ab^4 + b^5}$$

Let us now make some observations regarding the above expressions of $\boldsymbol{(a+b)^n}$ looking at the results, we see four things:

- The 'sum' of the powers is constant on each line for example $(\boldsymbol{a}+\boldsymbol{b})^{\mathbf{4}}$ total power on each term is 4.
- The power of $\boldsymbol{a}$ is decreasing from left to right example $\boldsymbol{a}^{\mathbf{4}}\boldsymbol{a}^{\mathbf{3}}\boldsymbol{a}^{\mathbf{2}}\boldsymbol{a}^{\mathbf{1}}$
- The power of b is increasing from left to right example $\boldsymbol{b}^{\mathbf{1}}\boldsymbol{b}^{\mathbf{2}}\boldsymbol{b}^{\mathbf{3}}\boldsymbol{b}^{\mathbf{4}}$
- The numbers form a triangle. The triangle is known as **the Pascal's triangle** which can be used to obtain binomial coefficients. The numbers are arranged as follows:

$n = 0$ 1

$n = 1$ 1 1

$n = 2$ 1 2 1

$n = 3$ 1 3 3 1

$n = 4$ 1 4 6 4 1

$n = 5$ 1 5 10 10 5 1

$n = 6$ 1 6 15 20 15 6 1

$n = 7$ 1 7 21 35 35 21 7 1

$n = 8$ 1 8 28 56 70 56 28 8 1

Starting with this triangle and then using the rules given in (1), (2) and (3) above, we can write down any expansion of$(\boldsymbol{a}+\boldsymbol{b})^{\boldsymbol{n}}$**.** For example expand $(\mathbf{3}\boldsymbol{a}+\mathbf{4}\boldsymbol{b})^{\mathbf{3}}$ we start with the line of Pascal's triangle for
$\mathbf{1}(\mathbf{3}\boldsymbol{a})^{\mathbf{3}}$+$3(\mathbf{3}\boldsymbol{a})^{2}(\mathbf{4}\boldsymbol{b})^{3}+(\mathbf{3}\boldsymbol{a})(\mathbf{4}\boldsymbol{b})^{\boldsymbol{b}}+\mathbf{1}(\mathbf{4}\boldsymbol{b})^{3}$
$n = 3$ which gives us 1, 3 , 3, 1 we get $\mathbf{27}\boldsymbol{a}^{\mathbf{3}}+\mathbf{108}\boldsymbol{a}^{\mathbf{2}}+\mathbf{144}\boldsymbol{a}\boldsymbol{b}^{\mathbf{2}}+\mathbf{64}\boldsymbol{b}^{\mathbf{3}}$

Binomial Expansion or the Binomial Theorem: A power of a Binomial can aslo be expanded by means of the Binomial Theorem.
$(\boldsymbol{a}+\boldsymbol{b})^{\boldsymbol{n}}=\boldsymbol{a}^{\boldsymbol{n}}+\frac{\boldsymbol{a}^{\boldsymbol{n}-\mathbf{1}}\boldsymbol{b}}{1!}+\frac{\boldsymbol{n}(\boldsymbol{n}-\mathbf{1})\boldsymbol{a}^{\boldsymbol{n}-\mathbf{2}}\boldsymbol{b}^{\mathbf{2}}+}{2!}+\frac{\boldsymbol{n}(\boldsymbol{n}-\mathbf{1})(\boldsymbol{n}-\mathbf{2})\boldsymbol{a}^{\boldsymbol{n}-\mathbf{3}}\boldsymbol{b}^{\mathbf{3}}}{1!}+\cdots$
For example, use the Binomial Theorem to expand $\mathbf{3}\boldsymbol{a}+\mathbf{4}\boldsymbol{b})^{\mathbf{3}}$

$$\mathbf{3}\boldsymbol{a}+\mathbf{4}\boldsymbol{b})^{\mathbf{3}}=\mathbf{3}\boldsymbol{a}^{\mathbf{3}}+\frac{\mathbf{3}(\boldsymbol{a})^{\mathbf{2}}\boldsymbol{b}}{1!}+\frac{\mathbf{3}(\mathbf{2})(\mathbf{3}\boldsymbol{a})(\mathbf{4}\boldsymbol{b})^{\mathbf{2}}+}{2!}+\frac{\mathbf{3}(\mathbf{2})(\mathbf{4}\boldsymbol{b})^{\mathbf{3}}}{3!}$$

= $\mathbf{27}\boldsymbol{a}^{\mathbf{3}}+\mathbf{108}\boldsymbol{a}^{\mathbf{2}}+\mathbf{144}\boldsymbol{a}\boldsymbol{b}^{\mathbf{2}}+\mathbf{64}\boldsymbol{b}^{\mathbf{3}}$
Which agrees with the answer in the previous example. You'll notice that the last term $\frac{\mathbf{3}(\mathbf{2})(\mathbf{1})}{3!}$ cancels out.

The Poisson distribution

A Poisson distribution is a probability distribution that is useful in describing the number of events that will occur in a specific time period or specific area or volume. This distribution is quite widely used in quality control. For example, it can be used to measure the number of defective spots in a car door panel. Other uses include:
Number of industrial accidents at a factory per month.
Number of errors per newspaper page.
Number of customers per hour at a certain bank teller.
The obvious difference between a binomial and a Poisson is the way the average for a Poisson process is expressed. The average is expressed as an average of an interval.

Characteristics of the Poisson distribution

The Poisson being a discrete distribution has the following characteristics:

1. The process has discrete random events.
2. The probability of the event does not change over time.
3. The probability of an event is independent of previous events.

Formula

The formula for a Poisson distribution is as follows:

Note that the symbol λ is used in some textbooks as opposed to .

Where: e = base of natural logs = 2.71828

or λ = mean number of occurrences

X = value of random variable

Example

λ = 3, find the Poisson probabilities $p(0), p(1), p(2), p(3)$ and $p(4)$.

$$\boldsymbol{p(x)=\frac{e^{-\lambda}\lambda^{x}}{x!}, x = 0, 1, 2, 3, 4}$$

$$\boldsymbol{p(0)=\frac{e^{-3}3^{0}}{0!} = e^{-3} = 0.0498}$$

$$\boldsymbol{p(1)=\frac{e^{-3}(3)^{1}}{1!} = 3e^{-3} = 0.1494}$$

$$\boldsymbol{p(2)=\frac{e^{-3}3^{2}}{2!} = \frac{9}{2}e^{-3} = 0.2241}$$

$$\boldsymbol{p(3)=\frac{e^{-3}3^{3}}{3!} = \frac{27}{6}e^{-3} = 0.2241}$$

$$\boldsymbol{p(4)=\frac{e^{-3}3^{4}}{4!} = \frac{81}{24}e^{-3} = 0.1681}$$

Sampling is a process of examining a representative number of items (people or things) out of the whole population or universe. To find a good sample is often not easy. For example, we sample the fruit at the top of a basket we may have no idea if there is a bad fruit in the middle or at the bottom of the basket. If we are studying college students' attitudes and we interview students on the steps of the School of business, we may never encounter an IT student. A sample only provides an estimated population measure while the accuracy of the estimate will depend on:

(i) The right sizes for a sample, the larger the sample size the greater is the probability that the sample is representative of the population.

(ii) Selecting the right sampling method so that the sample represents the population.

(iii) The extent of variability in the population.

Here are some of the definitions in sampling;

Population – a population is the set of all the individuals or objects which have a given characteristic, e.g. the set of all persons eligible to vote in a given country.

Sample – a sample is a subset of a population, e.g. the voters selected for questioning about their views.

Sampling – sampling is the process of taking a sample.

Sample Survey – the process of collecting the data from a sample is called a sample survey, e.g. asking the selected voters their political views is a sample survey.

Census – the process of collecting data from a whole population is called a census, e.g. a population census in which data about the entire population of a country is collected. (Note that the ten-yearly population census taken in the Zambia is one of the few questionnaires that the head of a household is compelled by law to complete.) You may ask a question why we sample. We sample because of the following reasons;

- **Time and cost** are probably the two most important reasons for sampling. If one wants to determine if customers in a given market will buy a product, one doesn't usually have the time or funds to interview all potential customers. For example, try interviewing everyone who uses Boom detergent paste in Zambia.
- **Testing** may prove destructive. If you want to test the durability of your product, stress tests on the entire output may leave you with no product to sell.
- **Accuracy** is another reason for sampling. One would think that a study of the entire population would be more accurate than a study of a sample. If one has a very large population and wishes to take a complete count, then one must hire a large number of inventory takers who must work at a rapid rate. As more personnel are hired, it is likely that they may be less efficient than the original employees. Thus, a limited number of skilled workers, studying a well defined sample, may provide more accurate results than a survey of the entire population.
- **Quality** - When only a few interviews are needed, it is easier to devote a greater degree of effort and control per interview than with a larger number of interviews. This will lead to better quality interviews and to a greater proportion of the questions being answered correctly without the necessity of a call-back. (A call-back is when an interviewer has to return to the respondent, if that is possible, in order to clarify the answer to a question.)
- **Control-** A sample survey is easier to control than a complete census. This greater control will lead to a higher response rate because it will be possible to interview every member of the sample under similar conditions. A comparatively small number of interviewers will be needed, so standardization of the interviews will be easier.
- **Industrial Application** - Sampling is not confined to surveys such as opinion polls which involve interviews and postal questionnaires; it is also very important in controlling industrial production. On an assembly line in a factory producing manufactured goods, it is necessary to check the standard of the product continuously. This is done by selecting a sample of the same product every day and testing to establish that each item in the sample meets the manufacturer's specifications. Sometimes this testing involves destroying the product. For example, in a tyre factory each tyre will be required to have a minimum safe life in terms of distance driven and to withstand a minimum pressure without a blowout. Obviously the whole population of tyres cannot be tested for these qualities. Even when the testing involves nothing more than measuring the length of a bolt or the pitch of a screw, a sample is used because of the saving in time and expense.

Sampling Methods –Refer to unit of this module

Statistical Inferences

Among the reasons for taking a sample is that the data collected from a sample can be used to infer information about the population from which the sample is taken. This process is known as **statistical inference**. The theory of sampling makes it possible not only to draw statistical inferences and conclusions from sample data, but also to make precise probability statements about the reliability of such inferences and conclusions.

Before we continue we must define some terms which are generally used in statistical inference:

Parameter – a constant measure used to describe a characteristic of a population. **Statistic** – a measure calculated from the data set of a sample. **Estimate** – the value of a statistic which, according to sampling theory, is considered to be close to the value of the corresponding parameter. **Sampling unit** – an item from which information is obtained. It may be a person, an organisation or an inanimate object such as a tyre. **Sampling frame** – is a list of all the items in a population. The sampling theory which is necessary in order to make statistical inferences is based on the mathematical theory of probability. We have already discussed probability in unit four and you may wish to revisit it. Statistical inference however, can be divided into two parts namely estimation and hypothesis testing. Firstly we deal with estimation which is the procedure or rules or formulae used to estimate population characteristics or parameters. Sample measures or statistics are used to estimate population measures (such as population means, μ the Greek symbol 'mu', population variance σ 2 the Greek symbol 'Sigma'). The corresponding sample measures are sample mean, ' x 'pronounced x-bar, and sample variance ' $S2$ ' respectively. Hypothesis testing is the process of establishing theory or hypothesis about some characteristics of the population and then draw information from a sample to see if the hypothesis is supported or not **Estimations;** There several properties of good estimators including the following**;**

(i) Unbiasness. An estimator is said to be unbiased if the mean of the sample mean x of all possible random samples of size n , drawn from a population of size N, equals the population mean (μ) . Therefore the mean of the distribution of the sample means will be the same as the population mean.
(ii) Consistency. An estimator is said to be consistent if, as the sample size increases, the accuracy of the estimate of the population parameter also increases.
(iii) Efficiency. An estimator is said to be efficient than any other if, it has the smallest variance among all the estimators.
(iv) Sufficiency. An estimator is said to be sufficient if it utilizes all the information about it to estimate the required population parameter.

In practical situation, it is not possible to have all the four qualities on one estimator. The researcher chooses which qualities he/she wants the estimator to have.

Distribution of Sample Means: Consider a population with mean = μ and standard deviation = σ . Samples of size n are taken from this population and the sample means x are found. The distribution of sample means has mean $\mu\ x=\mu$ and standard deviation (standard error) $\sigma\ x= \sigma n\surd$. If the population is normally distributed, or if the sample size is 'large' (i.e. $n>30$) , then the sample means is approximately normal. For example, A normal population has a mean = 500 and standard deviation = 125. Find the probability that a sample of 65 values has a mean greater than 538. We have $\mu=500$ *and* σ =125. The distribution of sample means has mean We have $\mu=500$ *and* σ =125. The distribution of sample means has mean $\mu\ x=500$ *and s* tan *dard error* $\sigma\ x\ =\ \sigma$ $\surd n=$ 125 $\surd 65$ =15.50 Therefore the probability that a sample mean greater than 538 is equal to the area shaded = 0.5000 – 0.4929 = 0.0071 from the tables. $\sigma\ x$=15.50 μx=500 538Sample means Z=538-500 15.50 =2.45 In another example, the daily output from a production line has a mean 7500 units with a standard deviation of 500 units. What is the probability that during the next 125 days the average output will be under 7400 units per day? We have μ=7500 *and* σ=500 . The population of sample means is pproximately Normal (since n is large) with μx=μ=7500 *and* $\sigma\ x = \sigma\ \sqrt{n}$= 500 √125=44 72 $\sigma\ x$=44 .72 7400 μx=7500 Sample means Z=7400-7500 44 .72 =-2 .2 P(*Sample ean*<7400)=*area shaded* =0.5000-0 .4875 =0.0125

Regression Analysis

Regression analysis attempts to establish the 'nature or type of relationship' between variables. That is, to study the functional relationship between the variables and thereby provide a mechanism for prediction or forecasts. Regression modelling is one of the most widely used techniques in business and academia to study the relationship between two variables. It is also one of the easiest models to construct! If an analyst for instance is trying to predict the share price of a particular sector there will be a whole range of independent variables to be considered. In this unit, we will restrict our attention to the particular case where a dependent variable y is related to a single independent variable x . The Regression Equation is such that when only one independent variable is used in making forecast, the technique used is called Simple Linear regression. The forecasts are made by means of a straight line using the equation $y=a+bx$ Where; *a=the y-intercept when x=0, b=slope=the amount that y changes with a unit change in x* The linear function is useful because it is mathematically simple and it can be shown to be reasonably close to the approximation of many situations. The first step to establish whether there is a relationship between variables is by means of a scatter diagram. This is a plot of the two variables on an *x- y* graph. Given that we believe there is a relationship between the two variables, the second step is to determine the form of this relationship.

For example consider the following data for a major Harry store. The daily high temperature and air conditioning units sold for 8 with randomly selected business days during the hot dry season.

Daily Temperature	**Number of Units**
27	5
35	6
18	2
20	1
46	6
36	4
26	3
23	3

Draw a scatter diagram

The distribution of points in the Scatter diagram suggests that a straight line roughly fits these points. The most straight forward method of fitting a straight line to the set of data points is ‘by eye’. The values of ‘a’ and ‘b’ can then be determined from the graph, ‘a’ is the intercept on the vertical axis and ‘b’ is the slope. The other method is that of semi averages. This technique consists of splitting the data into two equal

groups, plotting the mean point for each group and joining these points with a straight line.
Another example, using data in the Harry store example above, fit a straight line using the method of semi-averages. The procedure for obtaining the *y* on *x* regression line is as follows:
Step 1: Sort the data into size order by *x* - value**.**

x	y
18	2
20	1
23	3
26	3
27	5
35	6
36	4
46	1

Step 2: Split the data up into two equal groups, a lower half and upper half (if there is an odd of items drop the central one).

Lower Half		**Upper Half of data**	
x	y	x	y
18	2	27	5
20	1	35	6
23	3	36	4
26	3	46	1
Total : 87	**9**	**144**	**16**
Average: 21.75	**2.5**	**36**	**4**

Step 3: Calculate the mean point for each group Step 4: Plot the mean points in Step 3 on a graph within suitably scaled axes and joining them with a straight line. This is the required *y* on *x* regression line. **Least Square Line:** Let us consider a typical data point with coordinates $(\boldsymbol{x_i}, \boldsymbol{y_i})$ (See Figure 3.0). The error in the forecast (y coordinate of data point-forecasted coordinate as given by the straight line) is denoted by $\boldsymbol{e_i}$. The line which minimizes the value of $\boldsymbol{e_i}$ is called the "least square line" or the regression line. This can be shown by using calculus. Here we just give the 'best estimates of ' a ' and' $\boldsymbol{b}$ ' by the following formula

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

$a = \bar{y} - b\bar{x}$ where n is the number of data points
The values of $\boldsymbol{a}$ and $\boldsymbol{b}$ are then substituted into equation $\bar{\boldsymbol{y}} = \boldsymbol{a} + \boldsymbol{bx}$

The least squares line with the error term $\boldsymbol{e_i} \cdot \boldsymbol{x}$ If you fit the least squares line to the data in the above example, the table below shows the calculations for the estimates of $\boldsymbol{a}$ and $\boldsymbol{b}$.

x	y	x^2	y^2	xy
18	2	324	4	36
20	1	400	1	20
23	3	529	9	69
26	3	676	9	78
27	5	729	25	135
35	6	1225	36	210
36	4	1296	16	144
46	1	2113	1	46
$\boldsymbol{\sum x = 231}$	$\boldsymbol{\sum y = 25}$	$\boldsymbol{\sum x^2 = 7294}$	$\boldsymbol{\sum y^2 = 101}$	$\boldsymbol{\sum xy = 738}$

$\boldsymbol{n = 8}, \bar{y} = \mathbf{3.125}\,, \bar{x} = 23.1$

$$b = \frac{\sum xy - \frac{\sum x \sum y}{n}}{\sum x^2 - \frac{(\sum x)^2}{n}} \ldots\ b = \frac{738 - \frac{(231)(25)}{8}}{7295 - \frac{(231)^2}{8}} \ldots = \frac{16.125}{624.875} \ldots\ b = 0.0258$$

$$b = \bar{y} - b\bar{x} \ldots = 3.125 - 0.0258(28.875) \ldots = 2.38$$

Thus giving the equation for the regression line of

$$\boldsymbol{y = 2.28 + 0.0258}$$

Forecasting Using the Regression Line: Having obtained the regression line, It can be used to forecast the value of *y* for a given value of *x* . Suppose that we wish to determine the number of units sold if we have a daily temperature of 42*o c* .From the regression line the forecasted value *y* is ^ *y*=2 .38+0.0258(42)=3. 46 i.e. the expected number of units sold is 3. Now suppose that we wish to determine the number of units sold if the temperature is 49*oc*. The forecasted value of *y* is the given by ^ *y*=2 .38+0. 0258(49)=3 .64 i.e. the expected number of units sold is 4. The two examples differ due to the fact that the first *y* value was forecasted from an *x* value within the range of *x* values, while the second value outside the range of *x* values in the original data set. The first example is a case of **interpolation** and the second is that of **extrapolation.** With extrapolation, the assumption is that the relationship between the two variables continues to behave in the same way outside the given range of *x* values from which the least square line was computed.

Correlation Analysis

Correlation analysis is used to determine the degree of association between two variables. If two variables tend to change together in the same direction, they are said to be **positively correlated.** If they tend to change together in opposite directions, they are said to be **negatively correlated.** The degrees of correlation are **perfectly correlated** meaning that all points on the scatter graph lie exactly on the straight line. The other being **uncorrelated** meaning that the points on a scatter graph appear to be randomly scattered with no suggestion of any relationship. The **partly correlated** means that the points lie scattered in such a way that although they do not lie on a straight line, they do display a general tendency to be clustered around such a line. Having obtained the equation of the regression line, correlation analysis can be used to measure how well one variable is linearly related to another. The coefficient of correlation *r* can assume any value in the range -1 *to*+1 inclusive. A value of *r* is close to or equal to – 1, we have a negative correlation, while the value of r close to or equal to 1, we have a positive correlation. If a value r is close or equal to 0, we have no correlation. The sign of the correlation coefficient is the same as the sign of the slope of the regression line. The following scatter diagrams illustrate certain values of the correlation coefficient.

The method of investigating whether a linear relationship exists between two variables x and y is by calculating Pearson's product moment correlation coefficient (PPMCC) denoted by r and given by the formula:

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

For example, by calculating the PPMCC find the degree of association between weekly earnings and the amount of tax paid for each member of a group of 10 manual workers.

Weekly Wage($K'000$)	79	21	87	88	91	92	98	98	103	113
Income Tax ($K'000$)	10	8	14	14	17	12	18	22	21	24

The PMCC is calculated in the table below

x	y	x^2	y^2	xy
79	10	6241	100	790
81	8	6561	64	648
87	14	7569	196	1218
88	14	7744	196	1232
91	17	8281	289	1547
92	12	8464	144	1104
98	18	9604	324	11764
98	22	9604	484	2156
103	21	10609	441	2163
113	24	12769	576	2712
$\sum x = 930$	$\sum y = 160$	$\sum x^2 = 87446$	$\sum y^2 = 2814$	$\sum xy = 15334$

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

$$r = \frac{15334 - \frac{(930)(160)}{10}}{\left(\sqrt{7295} - \frac{(930)^2}{10}\right)\left(2814 - \frac{(160)^2}{10}\right)}$$

$$\frac{454}{\sqrt{(956)}(254)} = 0.921$$

r is near 1 and indicates a strong positive linear correlation between the two variables.

The Coefficient of Determination: The coefficient of determination is the square of the coefficient of correlation r. In words, it gives the proportion of the variation (in the y - values) that is explained (by the variation in the x - values). In Example 10, the correlation coefficient is $r = 0.839$. Therefore coefficient of determination: $r^2(0.839)^2 = 0.704$ (To 3 decimals) This means that only 70.4% of the variation in the variable y is due to the variation in the variable x . Note that the coefficient of etermination $r2$ is between 0 and +1 inclusive. **Spearman's Rank Correlation Coefficient:** An alternative method of measuring correlation is by means of the Spearman's rank correlation coefficient obtained by the formula.

$$r = 1 - \frac{6\sum d^2}{n(n^2-1)}$$

Where $d = difference\ between\ rankings$

In this example, two members of an interview panel have ranked seven applicants in order of preference for a specified post. Calculate the degree of agreement between the two members.

Applicant	**A**	**B**	**C**	**D**	**E**	**F**	**G**
Interviewer X	1	2	3	4	5	6	7
Interviewer Y	4	3	1	2	5	7	6

The differences in rankings below.

D	-3	-1	2	2	0	-1	1
d^2	9	1	4	4	0	1	1

$$\sum d = 0, \sum d^2 20 \ldots r = 1 - \frac{6\sum d^2}{n(n^2-1)} \ldots 1\ \frac{6(20)}{7(49^2-1)} \ldots 1 - \frac{120}{336}$$

$$= 1 - 0.3571 \ldots r = 0.6429$$

Activity

Conduct a multi-variant regression analysis for the data you collected in unit one: thereafter, determine and interpret the values of R^2 and adjusted R^2

UNIT FIVE

FINANCIAL MATHEMATICS

Introduction
Unit contents
Unit learning objectives

When we borrow money or invest it during financial dealings, *interest (I)* is paid to the lender by the borrower.
There are three factors that determine the interest being paid

1. The amount of money borrowed. The larger the amount borrowed, the larger the interest. We call the amount borrowed as the Principle (P) or capital amount (A). Let's refer to this amount as present value (P_v).
2. The length of *time* which it is borrowed. The longer the time, the larger the interest. It is calculated by dividing the total time into periods (days, weeks, months, quarters, years etc). Each period money gains interest. At the end of each period, the money is worth the original amount plus the interest. Let n = number of periods. New amount is referred to as *future value (F_v)*.
$$F_v = P_v + 1$$
3. The *rate* at which interest is changed. $\boldsymbol{i}, \boldsymbol{j}$ *and* $\boldsymbol{r}$ Are symbols commonly used to denote interest rate. The interest rate is the percentage or portion of the money borrowed which will constitute the interest for one period.
Example if K5, 000 is borrowed for a period of 1 year at an interest rate of 14% per year, the interest will be:
$$I = 5000 \cdot 0.14 = 700$$
Interest rates are expressed as percentages but also as decimal fractions.
$$i = 17\% = \frac{17}{100} = 0.17$$
Simple Interest
It is calculated on the initial amount borrowed or invested no matter how many interest periods are involved.
An amount of K100 is deposited into a bank for unspecified number of years, with an interest rate of 12% per annum(p.a) simple interest.
$$P_v = 100, \qquad i = 12\% p.a$$
The interest earned after 1 year:
$I = \ P_v \cdot i$
= 1000(0.12) = 12
The accumulated amount after 1 year:
$F_v = P_v + I$
=$P_v + P_v \cdot i = 112$
The interest for the second year would similarly amount to:
$I = P_v \cdot i$
=100(0.12)= 12
The interest earned after 2 years:
$I = P_v \cdot i \cdot 2$
=100(0.12) ·2
=24
The accumulated amount after 2 years:
$F_v = P_v + I$
=$P_v + P_v i \cdot 2$
= 124

Interest earned after n years:
$I = P_v \cdot i \cdot n$
The accumulated amount after n years
$F_v = P_v + I$
= $P_v + P_v \cdot i \cdot n$
F_v= $P_v(1 + i \cdot n)$

- Interest earned(I), Future Value F_v of the money invested for n periods at simple interest rate of i per period can be stated as

$$I = P_v \cdot i \cdot n$$
$$F_v = P_v(1 + i \cdot n$$

Compound Interest

If the interest accrued after an interest period added to the principle, so that for the next period based on this new principle amount (old principle plus interest), then the interest is said to have been compounded. That is, compound interest is paid on the initial principle and previously earned interest. In simple interest, the interest was a constant amount as calculated at the initial period for every interest period. In compound interest, it is calculated on the changing principle amount and the interest will change accordingly for each interest period.

Using the previous example the interest earned after year 1(I_1) =12, accumulated amount after 1 year (F_{v1}) =112.

The interest earned after 2 years
(I_2) = $F_{v1} \cdot i + I_1$
=$F_{v1} + P_v \cdot i$
$= 112(0.12) + 100(0.12)$
= 25.44
The Accumulated amount after 2 years
$F_{v2} = P_v + I_2$
=100 + 25.44
=125.44
These results could be repeated for n years.

NOMINAL AND EFFECTTIVE INTEREST RATES

The primary difference between simple and compound interest is that compound interest includes interest on the interest earned in the previous record, while simple does not. The nominal and effective interest rates have also the same basic relationship.

- The difference here is that the concepts of nominal and effective must be used when interest is compounded more than once each year.
- **Nominal interest rate, r,** is an interest rate that does not include any consideration of compounding

r = interest rate per period x number of periods

A nominal rate r may be stated for any time period, 1 year, 6 months, quarter, month, week, day, etc.

Effective interest rate is the actual rate that applies for a stated period of time. The compounding of interest during the time period of the corresponding nominal rate is accounted for by the effective interest rate.

- It is commonly expressed on an annual basis as the effective annual rate ia, but any time basis can be used.
- An effective rate has the compounding frequency attached to the nominal rate statement.

REAL INTEREST

Amortization Of Loans

Amortization is the repayment of a loan in periodic payments also known as in **installment loan**. Let's assume loans are paid in equal sizes of installments even though it's not an efficient way of paying them, it has become a common way of paying mortgages and car loans because majority of people have a fixed budget every month and require uniform payments. A mortgage is an agreement attached to a debt obligation that specifies the property or collateral that the borrower gives up interest in if they fail to make payments when promised

Sinking Fund

Sinking Fund is created for settling loans. When money accumulated in this fund at the end of a period, it must equal the total loan principle that must be paid. Meanwhile before the loan is due, only the interest is paid periodically on the agreed dates to the creditor. When theres a difference between the settlement fund and the principle at a certain time, it is called the loan book value. This is because the settlement fund is prepared specifically for the loan settlement. Therefore the annual payment that must be made by the debtor is the amount for the settlement fund and periodic interest.

Time Value of Money

An asset is tangible or intangible entity possessing monetary value. The value of an asset in today's currency is referred to as its Present value and the value of an asset at some point in the future is referred to as its future value.

The central idea of Time Value Analysis is to be able to recast the monetary value of some or more future cash flows into equivalent present monetary value.

Future Value

The face of value of an asset at some point in the future is referred to as its Future Value (FV). The future value of an asset after one period with Present Value and periodic Rate of return r is

$$FV_1 = PV + PVr = PV(1 + r)$$

The future value for the k^{th}period of the future value of the previous plus the rate of return for that period given by

$$FV_k = FV_{k-1}(1 + r)$$

The future value for the kth period may be expressed in terms of the present value as

$$FV_k = PV(1 + r)^k$$

Intuitive derivation of equation

The following is not a formal proof but it illustrates that the result is true by intuition

The Future Value after one year is – $FV_1 = PV(1 + r)$

The future value after two years is – $FV_2 = FV_1(1 + r) = PV(1 + r)^2$

The future value after three years is $FV_3 = FV_2(1 + r) = PV(1 + r)^3$

Continuing this fashion we conclude that the future value after k years is

$$FV_K = FV_{k-1}(1 + r) = PV(1 + r)^k$$

Example: consider one-time investment of K10, 000 for a term of 20 years with a nominal interest rate of 10% per year.

$PV = 10{,}000,$ $r = 0.1,$ $n = 20$

The future value of investment after 20 years is

$$FV = PV(1 + r)^n = 10000(1.1)^{20} = 67,275$$

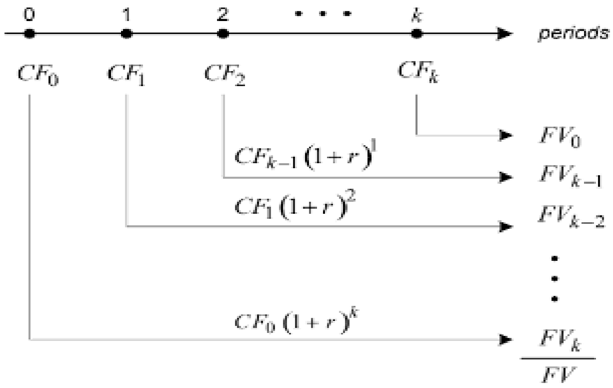
Future Value of cash flow stream

The future value of series n cash flows with periodic interest is given by he serios

Fv = $\sum_{i=0}^{k} CF_i(1 + r)^{k-i}$

= $CF_0(1 + r)^k + CF_1(1 + r)^{k-1} \ldots CF_n(1 + r)^0$

The future value of this series of cash flows is illustrated by the following cash flow diagram



Example Consider the series of cash flows: 100 in year 0, 200 in year 1 and 50 in year 2 with annual interest rate of 7%.

$r = .07, \qquad CF_0 = 100, \quad CF_1 = 200, \quad CF_2 = 50$

$$FV = \sum_{i=0}^{2} CF_i(1+r)^{2-i} = 100(1.07)^2 + 200(1.07) + 50 = 378.48$$

The present value

The present value with Discount rate r can be obtained from a future value that as accrued over a term of k years by

$$PV_k = \frac{FV_k}{(1+r)^k}$$

The following example illustrates the fact that discontinuing a future value to a present value is the inverse operation of fining the future value for a present value

Example: Assume a present value of K100,000 has accumulated after 8 years at kn interest rate of 7%.

$PV = 100{,}000, \qquad r = 0.1, \qquad n = 8$

The future after 8 years is

$$FV = PV(1+r)^k = 100{,}000(1.07)^8 = 107{,}181.8$$

Now consider you are presented with an opportunity to receive K107, 181.8, 8 years from today. What is that investment worth in today's Kwacha?

$$PV = FV(1+r)^{-k} = 10{,}718.2(1.07)^{-8} = 100{,}00$$

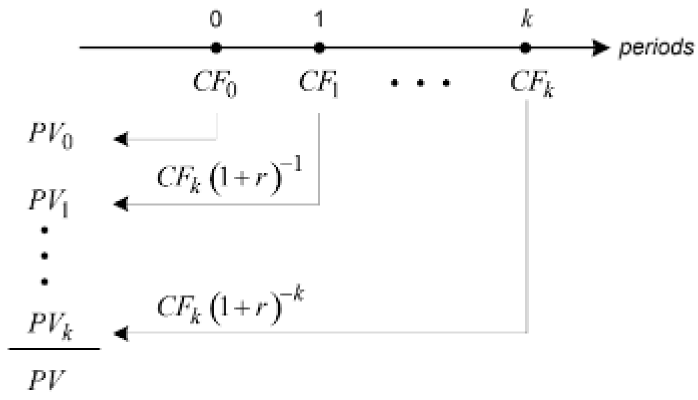
The present value of a cash flow stream

The present value of a stream of cash flows is found by taking the present value of each element in the series and is given by

$$PV = \sum_{i=0}^{2} \frac{CF_i}{(1+r)^i}$$

$$CF_0 + \frac{CF_1}{(1+r)} + \frac{CF_2}{(1+r)^2} \dots + \frac{CF_k}{(1+r)^k}$$

The present value of this series of cash flows is illustrated by the following cash flow diagram



Activity

Given a series of cash flows: 100 in year 0, 200 in year 1, 50 in year 2 with an annual interest rate of 7%:

$r = .07, \quad CF_0 = 100, \quad CF_1 = 200, \quad CF_2 = 50$

Present value of this cash flow stream after the third year and show that it reduces to:

:$PV = \sum_{i=0}^{2} CF_i (1+r)^{-i} = 100 + 200(1+1.07)^{-1} + 50(1+1.07)^{-2} = 330.58$

UNIT SIX

MATRIX ALGEBRA AND LINEAR PROGRAMMING

Introduction

A matrix is a collection of numbers ordered by rows and columns. The number of rows in a

matrix is usually specified by m and the number of columns by n and a matrix referred to as an 'm by n' matrix. For example $\begin{bmatrix} 5 & 8 & 2 \\ 1 & 0 & 7 \end{bmatrix}$ has two rows and three columns, so it is referred to as a '2 by 3' matrix.

There are special types of matrix such as:

1. **Vector Matrix** – This is a matrix which has one row called a vector row or one column called a vector column example

$$\begin{bmatrix} 5 \\ 0 \\ 7 \end{bmatrix} \text{ Or } \begin{bmatrix} 1 & 0 & 7 \end{bmatrix}$$

2. **Scalar Matrix** – is a matrix with only one row and one column. It is denoted by lower case italicized letters e.g. $\boldsymbol{x}$
3. **Square Matrix –** This a matrix with equal rows and columns. Example a '3 by 3' matrix
4. **Symmetric matrix** – is a square matrix in which $x_{ij} = x_{ji}$for all i and j. Matrix A is symmetric B is not symmetric

$$A = \begin{bmatrix} 0 & 1 & 5 \\ 1 & 6 & 2 \\ 5 & 2 & 7 \end{bmatrix}, \; B = \begin{bmatrix} 0 & 1 & 5 \\ 6 & 1 & 2 \\ 5 & 1 & 7 \end{bmatrix}$$

5. Diagonal Matrix is a symmetric matrix where all the off diagonal elements are 0. Matrix A is diagonal

$$\begin{bmatrix} 0 & 0 & 0 \\ 0 & 6 & 2 \\ 0 & 2 & 7 \end{bmatrix}$$

6. Identity(Unit) Matrix is one in which all elements of the leading diagonal (\) have a value of 1 and all other elements have value of 0. The identity matrix is denoted by I

$$I = \begin{bmatrix} 1 & 0 & 0 \\ 0 & 1 & 0 \\ 0 & 0 & 1 \end{bmatrix}$$

Just like there are rules for addition, subtraction, multiplication and division of numbers, these rules apply to matrices too.

(i) Addition of Matrices

Corresponding elements in two matrices can be added to form a single matrix.

A= $\begin{bmatrix} 2 & -1 \\ -7 & 4 \end{bmatrix}$ B= $\begin{bmatrix} -3 & 0 \\ 7 & -4 \end{bmatrix}$

$\begin{bmatrix} 2 & -1 \\ -7 & 4 \end{bmatrix} + \begin{bmatrix} -3 & 0 \\ 7 & -4 \end{bmatrix} = \begin{bmatrix} 2+(-3) & -1+0 \\ -7+7 & 4+(-4) \end{bmatrix}$

$= \begin{bmatrix} \mathbf{-1} & \mathbf{-1} \\ \mathbf{0} & \mathbf{0} \end{bmatrix}$

(ii) Subtraction

If A is a matrix and B is another matrix, then (A-B) is a single matrix formed by subtracting the elements of A.

Subtract A= $\begin{bmatrix} -3 & 0 \\ 7 & -4 \end{bmatrix}$ from B= $\begin{bmatrix} 2 & -1 \\ -7 & 4 \end{bmatrix}$

$\begin{bmatrix} 2 & -1 \\ -7 & 4 \end{bmatrix} - \begin{bmatrix} -3 & 0 \\ 7 & -4 \end{bmatrix} = \begin{bmatrix} 2-(-3) & -1-0 \\ -7-7 & 4-(-4) \end{bmatrix}$

$= \begin{bmatrix} 5 & -1 \\ -14 & 8 \end{bmatrix}$

(iii) Multiplication

When a matrix is multiplied by a number, called scalar multiplication, a single matrix results which each element of the original matrix has been multiplied by the number

Example:

If A = $\begin{bmatrix} -3 & 0 \\ 7 & -4 \end{bmatrix}$ B = $\begin{bmatrix} 2 & -1 \\ -7 & 4 \end{bmatrix}$ C= $\begin{bmatrix} 1 & 0 \\ -2 & -4 \end{bmatrix}$

Find 2A - 3B + 4C

For scalar multiplication each element is multiplied by the scalar quantity.

2A = 2 $\begin{bmatrix} -3 & 0 \\ 7 & -4 \end{bmatrix} = \begin{bmatrix} 6 & 0 \\ 14 & 8 \end{bmatrix}$

3B = 3 $\begin{bmatrix} 2 & -1 \\ -7 & 4 \end{bmatrix} = \begin{bmatrix} 6 & -3 \\ -21 & 12 \end{bmatrix}$

4C = $\begin{bmatrix} 1 & 0 \\ -2 & -4 \end{bmatrix} = \begin{bmatrix} 4 & 0 \\ -8 & -16 \end{bmatrix}$

$= \begin{bmatrix} 6 & 0 \\ 14 & 8 \end{bmatrix} - \begin{bmatrix} 6 & -3 \\ -21 & 12 \end{bmatrix} + \begin{bmatrix} 4 & 0 \\ -8 & -16 \end{bmatrix}$

$= \begin{bmatrix} -6-6+4 & 0-(-3)+0 \\ 14-(-21)+(-8) & -8-12+(-16) \end{bmatrix} = \begin{bmatrix} \mathbf{-8} & \mathbf{3} \\ \mathbf{27} & \mathbf{-36} \end{bmatrix}$

When a matrix A is multiplied by another matrix B, a single matrix results in which elements are obtained from the sum of the products of the corresponding rows of A and the corresponding columns of B. Two matrices A and B may be multiplied together, provided the number of elements in the rows of matrix A are equal to the number of elements in the columns of matrix B. In general terms, when multiplying a matrix of dimensions (m by n) by a matrix of dimensions (n by r), the resulting matrix has dimensions (m by r). Thus a 2 by 3 matrix multiplied by a 3 by 1 matrix gives a matrix of dimensions 2 by 1.

Example:

$$A = \begin{bmatrix} 3 & 4 & 0 \\ -2 & 6 & -3 \\ 7 & -4 & 1 \end{bmatrix} \quad B = \begin{bmatrix} 2 & -5 \\ 5 & -6 \\ -1 & -7 \end{bmatrix}$$

Let A x B = C where C = $\begin{bmatrix} C_{11} & C_{12} \\ C_{21} & C_{22} \\ C_{31} & C_{32} \end{bmatrix}$

C_{11} = is the sum of products of the first row elements of A and the first column elements of B, taken one at a time. $C_{11} = (3x2) + (4x5) + (0 + (-1))$

C_{12}= the sum of the products of the first row elements of A and the second column elements of B, taken one at a time. $C_{12=}\big(3x(-5)\big) + \big(4x(-6)\big) + (0x(-7))$

C_{21} = the sum of the products of the second row elements of A and the first column elements of B, taken one at a time $C_{21} = (-2x2) + (6x5) + (-3 + (-1))$

C_{22} = the sum of the products of the second row elements of A and the second column elements of B, taken one at a time $C_{22} = (-2x(-5) + \big(6x(-6)\big) + (-3x(-7))$

C_{31}= the sum of the products of the third row elements of A and the first column of elements of B, taken of B, taken one at a time $C_{31} = (7x2 + (-4x5) + (1x(-1))$

C_{32}= the sum of the products of the third row elements of A and the second column of elements of B, taken of B, taken one at a time

$$C_{31} = (7x - 5 + \left(-4x(-6)\right) + (1x((-7))$$

$$\begin{bmatrix} (3x2) + (4x5) + (0 + (-1)) & \left(3x(-5)\right) + \left(4x(-6)\right) + (0x(-7)) \\ (-2x2) + (6x5) + (-3 + (-1)) & (-2x(-5) + \left(6x(-6)\right) + (-3x(-7)) \\ (7x2 + (-4x5) + (1x(-1)) & (7x - 5 + \left(-4x(-6)\right) + (1x((-7)) \end{bmatrix}$$

$$= \begin{bmatrix} 26 & -39 \\ 29 & -5 \\ -7 & -18 \end{bmatrix}$$

In algebra, the commutative law of multiplication states that a x b = b x a. For Matrices, this law is only true in a few special cases and in general A X B is not equal to B X A.

The determinant of a 2 x2 matrix

A = $\begin{bmatrix} a & b \\ c & d \end{bmatrix}$ is defined as (ad-bc). The element are written between vertical lines thus $\begin{bmatrix} 3 & -4 \\ 1 & 6 \end{bmatrix}$ is written as $\begin{vmatrix} 3 & -4 \\ 1 & 6 \end{vmatrix}$ and is equal to [(3x6) – (-4 x 1)] = 22

The inverse or reciprocal of a 2 by 2 matrix

The inverse of matrix A is A^{-1} such that A X $A^{-1} = I$

For any matrix $\begin{bmatrix} a & b \\ c & d \end{bmatrix}$ the inverse may be obtained by:

i. Interchanging the positions of a and d

ii. Changing the signs of b and c

iii. Multiplying this new matrix by the reciprocal of the determinant

Let matrix A = $\begin{bmatrix} 3 & -2 \\ 7 & 4 \end{bmatrix}$

$$= \frac{1}{(3X4)-(7X(-2))} \begin{bmatrix} 4 & 2 \\ -7 & 3 \end{bmatrix}$$

$$= \frac{1}{26} \cdot \begin{bmatrix} 4 & 2 \\ -7 & 3 \end{bmatrix}$$

$$= \begin{bmatrix} \mathbf{\frac{2}{13}} & \mathbf{\frac{1}{13}} \\ \mathbf{\frac{-7}{26}} & \mathbf{\frac{3}{26}} \end{bmatrix}$$

The determinant of a 3 x 3 matrix

The minor of an element of a 3 by 3 matrix is the value of the 2 by 2 determinant obtained by covering up the row and column containing that element.

Let matrix A = $\begin{bmatrix} 1 & 4 & 3 \\ -5 & 2 & 6 \\ -1 & -4 & 2 \end{bmatrix}$

The minor element of 1 is obtained by covering the row $(1 \quad 4 \quad 3)$ and the column $\begin{pmatrix} \mathbf{1} \\ \mathbf{-5} \\ \mathbf{-1} \end{pmatrix}$ leaving the 2 by 2 determinant $\begin{vmatrix} 2 & 6 \\ -4 & 2 \end{vmatrix}$ = 28

The sign of a minor depends on its position within the matrix, the sgn pattern being $\begin{pmatrix} + & - & + \\ - & + & - \\ + & - & + \end{pmatrix}$**.** The signed minor of an element is called the **cofactor** of the element. The value of a 3 by 3 determinant is the sum of the products of the elements and

their cofactors of any row or any column of the corresponding 3 by 3 matrix. There are thus six different ways of evaluating a 3 ð 3 determinant — and all should give the same value.

Using the first row: $\begin{vmatrix} \mathbf{1} & \mathbf{4} & \mathbf{3} \\ \mathbf{-5} & \mathbf{2} & \mathbf{6} \\ \mathbf{-1} & \mathbf{-4} & \mathbf{2} \end{vmatrix}$

= 1$\begin{vmatrix} 2 & 6 \\ -4 & 2 \end{vmatrix} - 4\begin{vmatrix} -5 & 6 \\ -1 & 2 \end{vmatrix}$ + (-3) $\begin{vmatrix} -5 & 2 \\ -1 & -4 \end{vmatrix}$

= (4+24) – 4(-10+6) -3(20+2)

= 28 + 16 + 66 = **-22**

Using the second column

= -4$\begin{vmatrix} -5 & 6 \\ -1 & 2 \end{vmatrix}$ + 2$\begin{vmatrix} 1 & 3 \\ -1 & 2 \end{vmatrix}$ – (-4) $\begin{vmatrix} 1 & 3 \\ -5 & 6 \end{vmatrix}$

= -4 (-10 + 6) + 2(2 - 3) + 4(6 - 15)

= 16 – 2 – 36 = **-22**

The inverse or reciprocal of a 3 by 3 matrix

The adjoint of a matrix A is obtained by:

(i) forming a matrix B of the cofactors of A, and

(ii) **Transposing** matrix B to give B^T, where B^T is the matrix obtained by writing the rows of B as the columns of B^T. Then **adj** $\boldsymbol{A}$ = B^T. The **inverse of matrix** $\boldsymbol{A}$, A^{-1} is given by

$$\boldsymbol{A^{-1}} = \frac{\textbf{adj}\ \boldsymbol{A}}{|\boldsymbol{A}|}$$

Where adj A is the adjoint of matrix A and $|A|$ is the determinant of matrix A

Determine the inverse of the matrix $\begin{bmatrix} 1 & 5 & -2 \\ 3 & -1 & 4 \\ -3 & 6 & -7 \end{bmatrix}$

The inverse of matrix A is found by:

(i) Obtaining the matrix of the cofactors of the elements and

(ii) Transposing this matrix

The cofactor element of 1 is $+\begin{vmatrix} -1 & 4 \\ 6 & -7 \end{vmatrix}$ = - 17

The cofactor element of 5 is $-\begin{vmatrix} 3 & 4 \\ -3 & -7 \end{vmatrix}$ = 9

The cofactor element of -2 is $\begin{vmatrix} 3 & -1 \\ -3 & 6 \end{vmatrix}$ = 15 and so on

The matrix of cofactors is $\begin{bmatrix} -17 & 9 & 15 \\ 23 & -13 & -21 \\ 18 & -10 & -16 \end{bmatrix}$

The transpose of the matrix of cofactors i.e the adjoint of the matrix, is obtained by writing the rows as columns and is

$$\begin{bmatrix} -17 & 23 & 18 \\ 9 & -13 & -10 \\ 15 & -21 & -16 \end{bmatrix}$$

The determinant of $\begin{bmatrix} 1 & 5 & -2 \\ 3 & -1 & 4 \\ -3 & 6 & -7 \end{bmatrix}$

= 1(7-24) – 5(-21+12)-2(18 -3)

= 17 + 45 – 30 = -2

The inverse is $\frac{1}{-2}\begin{bmatrix} -17 & 23 & 18 \\ 9 & -13 & -10 \\ 15 & -21 & -16 \end{bmatrix} = \begin{bmatrix} 8.5 & -11.5 & -9 \\ -4.5 & 6.5 & -5 \\ -7.5 & 10.5 & 8 \end{bmatrix}$

Basic Linear Programming

A linear programming problem may be defined as the problem of maximizing or minimizing a linear function subject to linear constraints. The constraints may be equalities or inequalities. Here is a simple example. Find numbers x_1 and x_2 that maximize the sum $x_1 + x_2$ subject to the constraints $x_1 \geq 0, or\ x_2 \geq 0$ and

$$x1 + 2x2 \leq 4$$
$$4x1 + 2x2 \leq 12$$
$$-x1 + x2 \leq 1$$

There are two unknowns and five constraints.

Example

Given the payoff matrix $A = \begin{bmatrix} 6 & 0 & 3 \\ 0 & 10 & 3 \end{bmatrix}$,

evaluate the maxmin mixed strategy for player 1 and the minmax mixed strategy for player 2. Show your answer graphically.

Solution of $n \times m$ Games Using Linear Programming Methods

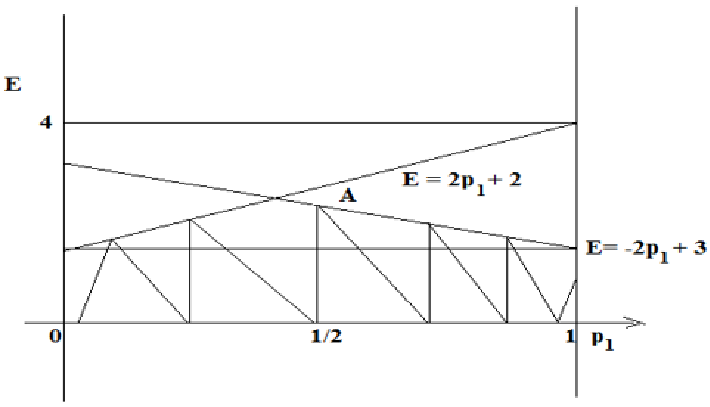
The graphical method is appropriate for solving a game problem when the payoff matrix is $2 \times n$ or $m \times 2$ in dimension. In some cases, we can be confronted by a $m \times n$ payoff matrix with m players and n strategies. In such situations linear programming has been used to solve such large games.

Formulation of the Game Problem as A Linear Programming Problem

Given a payoff matrix $\boldsymbol{A} = \begin{bmatrix} 4 & 1 \\ 2 & 3 \end{bmatrix}$ from the previous section, a graphical solution to the game gives

The new shading unlike the old shading delineates a convex set of points which is closed and bounded from the above. This convex set is the intersection of four closed half spaces defined by the inequalities:

$$E \leq 2p_1 + 2 \qquad (5.1)$$
$$E \leq -2p_1 + 3 \qquad (5.2)$$
$$p_1 \geq 0 \qquad (5.3)$$
$$p_1 \leq 1 \qquad (5.4)$$
$$E \geq 0 \qquad (5.5)$$



Viewed this ways, the search for the maximum point A, becomes a matter of finding the height possible horizontal supporting line to our convex set. If we regard the desired supporting line as a number of a family of horizontal lines generated by a linear objective function, we can express the problem as follows:

Maximise E

Subject to $E \leq 2p_1 + 2$ (5.6)

$E \leq -2p_1 + 3$ (5.7)

$0 \leq p_1 \leq 1$ (5.8)

$E \geq 0$ (5.9)

Since $p_1 + p_2 = 1$, we can transform equation 5.9 and have:

Maximise E

Subject to $E \leq 2p_1 + 2$ (5.10)

$E \leq -2p_1 + 3$ (5.11)

$p_1 + p_2 = 1$ (5.12)

$E \geq 0$ (5.13)

(5.14)

Since this introduces p_2 , we can modify the other two equations accordingly.
We have

$E_1 = [p_1 \quad 1 - p_1]\begin{bmatrix}4\\2\end{bmatrix} = 2p_1 + 2$

$E_2 = [p_1 \quad 1 - p_1]\begin{bmatrix}1\\3\end{bmatrix} = 3 - 2p_1.$

We can alter these equations and have:

$E_1 = [p_1 \quad p_2]\begin{bmatrix}4\\2\end{bmatrix}$

$= 4p_1 + 2p_2$

$E_2 = [p_1 \quad p_2]\begin{bmatrix}1\\3\end{bmatrix}$

$= p_1 + 3p_2$.

Substituting these into (5.12) and (5.13) we have:

$E \leq 4p_1 + 2p_2$ (5.15)

$E \leq p_1 + 3p_2$ (5.16)

and $p_1 + p_2 = 1$ (5.17)

(5.18)

Continues

Or

$-4p_1 - 2p_2 + E \leq 0$ (5.19)

$-p_1 - 3p_2 + E \leq 0$ (5.20)

$p_1 + p_2 - 1 = 0$ (5.21)

E has been introduced as a third choice variable. This compels us to subject it to the non negativity constant $E \geq 0$, if all the entries in the payoff matrix are non negative. If the payoff matrix contain one or more negative elements, the restrictions $E \geq 0$ is violated. We can resolve this problem by applying the transformation theorem, which assures us that the optimal strategy of a game is invariant with respect to the addition of a constant C to every element in the matrix.

Given that the payoff matrix $A = \begin{bmatrix} -3 & -1 \\ -2 & -5 \end{bmatrix}$, we can, by adding the constant sum 5 to every element, transform **A** to:

$$\boldsymbol{A}^* = \begin{bmatrix} 2 & 4 \\ 3 & 0 \end{bmatrix}$$

and then we can use the restriction $E^* \geq 0$, where $E^* = E + 5$ and formulate the problem in terms of E^*. The optimal strategies will come out to be the same, the only adjustment needed in interpreting the solution of the transformed program is to subtract 5 from E^* to get $\overline{E}$.

We can therefore rewrite the objective function as:

$$\prod = Op_1 + Op_2 + E$$

Therefore our linear program will be:

$$Maximise \prod = Op_1 + Op_2 + E$$

subject to $-4p_1 - 2p_2 + E \leq 0$

$-p_1 - 3p_2 + E \leq 0$

$p_1 + p_2 = 1$

$p_1, p_2, E \geq 0.$

This formulation of linear program differs from the standard maximization program because of the presence of the strict equality $p_1 + p_2 = 1$.

The simplex algorithm can now be used to solve the maximization game problem noting that no variables will be needed for the constant. This problem can be overcome by introducing an artificial variable to obtain an initial basic feasible solution.

Alternatively, we can define,

$$p_1 + p_2 \leq 1$$

instead of $p_1 + p_2 = 1$. This is justified on the grounds that should the expected payoff E be positive, then a player would not miss any opportunity of playing the game i.e. he/she will 100 cent of the time and will find $p_1 + p_2 = 1$ rather than $p_1 + p_2 < 1$. Should E be zero, a player will be indifferent as to whether to play or not to play so that he play 80% of the time (i.e. less than 100% of the time) with the result that $p_1 + p_2 < 1$. We can also by not counting 205 of the plays scale up p_1 and p_2 so that they will add up to 1. In summary, as long as $\bar{E} > 0$, the adjustment makes no difference to the game, whereas in the case of the complication and guarantee a solution that satisfies the condition $p_1 + p_2 = 1$.

Incorporating the last modification into the linear program, we have:

$$Maximise\ \prod = 0p_1 + 0p_2 + E$$

$$Subject\ to \begin{bmatrix} -4 & -2 & 1 \\ -1 & -3 & 1 \\ 1 & 1 & 0 \end{bmatrix} \begin{bmatrix} p_1 \\ p_2 \\ E \end{bmatrix} \leq \begin{bmatrix} 0 \\ 0 \\ 1 \end{bmatrix}$$

$$and\ p_1, p_2, E \geq 0.$$

Note:

(a) In the 3×3 coefficient matrix $\begin{bmatrix} -4 & -2 & 1 \\ -1 & -3 & 1 \\ 1 & 1 & 0 \end{bmatrix}$, the third row and the third column each of which contains a series of 1's followed by a 0 borders the 2×2 matrix

$$\begin{bmatrix} -4 & -2 \\ -1 & 3 \end{bmatrix}$$

-1 -3 that represents the negative of the transformed payoff matrix A multiplied by -1, i.e.

$$\begin{bmatrix} 4 & 2 \\ 1 & 3 \end{bmatrix}$$

This suggests a rule for deriving the coefficient matrix of the constraints once the payoff matrix **A** is given. i.e.

$$\begin{bmatrix} -4 & -2 & 1 \\ -1 & -3 & 1 \\ 1 & 1 & 0 \end{bmatrix} = \begin{bmatrix} -A^0 & 1 \\ 1 & 0 \end{bmatrix}$$

(b) The column vector of constants on the right of the constraints consists of a series of 0's followed by a unit element.

Summary of the Conversion Process

(a) Check whether the conversion matrix is non-negative. If so, proceed to (2). If not add an appropriate constant to each element to make the matrix non-negative.

(b) Write a linear objective function in which the relative frequency variables p_i are assigned zero coefficients. The expected payoff variable E is to be taken as the last variable and is given a unit coefficient. Since we are taking player 1's viewpoint, this objective function is to be maximized.

(c) In the constraints, the coefficient matrix on the left is the matrix $-A^0$, bordered with unit elements on the right and at the bottom and having a zero as the last element in the principal diagonal. The column vector to the right of the sign, on the other hand, is series of 0's followed by a single unit element.

(d) All the variables p_1 and E are restricted to be non-negative.

(e) In interpreting the solution, reverse the adjustment of the payoff in the first step.

Note:

The conversion process outlined for the 3×3 matrix is naturally applied to the general rectangular game with a $m \times n, m = n$ or $m \neq n$ payoff matrix,

Activity

A company wishes to increase its sales of a commodity by advertising on television and in the newspaper. Determine the net increase in profits for each television and newspaper advertisement given that the company's gross profits per commodity is $100, if each television advertisement increases sales by 65 000 units and each newspaper advertisement increases sales by 250 000 units given that a television advertisement costs $15 000 and a newspaper advertisement costs $24 000 and that the advertisement budget must involve no more than $75 000 in newspaper advertisement and no more than $360 000 in television advertisement. This must be achieved within an advertising budget of 1 million dollars.

UNIT SEVEN

DIFFERENTIAL AND INTEGRAL CALCULUS

Introduction

Calculus is a branch of mathematics involving or leading to calculations dealing with continuously varying functions. Calculus is a subject that falls into two parts:

(i) **Differential calculus** (or **differentiation**) and

(ii) **Integral calculus** (or **integration**).

Differentiation is used in calculations involving velocity and acceleration, rates of change and maximum and minimum values of curves.

In an equation such as $y = 3x^2 + 2x - 5$, y is said to be a function of x and may be written as y $y = f(x)$.

An equation written in the form $(x) = 3x^2 + 2x - 5$ is termed **functional notation**. The value of $f(x)$ when $x = 0$ is denoted by $f(0)$, and the value of $f(x)$when $x = 2$ is denoted by $f(0)$and so on. Thus when $f(x) = 3x^2 + 2x - 5$

$$f(0) = 3(0)^2 + 2(0) - 5 = -5$$

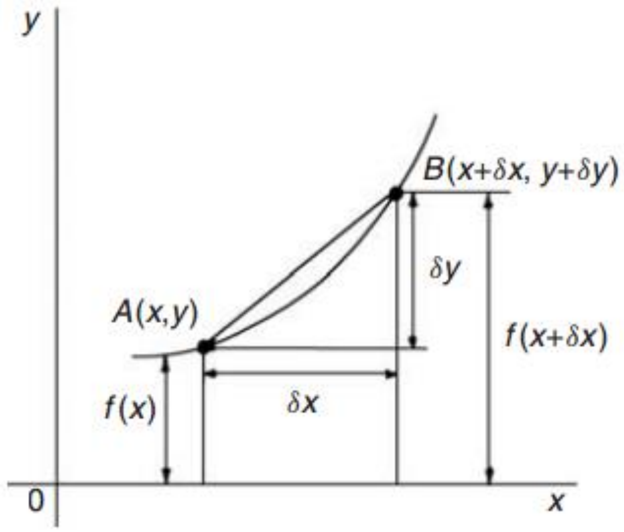
$$f(2) = 3(2)^2 + 2(2) - 5 = 11$$

Limit of a function

A given function $f(x)$ is said to have a limit M as x approaches c [in symbols, $\lim_{x \to c} f(x) = M$ if $f(x)$ can be made as close to M as we please for all values of $x \neq 0$ but sufficiently near to c, by having x get sufficiently close to c (approaching both from the left and right).

Differentiation from first principles

(i) The figure shows two points A and B close together on a curve delta x and delta y representing small increments in the x and y directions respectively



Gradient of a chord AB = $\frac{\delta y}{\delta x}$, However, $\frac{\delta y}{\delta x} = f(x + 8\delta) - f(x)$, Hence $\frac{\delta y}{\delta x} = \frac{f(x+8\delta)-f(x)}{\delta x}$

As δx approaches zero $\frac{\delta y}{\delta x}$ approaches a limiting value and the grairetnt of the chord approaches the gradient of the tangent at A.

(ii) When determining the gradient of a tangent to a curve there are two notations used. The gradient of the curve at A in the figure can be either written as

$$\lim_{\delta x \to 0} \frac{\delta y}{\delta x} \; or \; \lim_{\delta x \to 0} \{\frac{f(x + 8\delta) - f(x)}{\delta x}\}$$

In Leibniz notation , $\frac{dy}{dx} = \lim_{\delta x \to 0} \frac{\delta y}{\delta x}$

In functional notation

$$f'(x) = \lim_{\delta x \to 0} \{\frac{f(x + 8\delta) - f(x)}{\delta x}\}$$

(iii) $\frac{dy}{dx}$ is the same as $f'(x)$ and is called differential coefficient or the derivative. The process of finding the differential coefficient is called differentiation.

Example Differentiate from first principles $f(x) = x^2$ and determine the value of the gradient of the curve at x = 2.

- First find $f'(x)$ by using the expression

$$f'(x) = \lim_{\delta x \to 0} \{\frac{f(x + 8\delta) - f(x)}{\delta x}\}$$

$$f(x) = x^2$$

Substituting $f(x + \delta x) = (x + 8\delta\ ) = x^2 + 2x\delta x + \delta\, x^2$

$$f'(x) = \lim_{\delta x \to 0} \left\{ \frac{(x^2 + 2x\delta x + \delta\, x^2) - x^2}{\delta x} \right\}$$

$$= \lim_{\delta x \to 0} \left\{ \frac{2x\delta x + \delta\, x^2}{\delta x} \right\}$$

$$= \lim_{\delta x \to 0} \{2x\delta x + \delta x\}$$

As $\delta x \to 0, [2x\delta x + \delta x] \to [2x + 0]$. Thus $f'(x) = 2x$, i.e the differential coefficient of x^2 is $2x$. At $x = 2$, the gradient of the curve $f'(x) = 2(2) = 4$.

Find the derivative of $y = 8$

$y = f(x) = 8$ since there are no $x -$ values in the original equation, substituting $(x + \delta x)$ for x still gives $f(x + \delta x) = 8$.

$$\frac{dy}{dx} = f'(x) = \lim_{\delta x \to 0} \left\{ \frac{f(x + 8\delta) - f(x)}{\delta x} \right\}$$

$$= \lim_{\delta x \to 0} \left\{ \frac{8 - 8}{\delta x} \right\} = 0$$

When $y = 8$, $\frac{dy}{dx} = 0$

Finding the derivative means finding the gradient,, hence general for any line $y = k$ where k is a constant then $\frac{dy}{dx} = 0$

Find the differential coefficient of $y = 4x^2 + 5x - 3$ and determine the gradient of the curve $x = -3$.

Differentiation of $y = ax^n$ by the general rule.

From the differentiation by first principles, a general rule for differentiation ax^n emerges where a and n are any constants. This rule

If $y = ax^n$ then $\frac{dy}{dx} = anx^{n-1}$ or if $f(x) = ax^n$ then $f'(x) = anx^{n-1}$

When differentiating, results can be expressed in a number of ways

(i) If $y = 3x^2$ then $\frac{dy}{dx} = 6x$

(ii) If $f(x) = 3x^2$ then $f'(x) = 6x$

(iii) The differential coefficient of$3x^2$ is $6x$

(iv) The derivative of $3x^2$ is $6x$

(v) $\frac{d}{dx}(3x^2) = 6x$

Using general rule, differentiate the following with respect to x

(a) $y = 5x^7$ (b) $y = 3\sqrt{x}$ (c) $y = \frac{4}{x^2}$

(a) Using general rule $\frac{dy}{dx} = anx^{n-1}$, $a = 5$ and $n = 7$

(5)(7) $x^{7-1} = 35x^6$

(b) $y = 3\sqrt{x} = 3x^{\frac{1}{2}}$, $\quad a = 3, \ n = \frac{1}{2}$

$\frac{dy}{dx} = anx^{n-1} = (3)\frac{1}{2}x^{\frac{1}{2}-1}$

$= \frac{3}{2}x^{-\frac{1}{2}} = \frac{3}{2x^{\frac{1}{2}}} = \frac{\mathbf{3}}{\mathbf{2}\sqrt{\mathbf{x}}}$

(c)) $y = \frac{4}{x^2} = 4x^{-2}$ $\quad a = 4, \ n = -2$

$$\frac{dy}{dx} = anx^{n-1}$$

$$= (4)(-2)x^{-2-1}$$

$$= -8x^{-3}$$

$$= \frac{8}{x^3}$$

Summary of general derivatives

$F(x)$ function	$F'(x)$ Derivative
K — Constant	0
x^n	nx^{n-1}
$Sin(x)$	$cos(x)$
$cos(x)$	$-sin(x)$
$tan(x)$	$sec^2(x)$
e^{ax}	ae^{ax}

Derivative of a Sum (Difference) of Functions

The derivative of the sum (difference)of functions is the sum (difference) of the derivatives of those functions, that is, if $y = u + v$, where u and v are functions of x, Then

$$\frac{dy}{dx} = \frac{du}{dx} \pm \frac{dv}{dx}$$

Example

Differentiate $y = x^2 + tan(x)$ *with respect to x.*

Solution

To differentiate a sum, we differentiate each part or term, thus:

$$\frac{dy}{dx} = \frac{dx^2}{dx} + \frac{d(\tan(x))}{dx} = 2x + sec^2(x)$$

Derivative of a Product of Two Functions

Let y be a product of two functions that is $y = uv$, where u and v are functions of x, tlien we have

$$\frac{dy}{dx} = u\frac{dv}{dx} \pm \frac{du}{dx}$$

Example

Find the derivative of $y = x^5 \sin(x)$ *with respect to* x.

Solution

Let $y=uv$, *where* $u=x^5$ *and* $v=\sin(x)$. *Thus* $\frac{du}{dx}=5x^4$ *and* $\frac{dv}{dx} = \cos(x)$

Therefore $\frac{dy}{dx} = u\frac{dv}{dx} + v\frac{du}{dx}$

$= x^5 \cos(x) + 5x^4 \sin(x)$

Derivative of a Quotient of Two Functions

If y can be written as a quotient of two functions, that is $y = -$ where u and v are functions of x, then:

$$\frac{dy}{dx} = \frac{v\frac{du}{dx} - u\frac{dv}{dx}}{v^2}$$

This is called the gradient rule. It is often helpful to memorise a verbal statement of the quotient rule:

$$\frac{d(quotient)}{dx} = \frac{(denominator) - numerator\,\frac{d}{dx}\,(denominator)}{(denominator)}$$

Example *Find the derivative of* $y = \frac{tanx}{x^3}$ *with respect to x.*

Solution

Let $u = tan(x)$ *and* $v=x^3$

$$\frac{du}{dx} = \frac{x^3 \sec^2(x) - 3x^2\tan(x)}{x^6}$$

$$\frac{rsec^2(x) - 3\tan(x)}{x^4}$$

Derivative of Composite Functions

Suppose y is a function of u and u is a function of x, then we have:

$$\frac{dy}{dx} = \frac{dy}{du} \times \frac{du}{dx}$$

This is the famous chain rule.

Differentiate $y = (x^2 + 3x)^4$ *with respect to x,*

Solution

Let u = x2+3x. then $\frac{du}{dx} = \mathbf{2x + 3}$

and $\mathbf{y = u^4}$ $\frac{dy}{du} = \mathbf{4u^3}$

Hence $\frac{\mathbf{dy}}{\mathbf{dx}} = \mathbf{4u^3 \times (2x + 3)}$

= 4(x2+3x)3 (2x+3)

The Second Derivative

The second derivative is obtained by differentiating the first derivative. Thus $\frac{d}{dx}\frac{dy}{dx} = \frac{d^2y}{dx^2}$ is the second derivative.

Example *Given that* y = $3x^3 - 2x^2 + 4x - 2$, *find* $\frac{d^2y}{dx^2}$

Solution

$$\frac{dy}{dx} = 9x^2 - 4x + 4$$

$$\frac{d^2y}{dx^2} = \frac{d}{dx} = (9x^2 - 4x + 4) = 18x - 4$$

Implicit Differentiation

Implicit functions are those functions where the subject and object are combined in some terms. Given an equation involving x and y and assuming y is a differentiable function of x, we can find $\frac{dy}{dx}$ as follows;
Differentiate both sides of the equation with respect to x.

1. Collect all terms involving $\frac{dy}{dx}$ on the left side of the equation and move all other terms to the right side of the equation.

2. Factor $\frac{dy}{dx}$ out of the left of the equation.

3. Solve for $\frac{dy}{dx}$ by dividing both sides of the equation by the left-hand factor that

does not contain $\frac{dy}{dx}$

Example *Find* $\frac{dy}{dx}$ *given that* $y^3 = 4x$.

Solution

$$\frac{d(y^3)}{dx} = \frac{d(4x)}{dx}$$

$$\frac{d(y^3)}{dx} = 3y^2\frac{dy}{dx} = 4$$

$$\frac{dy}{dx} = \frac{4}{3y^2}$$

Application of Differentiation to Business Systems

In the following discussion, we shall show how derivatives are employed as a powerful means of extending our insight into real life problems and in solving these problems.

1.4.1 Maxima and Minima

Differentiation is handy for determining the point of maximum or minimum. Consider the graph of the function.

$y = -x^2 + 12x + 40$

Where is y a maximum? A function assumes its maximum value at a point where the derivative equals zero, that is at a stationary point.

Definition 1.2 *Stationary point*

A stationary point of a curve is a point at which the gradient of the curve is equal to zero. The reason why a point at which the gradient is zero is called a stationary point is the following: the gradient at a point measures the rate of change of the function at that point. If the gradient is zero then the rate of change is zero, in other words the function does not change at that point. For our example

$$\frac{dy}{dx} = \frac{dx}{dx}(-x^2 + 12x + 40)$$

$$= -2x+12$$

The point at which the function is maximum is obtained by setting ^ equal to zero. Therefore

$$-2x + 12 = 0 \qquad -2x = -12 \qquad x = 6.$$

Although it is necessary for the derivative to be zero at a maximum point, it is not *sufficient. There are other types of stationary points. It is, therefore, necessary to* analyse the gradient even further in the vicinity of stationary points.

Definition:*Maximum and Minimum*

A function $y = f[x)$, can achieve a finite maximum or minimum only at a point $x = a$ where the derivative $\frac{dy}{dx} = 0$; that is at a stationary point. Furthermore, the derivative $\frac{dy}{dx}$ is zero only at a maximum, a minimum and an inflection point.

1. If $\frac{dy}{dx} = 0$ and $\frac{d^2y}{dx^2} < 0$ at a point $x = a$ then $f(x)$ has a maximum point and the maximum is $f(a)$.

2. If $\frac{dy}{dx} = 0$ and $\frac{d^2y}{dx^2} > 0$ at a point $x = a$ then $f(x)$ has a minimum point and the minimum point is $f\ (a)$.

3. If $\frac{dy}{dx} = 0$ and $\frac{d^2y}{dx^2} = 0$ at a point x = a we determine the first (lowest order)

derivative such that $\frac{d^ky}{dx^2}, \neq 0$ that is $\frac{dy}{dx} = \frac{d^2y}{dx^2} = \cdots = \frac{d^{k-1}y}{dx^{k-i}} \neq 0$ but $\frac{d^ky}{dx^k} \neq 0$ it follows then that:

(i) If k is uneven, $x = a$ is an inflection point

(ii) If k is even and $\frac{d^k y}{dx^k} < 0$, $x = a$ is a maximum point

(iii) If k is even and $\frac{d^k y}{dx^k} > 0$, $x = a$ is a minimum point.

Example *A farmer wants to make a rectangular enclosure for his calves. If the farmer only has 200m of fencing available, what must the length and breath of the camp be for the area to be a maximum?*

Solution

Suppose the length of the enclosure is x metres and the area y square metres.

Therefore two lengths : 2 metres

Remaining fencing : 200-2x metres

Two breaths : 200-2x

Therefore breath $\frac{1}{2}(200 - 2x)$ = 100 - *x metres.*

Therefore Area = *length* x *br*

= x(100-x)

= $100x - x^2$

y = $100x - x^2$

$\frac{dy}{dx} = 100 - 2x$

Let $\frac{dy}{dx} = 0$

100-2x = 0

50 = x

$Now\ \frac{d^2y}{dx^2} = \frac{d}{dx}(100 - 2x)$

$= -2$

< 0

Therefore x = 50 is a maximum point. The area of the camp is a. maximum if length = 50 *m* and breath = — 50 = 50 *rn*. The camp must be a square for the area to be a maximum.

Example *A dealer finds that the sales of a certain item are closely linked to the price he asks per item. asks for x dollars per item, then he can expect to sell* 600 — 50a; *items. Each item costs him 5 dollars. much m.ust he ask per item to maximise his total profit?*

Solution

Let *y* represent his total profit. If he asks *x* dollars per item, then his profit per item sold = *x* — 5 dollars.

Number of sales = 600 — 50x

Total profit = (600 - 50x) $(x - 5)$

= $-50x^{-2} + 850x - 300$

$y = -50x^2 + 850x - 300$

$\frac{dy}{dx} = -100x + 850$

Let $\frac{dy}{dx} = 0$

-100x + 850 = 0

-100 x = -850

x = 8.5

$$\frac{d^2y}{dx^2} = -100$$

$$< 0$$

Therefore x = 8.5 is a maximum point. He must ask for $8.50 per item to maximize his total profit. The maximum total profit is: (600 - 50 x 8.5)(8.5 - 5) = $612.50.

Minimising Average Cost

We will illustrate this concept by means of an example.A manufacturer's total cost function is given by:

$$C = \frac{q^2}{4} + 3q + 400$$

where q is the number of units produced. At what level of output will the average cost per unit be a minim What is this minimum?

The quantity to be minimised is average cost $\bar{C} = \frac{c}{q}$

$$\frac{\frac{q^2}{4} + 3q + 400}{q}$$

$$\frac{q}{4} + 2 + \frac{400}{q}$$

Here q must be positive. To minimise $\bar{c}$ we differentiate, i.e.

$$\frac{d\bar{C}}{dq} = \frac{1}{4} - \frac{400}{q}$$

$$= \frac{q^2 - 1600}{4q^2}$$

Let $\frac{d\bar{C}}{dq} = 0$

$q^2 - 1600 = 0$

$(q - 40)(q + 40) = 0$

$q = 40$ (sin $q > 0$)

$$= \frac{d^2\bar{C}}{dq^2} = \frac{800}{q^3}$$

minimum average cost C = ^ + 3 + ^£ = 23.

which is positive for q = 40. $\bar{C}$ has a relative minimum when q = 40.

The minimum average cost $\bar{C} = \frac{40}{4} + \frac{400}{40} = 23$

INTRODUCTION TO INTEGRAL CALCULUS

Integral calculus is concerned with the reverse process of differentiation. We are given the derivative of a function and we must find the original function. The need for doing this arises in a natural way. For example, we may have a marginal revenue function and want to find the revenue function from it.

The Indefinite Integral Definition: *Anti-derivative*

If F is a function such that $F'(x) = f(x)$ then, F is called the anti-derivative of $f(x)$. An anti-derivative of $f(x)$ is simply a function whose derivative is $f(x)$. In differential notion we write $dF = f(x)dx$

For example, since the derivative of x^2 is 2x, x^2 is an anti-derivative of 2x. However it

is not the only anti-derivative of 2x. Since $\frac{d}{dx}(x^2 1) = 2x$

and $\frac{d}{dx}(x^2 - 5) = 2x$. Both $(x^2 + 1)$ and $(x^2 - 5)$ are also anti-derivatives of $2x$.
In fact it is obvious that because the derivative of a constant is zero, $x^2 + c$ is also an

anti-derivative of $2x$ for any constant c. Thus 2x has infinitely many anti-derivatives.

Since x^2 + c describes all anti-derivatives of $2x$. this is the most general anti-derivative of $2x$, denoted by $f\, 2xdx = x^2 + c$.

The symbol / is called the integral sign, $2x$ is the integrand and c is the constant of integration. The dx is part of the integral notation and indicates the variable involved. Here

x, is the variable of integration.

Basic Integration Formulae

1. $\int kdx = kx + c, k \text{ is a constant}$

2. $\int x^n dx = \frac{x^{n+1}}{n+1} + c, n \neq -1$

3 $\int e^x dx = e^x + c$

4 $\int kf(x)dx = k \int f(x)dx$, k is constant

5 $\int [f(x) \pm g(x)]dx = \int f(x)dx \pm \int g(x)dx$

6 $\int \mathbf{sin(x)dx = -cos(x) + c}$

7 $\int \mathbf{cos(x)dx = tan(x) + c}$

8 $\int \mathbf{sec^2(x)dx = tan(x) + c}$

Example Find $\int x^5 dx$.

Solution

$$\int x^5 dx = x^{5+1} + c$$
$$= x^6 + c$$

Example *Find* $\int 7xdx$

Solution

$$\int 7\,xdx = 7\int xdx$$

$$7\left[\frac{x^2}{2} + c_1\right]$$

$$\frac{7x^2}{2} + 7c_1$$

Since **$7c_1$** is just an arbitrary constant, for simplicity we shall replace it by c thus:

$$\int 7xdx = \frac{7x^2}{2} + c$$

Example *Evaluate* $\int (x2\ +\ 2x)dx$

$$\int (x^2 + 2x)dx = \int x^2 dx + \int 2xdx$$

$$= \frac{x^{2=1}}{2+1} + c_1 + 2\left[\frac{x^{1+1}}{1+1} + c_2\right]$$

$$= \frac{x^3}{3} + c_1 + x^2 + 2c_2$$

$$= \frac{x^3}{3} + x^2 + c_1 + 2c_2$$

$+ c$

Solution

For convenience, we shall replace the constant $c_1 + 2c_2$ by c. **Thus**

$$\int (x^2 + 2x)dx = \frac{x^3}{3} + x^2| + c$$

ExampleFind $\int \frac{(2x-1)(x+3}{6} dx$

Solution

By *factoring out the constant $ and multiplying out the binomials, we get:*

$$\int \frac{(2x-1)(x+3}{6} dx = \frac{1}{6}\int (2x^2 + 5x - 3)dx$$

$$= \int \frac{1}{6}\left[\frac{2x^3}{3}\frac{5x^2}{2} - 3x\right] + c$$

$$= \frac{x^3}{9} + \frac{5x^2}{12} - \frac{x}{2}$$

Example *Find* $\int \frac{x^{3}-1}{x^{2}} dx$

Solution

We can break up the integrand into fractions by dividing each term in the numerator by the denominator:

$$\int \frac{x^{3-1}}{x^{2}} dx = \int \left(\frac{x^{3}}{x^{2}} - \frac{1}{x^{2}}\right) dx$$

$$= \int (x - x^{-2}) dx$$

$$= \frac{x^{2}}{2} + \frac{x-1}{-1} + c$$

$$\frac{x^{2}}{2} + \frac{1}{x} + c$$

Find the general indefinite integral

$$\int \left(10x^{3} - 2\sec^{2} x\right) dx$$

Solution:

$$\int \left(10x^{3} - 2\sec^{2} x\right) dx = 10\int x^{3} - 2\int \sec^{2} x \, dx$$

$$= 10\frac{x^{4}}{4} - 2\tan x + C$$

$$= \frac{5x^{4}}{2} - 2\tan x + C$$

The following is a table of formulas of the commonly used Indefinite Integrals. You can verify any of the formulas by differentiating the function on the right side and obtaining the integrand.

Table of Indefinite Integral Formulas

$$\int cf(x)dx = c\int f(x)dx$$

$$\int [f(x)+g(x)]dx = \int f(x)dx + \int g(x)dx$$

$$\int k\,dx = kx + C$$

$$\int x^n dx = \frac{x^{n+1}}{n+1} + C \quad (n \neq 1)$$

$$\int \sin x\,dx = -\cos x + C$$

$$\int \cos x\,dx = \sin x + C$$

$$\int \sec^2 x\,dx = \tan x + C$$

$$\int \csc^2 x\,dx = -\cot x + C$$

$$\int \sec x \tan x\,dx = \sec x + C$$

$$\int \csc x \cot x\,dx = -\csc x + C$$

Definite Integrals

Activity

Login to a youtube link of your choice and follow video lectures on various techniques employed in the calculations of definite integral calculus.

UNIT EIGHT

TIME SERIES ANALYSIS

Introduction

A time series is a sequence of measurements or observations made on a variable at regular time intervals over a relatively long period of time. Many business variables have observations made on them at regular time intervals. Examples of these variables are daily sales, monthly payroll, annual exports, annual profits and so on. For instance, a country's annual exports recorded over 30 years constitute time series data. Times series data are important in that they help business managers to review past performance and they provide a basis for predicting future values of the time series.

In this unit we make use of time series charts to describe the different components of a time series. We use the method of least squares to fit a trend line to time series data. We demonstrate how to use the moving average method to smoothen a time series.

Objectives

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

- define a time series
- describe the components of a time series
- draw time series charts
- carryout a trend analysis in time series data using the least squares method and the moving-average method
- deseasonalise data using the Ratio-to-Moving Average method
- predict future values of a time series

Components of a Time Series

A time series is best analysed by decomposing it into different components. The components of a time series are:

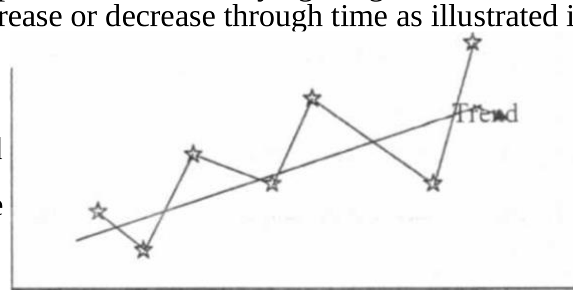
- Trend component (T)
- Seasonal component (S)
- Cyclical component (C)
- Irregular Component (I)

Trend component

The trend component is an underlying longer-term movement in the series showing a steady tendency of increase or decrease through time as illustrated in Figure 9.1.

Time in years

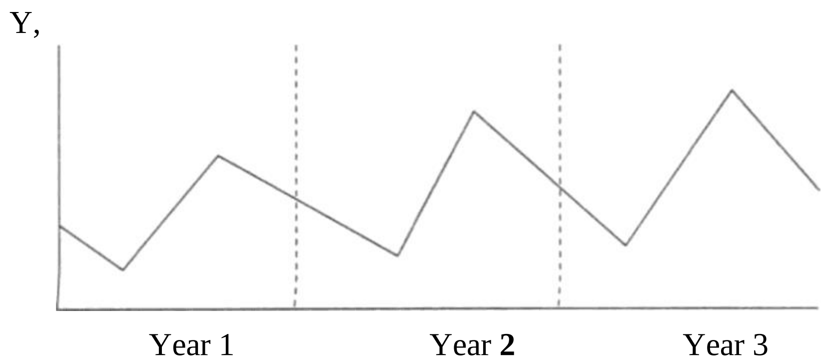
The trend shows the overall



Long Term Trend

movement in the series. **9.2.2 Seasonal component**

The seasonal component is a short-term recurrent component, which may be daily, weekly, monthly, or quarterly. The type of 'seasonal' component thus depends on how regularly the data are collected. However, seasonal variation is usually a feature of data collected quarterly. The variation follows a complete cycle throughout a whole year, with the same general pattern repeating itself year after year as illustrated by Figure 9**.2.**



Seasonal Component

Examples of time series variables (Y_t) that display seasonal variation are:

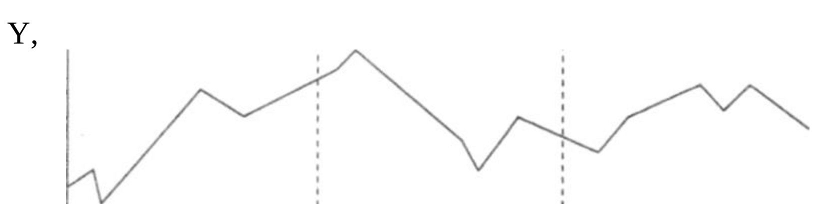
- Sales of seasonal items such as blankets/jerseys, school uniforms, umbrellas, fruits
- Credit card spending which is generally high towards and during the festive season
- Electricity consumption which varies depending on time of the day

Cyclical component

The cyclical component is a long-term recurrent component that repeats over several years. Cyclical movements differ in intensity and also vary in lengths usually lasting from **2** to 10 years. In business, cyclical behaviour is often referred to as the business cycle characterised by troughs and peaks of business activity.

Cyclical Variation Irregular component

The irregular component accounts for variation which is of a random nature and is not part of either the trend or the recurrent components. It does not contain any obviously predictable pattern. The variation is due to sporadic forces such as natural disasters (floods, drought, cyclones) or man-made disasters such as civil wars, strikes, boycotts. Figure 9.4 is an illustration of the irregular component.



Year 1 Year 2 Year 3

Irregular Trend

Time Series Models

The relationship between the components of a time series can be described by two models of a time series which are the Additive Model and the Multiplicative Model.

Additive Model
The model assumes that the components are added together with each observation Y_t being the sum of a set of components:

$$Y_t = T_t + S_t + C_I + I_t$$

The model is appropriate for series that have regular and constant fluctuations around a trend. To decompose an additive time series you have to subtract the components from each other.

Multiplicative Model
The model assumes that the observed time series values are a product of the four components, when all exist. This model is more commonly used than the additive model because it is found to describe more appropriately time series in a wide range of applications. It is more appropriate for series that have regular but not constant fluctuations around a trend. To decompose a time series which is assumed to be multiplicative, we divide the components.

Moving average method
A moving average (MA) of a time series is an average of a fixed number of observations that moves as we progress down the series. The moving averages smoothes out peaks and valleys in the original series to leave out a relatively smooth trend. The moving averages are therefore estimates of the trend at different stages of the series.

The moving averages are centred at the middle of the observations from which it has been calculated. The term of the moving average series is meant to coincide with the periodicity of the original series. For example, a four-point moving average will be appropriate for quarterly data.

Example
The daily sales of an airtime vendor over 12 days are recorded below:
37 24 62 80 77 95 94 133 148 155 128 161 Calculate a 3- point moving average of the sales.

Solution
A 3-point moving average requires you to find the average of three sets of observations at a time.
The first MA = (37 + 24 +62)/3 = 41.00
The second MA = (24 + 62 + 80)/3 = 55.33
The third MA = (62 + 80 + 77)/3 = 73.00 and so on.

Note that the moving averages are centred at the middle of the data used to calculate it so that we lose two observations one at the start and the other at the end. Centering is problematic with even terms because, for an even term, the moving averages are 'out of phase' with the time series observations. To centre the MA series, a further 2-point MA is found by averaging every consecutive pair as illustrated in Example.

Example

The data below shows the sales ($000s) of a seasonal good at a retail outlet over three years.

a) Calculate four-point moving averages for the series

Solution

a) **Table 9.1 A 4-Point Centred MA of Sales**

Year	Quarter	Sales (Y,)	Uncentred 4-point MA	Centred 4-point MA(T)
			21.25	
				21.500
			21.75	
				22.125
			22.50	
				22.875
			23.25	
				23.375
			23.50	
				23.375
			23.25	

Activity

Plot the four-point MA series on the same graph as the original series.

UNIT NINE

INDEX NUMBERS

Introduction

An index number is a number that measures the relative change in a set of measurements over time. Index numbers show changes over time by expressing the new value, V_n as a percentage of some existing value, V_0 called the base value.

Index number $= \frac{V}{v} \times 100$

The base value is the value of the variable at some reference point in the past called the base period. By convention, the index number of the base period is always assumed to be 100.

In this unit, you will be taught how to construct price and quantity indices for both weighted and unweighted indices. We will also look at the problems that are encountered in the construction of index numbers.

Learning Objectives

By the end of the unit, you should be able to:

- define index numbers
- calculate simple index numbers and weighted index numbers
- change the base from one period to another
- compare base weighting with current weighting
- state the purpose of index numbers
- use the CPI to adjust for inflation
- describe the challenges that are encountered in the construction of index numbers

Types of Index Numbers

There are three major categories of index numbers which are:

1. Price indices,
2. Quantity indices, and
3. Value indices.

Price indices

Price indices measure changes in price over time. Some examples of price indices are:

- Consumer Price Index (CPI) which measures the overall price change, from month to month, of a representative selection of goods and services that are relevant to a typical household. The CPI is used to calculate the rate of inflation.
- Producer Price Index (PPI) which measures the average change over time in selling prices received by domestic producers for their output.

Quantity indices

Quantity indices measure how much of a commodity is produced or consumed over time. Some examples of quantity indices are:

- Industrial index which gives a measure of change in industrial output now compared to a past reference point

- Mining index which gives a measure of change in minerals production now compared to a specified base period

Value indices

Value indices measure changes in total monetary worth of say exports (export index) or imports (import index) of an economy between two time periods.

Simple Index Numbers

The word 'simple' implies the measurements are for a single variable. A simple index number is the ratio of two values of a variable, expressed as a percentage. The most commonly referred to simple indices are: the Simple Price Index and the Simple Quantity Index.

Simple price index

The Simple Price Index (SPI), sometimes known as a price relative, measures changes in the price of a commodity. It shows the effect of a price change on a single product. The current price is expressed as a percentage of the price at base period,

Activity

The following data give the prices and quantities of two products sold in 2016 and 2017

product	2016		2017	
	Price	quantity	Price	quantity
X	30	3 000	50	1 500
Y	15	400	20	250

Use 2016 as the base period, calculate the:

a) Simple Price Index for product X
b) Simple Quantity Index for product Y

UNIT TEN

BUSINESS DECISION MAKING

DECISION THEORY

Steps involved in decision theory approach:

- Determine the various alternative courses of actions from which the final decision has to be made.
- Identify the possible outcomes, called the states of nature or events for the decision problem.
- Construct a pay off table.
- The decision maker chooses the criterion which results in largest pay off.

Decision making environments:

- Decision under certainty
  Whenever there exists only one outcome for a decision, we are dealing with this category
- Decisions under uncertainty:
  When more than one outcome can result from any single decision i.e. more than one state of nature exists.
- Decision under risk:
  The decision maker chooses from among several possible outcomes where the probability of occurrence can be stated objectively from the past data.
- Decision under conflict:
  Neither states of nature are completely known nor are they completely uncertain.

DECISION UNDER UNCERTAINTY:

There are five criterion on the basis of which rules for making a decision is Formulated:

Criterion of pessimism:

- Minimax or Maximin
- Maximin is a conservative approach to assume worst possible outcomes
- Steps involved:
  Find minimum assured pay off for each alternative
  Choose the maximum of minimum values.
- Minimax involves two steps:
  Determine maximum possible cost for each alternative
  Choose the alternative minimum of above costs

Criterion of optimism:

•Mimimin or maximax

•Two extreme optimism

•Decision makes ensures that he should not miss the opportunity to achieve the the greatest possible pay off or lowest possible cost

•Steps involved:
  Determine maximum possible payoff
  Select a alterative which corresponds to maximum of above maximum pay off

•Minimin of cost is done in similar manner

Laplace criterion:

• It is assumed that all states of nature will occur with equal probability

•Probabilities of each state of nature is given by 1/(number of states of nature)

•Steps involved:

i. Assign equal probabilities to each payoff of a strategy
ii. Determine the expected pay off value for each alternative.
iii. Select the alternative which corresponds to the maximum payoff or minimum cost

Criterion of realism or Hurwicz criterion:

•Coefficient of optimism α

•$0<\alpha<1$ where 0 signifies total pessimism and 1 total optimism

•Steps involved:

i. Decide the coefficient of optimism and the coefficient of pessimism
ii. Determine the maximum as well as minimum pay off for each alterative

h= α x maximum for each alterative + $(1-\alpha)$ x minimum for each alterative

iii. Select the alternative with highest value of h.

Example:

A farmer wants to decide which of the three crops he should plant on his 100 Acre farm. The profit from each is dependent on the rainfall during the growing seasons. The farmer has categorized the amount of rainfall as high, medium, low. His estimated profit for each is show in the table:

Rainfall	Crop A	Crop B	Crop C
High	8000	3500	5000
Medium	4500	4500	5000
Low	2000	5000	4000

If the farmer wishes to plant only one crop, decide which will be his choice using

•Maximax criterion
•Maximin criterion
•Hurwicz criterion
•Laplace criterion

Rainfall	Crop A	Crop B	Crop C
High	8000	3500	5000
Medium	4500	4500	5000
Low	2000	5000	4000

i. Maximax criterion:
From table we observe that maximum pay off for each alternative are 8000, 5000 ad 5000 respectively. Maximum among these is 8000 corresponding to crop A. So this strategy chooses crop A .

ii. Maximin criterion selects crop C

iii. Hurwicz criterion:
Assuming degree of optimism $\alpha = 0.6$ ad therefore $1-\alpha = 0.4$, the value of h is calculated in the table:

Alternativ e	Maximum pay off	Minimum pay off	h
Crop A	8000	2000	5600
Crop B	5000	3500	4400
Crop C	5000	4000	4600

The maximum value is 5600 so this criterion selects crop A.

iv. Laplace criterion:

Assign equal probabilities i.e. 1/3. The expected pay off is calculated for each alterative:

E (Crop A)=1/3(8000)+1/3(4500)+1/3(2000)= 4833
E (Crop B)=4333
E (Crop C)=4666

Hence this criterion also selects crop A.

Activity

Consider the payoff table below which shows decision alternatives (d_1; d_2 and d_3) that a company is faced with under three different states of nature (S_1; S_2 and S_3).

d_i/s_j	S_1	S_2	S_3
d_1	26	12	34
d_2	14	8	16
d_3	35	19	10
$P(S_j)$	0.4	0.5	0.1

Determine the appropriate decision that the company should make using the following approaches:

i. Optimistic approach (maximax)
ii. Conservative approach maximin
iii. Minimax Regret Approach
iv. Expected Monetary Value (EMV) approach.

THE END

REMEMBER TO READ LATEST PUBLICATIONS SO THAT YOU REMAIN RELEVANT TO YOUR CAREER DEVELOPMENT AND CONTRIBUTE POSITIVELY TO SOCIETY

ENJOY SCHOOL AND ADD PASSION TO YOUR STUDIES

You can also access some published business articles on this website (http://www.williamphiriabwino.com)

You can equally consult other websites that will add value to your studies

REFERENCES

George B. Thomas and Wross L. Finney, (1983). *Calculus and Analytic* ***geometry, S****th* Edition, Edson Wesley, 1983.

Hamdy A. Taha, (1997) *Operations Research*, *6th* edition, Prentice Hall,

http://www.williamphiriabwino.com

Kachepa, A. (2017). Operations Research. Harare; Mount Pleasant

Lawrence L. Lapin, (1994) *Quantitative Methods for Business Decisions with Cases* Sixth Edition, Duxbury Press Belmont, California.

Munzara, A. T. (2014). Statistics for Managers. Harare; Mount Pleasant

Phiri. W. (2016). Applied Descriptive Statistics and ICT Education. Lusaka; MP Publishers