THE REPUBLIC OF UGANDA

Ministry of Education and Sports

HOME-STUDY LEARNING

SENIOR 5

MATHEMATICS

August 2020

NCDC
NATIONAL CURRICULUM DEVELOPMENT CENTRE

UGANDA NATIONAL EXAMINATIONS BOARD

NCDC
NATIONAL CURRICULUM
DEVELOPMENT CENTRE

Published 2020

ISBN: 978-9970-00-193-4

This material has been developed as a home-study intervention for schools during the lockdown caused by the COVID-19 pandemic to support continuity of learning.

National Curriculum Development Centre
P.O. Box 7002,
Kampala- Uganda
www.ncdc.go.ug

FOREWORD

Following the outbreak of the COVID-19 pandemic, government of Uganda closed all schools and other educational institutions to minimize the spread of the coronavirus. This has affected more than 36,314 primary schools, 3129 secondary schools, 430,778 teachers and 12,777,390 learners.

The COVID-19 outbreak and subsequent closure of all has had drastically impacted on learning especially curriculum coverage, loss of interest in education and learner readiness in case schools open. This could result in massive rates of learner dropouts due to unwanted pregnancies and lack of school fees among others.

To mitigate the impact of the pandemic on the education system in Uganda, the Ministry of Education and Sports (MoES) constituted a Sector Response Taskforce (SRT) to strengthen the sector's preparedness and response measures. The SRT and National Curriculum Development Centre developed print home-study materials, radio and television scripts for some selected subjects for all learners from Pre-Primary to Advanced Level. The materials will enhance continued learning and learning for progression during this period of the lockdown, and will still be relevant when schools resume.

The materials focused on critical competences in all subjects in the curricula to enable the learners to achieve without the teachers' guidance. Therefore effort should be made for all learners to access and use these materials during the lockdown. Similarly, teachers are advised to get these materials in order to plan appropriately for further learning when schools resume, while parents/guardians need to ensure that their children access copies of these materials and use them appropriately.
I recognise the effort of National Curriculum Development Centre in responding to this emergency through appropriate guidance and the timely development of these home study materials. I recommend them for use by all learners during the lockdown.

Alex Kakooza
Permanent Secretary
Ministry of Education and Sports

ACKNOWLEDGEMENTS

National Curriculum Development Centre (NCDC) would like to express its appreciation to all those who worked tirelessly towards the production of home–study materials for Pre-Primary, Primary and Secondary Levels of Education during the COVID-19 lockdown in Uganda.

The Centre appreciates the contribution from all those who guided the development of these materials to make sure they are of quality; Development partners - SESIL, Save the Children and UNICEF; all the Panel members of the various subjects; sister institutions - UNEB and DES for their valuable contributions.

NCDC takes the responsibility for any shortcomings that might be identified in this publication and welcomes suggestions for improvement. The comments and suggestions may be communicated to NCDC through P.O. Box 7002 Kampala or email admin@ncdc.go.ug or by visiting our website at http://ncdc.go.ug/node/13.

Grace K. Baguma
Director,
National Curriculum Development Centre

ABOUT THIS BOOKLET

Dear learner, you are welcome to this home-study package. This content focuses on critical competences in the syllabus.

The content is organised into lesson units. Each unit has lesson activities, summary notes and assessment activities. Some lessons have projects that you need to carry out at home during this period. You are free to use other reference materials to get more information for specific topics.

Seek guidance from people at home who are knowledgeable to clarify in case of a challenge. The knowledge you can acquire from this content can be supplemented with other learning options that may be offered on radio, television, newspaper learning programmes. More learning materials can also be accessed by visiting our website at www.ncdc.go.ug or ncdc-go-ug.digital/. You can access the website using an internet enabled computer or mobile phone.

We encourage you to present your work to your class teacher when schools resume so that your teacher is able to know what you learned during the time you have been away from school. This will form part of your assessment. Your teacher will also assess the assignments you will have done and do corrections where you might not have done it right.

The content has been developed with full awareness of the home learning environment without direct supervision of the teacher. The methods, examples and activities used in the materials have been carefully selected to facilitate continuity of learning.

You are therefore in charge of your own learning. You need to give yourself favourable time for learning. This material can as well be used beyond the home-study situation. Keep it for reference anytime.

Develop your learning timetable to ca ter for continuity of learning and other responsibilities given to you at home.

Enjoy learning

Topic 1: Descriptive Statistics

Lesson 1: Representation of Data

Competences

In this lesson, you should be able to learn how to:

(i) form a frequency table from a set of raw data.
(ii) draw a histogram for equal and unequal classes and use the histogram to determine the mode.
(iii) draw a frequency polygon on its axes and on a histogram.
(iv) draw an Ogive and use it to determine the median.

Introduction

Statistics is a branch of science that deals with the collection, interpretation, presentation and analysis of data. When numerical information is collected, it's called raw data before it is arranged.

After collection, data is arranged in a table which shows the different scores and their number of occurrences (frequencies). This is called a **frequency table**. Frequency tables are of two types; **Ungrouped** and **Grouped** frequency table.

In O-level you saw how frequency tables are generated. The ungrouped frequency table is formed when the spread of data is small. For example, if a class is given a test marked out of 10 marks, each of the scores 0 to 10 is taken individually in the frequency table.

Ungrouped frequency table

The table shows the marks of students in a class in a certain test.

Mark	Freq.
0	3
1	5
2	4
3	4
4	6
5	5
6	3
7	4
8	3
9	2
10	1

3 students got 0 mark, 5 students got 1 mark, 4 students got 2 marks etc.

While a grouped frequency table is formed when the spread of data is large i.e. scores ranging from 1 to 100. If a test is marked out of 100, the marks will be grouped in classes of our convenience and each mark counted in its respective class.

A Grouped frequency table

The marks of students in a class with many students can be grouped for convenience as shown in the table.

Class	Freq.
10–19	4
20–29	3
30–39	7
40–49	10
50–59	6
70–79	5
80–89	3
90–99	2

When a grouped frequency table is made, some other concepts are generated; these include:

(i) **Class limits:** are the end marks in a given class. They are of two types: The lower-class limit and the upper-class limit. For example, in the class 20–29, 20 is the lower-class limit and 29 is the upper-class limit.

(ii) **Mid mark (x):** is the mark taken to represent a given class. It is got by adding the class limits and divide by 2 (i.e.) in the class 30–34, $x = \frac{30+34}{2} = 32$

(iii) **Class interval:** is the size of a given class. It is denoted by letter i and got by subtracting the lower-class limit from the upper-class limit and adding 1 i.e. in the class 80–89, i = (89 – 80) + 1 = 9 + 1 = 10.

(iv) **Class boundaries:** are the real class limits of a given class. They are got from the class limits by subtracting 0.5 from the lower-class limit and adding 0.5 to the upper-class limit to get the lower-class boundary and upper-class boundary respectively. That is, in the class 65–69, the class boundaries are (65 – 0.5, and 69 + 0.5) = 64.5 – 69.5.

A grouped frequency table

MARKS	FREQUENCY	MID MARK (X)	CLASS BOUNDARIES
30–34	11	32	29.5–34.5
35–39	17	37	34.5–39.5
40–44	13	41	39.5–44.5
45–49	24	47	44.5–49.5
50–54	22	52	49.5–54.5
55–59	15	57	54.5–59.5
60–64	8	62	59.5–64.5
65–69	4	67	64.5–69.5

A frequency table can be represented by class boundaries instead of class limits.

Marks	20 - < 30	30 - < 40	40 - < 50	50 - < 60	60 - < 70
Frequency	5	8	15	12	6

Marks	20 - 30	30 - 40	40 - 50	50 - 60	60 – 70
Frequency	5	8	15	12	6

Weight	< 20	< 30	< 40	< 45	< 55
Frequency	11	15	8	9	10

After representing data in a tabular form, it can further be represented graphically in different forms. Among them, we have bar graph, histogram, ogive, frequency polygon, pie chart etc.

Histogram

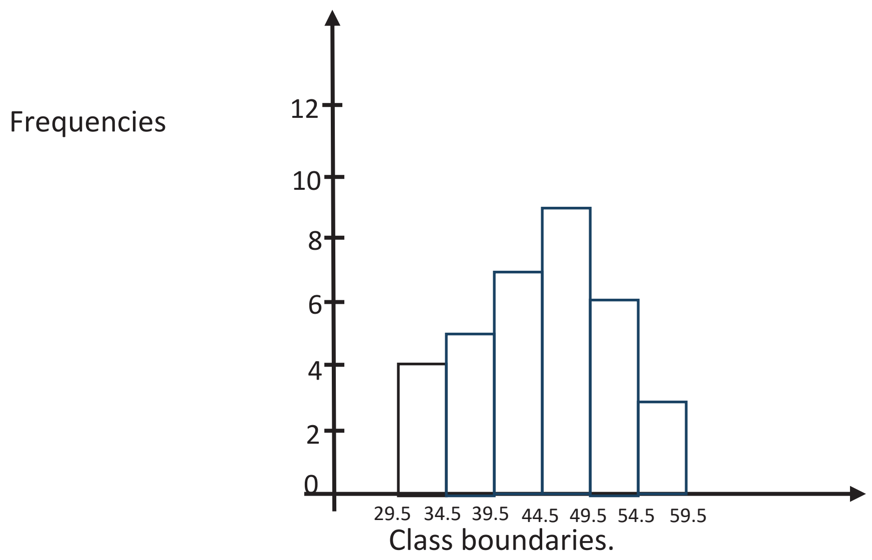
Histogram is a graph where class frequencies are plotted against class boundaries. Histograms are of two types. One of equal class interval and another of unequal class interval.

Histogram with Equal Class Interval

Example

Draw a histogram for the data below.

Class	Freq.
30–34	4
35–39	5
40–44	7
45–49	9
50–54	6
55–59	3



Histogram with Unequal Class Width

Here histogram is drawn with:

a) frequency density on the vertical axis.
b) class boundaries on the horizontal axis.

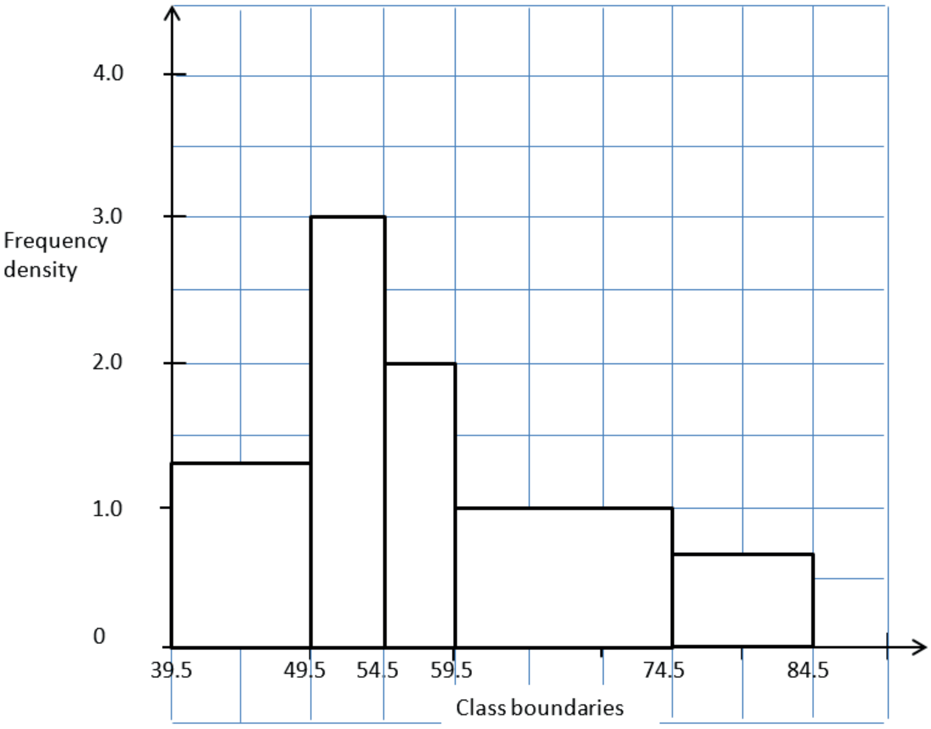
Frequency density is obtained by dividing frequency of that class with its class width.

Example

Draw a histogram for the frequency table below.

Class	Frequency	Class width, i	Frequency density ($\frac{f}{i}$)	Class boundaries
40–49	13	10	1.3	39.5–49.5
50–54	15	5	3.0	49.5–54.5

Class	Frequency	Class width, i	Frequency density $(\frac{f}{i})$	Class boundaries
55–59	10	5	2.0	54.5–59.5
60–74	15	15	1.0	59.5–74.5
75–84	7	10	0.7	74.5–84.5



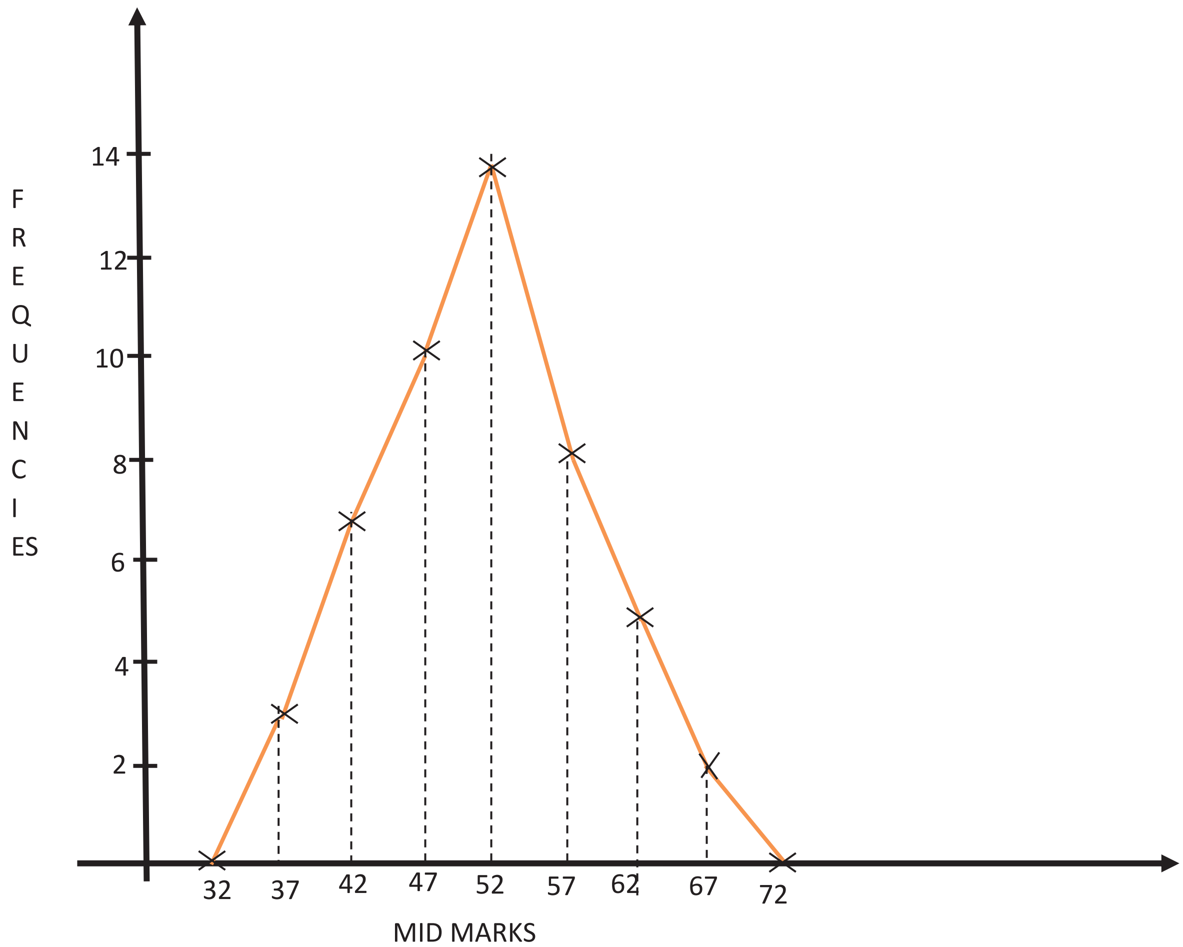
Frequency Polygon

Frequency polygon is drawn with frequencies on the vertical axis and mid marks on the horizontal. It can also be put on the same graph with the histogram.

Example

Draw a frequency polygon for the data below.

Class	**Frequency**	**Mid mark**
35–39	3	37
40–44	7	42
45–49	10	47
50–54	14	52
55–59	8	57
60–64	5	62
65–69	2	67



Cumulative Frequency Curve (Ogive)

The cumulative frequency curve is a curve drawn with cumulative frequency on the vertical axis and class boundaries on the horizontal axis. We plot upper class boundaries and the cumulative frequencies and join those points with a smooth curve.

Example 3

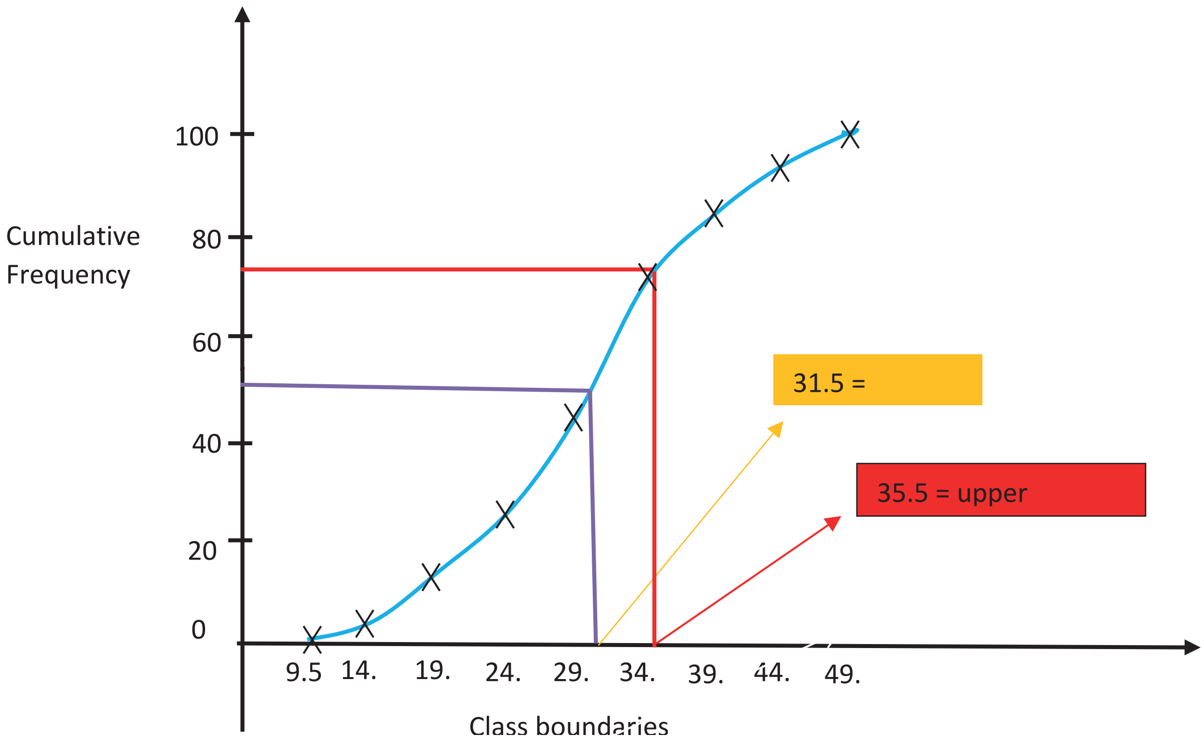
The frequency distribution table shows the weights of 100 children measured to the nearest kilogram.

Weight	Frequency
10–14	5
15–19	9
20–14	12
25–29	18
30–34	25
35–39	15
40–44	10
45–49	6

Draw a cumulative frequency curve for the data.

Solution

Class Boundaries	Frequency	Cumulative Frequency
9.5	0	0
14.5	5	5
19.5	9	14
24.5	12	26
29.5	18	44
34.5	25	69
39.5	15	84
44.5	10	94
49.5	6	100



Exercise

1. The lengths in cm of 40 metal rods were as follows:

Lengths	Frequency
30 – < 35	8
35 – < 40	5
40 – < 55	12
55 – < 60	9
60 – < 65	6

(a) Calculate the:
   (i) mean length.
   (ii) upper quartile.
(b) Display the data on a histogram and use it to estimate the mode.

2. The times (minutes) taken by the taxi to move from Kampala to Gomba were recorded over a certain period of time and grouped as follows:

Time (minutes)	Frequency (f)
80–84	10
85–89	15
90–94	35
95–99	40
100–104	28
105–109	15
110–114	4
115–119	2
120–124	1

(a) Calculate the mean time of travel from Kampala to Gomba by taxi.
(b) Draw a cumulative frequency curve for the data. Use it to estimate the:
   (i) median time for the journey.
   (ii) semi interquartile range of time of travel from Kampala to Gomba.

Lesson 2: Measures of Central Tendency and Measures of Dispersion

Competences

In this lesson, you will learn how to:

(i) calculate the mean, mode and median.
(ii) calculate the mean deviation, variance and standard deviation.

In O-Level, you saw the measures of central tendency by calculating the mean, mode and median. You used the formulae:

Mean = $\frac{\Sigma f(x)}{\Sigma f}$, or using assumed mean. Mean = $X_A + \frac{\Sigma f(d)}{\Sigma f}$

Mode = $L_o + \left(\frac{\Delta_1}{\Delta_1 + \Delta_2}\right)C$ and Median = $L_o + \left(\frac{\frac{N}{2} - F_b}{f_m}\right)C$.

Measures of Dispersion

These are measures used to find out how the observations are spread out from the average. These include: mean deviation, quartile range, variance and standard deviation.

(a) Mean deviation

The mean deviation of a set of numbers is given by:

$\sum_1^n \frac{|x_{i-M}|}{n}$ where M is the mean

Find the mean deviation of the set of numbers 24, 43, 38, 28, 36, 40, 26, 37.

$M = \frac{\sum x}{n} = \frac{24+43+38+28+36+40+26+37}{8} = 34$

Mean deviation $= \frac{|24-34|+|43-34|+|38-34|+|28-34|+|36-34|+|40-34|+|26-34|+|37-34|}{8}$

$= \frac{10+9+4+6+2+6+8+3}{8} = \frac{48}{8}$

= 6.

(b) Variance

Variance is the sum of the mean deviations squared divided by the number of observations.

Variance $= \frac{\sum_1^n (x_{i-M})^2}{n}$.

If the observations appear with the frequencies (repeated values), then variance is given by:

Variance $= \frac{\sum_1^n f(x_{i-M})^2}{n}$.

A simplified form of the formula for variance used for computations is:

Variance $= \frac{\sum x^2}{n} - \left(\frac{\sum x}{n}\right)^2$ or Variance $= \frac{\sum fx^2}{\sum f} - \left(\frac{\sum fx}{\sum f}\right)^2$

Standard deviation is the positive square root of the variance.

i.e. standard deviation $= \sqrt{\frac{\sum fx^2}{\sum f} - \left(\frac{\sum fx}{\sum f}\right)^2}$.

Example

The frequency distribution table shows the marks scored by students in a certain school.

x	55	63	65	66	70	72	75	80	90
f	2	2	3	1	2	2	4	3	1

Calculate the standard deviation of the data.

x	f	x^2	fx	fx^2
55	2	3025	110	6050
63	2	3969	126	7938
65	3	4225	195	12675
66	1	4356	66	4356
70	2	4900	140	9800

72	2	5184	144	10368
75	4	5625	300	22500
80	3	6400	240	19200
90	1	8100	90	8100
	$\sum f = 20$		$\sum fx = 1411$	$\sum fx^2 = 100987$

Variance $= \frac{\sum fx^2}{\sum f} - \left(\frac{\sum fx}{\sum f}\right)^2$

$= \frac{100987}{20} - \frac{1411}{20}$

$= 72.05$

Standard deviation $= \sqrt{72.05}$

$=$ **8.488**

Exercise

1. The table below shows the distribution of marks of students in a test.

Score	Frequency
$20 \leq x < 30$	4
$x < 45$	3
$x < 50$	9
$x < 65$	21
$x < 75$	3
$x < 80$	5
$x < 100$	14

(a) Draw a histogram and use it to estimate the modal mark.

(b) Calculate the: (i) mean score.
(ii) standard deviation score.

2. The heights of a group of workers in a factory were recorded as shown in the frequency table below.

Height (cm)	Frequency
170 – 175	19
175 – 180	36
180 – 185	70
185 – 190	64
190 – 195	39
195 - 200	22

(a) Estimate the mean and standard deviation of the worker's height.
(b) Plot an ogive.
(c) Use the ogive to estimate the median and interquartile range for the data.

TOPIC 2: INDEX NUMBERS

LESSON 1: SIMPLE AGGREGATE PRICE INDEX

In this lesson, you should be able to:
- (i) identify index numbers.
- (ii) calculate simple average relative indices.
- (iii) calculate simple aggregate index numbers.

Index numbers are statistical measurements which show the changes in variables or group of variables with respect to time.

An index number is a statistic indicating the relative change occurring in each successive period of time in the price, volume or value of a commodity or in a general economic variable or gross output, with reference to a previous base period conventionally given the number 100.

TYPES OF INDEX NUMBERS

These include:
1. Relative indices
2. Simple index numbers/simple average relative indices
3. Simple aggregate index numbers
4. Weighted indices.
    - (a) Weighted average indices
    - (b) Weighted aggregate indices (composite indices)
    - (c) Laspeyre aggregate index
    - (d) Paasche price index
5. Value index numbers

BASE YEAR

This is the year or period against which all the years or periods are compared. The base year is normally given a standard statistical value or measure of 100. Sometimes, the previous year is taken as the base year.

Price in the base year is denoted by P_o,
Quantity in base year is denoted by q_o,
Wage in base year is denoted by W_o etc.

CURRENT YEAR

This is the year or period for which the index is to be calculated (computed) it is also known as **given year**.
Price in the current year is denoted by P_1
Quantity in the current year is denoted as q_1
Wage in the current year id denoted as W_1
Throughout this chapter we shall use P_1, Q_1, W_1, etc. to denoted variables in the current year.

RELATIVE INDICES/SIMPLE INDEX NUMBERS

These include the following among others

1. Price indices
2. Quantity (quantum) indices
3. Wage indices

PRICE INDICES/RELATIVES

These measure the changes in prices of items between the base year and the current year.

Example 1

An article cost Shs 500 in 1990 and Shs 800 in 1994. Taking 1990 as the base year, find the price relative in 1994.

Solution;

Price index or price relative = $\frac{P_1}{P_0}$X100

$=\frac{800}{500}$X100

=160

Therefore 160-100=60% increment

Comment: The price of the article increased by 60% in 1994

Example 2

The wages of nurses in Uganda in 1995 was Shs 20,000/- . The wages of the same nurses in 1997 was increased by Shs 25,000/-. Using 1995 as the base year, calculate the nurses' wage index for 1997

Solution:

W_1=20000 +25000

W_1= 45,000/-

W_0 =20,000/-

Wage index (wages relative) = $\frac{W_1}{W_0}$X100

$=\frac{45000}{20000}$X100

= 225

Therefore 225-100 = 125% increment

Comment: The nurses' wages increased by 125% in 1997

Example 3

In the year 2000, the price of the commodity using 1999 as the base year was 166. In 2006, the index using 2000 as the base year was 123. What is the index in 2006 using 1999 as the base year?

SOLUTION

$\frac{P_{2000}}{P_{1999}} \times 100 = 166$

$\frac{P2000}{P1999} = \frac{166}{100}$

Therefore: $P_{2000} = P_{1999} \times \frac{166}{100}$ ---- (i)

And

$\frac{P_{2006}}{P_{2000}} \times 100 = 123$

$\Rightarrow \frac{P2006}{P2000} = \frac{123}{100}$

Therefore: $P_{2000} = P_{2006} \times \frac{100}{123}$ ----(ii)

Comparing, (i) = (ii)

$\frac{166}{100}$x P_{1999} = $\frac{100}{123}$ x P_{2006}

$\frac{166}{100}$ x $\frac{123}{100}$ = $\frac{P2006}{P1999}$

But index in 2006 using 1999 as the base year is given by P.R = $\frac{P2006}{P1999}$x100

Therefore: P.R = ($\frac{166}{100}$x$\frac{123}{100}$) x100

The index is 204.18

3. **SIMPLE PRICE INDEX**

- Simple price index for each commodity is also known as the price relative (price index) for each commodity.

S.P.I = $\frac{P1}{Po}$x100

A comment should be given for every commodity.

- Simple price index for the overall commodities combined is also known as the simple average price index/relative.

S.P.I = $\frac{\Sigma(\frac{P1}{Po})}{n}$x100 where **n** = number of items.

This also works when the units of the items are not uniform.

A comment is necessary.

Example 1

The table below shows the price and quantities of selected items consumed in the period 1980 – 1982.

Items	**Price (shillings per kilogram)**			**Quantities(kilograms)**		
	1980	**1981**	**1982**	**1980**	**1981**	**1982**
Salt	100	150	80	200	80	70
Sugar	500	450	600	300	70	62
Millet	400	450	500	250	100	70
Beans	300	450	550	150	120	80

(a) By using 1980 as the base year, find

(i) the price relative for each of the items in 1981

(ii) the simple quantum index for each of the items in 1981

(b) By using 1980 – 1981 as the base year, find the

(i) simple price index for each of the items in 1982.

(ii) simple average price index for all the items.

SOLUTION

(a)

(i) Price Relative = $\frac{P81}{P80}$ x 100

For salt, P.R = $\frac{150}{100}$ x 100

= 150

Therefore 150 – 100 = 50%

The price of salt increased by 50% in 1981

For Sugar, P.R = $\frac{450}{500}$ x 100

= 90

Therefore 90 – 100 =- 10%

The price of sugar reduced by 10% in 1981

For millet: P.R = $\frac{450}{400}$ x 100

= 112.5

Therefore 112.5 – 100 =- 12.5%

The price of millet increased by 12.5% in 1981

For beans: P.R = $\frac{500}{300}$ x 100

= 166.667

Therefore 166.667 – 100 =- 66.667

The price of beans increased by 66.667% .

(ii) Simple quantum index

= $\frac{Q81}{Q80}$x100

For salt, Q.I = $\frac{80}{200}$x100

=40

The quantity of salt consumed reduced by 60% in 1981

For sugar, Q.I = $\frac{70}{300}$x100

=23.3

The quantity of sugar consumed decreased by 76.7%.

For millet, Q.I = $\frac{100}{250}$x100

=40

The quantity of millet consumed in 1981 decreased by 60%

For beans, Q.I = $\frac{120}{150}$x100

=80

The quantity of beans consumed in 1981 decreased by 20%.

(b) (i)

Items	**Po (average of 1980 and 1981)**	**P1 (1982)**	$\frac{P_1}{P_0}$	$\frac{P1}{Po}$**x1oo**
Salt	(100+150)/2=125	80	0.64	64
Sugar	(500+450)/2=475	600	1.2632	126.32
Millet	(400+450)/2=425	500	1.1765	117.65
Beans	(300+500)/2=400	550	1.375	137.5
Total			**4.4547**	

Comments for each commodity.

- For salt: S.P.I = 64

There was a decrease of 36% in the price of salt in 1982 as compared to the price of salt in the period of 1980-1981.

- For sugar: S.P.I = 126.32

There was an increase of 26.32% in the price of sugar in 1982 as compared to the price of sugar in the period of 1980-1981.

There was an increase of 17.65% and 37.5% in the price of millet and beans respectively in 1982 as compared to their prices in the period of 1980 – 1981

(ii)

Simple price index for the overall commodities combined (simple average price index)

S.P.I = $\frac{\Sigma(\frac{P_1}{P_0})}{n}$x100

= $\frac{4.4547}{4}$ x 100

= 111.3675

Comment

On average there was an increase in the overall prices of all commodities in 1982 as compared to the prices of commodities in the period of 1980-1981

SIMPLE AGGREGATE INDEX NUMBERS

This refers to the changes in all the variables as a whole between the base and the current year.

Simple aggregate price index =$\frac{\Sigma P_1}{\Sigma P_0}$**x 100**

Simple aggregate quantity index =$\frac{\Sigma Q_1}{\Sigma Q_0}$ **x100**

Simple aggregate wages index =$\frac{\Sigma W_1}{\Sigma W_0}$ **x 100**

Simple aggregate price index is abbreviated as S.A.P.I

Example 2:

The table below shows the price in pounds of maize and beans in 1999 and 2004.

Items	1999	2004
Maize (bag)	80	120
Beans (bag)	10	15

Determine the simple aggregate price index for the total cost of one bag of maize and beans.

SOLUTION:

Items	**1999**	**2004**
Maize (bag)	80	120
Beans (bag)	10	15
Total	90	135

S.A.P.I= $\frac{\Sigma P_1}{\Sigma P_0}$x 100

= $\frac{135}{90}$x 100

= 150

Therefore 150 – 100 = 50

The price of the two commodities increased by 50% in 2004

EXERCISE 1

1. One litre of petrol cost Shs 1200 in 1998 and Shs.1800 in 2004. Taking 1998 as the base year, find the price relative in 2004.
2. The wages of a certain group of workers in 2002 was 130,000/- and in 2006 it was increased by 70,000/-. Using 2002 as the base year, calculate the workers wage index in 2006.
3. In 2003, the price of rice using 2001 as the base year was 132. In 2006, the index using 2003 as the base year was 105. What is the index in 2006 using 2001 as the base year?
4. The cost of making a cake is calculated from the cost of baking flour, sugar, milk and eggs. All these ingredients are measured in kgs. The table below gives the cost of each of these items in 1985 and 1986

Items (kgs)	**Price**	
	1985	**1986**
Flour	60	78
Sugar	50	40
Milk	25	30
Eggs	10	15

Using 1985 as the base year, calculate the price relatives for each item. Hence find simple price index for the cost of making a cake.

5. A company imported the following commodities A,B,C form Kenya, U.S.A and England respectively.

Commodity	**Price**	
	1992	**1996**
A	Ksh 20,000	Ksh 25000
B	$ 4,000	$ 4800
C	£ 10,000	£ 10,000

Calculate the simple price index for 1996

6. The table below shows the amount spent per month by an average family in Gulu in the years 2000 and 2001.

Item	Amount in Uganda Shs.	
	2000	2001
Clothing	100,000	80,000
Housing	50,000	40,000
Electricity	40,000	50,000
Water	20,000	25,000
Transport	50,000	60,000
Food	150,000	160,000

Using 2000 as the base year, calculate the simple aggregate expenditure index in 2001

LESSON 2: WEIGHTED INDICES

In this lesson, you will learn how to:

(i) Calculate weighted price index

(ii) Calculate Laspeyre's and Paasches's aggregate price index

(iii) Calculate value index

1. **Weighted price index** also known as weighted average price index and is denoted by

W.P.I = $\frac{\Sigma(\left(\frac{P1}{Po}\right)xW)}{\Sigma W}$ x 100 where **W** = weights of commodities given

This can be used when the units of the items are not uniform.

2. **Weighted aggregate price index** also known as composite index or cost of living index or just weighted price index depending on the question setting.

a) **W.A.P.I** = $\frac{\Sigma P1W}{\Sigma PoW}$ x 100

where W= weights of the commodities. This is used when units of the items are uniform

b) or **W.A.P.I** = $\frac{\Sigma PW}{\Sigma W}$ where P= price relatives, P = $\frac{P1}{Po}$x100

NOTE.

2(b) is used when price relatives are given or can be calculated.

1 and **2(a)** are used when prices of the items and weights attached are given.

3a) **Laspeyre's aggregate price index** is when the base year quantities are used as weights.

Therefore 2a is modified as L.A.P.I= $\frac{\Sigma P_1 Q_0}{\Sigma P_0 Q_0}$ x100

3b) **Laspeyre's aggregate quantity index** is where base year prices are used as weights.

Therefore **2a** is modified as L.A.Q.I = $\frac{\Sigma Q_1 P_0}{\Sigma Q_0 P_0}$

NOTE: Laspeyre's index shows how the price or quantity of the base year's bill of goods change.

4. **Paasche's aggregate price index** is where current year quantities are used as weights.

Therefore **2a** is modified as **P.A.P.I** =$\frac{\Sigma Q_1 P_0}{\Sigma Q_0 P_0}$x100

NOTE: Paasche's index shows how the given year's bill of goods or expenditure changes.

5. **Value index number**, is the total expenditure incurred in a current year expressed as the fraction of the total expenditure incurred in the base year. It is denoted by:

V.I = $\frac{\Sigma P_1 Q_1}{\Sigma P_0 Q_0}$ x 100

Example 1

The table below shows the prices of items and their corresponding weights in the years 2000 and 2004.

Items	Price (U Shs)		weight
	2000	2004	
Food	55000	60000	4
Housing	48000	52000	2
Transport	16000	20000	1

Using 2000 as the base year, calculate the weighted price index for the items in 2004.

SOLUTION

Since the units of food, housing and transport are not uniform, so we use

W.P.I = $\frac{\Sigma(\frac{P1}{Po})xW}{\Sigma W}$ x100

Item	P1	$\frac{P1}{Po}$	W	$\frac{P1}{P0}$ x W
Food	60,000	1.0909	4	4.3636
Housing	52,000	1.0833	2	2.1666
Transport	20,000	1.25	1	1.25
Total			7	7.7802

W.P.I = $\frac{7.7802}{7}$ x 100

= 111.1457

≈111

Or W.P.I = $\frac{Total\ cost\ of\ all\ items}{Total\ number\ of\ items}$

W.P.I $_{2004}$ = $\frac{60{,}000+52{,}000+20{,}000}{4+2+1}$

= $\frac{132{,}000}{7}$

= 18857.1429

W.P.I $_{2000}$ = $\frac{55000+48{,}000+16{,}000}{7}$

= $\frac{11900}{7}$

= 17,000

Now W.P.I for items = $\frac{W.P.I\ 2004}{W.P.I\ 2000}$ x 100

= $\frac{18857.1429}{17{,}000}$ x100

= 110.9244

≈ 111

Therefore for this example, when the units of items are not uniform, the weighted aggregate price index is = $\frac{\textbf{Weighted average price index in 2004}}{\textbf{weighted average price index in 2000}}$ **x100**

W.A.P.I= $\frac{18857.1429}{17{,}000}$ x100

= 110.9244

≈ 111

The purchasing power of the shilling is reduced by (111-100) = 11%

Example 2

The price s per unit (in Uganda shillings,Ushs) of four foodstuffs A , B, C and D in December 2004 and December 2005 are as shown in the table below.

Food stuff	Price in Ushs.	
	2004	2005
A	635.0	887.5
B	720.0	815.0
C	730.0	1045.0
D	362.0	503.0

The weights of the foodstuffs A, B, C and D are 6, 4, 3 and 7 respectively. Taking 2004 as the base year, calculate for 2005 the:

a) Weighted aggregate index
b) Price of the food staff costing Shs 500 in December 2004, using the weighted aggregate index in (a) above.

SOLUTION

a) Note that Foodstuff may have same units e.g. kgs. Also the question is clearly asking weighted aggregate index.

Therefore weighted price index (Weighted aggregate price index) = $\frac{\Sigma P1W}{\Sigma PoW}$ x 100

Foodstuff	**W**	**Po**	**PoW**	**P_1**	**P1W**
A	6	635.0	3810	887.5	5325
B	4	720.0	2880	815.0	3260
C	3	730.0	2190	1045.0	3135
D	7	362.0	2534	503.0	3135
TOTAL			11414		14855

W.A.P.I = $\frac{14855}{11414}$ X100

= 130.1472

The total price of food stuffs increased by 130-100 = 30% in 2005.

b) $\frac{p2005}{P2004}$x 100 = 130.1472

$\frac{P2005}{500}$x 100 = 130. 1472

P2005 = 130.1472 x 6

= 650.736

Therefore, the price of a food stuff costing Shs 500 in December 2004 will be Shs 650.7 in December 2005

Example 3

The data below shows items with their corresponding price relatives and weights.

ITEMS	PRICES RELATIVES	WEIGHT
FOOD	120	172
CLOTHING	124	160
HOUSING	125	170
TRANSPORT	135	210
OTHERS	104	140

a) Find the cost of living index
b) Comment on your results

SOLUTION

If price relatives are given.

C.L.I = W.P.I (W.A.P.I) = $\frac{\Sigma PW}{\Sigma W}$

Where P= Price relatives/price indices

ITEMS	P	W	PW
FOOD	120	172	20640
HOUSING	124	160	19840
HOUSING	125	170	21250
TRANSPORT	135	210	28350

ITEMS	P	W	PW
OTHERS	104	140	14560
TOTAL		**852**	**104,640**

Therefore C.L.I = $\frac{104,640}{852}$

= 122.8169

(122.8169-100)

= 22.8169%

Comment

It is 22.8169% more expensive to live in this area (current area) where the above data was got than in the base area (though the base area is not shown)

Example 4

The table below shows the consumption of some commodities.

	2003		**2006**	
Items	**Price (Shs)**	**Quantity (Kgs)**	**Price (Shs)**	**Quantity (Kgs)**
Posho	700	50	1000	57
Beans	600	25	900	30
Rice	1100	15	1300	18
Salt	350	12	400	16
Meat	3500	14	4200	17

Calculate the

(i) Laspeyre's price index

(ii) Paasche's price index

(iii) Value index using 2003 as the base

SOLUTION

Items	P_o	Q_o	P_oQ_o	P_1	Q_1	P_1Q_1	P_oQ_1	P_1Q_o
Posho	700	50	35000	1000	57	57000	39900	50000
Beans	600	25	15000	900	30	27000	18000	22500
Rice	1100	15	16500	1300	18	23400	19800	19500
Salt	350	12	4200	400	16	6400	5600	4800
Meat	3500	14	49000	4200	17	71400	59500	58800

Items	P_o	Q_o	P_oQ_o	P_1	Q_1	P_1Q_1	P_oQ_1	P_1Q_o
Total			119700			185200	142800	155600

(i) L.A.P.I = $\frac{\Sigma P1Qo}{\Sigma PoQo}$ x 100

$= \frac{155600}{119700}$ x 100

= 129.9916

≈ 130

Comment

The bill of goods or expenditure on goods in 2003 increased by 30% i.e. (130-100 = 30)

(ii) P.A.P.I = $\frac{\Sigma P1Q1}{\Sigma PoQ1}$ x 100

$= \frac{185200}{142800}$ x 100

= 129.6919

≈ 130

Comment

The bill of goods or expenditure on goods in 2006 increased by 30%.i.e (130-100 = 30)

NOTE: In most cases L.A.P.I and P.A.P.I give different answers

And L.A.Q.I = $\frac{\Sigma Q1Po}{\Sigma QoPo}$ x 100

$= \frac{142800}{119700}$ x 100

= 119.2982

≈ 119

Comment

The quantity of goods purchased or bought in 2003 increased by (119-100) =19%

(iii) Value Index = $\frac{\Sigma P1Q1}{\Sigma PoQo}$ x 100

$= \frac{185200}{119700}$ x 100

= 154.7201

≈ 155

Comment

The prices or expenditure in 2006 were higher than those prices or expenditure in 2003 by (155-100)=55%

EXERCISE 2

1. Calculate the weighted price and quantity index numbers to measure change in domestic consumption of the indicated food items(use 1991 as base year)

		Prices in Shillings		Quantity	
Commodity	**Units**	**1991**	**1995**	**1991**	**1995**
A (Banana)	Kg	180	150	1500	2500
B (bread)	Units	500	700	80	100
C (milk)	Litre	400	700	60	60
D (Vegetable)	Kg	1000	800	45	60
E (Fruit)		900	600	120	200

2. The table shows the prices and quantities of some four commodities A, B, C, and D for the years 1986 and 1987.

ITEMS	**PRICE PER UNIT U(Shs.)**		**QUANTITIES**	
	1986	1987	1986	1987
A	100	120	36	42
B	110	100	69	88
C	50	65	10	12
D	80	85	11	10

Using 1986 as the base year

(i) Calculate the price index number for each item for 1987.

(ii) Calculate the simple aggregate price index number for 1987

(iii) Calculate the weighted aggregate price index number or 1987.

(iv) If A, B, C, and D were ingredients needed to make a loaf of bread and in 1986 a loaf cost shillings 60, using the index in (ii) above, estimate the cost of a loaf in 1987.

Topic 3: SCATTER GRAPHS AND CORRELATIONS

1. Identify correlations
2. Draw scatter diagrams
3. Find coefficients of correlations and comment

LESSON 1: SCATTER GRAPHS

In this lesson you should be able to:

(i) state the different types of scatter graphs.

(ii) draw scatter graphs and a line of best fit.

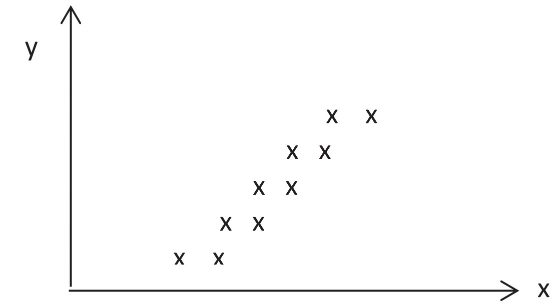
A scatter graph is one that shows the relationship between two variables diagrammatically. One variable is represented on the horizontal (x – axis) and the other on the vertical (y – axis).

When a scatter diagram is drawn, it shows the relationship between the two variables plotted. The degree of this relationship is measured by what is known as correlation coefficient.

TYPES OF CORRELATTION (SCATTER GRAPHS)

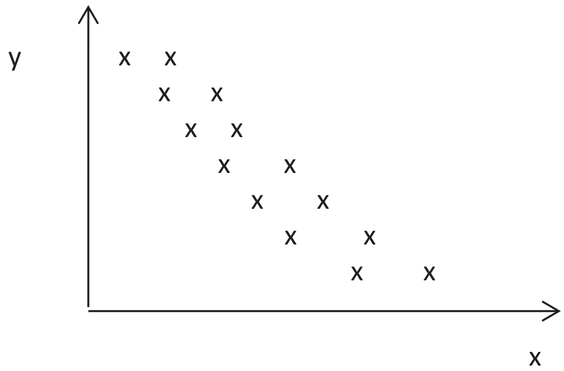
The spread of points on the scatter diagram gives the type of correlation.

(a) Positive correlation



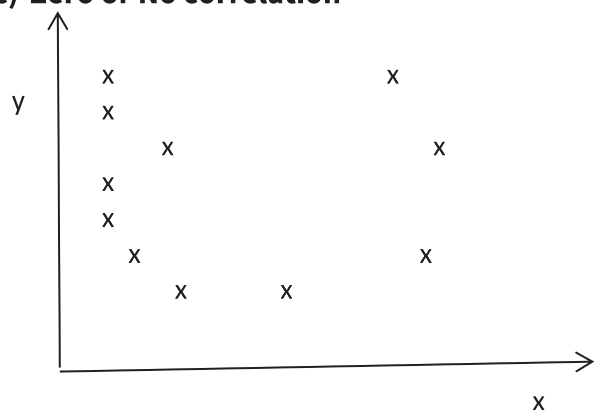
When the value of x increases, the value of y also increases. The correlation coefficient is between 0 and 1.

(b) Negative correlation



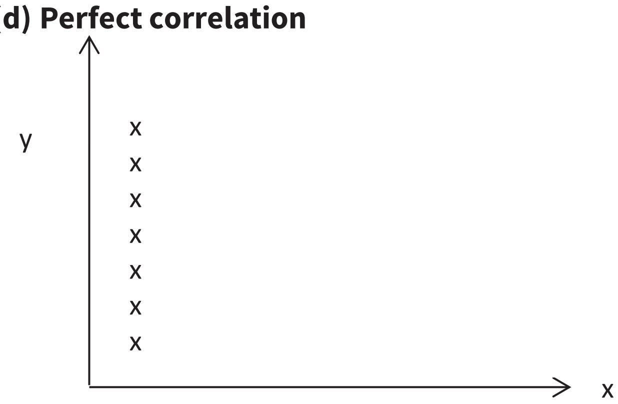
When the x value increases, the y value decreases. For example, the variable on the x – axis may be number of students and on y we have days the food in the store will last. The correlation coefficient is between -1 and 0.

(c) Zero or No correlation



There is no correlation between the variables. Correlation coefficient is 0.

(d) Perfect correlation



When a change in one variable corresponds to a change in equal degree of the other. The correlation coefficient is 1.

Line of best fit

This is the line judged to fit best the pattern of points. It is drawn so as to pass centrally through the graph of points. Since all the points can lie to the line, the objective here is to minimize the total divergence of points from the line hence the best to predict the average of **y** on **x** or the average of **x** on **y.**

NOTE:

- To draw a line of best fit, we first calculate the mean $(\bar{X}, \bar{Y})$ of the variables given.
- Then draw the line through the mean by use of eye to see that the points are evenly distributed on either side and not all points may lie on the line.
- To comment, it depends on the slope of the line of best fit. if it has a positive gradient it means as one variable increases the other also increases and vice versa.
- For a scatter graph, it is not guarantee that you must start from zero on both axes, it rather depends on the values given.

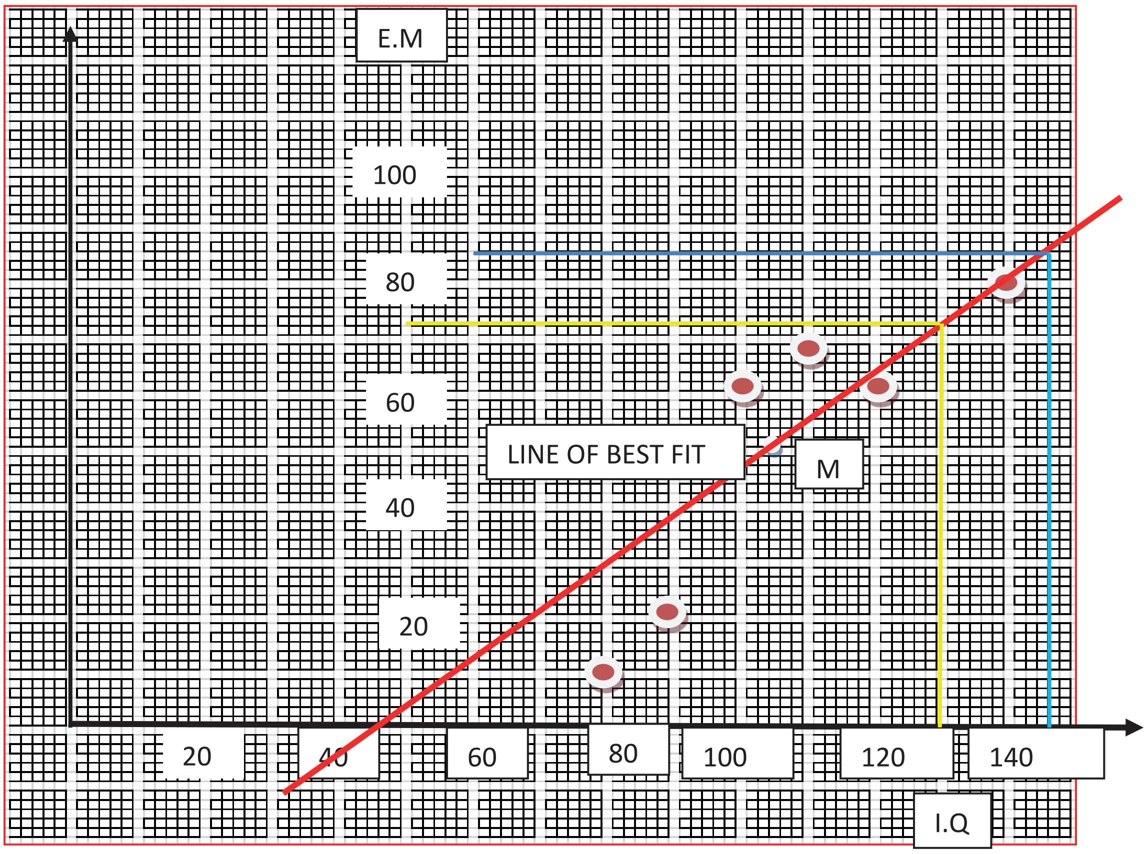
Example

1. The I. Qs of a group of six students were measured and sat an exam. The I. Qs and the examination marks were as follows.

student	I.Q	Examination marks (E.M)
A	110	70
B	100	60
C	140	80
D	120	60
E	80	10
F	90	20

(a) Represent the above data on a scatter graph and on it draw a line of best fit.
(b) Estimate the I.Q of a student whose marks are 85.
(c) What marks will a student with an I.Q of 130 get?

A SCATTER DIAGRAM SHOWING THE I.Qs OF SIX STUDENTSAND THEIR MARKS



(b) When the marks are 85, the I.Q of a student is 146

(c) A student with an I. Q of 130 will get 72 marks.

Exercise

1. The table below shows the marks obtained by 10 students in a physics and chemistry test.

Physics	80	75	65	90	95	98	78	65	54	60
Chemistry	70	85	70	90	92	88	76	70	73	76

(a) Represent the above data on a scatter diagram and on it draw a line of best fit.
(b) If a student scored 76% in physics, predict his score in chemistry.
(c) If $y = 81$, predict the corresponding value of x.

2. The price of matooke is found to depend on the distance the matooke is away from the nearest town. The table below gives the average price of matooke for markets around Kampala city.

Dist d	40	8	17	20	24	30	10	28	16	36
Price p	120	160	140	130	135	125	150	130	145	125

(i) Plot this data on a scatter graph
(ii) Draw the line of best fit on your diagram and use it to find the price of matooke when the distance is 38
(iii) Calculate the rank correlation coefficient and comment on your answer.

LESSON 2: RANK CORRELATION COEFFICIENT

In this lesson you should learn how to:
(i) calculate spearman's rank correlation coefficient.
(ii) comment on the relationship between the two variables.

We can calculate to measure the degree of the relationship between any two variables. The variable you get will be interpreted as follows:

INTERPRETATION OF THE MAGNITUDE OF THE CORRELATION COEFFICIENT

Correlation coefficient	Interpretation /comment
0-0.19	Negligible (very low correlation)
0.2-0.39	Slight (low correlation)
0.3-0.59	Moderate correlation
0.6-0.79	Substantial (high) correlation
0.8-1.0	Very high correlation

All the above are positive correlation. If the coefficient is like -0.6429 then we will say that "there is a substantial (high) negative correlation between x and y."

The methods for measuring the degree of the relationship include;

1. Spearman's rank correlation coefficient ($\wp$)
2. Kendall's rank correlation coefficient (t)

SPEARMAN'S RANK CORRELATION

It is given by $\wp = 1 - \frac{6\Sigma d^2}{n(n^2-1)}$

Where **d,** is the rank difference.
n is the number of pairs

Example:
Compute the spearman's rank correlation coefficient from the data below;

X	12	4	10	10	11	6	7	13
y	5	2	5	8	6	3	3	5

Solution

- We rank by assigning 1 to the lightest value.
- In case of a tie like 10 and 10 on the x values where one is in position 4 and another in position 5, it will be $\frac{4+5}{2}$ = 4.5. This implies each takes on position 4.5. In this case position 5 will be skipped.

X	Y	R_x	R_y	d = R_x- R_y	d_2
12	5	2	4	-2	4
4	2	8	8	0	0
10	5	4.5	4	0.5	0.25
10	8	4.5	1	3.5	12.25
11	6	3	2	1	1.00
6	3	7	6.5	0.5	0.25
7	3	6	6.5	-0.5	0.25
13	5	1	4	-3	9.00
TOTAL					**27**

Number of pairs is 8

Using $\wp = 1-\frac{6\Sigma d^2}{n(n^2-1)}$

$= 1 - \frac{6x27}{8(64-1)}$

$= 1 - \frac{162}{504}$

$\wp = 0.6786$

Comment: There is a substantial positive correlation between **X** and **Y.**

SIGNIFICANCE TEST FOR r_s

r_s is the rank correlation coefficient worked out from the given data. e.g. spearman's, Kendall's and others.

r_T is the rank correlation coefficient given in the mathematical table for significance levels for correlation coefficients on page 57 (A' level UNEB log book)

Steps followed

1. State the null hypothesis(H_o) e.g. H_o: r = 0 (there is no correlation between the two variables in question)
2. State the alternative hypothesis (H_1) e.g. H_1: r > 0 (there is positive correlation between the two variables.
   Or H_1: r < 0 (there is negative correlation between the two variables).
   Or H_1: $r \neq 0$ (there is correlation between the two variables.
   You can state any of the three H_1sabove
3. State the significance level ($\propto$) e.g. 1% level, 5% level or 10% level of significance.
4. State the critical region e.g. if number of pairs (n)=8, r_t= 0.86 at 1% level of significance (r_t=0.86 is for spearman)

Exercise:

1. The table below shows the number of female and male students admitted at a university to offer ten different courses A, B, C, ….., J.

Course	A	B	C	D	E	F	G	H	I	J
Female	58	54	60	70	62	54	70	58	80	58
Male	50	38	54	68	60	38	68	70	68	64

Calculate a rank correlation coefficient for the above data and comment on your answer.

2. The table below shows the percentage of sand (y) in the soil at different depths (x) in cm.

Soil depth, (x)cm	35	65	55	25	45	75	20	90	51	69
% of sand, (y) cm	86	70	84	92	84	68	96	84	86	77

(a) Draw a scatter graph for the above data and draw a line of best fit
(b) Use the line of best fit to estimate the percentage of sand at a depth of 12cm
(c) Calculate the spearman's rank correlation coefficient between the percentages of sand in the soil and the depth of the soil.

Topic 4: PROBABILITY THEORY

LESSON 1: LAWS OF PROBABILITY

In this lesson, you should be able to:

(i) state the laws of probability.

(ii) apply the laws of probability to solve problems.

Definition: Probability is an attempt to use mathematics to estimate by means of a numerical answer the chance that some event will happen.

Event: This is an occurrence in a defined context

Trial: This is a single attempt to obtain a defined event

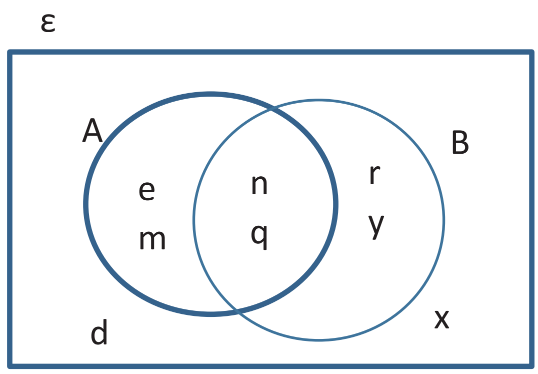
Sample space: This is the total or summation of all the events. Generation of a sample space include; **Table of outcomes** and **Tree diagrams**.

LAWS OF PROBABILITY

Consider an experiment whose sample space is (S). For each event (E) of the sample space to occur, we assume that the probability of that event satisfies the following laws;

1. $0 \leq P(E) \leq 1$. Probability of an event E to occur it can't be negative and can't be greater than 1
2. P(S) = 1. Probability of the sample space is 1
3. P(A ∩ B) = P(B ∩ A). also P(A U B) = P(B U A)

Consider $A = \{ m, n, q, e\}$ and $B = \{y, r, n, q\}$



A ∩ B = $\{n , q\}$

B ∩ A = $\{n , q\}$

BUA = A U B= $\{e, m, n, q, y, r\}$

$P(A \cap B) = \frac{2}{8} = \frac{1}{4}$

$P(A \cup B) = \frac{6}{8} = \frac{3}{4}$.

4. $P(A \cap B)^I = P(A^I \cup B^I)$:

   $A \cap B = \{n, q\}$ $(A \cap B)^I = \{e, m, y, r, d, x\}$

   $A^I = \{r, y, d, x\}$ $B^I = \{e, m, d, x\}$

   $A^I \cup B^I = \{r, y, d, x, e, m\}$

   Therefore; $(A \cap B)^I = A^I \cup B^I$.

   $P(A \cap B)^I = P(A^I \cup B^I) = \frac{6}{8} = \frac{3}{4}$.

5. $P(A \cup B)^I = P(A^I \cap B^I)$:

   $A \cup B = \{e, m, q, n, r, y\}$ $(A \cup B)^I = \{d, x\}$

   $A^I = \{r, y, d, x\}$ $B^I = \{e, m, d, x\}$

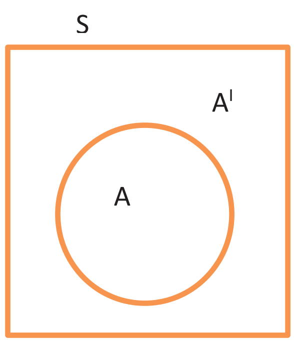
   $(A^I \cap B^I) = \{d, x\}$

   Therefore ;$P(A \cup B)^I = P(A^I \cap B^I) = \frac{2}{8} = \frac{1}{4}$.

   Note: Rules 4 and 5 are called the **Demorgan's laws.**

ADDITION RULES OF PROBABILITY

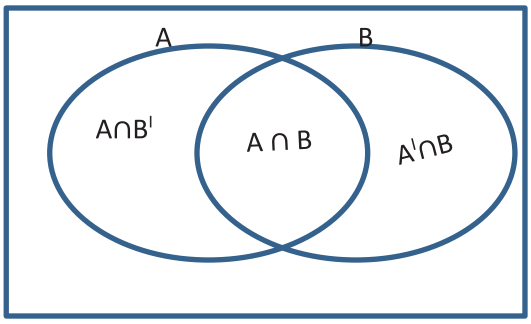
1. $P(A) + P(A^/) = 1$. Where $A^/$ means A complement (outside of set A)



   $P(A) + P(A^/) = P(S)$. but $P(S) = 1$

   $\therefore P(A) + P(A^/) = 1$ or $P(A) = 1 - P(A^/)$

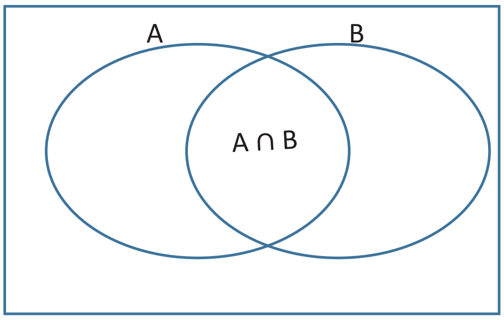
2. $P(A) = P(A \cap B) + P(A \cap B^/)$



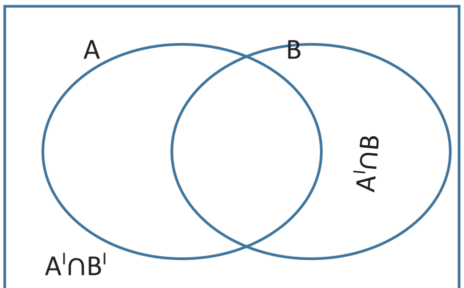
Similarly;

$P(B) = P(A \cap B) + P(A' \cap B)$

3. $P(A \cup B) = P(A) + P(B) - P(A \cap B)$



4. $P(A') = P(A' \cap B) + P(A' \cap B')$



$A' = (A' \cap B) \cup (A' \cap B')$

$P(A') = P(A' \cap B) \cup P(A' \cap B')$

Similarly;

$P(B') = P(A \cap B') \cup P(A' \cap B')$

Contingency table

	A	**A'**	
B	**P(A ∩ B)**	**P(A' ∩ B)**	**P(B)**
B'	**P(A ∩ B')**	**P(A' ∩ B')**	**P(B')**
	P(A)	**P(A')**	**1**

ADD

ADD

This contingency table is a table which Summarises the addition laws of probability.

P(A) + P(A^I) = 1,

P(A ∩ B) + P(A ∩ B^I) = P(A),

P(A ∩ B) + P(A^I ∩ B) = P(B),

P(A^I ∩ B) + P(A^I ∩ B^I) = P(A^I)

and P(A ∩ B^I) + P(A^I ∩ B^I) = P(B^I).

Example:

1. Given that P(AUB) = $\frac{7}{8}$, P(A∩B) = $\frac{1}{4}$ and P(A/) = $\frac{5}{8}$. Find;
   (i) P(A)
   (ii) P(B)
   (iii) P(A∩B/)
   (iv) P(A/UB/)

Solution:

(i) P(A) = 1 – P(A/)

$= 1 - \frac{5}{8}$

$= \frac{3}{8}$

(ii) P(A∪B) = P(A) + P(B) – P(A∩B)

$\frac{7}{8} = \frac{3}{8} + P(B) - \frac{1}{4}$

$\frac{7}{8} - \frac{3}{8} + \frac{1}{4} = P(B)$

$\frac{3}{4} = P(B).$

(iii) P(A) = P(A∩B) + P(A∩B/)

$\frac{3}{4} = \frac{1}{4} + P(A\cap B^/)$

$\frac{3}{4} - \frac{1}{4} = P(A\cap B^/)$

$\frac{1}{8} = P(A\cap B^/)$

(iv) $P(A^{/} \cup B^{/}) = P(A \cap B)^{I}$

$= 1 - P(A \cap B)$

$= 1 - \frac{1}{4} = \frac{3}{4}$.

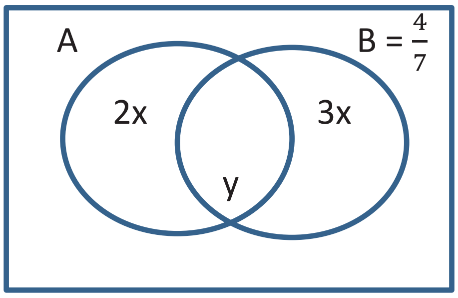
Example 2:

Two events A and B are such that $P(A^{/} \cap B) = 3x$, $P(A \cap B^{/}) = 2x$, $P(A^{/} \cap B^{/}) = x$ and $P(B) = \frac{4}{7}$. Using the Venn diagram, find the values of;

(i) x

(ii) $P(A \cap B)$

(i) Let y be the intersection



$P(B) = P(A \cap B) + P(A^{/} \cap B)$

$\frac{4}{7} = y + 3x$

$\frac{4}{7} - 3x = y$ …………..(i)

$P(\varepsilon) = 1$

$2x + y + 3x + x = 1$

$6x + (\frac{4}{7} - 3x) = 1$

$3x = 1 - \frac{4}{7}$

$3x = \frac{3}{7}$

$x = \frac{1}{7}$.

(ii) $P(A \cap B) = y$

$y = \frac{4}{7} - \frac{3}{7}$

$y = \frac{1}{7}$.

EXERCISE:

1. Two events M and N are such that $P(M) = 0.7$, $P(M \cap N) = 0.45$ and $P(M^{/} \cap N) = 0.18$.
   Find; (i) $P(N^{/})$
   (ii) P (M U N).

2. The probability that that Peter reads the New vision is 0.75. And the probability that he reads the New vision but not the Monitor is 0.65. The probability that he reads neither of the two papers is 0.15. Find the probability that he reads the Monitor.

3. The probability that a student passes Mathematics is $\frac{2}{3}$, the probability that he passes Physics is $\frac{4}{9}$. If the probability that he passes at least one of them is $\frac{4}{5}$. Find the probability that he passes both subjects.

Lesson 2: Probability of the 'OR', 'AND' situations

In this lesson you should be able to:

(i) define the 'or' , 'and' situations

(ii) define mutually exclusive and independent events

(iii) calculate probabilities using mutually exclusive and independent events

THE OR SITUATION

If A and B are two events, the probability that either event A or B or both occur is denoted by P(AUB).

THE AND SITUATION

For two events A and B, the probability that both events A and B occur together is P(A∩B).

Example:

Two dice are thrown. What is the probability of scoring either a double or a sum greater than 8?

Solution;

Table of outcomes can be used to generate the sample space.

	First die					
Second die	1,1	1,2	1,3	1,4	1,5	1,6
	2,1	2,2	2,3	2,4	2,5	2,6
	3,1	3,2	3,3	3,4	3,5	3,6
	4,1	4,2	4,3	4,4	4,5	4,6
	5,1	5,2	5,3	5,4	5,5	5,6
	6,1	6,2	6,3	6,4	6,5	6,6

Table of sums

2	3	4	5	6	7
3	4	5	6	7	8
4	5	6	7	8	9
5	6	7	8	9	10
6	7	8	9	10	11
7	8	9	10	11	12

Let A represent a double and B a sum greater than 8

A = $\{(1,1)(2,2)(3,3)(4,4)(5,5)(6,6)\}$

B = $\{(3,6)(4,5)(4,6)(5,4)(5,5)(5,6)(6,3)(6,4)(6,5)(6,6)\}$

n(A) = 6, n(B) = 10, n(A∩B) = 2

P(A∪B) = P(A) + P(B) – P(A∩B)

$= \frac{6}{36} + \frac{10}{36} - \frac{2}{36} = \frac{7}{18}$.

∴P(Double or sum greater than 8) = $\frac{7}{18}$.

Mutually Exclusive events

Two events A and B are said to be mutually exclusive if they cannot occur at the same time. i.e. the probability that they both occur is zero.

If A and B are mutually exclusive then P(A∩B) = 0

P(AUB) = P(A) + P(B).

Example:

1. Given that A and B are mutually exclusive events such that P(A) = 0.5 and P(AUB) = 0.9 find; (i) P(B)
   (ii) $P(A^{/}\cap B^{/})$

Solution:

For mutually exclusive events;

P(AUB) = P(A) + P(B)

P(B) = P(AUB) – P(A)

= 0.9 – 0.5

= 0.4

(ii) $P(A^{/}\cap B^{/}) = P(AUB)^{/}$

= 1 – P(AUB)

= 1 – 0.9

= 0.1

2. In a race, the probability that Grace wins is 0.4, the probability that Mahad wins is 0.2 and the probability that Denis wins is 0.3. Find the probability that;
   (i) Grace or Denis wins
   (ii) Neither Denis nor Mahad wins

Solution:

(i) P(GUD) = P(G) + P(D) $[Since\ G, U\ \&D\ are\ mutually\ exclusive\ events\ ]$

$= 0.4 + 0.3$

$= 0.7$

(ii) $P(D^/ \cap M^/) = P(DUM)^I$

$= 1 - P(DUM)$

$= 1 - [P(D) + P(M)]$

$= 1 - (0.3 + 0.2)$

$= 1 - 0.5$

$= 0.5$

INDEPENDENT EVENTS

Events are said to be independent if and only if the occurrence of one event does not influence the occurrence of the other event. Or the nonoccurrence of one does not influence the nonoccurrence of the other.

For independent events;

P(A and B) = P(A ∩ B) = P(A) x P(B).

Example1:

Given that A and B are independent events such that P(A) = $\frac{2}{5}$, P(AUB) = $\frac{4}{5}$.

Find; (i) P(B)

(ii) $P(A^/UB^i)$.

Solution:

(i) P(AUB) = P(A) + P(B) – P(A∩B)

Since A and B are independent events,

Then P(A∩B) = P(A) x P(B)

$\Rightarrow \frac{4}{5} = \frac{2}{5} + P(B) - \frac{2}{5}P(B)$

$\frac{4}{5} - \frac{2}{5} = P(B)\left[1 - \frac{2}{5}\right]$

$\frac{2}{5} = \frac{3}{5}P(B)$

$P(B) = \frac{2}{3}$.

(ii) $P(A^/UB^/) = P(A \cap B)^I$

$= 1 - P(A \cap B)$

$= 1 - \frac{2}{5} x \frac{2}{3}$

$= 1 - \frac{4}{15} = \frac{11}{15}$

Example 2:

Given that A and B are independent events, show that A^I and B^I are also independent.

Solution:

For independent events; P(A∩B) = P(A) x P(B).

We need to show that $P(A^/ \cap B^/) = P(A^/) \times p(B^/)$

$P(A^{/}\cap B^{/}) = P(AUB)^{I}$

$= 1 - P(AUB)$

$= 1 - [P(A) + P(B) - P(A \cap B)]$

$= 1 - [P(A) + P(B) - P(A) \times P(B)]$

$= 1 - P(A) - P(B) + P(A)xP(B)$

$= 1 - P(A) - P(B)[1 - P(A)]$

But; $1 - P(A) = P(A^{/})$

$\Rightarrow P(A^{/}\cap B^{/}) = P(A^{/}) - P(B)P(A^{/})$

$= P(A^{/})[1 - P(B)]$

Also; $1 - P(B) = P(B^{/})$

$\therefore P(A^{/}\cap B^{/}) = P(A^{/})P(B^{/})$. Hence they are independent.

Exercise:

1. Events A and B are independent and P(A) = 0.3, P(AUB) = 0.6. Find;
   (i) P(A∩B)
   (ii) P(B)

2. If A and B are mutually exclusive events that P(A) = 0.2, $P(AUB)^{I}$ = 0.3. Find;
   (i) P(B)
   (ii) $P(A^{/}\cap B)$
   (iii) $P(A\cap B^{/})$

3. Events A and B are independent such that P(A∩B) = $\frac{1}{4}$, P(AUB) = $\frac{3}{4}$. Find;
   (i) P(A)
   (ii) P(B)

4. The probability that two independent events occur together is $\frac{2}{15}$. The probability that either or both events occur is $\frac{2}{3}$. Find the individual probabilities of the two events.

LESSON 3: CONDITIONAL PROBABILITY

Learning Outcomes

Learners should be able to:

(i) define conditional probability.
(ii) draw tree diagrams.
(iii) apply conditional probability and tree diagrams to solve probabilities.

A conditional probability is a situation where an event will occur on the condition that another event has occurred. If A and B are events then the conditional probability of A given B is;

$$P\left(A/_{B}\right)=\frac{P(A\cap B)}{P(B)}.$$

This is provided $P(B) \neq 0$.

Example 1.

Given that A and B are events such that $P(B) = \frac{1}{6}$, $P(A\cap B) = \frac{1}{12}$ and $P\left(B/_{A}\right) = \frac{1}{3}$.

Find: (i) P(A)

(ii) $P\left(A/_{B}\right)$

(iii) $P\left(A/_{B^I}\right)$

Solution: (i) $P\left(B/_{A}\right) = \frac{P(B\cap A)}{P(A)}$

$$\frac{1}{3} = \frac{\frac{1}{12}}{P(A)}$$

$$P(A) = \frac{1}{4}.$$

(ii) $P\left(A/_{B}\right) = \frac{P(A\cap B)}{P(B)}$

$$= \frac{1}{12} \div \frac{1}{6}$$

$$= \frac{1}{2}.$$

(iii) $P\left(A/_{B^I}\right) = \frac{P(A\cap B^I)}{P(B^I)}$

$$= \frac{P(A) - P(A\cap B)}{1 - P(B)}$$

$$= \left(\frac{1}{4} - \frac{1}{12}\right) \div \left(1 - \frac{1}{6}\right)$$

$$= \frac{1}{6} \div \frac{5}{6}$$

$$= \frac{5}{6}.$$

Example 2:

Given that A and B are two events such that $P(A) = \frac{1}{3}$, $P(B) = \frac{1}{4}$, $P(AUB) = \frac{1}{2}$

Find:

(i) $P\left(A/_{B}\right)$

(ii) $P\left(A/_{B^I}\right)$

Solution:

(i) $P\left(A/B\right) = \frac{P(A\cap B)}{P(B)}$

$P(A\cup B) = P(A) + P(B) - P(A\cap B)$

$\frac{1}{2} = \frac{1}{3} + \frac{1}{4} - P(A\cap B)$

$P(A\cap B) = \frac{7}{12} - \frac{1}{2}$

$= \frac{1}{12}$

$P\left(A/B\right) = \frac{1}{12} \div \frac{1}{4}$

$= \frac{1}{3}.$

(ii) $P\left(A/B^I\right) = \frac{P(A\cap B^I)}{P(B^I)}$

$= \frac{P(A) - P(A\cap B)}{1 - P(B)}$

$= \left(\frac{1}{3} - \frac{1}{12}\right) \div \left(1 - \frac{1}{4}\right)$

$= \frac{1}{4} \div \frac{3}{4}$

$= \frac{1}{3}.$

Tree Diagrams

To draw a tree diagram, we begin from a single point and produce a pair of branches for the first trial. We continue drawing branches for each of the subsequent outcomes.

To obtain the required probability look at the branches and trace out the path associated with events of interest then;

- Take product of the probabilities along that path
- Do the same for other similar paths of interest
- Sum the products

Note: At every junction where branches meet, the probabilities sum up to one.

Example 1

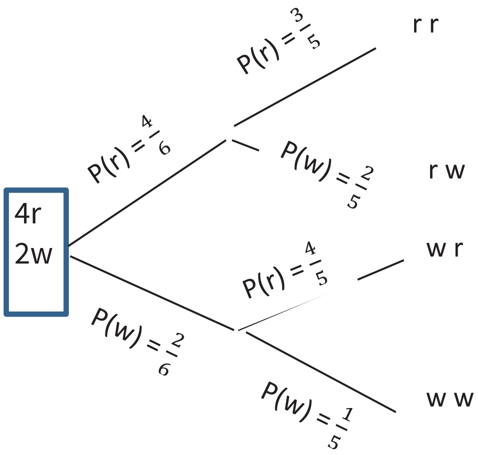
A box contains 4 red and 2 white balls. If two balls are picked from the box one at a time without replacement. Find the probability that;

(i) The second ball is red

(ii) Both balls are white

(iii) Both balls are of the same colour

Solution:



(i) The second ball is red

$= P(rr) \text{ or } P(wr)$

$= \left(\frac{4}{6} x \frac{3}{5}\right) + \left(\frac{2}{6} x \frac{4}{5}\right)$

$= \frac{2}{5} + \frac{4}{15}$

$= \frac{2}{3}$

(ii) Both balls are white

$= P(ww)$

$= \frac{2}{6} x \frac{1}{5} = \frac{1}{15}$.

(iii) Both balls are of same colour

$= P(rr) \text{ or } P(ww)$

$= \left(\frac{4}{6} x \frac{3}{5}\right) + \left(\frac{2}{6} x \frac{1}{5}\right)$

$= \frac{2}{5} + \frac{1}{15}$

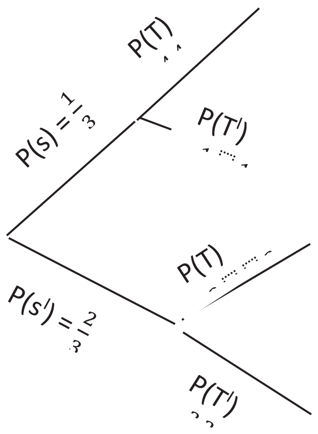
$= \frac{7}{15}$

Example 2:

The probability that it will be sunny tomorrow is $\frac{1}{3}$. If it is sunny, the probability that Simon plays tennis tomorrow is $\frac{4}{5}$. If it is not sunny, the probability that he plays tennis is $\frac{2}{5}$. Find the probability that Simon play tennis tomorrow.

Solution:

Let S represent sunny, and T represent playing tennis.



$$P(T) = P(S \cap T) + P(S' \cap T)$$
$$= \left(\frac{1}{3} \times \frac{2}{5}\right) + \left(\frac{2}{3} \times \frac{2}{5}\right)$$
$$= \frac{2}{15} + \frac{4}{15}$$
$$= \frac{6}{15} = \frac{2}{5}.$$

Exercise:

1. Show that $P\left(A/_B\right) + P\left(A^I/_B\right) = 1$
2. The events A and B are independent with P(A) = $\frac{1}{2}$ and P(A U B) = $\frac{2}{3}$, find;
   (i) P(B)
   (ii) $P\left(A/_B\right)$
   (iii) $P\left(B^I/_A\right)$

3. (a) A bag contains 30 white (W), 20 blue (B) and 20 red (R) balls. Three balls are drawn at random one after the other without replacement. Determine the probability that the first ball is white and the third ball is also white.

   (b) Events A and B are such that P(A) = $\frac{4}{7}$, P(A U B/) = $\frac{1}{3}$ and $P\left(A/_B\right) = \frac{5}{14}$.
   Find; (i) P(B). (ii) $P(A' \cap B')$.

4. A and B are two identical boxes, Box A contains one diamond ring and two gold rings. Box B has 3 diamond and 4 gold rings. A box is chosen at random and from it one ring is randomly taken and put into the other box. And a ring is then randomly drawn from the later box.

Determine the probability that;

(i) Both rings are diamond

(ii) The first ring is gold

(iii) The first ring is diamond given that the 2^{nd} is gold.

5. A box contains two types of balls, red and black. When a ball is picked from the box, the probability that it is red is $\frac{7}{12}$. Two balls are selected at random from the box without replacement.

   Find the probability that;

   (i) The second ball is black

   (ii) The first ball is red given that the second is black.

6. An interview involves written, oral and practical tests. The probability that an interviewee passes the written test is 0.8, the oral test is 0.6 and the practical test is 0.7.

   What is the probability that the interviewee will pass

   (i) The entire interview

   (ii) Exactly two of the interview tests.

Topic 5: Discrete Probability Distribution

Lesson 1. Random Variables

In this lesson you should be able to learn how to:

(i) describe a random experiment.
(ii) state the properties of a discrete random variable.
(iii) generate a probability distribution function of a discrete random variable from a given experiment.

When an experiment is performed, it is common that the main interest is some function of the outcome as opposed to the actual outcome itself.

For example, in tossing a pair of dice, we are often interested in the sum of the two dice and not really about the separate values of the dice.

We may be interested in knowing that the sum is 8 but not concerned about the actual outcomes i.e. (2,6), (3,5), (4,4), (5,3) or (6,2).

Random variable is a real valued function representing the outcome of a random experiment. X is used to denote the random variable and x denotes its outcomes.

Examples

Random Experiment	Random Variable
Tossing a coin 3 times	X = number of heads obtained
Rolling a pair of dice	X = sum, is 8
Measure the height of students	X = height is greater than 140 cm

Discrete Random Variable

A discrete random variable is one which takes on a countable number of values (outcomes). It assumes that each value has a certain probability.

Examples of a discrete random experiment and corresponding set of outcomes may be:

(a) a score when a die is tossed: X = 1, 2, 3, 4, 5, 6.
(b) the number of boys in a family of five children: X = 0, 1, 2, 3, 4, 5.
(c) the number of heads when two coins are tossed: X = 0, 1, 2.

If X is a discrete random variable, then the function $f(x) = P(X = x)$. The function is called the probability distribution of x.

Properties of a discrete random variable

1. $P(X = x) = f(x) \geq 0$. For all values of x
2. Summation of the probability for all values of x is 1

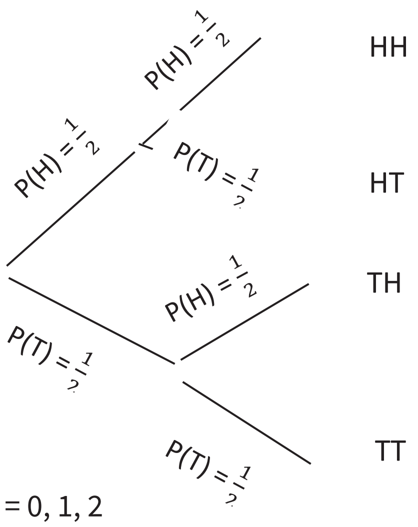
   i.e. $\sum_{all\,x} P(X = x) = 1$ **or** $\sum_{all\,x} f(x) = 1$

   Total probability is 1

Example 1

If a random variable X is the number of heads obtained when an unbiased coin is tossed twice:

(i) form a sample space.
(ii) find the probability distribution.



Number of heads = 0, 1, 2

When x = 0: $P(X = 0) = P(TT)$

$= \frac{1}{2} \times \frac{1}{2} = \frac{1}{4}$

When x = 1: $P(X = 1) = P(HT)$ or $P(TH)$

$= \frac{1}{2} \times \frac{1}{2} + \frac{1}{2} \times \frac{1}{2}$

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

$= \frac{1}{2}$

When x = 2: $P(X = 2) = P(HH)$

$= \frac{1}{2} \times \frac{1}{2} = \frac{1}{4}$

∴ The probability distribution is

X	0	1	2
P(X = x)	$\frac{1}{4}$	$\frac{1}{2}$	$\frac{1}{4}$

Exercise

1. Write the probability distribution of the score when an ordinary die is thrown.
2. Two discs are drawn at random without replacement from a bag containing 3 blue and 4 yellow discs. If X is a random variable for the number of blue discs drawn, construct a probability distribution for X
3. In a bag there are 3 green counters, 4 black counters and 2 red counters. Two counters are picked at random from the bag: one after the other without replacement. Find the probability distribution for the green counters.

Lesson 2: Use Properties of Discrete Probability Distribution

In this lesson you should be able to apply the laws of discrete random variable to solve problems.

Example

1. A discrete random variable X has a probability distribution function given by:

$$f(x) = \begin{cases} \frac{a}{2x}; x = 1,2,3,4 \\ 0,..elswhere \end{cases}$$

Determine:

(i) The value of a

(ii) P(X ≤ 3)

Solution

(i) For a discrete random variable,

$\sum_{all\ x} P(X = x) = 1$

P(X = 1) + P(X = 2) + P(X = 3) + P(X = 4) = 1

$\frac{a}{2\ x\ 1} + \frac{a}{2\ x\ 2} + \frac{a}{2\ x\ 3} + \frac{a}{2\ x\ 4} = 1$

$\frac{a}{2} + \frac{a}{4} + \frac{a}{6} + \frac{a}{8} = 1$

12a + 6a + 4a + 3a = 24

25a = 24

a = $\frac{24}{25}$.

$$f(x) = \begin{cases} \frac{12}{25x}; x = 1,2,3,4 \\ 0,..elswhere \end{cases}$$

(ii) P(X ≤ 3)

P(X = 1) + P(X = 2) + P(X = 3)

$\frac{12}{25\ X\ 1} + \frac{12}{25\ X\ 2} + \frac{12}{25\ X\ 3}$

$\frac{12}{25} + \frac{6}{25} + \frac{4}{25}$

$\frac{22}{25}$

2. A discrete random variable has a probability distribution function as shown below;

X	1	2	3	4	5
P(X = x)	0.2	0.25	0.4	e	0.05

Determine:

(i) the value of e

(ii) P(1 ≤ X < 3)

(iii) P(X > 2)

Solution

(i) $\sum_{all\ x} P(X = x) = 1$

0.2 + 0.25 + 0.4 + e + 0.05 = 1

0.9 + e = 1

e = 1 – 0.9

e = 0.1.

(ii) P(1 ≤ x < 3)

P(X = 1) + P(X = 2)

0.2 + 0.25

0.45

(iii) P(X > 2)

P(X = 3) + P(X = 4) + P(X = 5)

0.4 + 0.1 + 0.05

0.55

3. The probability distribution function of a random variable X is given by:

$$P(X = x) = \begin{cases} kx + d : x = -2,-1,0\,1,2 \\ 0,..elswhere \end{cases}$$

Where **k** and **d** are constants, if P(x = 2) = 2P(x = -2).
Determine:
(i) The value of **k** and **d**.
(ii) P(X $\geq$ 0)

Solution:

(i) $\sum_{all\ x} P(X = x) = 1$

P(X=-2) + P(X=-1) + P(X=0) + P(X=1) + P(X=2) = 1

(-2k+d) + (-k+d) + (d) + (k+d) + (2k+d) = 1

-2k+d –k+d + d + k+d +2k+d = 1

5d = 1

d $=\frac{1}{5}$

(ii) P(X=2) = 2P(X=-2)

2k + d = 2(-2k+d)

2k + d = -4k + 2d

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

6k = $\frac{2}{5}$ - $\frac{1}{5}$

6k = $\frac{1}{5}$

k = $\frac{1}{30}$

(iii) P(X $\geq$ 0)

= P(X=0) + P(X=1) + P(X=2)

= d + k + d + 2k + d

= 3d + 3k

$= \frac{3}{5} + \frac{3}{30}$

$= \frac{7}{10}$

Lesson 3: Mean or Expectation and Variance of Discrete Random Variable

In this lesson, you should be able to learn how to:
(i) calculate the mean of a discrete random variable.
(ii) calculate the variance of a discrete random variable.

The mean, E(x) or μ of a discrete random variable is given by:

E(X) = $\sum_{all\ x} \boldsymbol{xP(X = x)}$

And the variance is given by:

Var(x) = E(x²) - $\boldsymbol{(E(x))^2}$**.**

Where E(x^2) = $\sum_{all\ x} \boldsymbol{x^2 P(X = x)}$

And Standard deviation (Sigma) $\delta$is given by $\sqrt{Var(x)}$.

i.e. Standard deviation = $\sqrt{\mathbf{E}(x^2) - (E(x))^2}$

Example

A random variable X has the following distribution: P(X=0) = P(X=1) = 0.1, P(X=2) = 0.2, p(X=3) = P(X=4) = 0.3. Find:

(i) the mean
(ii) the variance
(iii) standard deviation

Solution

(i) E(X) = $\sum_{all\ x} xP(X = x)$

x	P(X=x)	xP(X=x)	x^2	$x^2P(X=x)$
0	0.1	0	0	0
1	0.1	0.1	1	0.1
2	0.2	0.4	4	0.8
3	0.3	0.9	9	2.7
4	0.3	1.2	16	4.8
		$\sum = \mathbf{2.6}$		$\sum = \mathbf{8.4}$

E(x) = $\sum_{all\ x} xP(X = x)$
= 2.6

(ii) Var(x) = E(x^2) - $(E(x))^2$
= 8.4 - $(2.6)^2$
= 8.4 - 6.76
= 1.6

(iii) Standard deviation = $\sqrt{Var(x)}$
= $\sqrt{1.6}$
= 1.265

Properties of the mean

(i) E(a) = a where **a** is a constant.
(ii) E(ax) = aE(x)
(iii) E (ax + b) = aE(x) + b where **a** and **b** are constants.

Properties of the variance

(i) Var(a) = 0
(ii) Var(ax) = a^2Var(x)
(iii) Var (ax + b) = a^2Var(x) + 0 where a and b are constants.

Exercise

1. A discrete random variable X takes integral values from 0 to 5 with probabilities given by:

$$P(X=x)=\begin{cases} k(2x+1).; x=0,1.2.3 \\ k(11-2x).; x=4,5. \\ 0,.elswhere \end{cases}$$

Find**:**
(i) The value of k
(ii) $P(2 < x \leq 4)$
(iii) The expectation of X
(iv) The variance of X

2. A discrete random variable X has a probability density function given by:

X	-1	0	1	2
P(X = x)	0.25	0.10	0.45	0.20

If Y is a random variable defined by Y = (0.5x + 3).

Determine
(i) E(x)
(ii) E(y)
(iii) Var(y)

3. Box A contains 4 red sweets and 3 green sweets. Box B contains 5 red sweets and 6 green sweets. Box A is twice as likely to be picked as box B. if a box is chosen at random and two sweets are removed from it one at a time without replacement:
   (a) find the probability that the two sweets removed are of the same colour.
   (b) (i) construct a probability distribution table for the number of red sweets removed.
   (ii) find the mean number of red sweets removed.

Lesson 4: Mode and Median of a Discrete Random Variable

In this lesson you should be able to:
(i) determine the mode of a discrete random variable.
(ii) generate a cumulative mass function (c.d.f) of a discrete random variable.
(iii) determine the median of a discrete random variable.

For a discrete random variable, mode is the value of **x** which has the highest probability.

Example

A discrete random variable X has a probability distribution defined by:

X	0.5	1	1.5	2
P(X = x)	M	m^2	$2m^2$	M

(i) Find the value of m.
(ii)Find the mode.
(iii) Find $P(1 \leq x < 2)$.

Solution

(i) For a P.D.F of a discrete random variable,

$\sum_{all\ x} P(X = x) = 1$

$m + m^2 + 2m^2 + m = 1$

$3m^2 + 2m - 1 = 0$

$(3m - 1)(m + 1) = 0$

$3m - 1 = 0$ or $m + 1 = 0$

$m = \frac{1}{3}$ or $m = -1$

We take the positive value $m = \frac{1}{3}$.

Hence the p.d.f:

X	0.5	1	1.5	2
P(X = x)	$\frac{1}{3}$	$\frac{1}{9}$	$\frac{2}{9}$	$\frac{1}{3}$

(ii) $P(X = 0.5) = \frac{1}{3}$, and $P(X = 2) = \frac{1}{3}$. is the highest probability, therefore the mode is

0.5 and 2. The distribution is a Bi-modal (it has two modes).

(iii) $P(1 \leq x < 2)$

$P(X = 1) + P(X = 1.5)$

$\frac{1}{9} + \frac{2}{9} = \frac{1}{3}$.

Cumulative Mass Function (C.M.F) F(X)

The cumulative distribution function of a discrete random variable is given by:

$$F(X) = P(X \leq x)$$

Note: Cumulative is obtained by adding probabilities of the corresponding values of x.

Example

Consider the following random variable which the probability distribution as:

X	1	2	3	4	5	6
P(X=x)	$\frac{1}{6}$	$\frac{1}{6}$	$\frac{1}{6}$	$\frac{1}{6}$	$\frac{1}{6}$	$\frac{1}{6}$

Obtain the cumulative distribution function.

Solution

$F(1) = P(X \leq 1)$, $= \frac{1}{6}$.	$F(2) = P(X \leq 2)$, $= P(X = 1) + P(X = 2)$ $= \frac{1}{6} + \frac{1}{6} = \frac{1}{3}$	$F(3) = P(X \leq 3)$ $= F(2) + P(3)$ $= \frac{1}{3} + \frac{1}{6} = \frac{1}{2}$
$F(4) = P(X \leq 4)$, $= F(3) + P(4)$ $= \frac{1}{2} + \frac{1}{6} = \frac{2}{3}$	$F(5) = P(X \leq 5)$, $= F(4) + P(5)$ $= \frac{2}{3} + \frac{1}{6} = \frac{5}{6}$	$F(6) = P(X \leq 6)$ $= F(5) + P(6)$ $= \frac{5}{6} + \frac{1}{6} = 1.$

Note: Always the last value of F(X) = 1.

Hence, the cumulative distribution function will be:

x	**P(X = x)**	**F(X)**
1	$\frac{1}{6}$	$\frac{1}{6}$
2	$\frac{1}{6}$	$\frac{1}{3}$
3	$\frac{1}{6}$	$\frac{1}{2}$
4	$\frac{1}{6}$	$\frac{2}{3}$
5	$\frac{1}{6}$	$\frac{5}{6}$
6	$\frac{1}{6}$	1

Median of a Discrete Random Variable

The median of a probability distribution function of a discrete random variable X is the smallest value for which F(X) is at least 0.5.

i.e. if m is the median then: (i) $F(m) \geq 0.5$

(ii) $1 - F(m - 1) \geq 0.5$

Example

A random variable X has the following probability distribution;

X	1	2	3	4
P(X = x)	0.1	0.4	0.3	0.2

Find:
(i) the cumulative distribution function.
(ii) the median.

Solution

(i) C.D.F $= F(X) = P(X \leq x)$

X	P(X = x)	P(X ≤ x)
1	0.1	0.1
2	0.4	0.5
3	0.3	0.8
4	0.2	1.0

(ii) The median
$F(m) \geq 0.5$, where m is the median
From the table, F (2) = 0.5
Therefore median = 2.

Exercise

1. A bag contains 7 ripe mangoes and 8 raw mangoes. Two mangoes are picked in succession with replacement. Find the most likely number of ripe mangoes picked.
2. The discrete random variable has cumulative mass function F(X) given by:

$$F(X) = 1 - \left[1 - \frac{x}{4}\right]^x . \text{ For } x = 1, 2, 3, 4.$$

(a) Show that F(3) = $\frac{63}{64}$ and F(2) = $\frac{3}{4}$.
(b) Obtain the probability distribution function of X.
(c) Find E(X) and Var(X).
(d) Find P$(X > E(X))$.

3. A discrete random variable X has probability function given by:

$$P(X = x) = \begin{cases} \frac{x}{k} : x = 1,2,3,...n \\ 0,..elswhere \end{cases}$$

Where k is a constant.
Given that the expectation of X is 3.
Find:
(i) The value of n and the constant k
(ii) The median and the variance of X
(iii) P$\left(X = 2/X \geq 2\right)$.

TOPIC 6: Continuous Random Variable

Lesson 1: Properties of a Continuous Random Variable X

Competences

In this lesson, you should be able to learn how to:
(i) define continuous random variables.
(ii) state the properties of a continuous random variable.
(iii) apply the properties of random variables to solve probability problems.

A continuous random variable is a theoretical representation of continuous variables such as weight, temperature, time, distance, mass, height. A probability distribution function f(x) of random variable x is said to be continuous if it has a continuous domain.

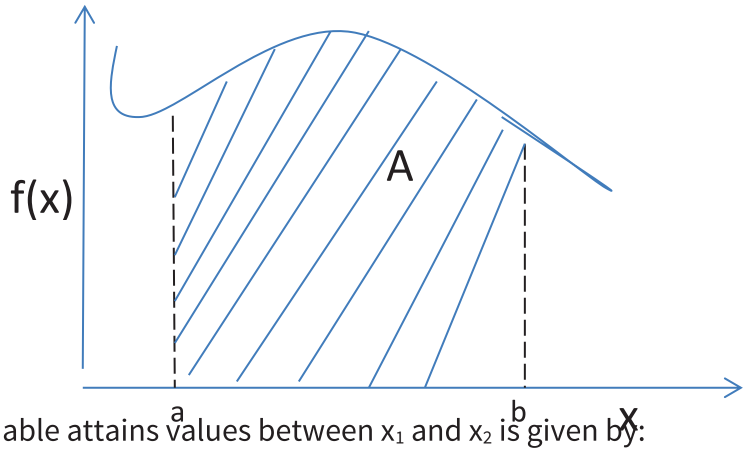
Properties of Continuous Random Variable f(x)

1. f(x) $>$ 0

2. $\int_{all\,x} f(x)dx = 1$ or $\int_{-\infty}^{\infty} f(x)dx = 1$

   i.e. the total area under a curve is 1.

   A = $\int_a^b f(x)dx = 1$



How to Obtain Probabilities

The probability that a random variable attains values between x_1 and x_2 is given by:

$$P(x_1 \le x \le x_2) = \int_{x_1}^{x_2} f(x)dx.$$

Example

1. A continuous random variable has a probability distribution function given by:

$$f(x) = \begin{cases} kx.;.for 0 \le x \le 4 \\ 0.;.elsewhere \end{cases}$$

Determine

(a) The value of k
(b) $P(1 \le x \le 2)$
(c) Sketch the graph of f(x)

Solution

(a) For a p.d.f, $\int_{-\infty}^{\infty} f(x)dx = 1$

$$\int_0^4 k(x)dx = 1$$

$$\left[\frac{kx^2}{2}\right]_0^4 = 1$$

$$\frac{k(4)^2}{2} - \frac{k(0)^2}{2} = 1$$

$$\frac{16k}{2} = 1$$

$$k = \frac{1}{8}.$$

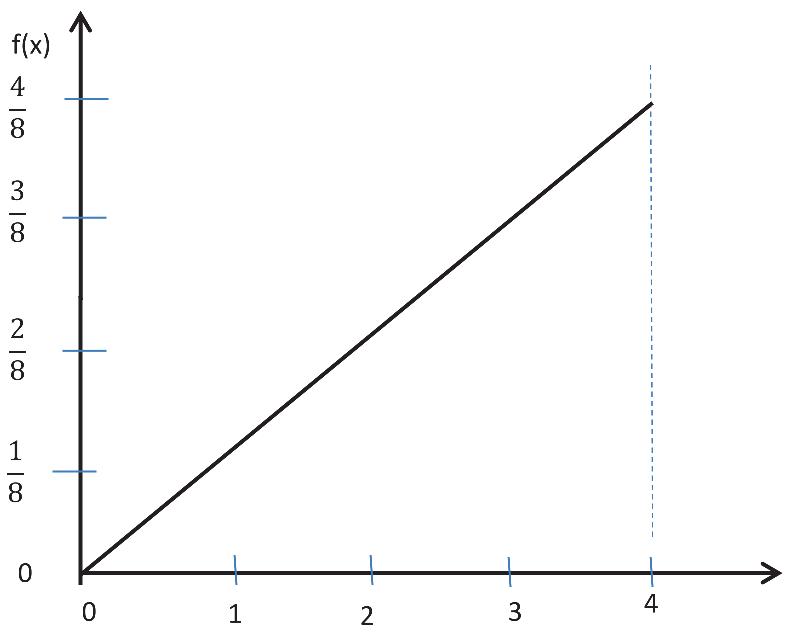
$$f(x) = \begin{cases} \frac{1}{8}x.;. for 0 \le x \le 4 \\ 0.;. elsewhere \end{cases}$$

(b) P(1 ≤ x ≤ 2) = $\int_1^2 \frac{1}{8}(x)dx$

$= \left[\frac{x^2}{16}\right]_1^2$

$= \frac{2^2}{16} - \frac{1^2}{16} = \frac{4}{16} - \frac{1}{16} = \frac{3}{16}.$

(c) Graph of f(x)



Example 2

A continuous random variable has a probability distribution function where:

$$f(x) = \begin{cases} kx.;. for 0 \le x \le 2 \\ k(4-x).;. for 2 \le x \le 4 \\ 0.;. elsewhere \end{cases}$$

Determine

(i) The value of the constant k.
(ii) P(1 ≤ x ≤ 3).
(iii) Sketch the graph of f(x).

Solution

(i) $\int_{-\infty}^{\infty} f(x)dx = 1$

$\int_0^2 kx\,dx \;+\; \int_2^4 k(4-x)dx = 1$

$\left[\frac{kx^2}{2}\right]_0^2 \;+\; k\left[4x - \frac{x^2}{2}\right]_2^4 = 1$

$k\left[\frac{4}{2} - \frac{0}{2}\right] \;+\; k[(16-8)-(8-2)] \;=\; 1$

$2k \;+\; 2k \;=\; 1$

$4k \;=\; 1$

$k \;=\; \frac{1}{4}.$

$$f(x) = \begin{cases} \frac{1}{4}x.;.for 0 \le x \le 2 \\ \frac{1}{4}(4-x).;.for 2 \le x \le 4 \\ 0.;.elsewhere \end{cases}$$

(ii) $P(1 \le x \le 3)$

$= \int_1^3 f(x)\,dx$

$= \int_1^2 \frac{1}{4}x\,dx \;+\; \int_2^3 \frac{1}{4}(4-x)dx$

$= \left[\frac{x^2}{8}\right]_1^2 \;+\; \frac{1}{4}\left[4x - \frac{x^2}{2}\right]_2^3$

$= \left(\frac{4}{8} - \frac{1}{8}\right) \;+\; \frac{1}{4}\left[\left(\frac{12}{1} - \frac{9}{2}\right) - (8-2)\right]$

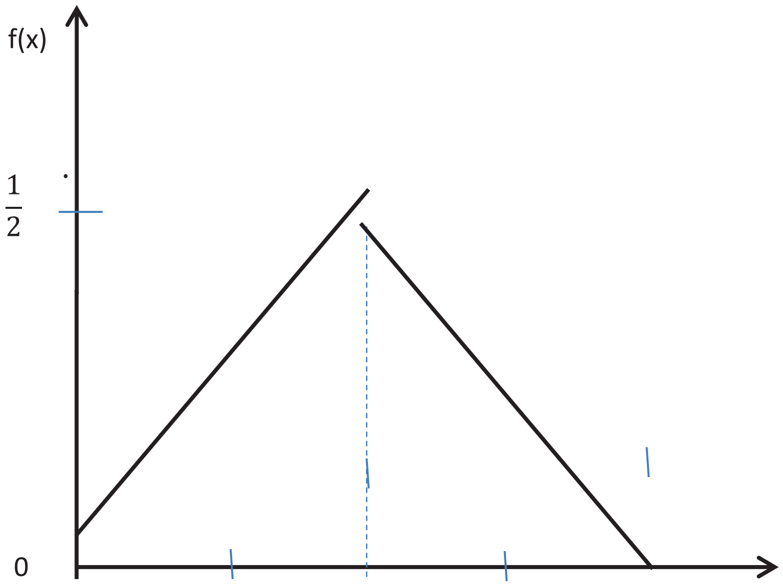
$= \frac{3}{8} \;+\; \frac{1}{4}\left[\left(\frac{24-9}{2}\right) - 6\right]$

$= \frac{3}{8} \;+\; \frac{1}{4}\left[\frac{15}{2} - \frac{6}{1}\right]$

$= \frac{3}{8} \;+\; \frac{1}{4} \times \frac{3}{2}$

$= \frac{3}{8} \;+\; \frac{3}{8} \;=\; \frac{6}{8} = \frac{3}{4}$

(c) Graph of f(x);



Exercise

1. A continuous random variable has a probability distribution function given by;

$$f(x) = \begin{cases} k \quad for\ 0 \le x \le 2 \\ k(2x-3) \quad for\ 2 \le x \le 3 \\ 0 \quad elsewhere \end{cases}$$

Determine

(a) The value of the constant k.

(b) P(1 $\le$ x $\le$ 2.5).

(c) Sketch the graph of f(x).

Lesson 2: Expectation (Mean) and Variance of a Continuous Random Variable

Competence

In this lesson, you should be able to learn how to calculate the mean and variance of a continuous random variable.

The mean of a continuous random variable is given by:

$$E(x) = \int_{-\infty}^{\infty} x f(x) dx$$

And the variance is given by:

$$Var(x) = E(x2) - [E(x)]^2$$

Where, E(x2) $=\int_{-\infty}^{\infty} x^2 f(x)dx.$

Example

1. A continuous random variable has a probability distribution function f(x) where:

$$f(x) = \begin{cases} \frac{x}{8}; for 0 \leq x \leq 4 \\ 0; elsewhere \end{cases}$$

Determine:

(i) E(x)
(ii) Var(x)
(iii) Standard deviation

Solution

(i) $E(X) = \int_0^4 \frac{x^2}{8}\, dx$

$= \left[\frac{x^3}{24}\right]_0^4$

$= \left(\frac{64}{24}\right) - \left(\frac{0}{24}\right)$

$= \frac{8}{3}.$

(ii) $Var(x) = E(x^2) - [E(x)]^2$

$E(x^2) = \int_{-\infty}^{\infty} x^2 f(x)dx$

$= \int_0^4 \frac{x^3}{8}\, dx$

$= \left[\frac{x^4}{32}\right]_0^4$

$= \frac{256}{32} - \frac{0}{32}$

$= 8.$

$Var(x) = 8 - \left(\frac{8}{3}\right)^2$

$= 8 - \frac{64}{9}$

$Var(x) = \frac{8}{9}.$

(iii) Standard deviation

S.D $= \sqrt{Var(x)}$

$$= \sqrt{\frac{8}{9}}$$
$$= 0.943$$

Example 2

A continuous random variable has a probability distribution function given by:

$$f(x) = \begin{cases} ak; for 0 \leq x \leq 10 \\ a(20-x); for 10 \leq x \leq 20 \\ 0; elsewhere \end{cases}$$

Find:

(i) The value of a

(ii) E(x)

(iii) Var(x)

Solution

(i) $\int_{-\infty}^{\infty} f(x)dx = 1$

$$\int_0^{10} ax\, dx + \int_{10}^{20} a(20-x)dx = 1$$

$$\left[\frac{ax^2}{2}\right]_0^{10} + a\left[20x - \frac{x^2}{2}\right]_{10}^{20} = 1$$

$$\left[\frac{100a}{2} - \frac{0}{2}\right] + a[(400-200)-(200-50)] = 1$$

$$50a + 50a = 1$$

$$100a = 1$$

$$a = \frac{1}{100}.$$

$$f(x) = \begin{cases} \frac{1}{100}k; for 0 \leq x \leq 10 \\ \frac{1}{100}(20-x); for 10 \leq x \leq 20 \\ 0; elsewhere \end{cases}$$

(ii) E(x) = $\int_{-\infty}^{\infty} xf(x)dx$

$$= \frac{1}{100}\int_0^{10} x^2\, dx + \frac{1}{100}\int_{10}^{20}(20x - x^2)dx$$

$= \frac{1}{100}\left[\frac{x^3}{3}\right]_0^{10} + \frac{1}{100}\left[\frac{20x^2}{2} - \frac{x^3}{3}\right]_{10}^{20}$

$= \frac{1}{100}(\frac{1000}{3} - \frac{0}{3}) + \frac{1}{100}\left[\left(4000 - \frac{8000}{3}\right) - (1000 - \frac{1000}{3}\right]$

$= \frac{10}{3} + \frac{1}{100}\left[\frac{4000}{3} - \frac{2000}{3}\right]$

$= \frac{10}{3} + \frac{1}{100} x \frac{2000}{3}$

$= \frac{10}{3} + \frac{20}{3}$

$= \frac{30}{3} = 10.$

(iii) Var(x) = $E(x^2)$ - $[E(x)]^2$

$E(x^2) = \int_{-\infty}^{\infty} x^2 f(x)dx$

$= \frac{1}{100}\int_0^{10} x^3 \, dx + \frac{1}{100}\int_{10}^{20}(20x^2 - x^3)dx$

$= \frac{1}{100}\left[\frac{x^4}{4}\right]_0^{10} + \frac{1}{100}\left[\frac{20x^3}{3} - \frac{x^4}{4}\right]_{10}^{20}$

$= \frac{1}{100}\left[\frac{10000}{4} - \frac{0}{4}\right] + \frac{1}{100}\left[\left(\frac{160000}{3} - \frac{160000}{4}\right) - (\frac{20000}{3} - \frac{10000}{4})\right]$

$= \frac{1}{100}(2500) + \frac{1}{100}\left[\frac{40000}{3} - \frac{12500}{3}\right]$

$= 25 + \frac{1}{100}(\frac{27500}{3})$

$= 25 + \frac{275}{3} = \frac{350}{3}.$

$\therefore$ Var(x) $= \frac{350}{3} - 10^2$

$= \frac{50}{3}.$

Exercise

1. A continuous random variable has probability distribution function given by;

$$f(x) = \begin{cases} kx^2; for 0 \le x \le 2 \\ k(6-x); for 2 \le x \le 6 \\ 0; elsewhere \end{cases}$$

Find the

(a) Value of k

(b) $P(1 \le x \le 4)$

(c) E(x)
(d) Var(x).

Lesson 3: Cumulative Distribution Function F(X)

Competence
In this lesson you should be able to learn how to obtain the cumulative mass function F(X) and also how to get probability distribution function f(x) when given F(X).

The cumulative distribution function C.D.F F(X) is given by:

$$F(X) = \int_{-\infty}^{x} f(t)dt$$

Example

Given that:

$$f(x) = \begin{cases} \frac{x}{8}; for 0 \le x \le 4 \\ 0; elsewhere \end{cases}$$

Find F(X) or the cumulative distribution function.

Solution

Over the interval; $0 \le x \le 4$

$$F(X) = \int_{-\infty}^{x} f(t)dt$$
$$= \int_{0}^{x} \frac{1}{8} t \, dt$$
$$= \frac{1}{8}\left[\frac{t^2}{2}\right]_0^x = \frac{x^2}{16}$$

$F(0) = 0$, $F(4) = 1$ and 1 over the interval $x \ge 4$

Therefore,

$$F(x) = \begin{cases} 0; for..x \le 0 \\ \frac{x^2}{16}; for.0 \le x \le 4 \\ 1;. for.x \ge 4 \end{cases}$$

Example 2

A continuous random variable has a probability distribution function given as below:

$$f(x) = \begin{cases} \frac{x}{3}; for. 0 \le x \le 2 \\ 2 - \frac{2x}{3}; for. 2 \le x \le 3 \\ 0;.elsewhere \end{cases}$$

Determine F(X)

$$F(X) = \int_{-\infty}^{x} f(t)dt$$

Over the interval $0 \le x \le 2$

$$F(X) = \int_0^x \frac{t}{3}dt = \left[\frac{t^2}{6}\right]_0^x = \frac{x^2}{6}$$

When x = 2, $F(2) = \frac{2}{3}$.

Over the interval $2 \le x \le 3$

$$F(X) = F(2) + \int_2^x \left(2 - \frac{2t}{3}\right) dt$$

$$= \frac{2}{3} + \left[2t - \frac{t^2}{3}\right]_2^x$$

$$= \frac{2}{3} + (2x - \frac{x^2}{3}) - (4 - \frac{4}{3})$$

$$= \frac{2}{3} + 2x - \frac{x^2}{3} - \frac{8}{3}$$

$$= \frac{6x - x^2 - 6}{3}$$

$$F(3) = \frac{18 - 9 - 6}{3} = 1$$

Therefore,

$$F(x) = \begin{cases} 0; for..x \leq 0 \\ \frac{x^2}{6}; for.0 \leq x \leq 2 \\ \frac{6x - x^2 - 6}{3};. for2 \leq x \leq 3 \\ 1; for.x \geq 3 \end{cases}$$

Exercise

A continuous random variable x has probability distribution function given by;

$$f(x) = \begin{cases} kx(3-x); for.0 \leq x \leq 2 \\ k(4-x); for.2 \leq x \leq 3 \\ 0;.elsewhere \end{cases}$$

Determine

(i) The value of k

(ii) F(X)

(iii) P(1 ≤ x ≤ 3).

Lesson 4: Mode and Median of a Continuous Random Variable

Competence

In this lesson, you should be able to learn how to calculate the mode and median of a continuous random variable.

Mode of a Continuous Random Variable

This is the value of x for which f^I(x) or $\frac{d}{dx}$(fx) = 0.

Example

A continuous random variable has probability distribution function f(x) where:

$$f(x) = \begin{cases} \frac{3}{80}(2+x)(4-x); for.0 \leq x \leq 4 \\ 0;.elsewhere \end{cases}$$

Find the mode.

Solution

f(x) = $\frac{3}{80}(8 + 2x - x^2)$

$\frac{d}{dx}$f(x) = $\frac{3}{80}(2 - 2x$

$f^{/}(x) = \frac{3}{40}(1 - x) = 0$

$1 - x = 0$

$x = 1.$

Note: If you get more than one value, check for the maximum by second derivative test $f^{//}(x) < 0$.

Example 2:

A continuous random variable has probability distribution function given by:

$$f(x) = \begin{cases} \frac{x}{108}(6-x)^2; for.0 \le x \le 6 \\ 0;.elsewhere \end{cases}$$

Obtain the mode.

f(x) = $\frac{x}{108}(36 - 12x + x^2) = \frac{1}{108}(36x - 12x^2 + x^3)$

$\frac{d}{dx}$f(x) = $\frac{1}{108}(36 - 24x + 3x^2) = \frac{1}{36}(12 - 8x + x^2)$

For mode; $\frac{d}{dx}$f(x) = 0

$\frac{1}{36}(12 - 8x + x^2) = 0$

$x^2 - 8x + 12 = 0$

$(x - 2)(x - 6) = 0$

Either x = 2 or x = 6.

$f^{/}(x) = \frac{1}{36}(12 - 8x + x^2)$

$f^{//}(x) = \frac{1}{36}(2x - 8) = \frac{1}{18}(x - 4)$

when x = 2, $f^{//}(2) = \frac{-1}{9}$ when x = 6, $f^{//}(6) = \frac{1}{9}$

.

for mode, $f^{//}(x) < 0$: ∴ mode = 2.

Median

The median divides the area under a curve into two equal parts. Hence the median of a continuous random variable is given by:

$\int_{-\infty}^{m} f(x)dx = \frac{1}{2}$ or $\int_{m}^{\infty} f(x)dx = \frac{1}{2}$ or m is the value for which F(X) $\geq \frac{1}{2}$.

Example

The probability distribution function of a continuous random variable is given by:

$$f(x) = \begin{cases} \frac{x}{6}; for.0 \leq x \leq 3 \\ 2 - \frac{1}{2}x; for.3 \leq x \leq 4 \\ 0;.elsewhere \end{cases}$$

Obtain the median.

Solution

For median $\int_{-\infty}^{m} f(x)dx = \frac{1}{2}$

Over the interval $0 \leq x \leq 3$,

$\frac{1}{6}\int_{0}^{3} xdx = \frac{1}{6}\left[\frac{x^2}{2}\right]_0^3 = \frac{1}{12}(9-0) = \frac{3}{4}$.

Since $\frac{3}{4}$ is $> \frac{1}{2}$ then median lies in the interval $0 \leq x \leq 3$

$\frac{1}{6}\int_{0}^{m} xdx = \frac{1}{2}$

$\frac{1}{12}[x^2]_0^m = \frac{1}{2}$

$\frac{m^2}{12} = \frac{1}{2}$

$m^2 = 6$

$m = \pm 2.45$

m = 2.45 (since it lies in the interval $0 \leq x \leq 3$)

Exercise

1. A continuous random variable X has a probability distribution function given by:

$$f(x) = \begin{cases} 2a(x+1); for. -1 \leq x \leq 0 \\ a(2-x); for.0 \leq x \leq 2 \\ 0;.elsewhere \end{cases}$$

Determine

(i) The value of a.

(ii) The median.

2. A continuous random variable has probability distribution function given by;

$$f(x)=\begin{cases}\beta;\ for.2\le x\le 3\\ \beta(x-2);\ for.3\le x\le 4\\ 0;.elsewhere\end{cases}$$

(i) Sketch the graph of f(x).
(ii) Find the value of β.
(iii) Find the median; m.
(iv) Find P(2.5 $\le$ x $\le$ 3.5) .

Lesson 4: Uniform Distribution

Competences

In this lesson, you should be able to learn how to:

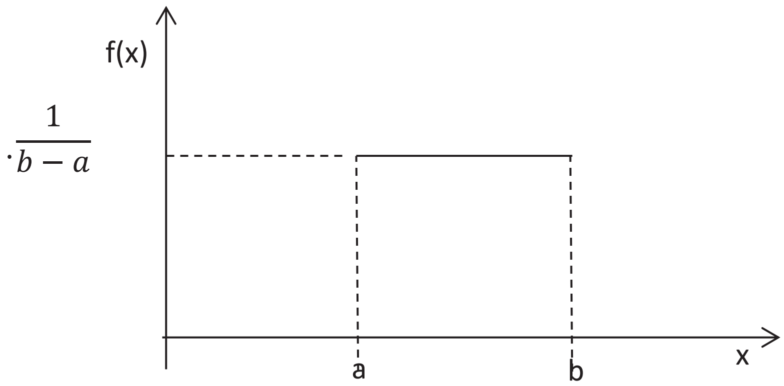
(i) define and sketch a uniform distribution.
(ii) calculate the mean and variance of a uniform distribution.

This is a continuous random variable whose density function is defined by:

$$f(x)=\begin{cases}\dfrac{1}{b-a};\ for.a\le x\le b\\ 0;.elsewhere\end{cases}$$

Where a and b are constants and a $\le$ b.

Graph of f(x) for a Uniform Distribution



Because of its shape, this function is known as a rectangular distribution.

Example

Find the mean and variance of the continuous random variable X which is uniformly distributed over the interval;

(a) 0 to 1

(b) 2 to k.

Solution

(a) Since it's a uniform distribution, then:

$$f(x) = \begin{cases} \frac{1}{b-a}; for.a \le x \le b \\ 0;.elsewhere \end{cases}$$

Where a = 0 and b = 1

Therefore,

$$f(x) = \begin{cases} 1; for.0 \le x \le 1 \\ 0;.elsewhere \end{cases}$$

E(x) $= \int_{allx} xf(x)dx$

$= \int_0^1 xdx = \left[\frac{x^2}{2}\right]_0^2 = \frac{1}{2}$

Var(x) = E(X^2)- $[E(X)]^2$

E(X^2) $= \int_{allx} xf(x)dx = \int_0^1 x^2dx = \left[\frac{x^3}{3}\right]_0^1 = \frac{1}{3}$.

Var(x) $= \frac{1}{3} - \left(\frac{1}{2}\right)^2 = \frac{1}{3} - \frac{1}{4} = \frac{1}{12}$.

(a) a = 2 and b = k

E(X) $= \int_2^k \frac{x}{k-2}dx = \frac{1}{k-2}\left[\frac{x^2}{2}\right]_2^k$

$= \frac{1}{k-2}(k^2 - 2^2)$

$= \frac{k+2}{2}$

E(X^2) $= \int_2^k \frac{x^2}{k-2}dx = \frac{1}{k-2}\left[\frac{x^3}{3}\right]_2^k$

$= \frac{1}{k-2}\left(\frac{k^3 - 2^3}{3}\right)$

$(a - b)^3 = a^3 - 3a^2b + 3ab^2 - b^3$

Thus, $\Rightarrow k^3 - 2^3 = (k - 2)^3 + 6k(k - 2)$

$$E(X^2) = \frac{k-2}{k-2}\left(\frac{(k-2)^2+6k}{3}\right)$$
$$= \left(\frac{k^2+2k+4}{3}\right)$$
$$Var(X) = \left(\frac{k^2+2k+4}{3}\right) - \left(\frac{k+2}{2}\right)^2$$
$$= \left(\frac{k^2+2k+4}{3}\right) - \frac{k^2+4k+4}{4}$$
$$= \frac{4k^2+8k+16-3k^2-12k-12}{12}$$
$$= \frac{k^2-4k+4}{12}$$
$$= \frac{(k-2)^2}{12}.$$

Note: For a uniform distribution:

$E(X) = \frac{a+b}{2}$ and $Var(X) = \frac{(b-a)^2}{12}$

Where a and b are the lower and upper limits.

Exercise

1. During rush hours it was observed that the number of vehicles departing for Entebbe from Kampala Old Taxi park is a random variable x with a uniform distribution over the interval $[x_1, x_2]$. If in one hour, the expected number of vehicles leaving the stage is 12, with variance of 3, calculate the:
   (i) values of x_1 and x_2.
   (ii) probability that at least 11 vehicles leave the stage.

2. A random variable X has a probability distribution function given by:

$$f(x) = \begin{cases} kx; for.2 \le x \le 3 \\ (k - x^2); for.3 \le x \le 4 \\ 0;.elsewhere \end{cases}$$

   (i) Find the constant k and sketch the graph of f(x).
   (ii) Determine E(X).
   (iii) Find the cumulative distribution function F(X).

NCDC
NATIONAL CURRICULUM DEVELOPMENT CENTRE

National Curriculum
Development Centre,
P.O. Box 7002,
Kampala.

www.ncdc.go.ug