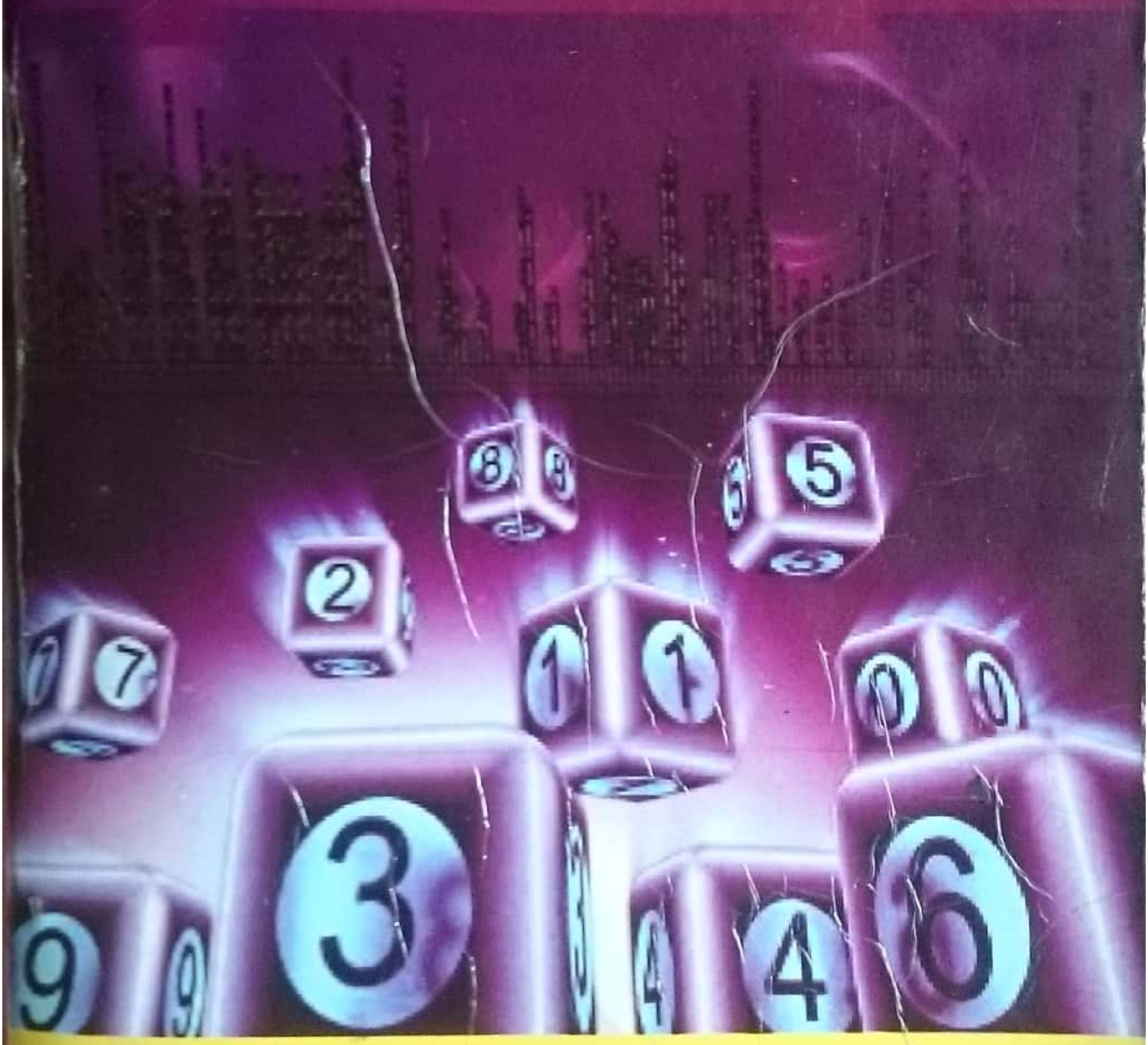
A COMPREHENSIVE APPROACH TO ADVANCED LEVEL STATISTICS AND NUMERICAL METHODS



MUKOSE MUHAMMED

THIRD EDITION

2022

A COMPREHENSIVE APPROACH TO ADVANCED LEVEL STATISTICS AND NUMERICAL METHODS

2022

THIRD EDITION 2008

M.MUKOSE BSC/ED (HONS), M.A
DEMOGRAPHY (MUK)
MATHEMATICS DEPARTMENT
NABISUNSA GIRLS' SCHOOL

CONTACTS:
TEL. NO. 0772429998 OR 0701429988
EMAIL mukosm@yahoo.com

ACKNOWLEDGEMENT

I hereby extend my sincere appreciation to the Headmistress, Nabisunsa Girls' School Hajati Aisha Lubega, the Deputy Headmistresses Hajati Sengendo and Hajati Kyagaba; the Director of studies, Mr. Miuwa Mustafa; credit also goes to my fellow staff members for their warm company and contribution towards this compilation.

To my dear parents Mr. Mutengu Abdu and Mrs. Sarah. Mutengu whose tireless efforts in educating me have reaped these results.

I offer my gratitude. Heartfelt thanks too, go to my dear collegues Mrs. Ssebowa . A., Hajati Matovu Sarah, Mr. Wamala Mohammed, Mr. Nasimolo. P, Mr. Kaddu Juma, Sheikh. Ssali Ebraheem, Mr. Kiganda Abdu, Mr. Kasoma Abdu, Sheikh Kalema. H, Mr. Kaggwa Ronald, Mr. Mbajja. B, Mr. Ssemujju J, Mr. Buyinza (Buddo), Mr. Isanga (Kiira) Mr. Otigo E(Headmaster), Mr. Kigozi.K, Mr. Kakaire.M. Mr. Turinawe. J, Mr. Kateregga.K, Mr. Iwalwa.F, Mr. Ddungu R, Mr. Sairo , Mr Ssendawula.I, Mr. Nabiso, Mr. Kigozi. W, Ms. Mutonyi J, Mrs. Ebal J, Twaha. S and Sooka Karim for all the invaluable assistance given.

Mr. Heri Muhamed, Ms. Fatuma Kirunda, Mrs Juliet Mugalu and Bukenya Grace are credited with typing and type setting the manuscript.

My family and Uncle Nsambu R for financial assistance and time. Without the useful contribution to the compilation of this book by both my past and present Mathematics students, what a formidable task it would have been!

MUKOSE MUHAMMED
MATHEMATICS DEPARTMENT
NABISUNSA GIRLS' SCHOOL
P.O. BOX 3027 KAMPALA

INTRODUCTION

2022

I have written this book purposely to meet the needs for students offering Mathematics at advanced level. This book is based on the UACE syllabus for paper two which include Statistics, Mechanics and Numerical methods. However, in this book Statistics and Numerical Methods have been catered for. The topics there in have been comprehensively covered to cater for the increasing demand by most schools to cover the different papers concurrently. Accordingly, topical issues such as combinations and permutations have been looked at to help in the student's understanding for probability and binomial distribution.

Bearing in mind that students will find it useful as study aid, I designed exercises at the end of each topic for discussion, to stimulate the student's understanding of the subject matter. I hope students offering the following subject – combinations MEG, PEM, PCM, BCM, PCB, PCB/M will find this book extremely useful. So will those undergrates in higher institutions of learning.

As for the teachers, this book is designed to serve as a key teaching – aid.

CONTENT

CHAPTER ONE

DESCRIPTIVE STATISTICS

1.1. Introduction 1
1.2. Presentation of Data 4
1.3. Measures of Central Location........................ 22
1.4. Measures of Disperson 26
1.5. Percentiles, Deciles and Quartiles 37
Exercise... 50
1.6. Index Numbers 54
1.7. Permutation and Combination 60

CHAPTER TWO

PROBABILITY THEORY

2.1 Introduction 87
2.2 Interaction with Set Theory 90
2.3 The Contingency Table 93
2.4 Probability Situations.............................. 96
2.5 Independent Events 101
2.6 Conditional Probability 104
Exercise 108
2.7 Probability Tree Diagrams.......................... 110
2.8 Bayes' Tree Diagrams.............................. 118

2022

2.9 Calculations involving Combination & Probability.... 121

2.10 Other examples 123

Exercise 127

CHAPTER THREE

DISCRETE PROBABILITY DISTRIBUTION

3.1 Introduction 135

3.2 Mean or Expected value 139

3.3 Variance 140

3.4 Properties of Mean 145

3.5 Properties of Variance 146

3.6 Mode 150

3.7 Cumulative Mass Function 153

3.8 Median 155

3.9 Graph of f(x) 155

3.10 Graph of F(x) 155

Exercise 158

CHAPTER FOUR

BINOMAL DISTRIBUTION

4.1 Introduction 164

4.2 Properties of Binomial experiment 165

4.3 Formula for Binomial 165

4.4 Format for the Table of sums 172

4.5 Expectation and variance 175

4.6 The most likely value of X 179

Exercise 181

CHAPTER FIVE

CONTINUOUS PROBABILITY DENSITY FUNCTION

5.1 Introduction 184

5.2 How to Obtain probabilities 184

5.3 Graph of f(x) 186

5.4 How to obtain f(x) from the Sketch 195

5.5 Mean and Variance 198

5.6 The median 203

5.7 Mode 206

5.8 The cumulative Distribution Function F(x)............ 207

5.9 Probabities, Mean and p.d.f (fx) from F(x) 211

5.10 Uniform Distribution 219

Exercise 223

CHAPTER SIX

THE NORMAL DISTRIBUTION

6.1 Introduction 238

6.2 Properties of the Normal Curve 239

6.3 Standardisation of Normal Distribution 239

6.4 How to Read Tables.. 240

6.5 How to Obtain Z-values 247

6.6 Distribution of Sample Mean 253

Exercise 255

6.7 Normal Approximation to Binomal Distribution ... 260

Exercise 265

CHAPTER SEVEN

ESTIMATION

7.1 Introduction 267

7.2 Point Estimation 267

7.3 Interval Estimation 267

7.4 Confidence Interval 268

7.5 Confidence Coefficient 268

7.6 Estimation of Mean 268

CHAPTER EIGHT

SCATTER GRAPHS AND CORRELATION

8.1 Introduction – Scatter graphs............................ 275

8.2 Regression Line 276

8.3 Correlation 279

8.4 Interpretation of magnitude of correlation coefficient. 281

8.5 Spearman's Rank Correlation Coefficient 281

8.6 Kendall's Rank correlation 285

Exercise 289

CHAPTER NINE

ERRORS

9.1 Introduction 292

9.2 Types of Errors 292

9.3 Limits of Accuracy 296

9.4 Properties of Rounding Errors 301

Exercise 314

CHAPTER TEN

LINEAR INTERPOLATION

10.1 Linear Interpolation 318

10.2 Expression of Linear Interpolation.................. 319

10.3 Inverse Interpolation 323

10.4 Extrapolation 325

CHAPTER ELEVEN

APPROXIMATE NUMERICAL METHODS

11.1 Location of roots f(x) = 0 328

11.2 Approximations / Iterations 334

11.2.1 Interpolation.. 334

11.2.2 Newton Raphson Method 336

11.2.3 General Iterative Method 343

11.2.4 Test for Convergence 346

11.2.5 Trapezium rule 348

Exercise 352

CHAPTER TWELVE

FLOW CHARTS

12.1 Introduction 356

12.2 Flow charts and Computer based programs...... 359

12.3 Dry Run 361

12.4 Constructing Flow Charts 370

Exercise 387

FORMULAE 392

STATISTICAL TABLES 395

CHAPTER ONE

DESCRIPTIVE STATISTICS

1.1 INTRODUCTION

This is the branch of mathematics dealing with collection, interpretation, presentation and analysis of data. Data refers to facts and figures collected for a purpose.

STATISTICAL DATA

Data can be categorized into two, namely Qualitative and Quantitative.

Qualitative data measures attributes such as sex, colour and so on. While quantitative data can be represented by numerical quantity. For instance, height, mass, time and so on. In this part of statistics it is only the quantitative data that will be considered. Quantitative data is mainly of two forms either continuous or discrete.

(a) Discrete and Continuous data

Statistical data can be of two forms; Discrete or Continuous. Information collected by counting is discrete and usually takes integral values e.g. numbers of students in a class, school, etc.. Continuous data can take on any value, for instance height, weight, time, distance, mass, temperature.

The quantity, which is counted or measured, is called the variable. Other terms used include

(a) Crude and Classified data

Crude data are individual values of variable in no particular order of magnitude, written down as they occurred or were measured. When the numbers have been arranged in order and grouped in a small number of classes, they are called classified or grouped data

(b) Population and Sample

A population is the total set of items under consideration and is defined by some characteristics of these items. A sample is a finite subset of a population.

Frequency

This refers to the number of times an item occurs. In most cases many items occur more than once, therefore a frequency distribution table is used to remove repetitions.

Example

The raw data gives the marks of 10 students in class.

35, 40, 45, 50, 60, 35, 40, 35, 45, 70

Represent the above data using frequency table

A frequency table can be used to represent the data

marks	35	40	45	50	60	70
frequency	3	2	2	1	1	1

This kind of representation is commonly used for ungrouped data. However if the data is large then grouping is necessary.

Grouped data

Weight (kg)	Number of students
40 - 44	3
45 - 49	10
50 - 54	16
55 - 59	10
60 - 64	4

The first column represents the classes. So 40 – 44 is a class. The second column represents the class frequencies. For any given class (50 – 54), the lower value is called lower class limit and upper value is upper class limit.

Class boundaries are the true class limits. For the class 60 -64, the lower class boundary is 59.5 and the upper is 64.5.

Note that the class boundaries depend on degree of accuracy.

For example (7.0 – 7.4). Lower class boundary is 6.95 and upper class boundary is 7.45.

Class width is difference between upper class boundary and lower class boundary.

For 50 – 54. The class width = 54.5 – 49.5 = 5

For 7.0 – 7.4. The class width = 7.45 – 6.95 = 0.5

Note: that the class width , size and length are the same.

Class mark is the average of lower and upper limit

Example (40 – 44)

Class mark = (40 + 44) /2 = 42

Note: Tallies are obtained when forming the classes.

1.2 PRESENTATION OF DATA

Ways of presenting data include

i. Bar graphs
ii. Histogram
iii. Frequency polygon
iv. The Ogive
v. Pie chart

BAR GRAPH

A bar graph or bar chart is a graph where the class frequencies are plotted against class limits.

HISTOGRAM

A histogram is a graph where the class frequencies are plotted against class boundaries.

Example

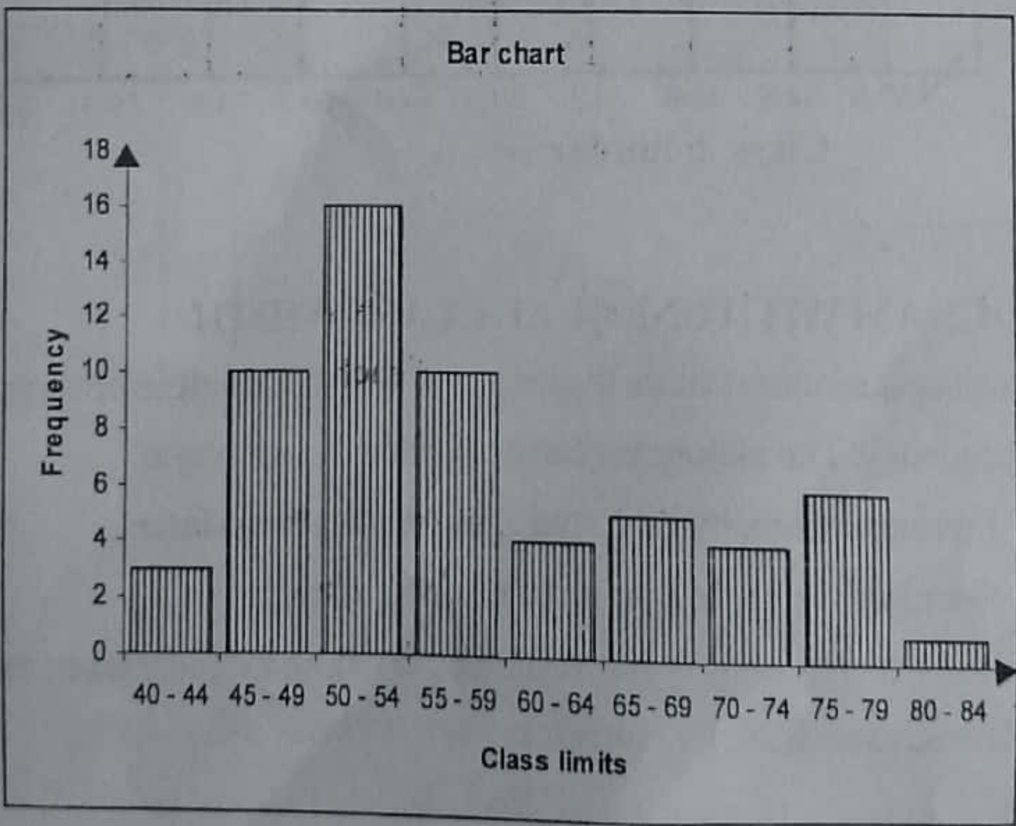
The table below shows the weights of some freshers in 2000/2001 academic year who under went medical examination at the sick bay.

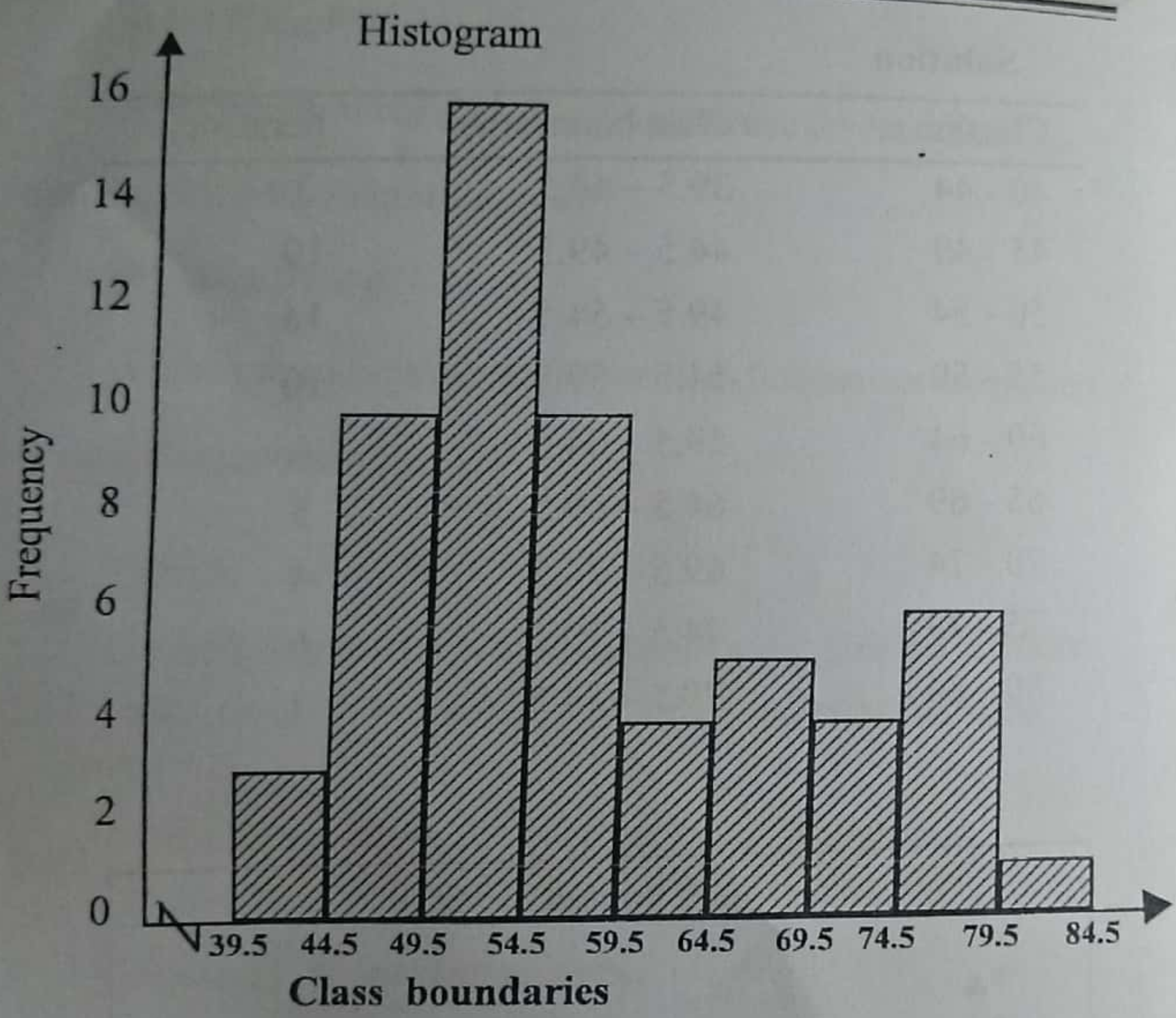
Use the data given to plot a bar graph and histogram.

Weight (kg)	Number of students
40 – 44	3
45 – 49	10
50 – 54	16
55 – 59	10
60 – 64	4
65 – 69	5
70 – 74	4
75 – 79	6
80 - 84	1

Solution

Class limits	Class boundaries	frequency
40 - 44	39.5 – 44.5	3
45 - 49	44.5 – 49.5	10
50 - 54	49.5 – 54.5	16
55 - 59	54.5 – 59.5	10
60 - 64	59.5 – 64.5	4
65 - 69	64.5 – 69.5	5
70 - 74	69.5 – 74.5	4
75 - 79	74.5 – 79.5	6
80 - 84	79.5 – 84.5	1





HISTOGRAM WITH UNEQUAL CLASS WIDTH

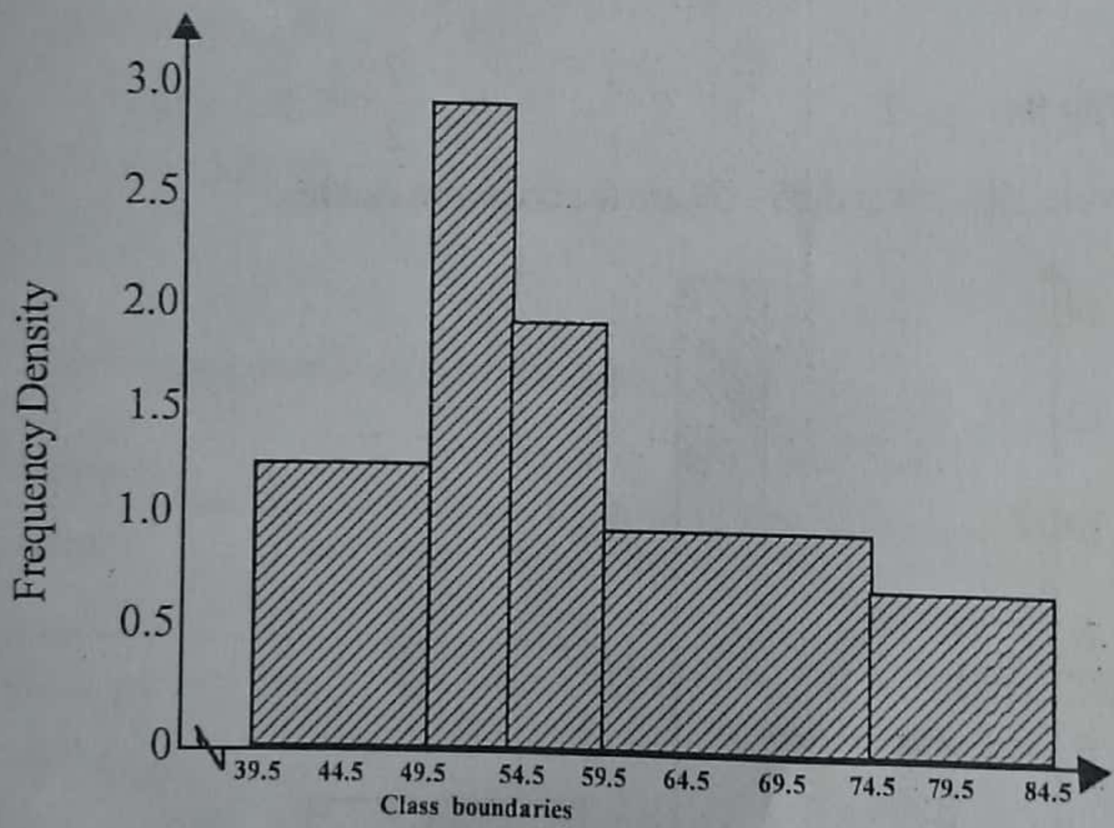
In case of unequal class width, the height of each rectangle is not equal to the frequency. The histogram can be plotted in two ways:

- Frequency density is plotted against class boundaries.
- Standard frequency against class boundaries.

Where the frequency density is obtained by dividing frequencies of that class – by – class width.

Example

Class boundaries	Frequency	Frequency density
39.5 – 49.5	13	1.3
49.5 – 54.5	15	3.0
54.5 – 59.5	10	2.0
59.5 – 74.5	15	1.0
74.5 – 84.5	7	0.7

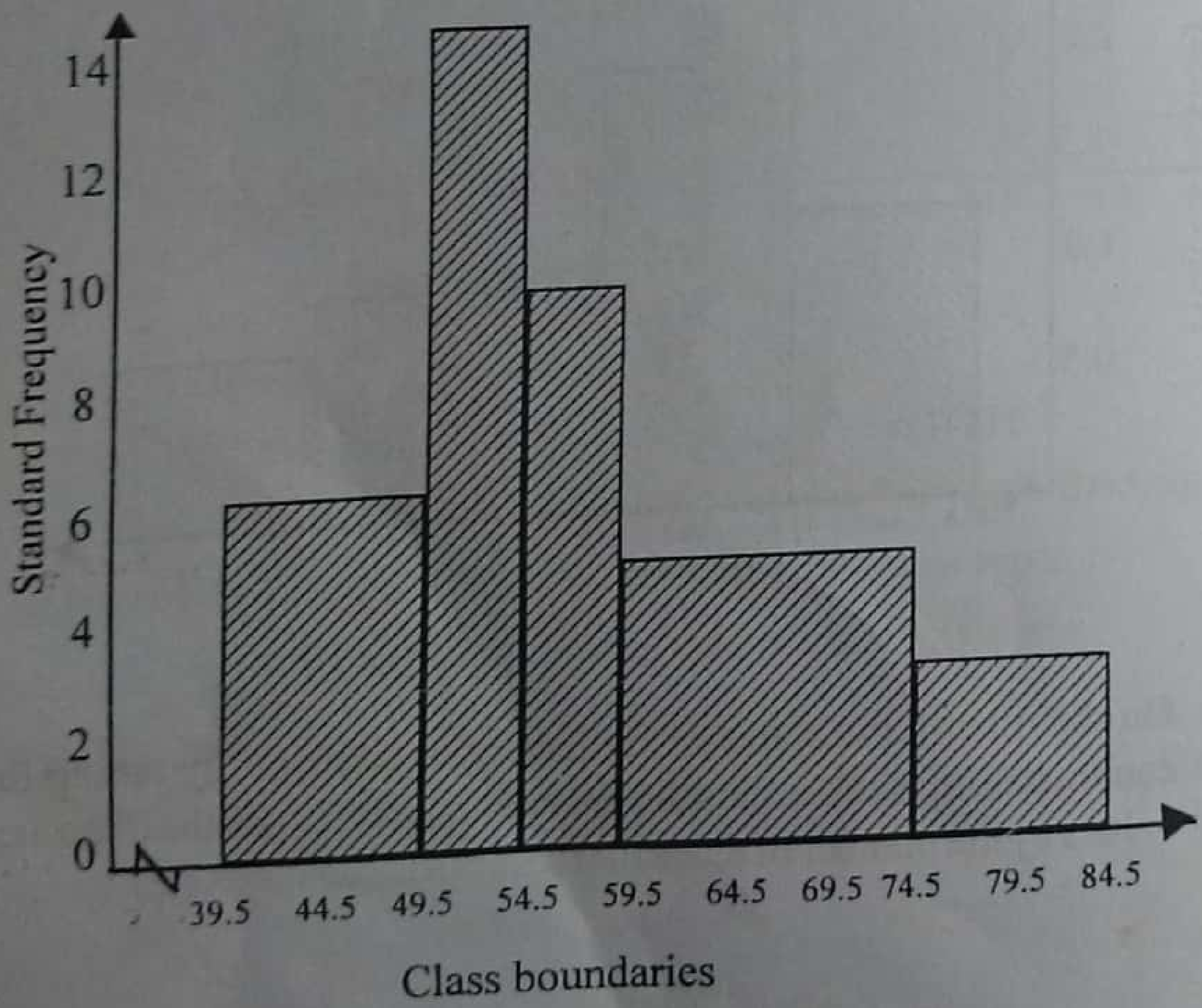


On the other hand standard frequency is obtained by taking the common class width as the standard one and then divide the frequency values by the number of times the class is the standard one.

Example

Class limits	No. of times	Frequency	*Standard frequency*
40 - 49	2	13	6.5
50 - 54	1	15	$\frac{15}{1}=15$
55 - 59	1	10	10
60 - 74	3	15	$\frac{15}{3}=5$
75 - 84	2	7	$\frac{7}{2}=3.5$

Note: 50 – 54 and 55 – 59 are the common classes.



Example

The table below shows the population of Cairo in millions for the different age groups.

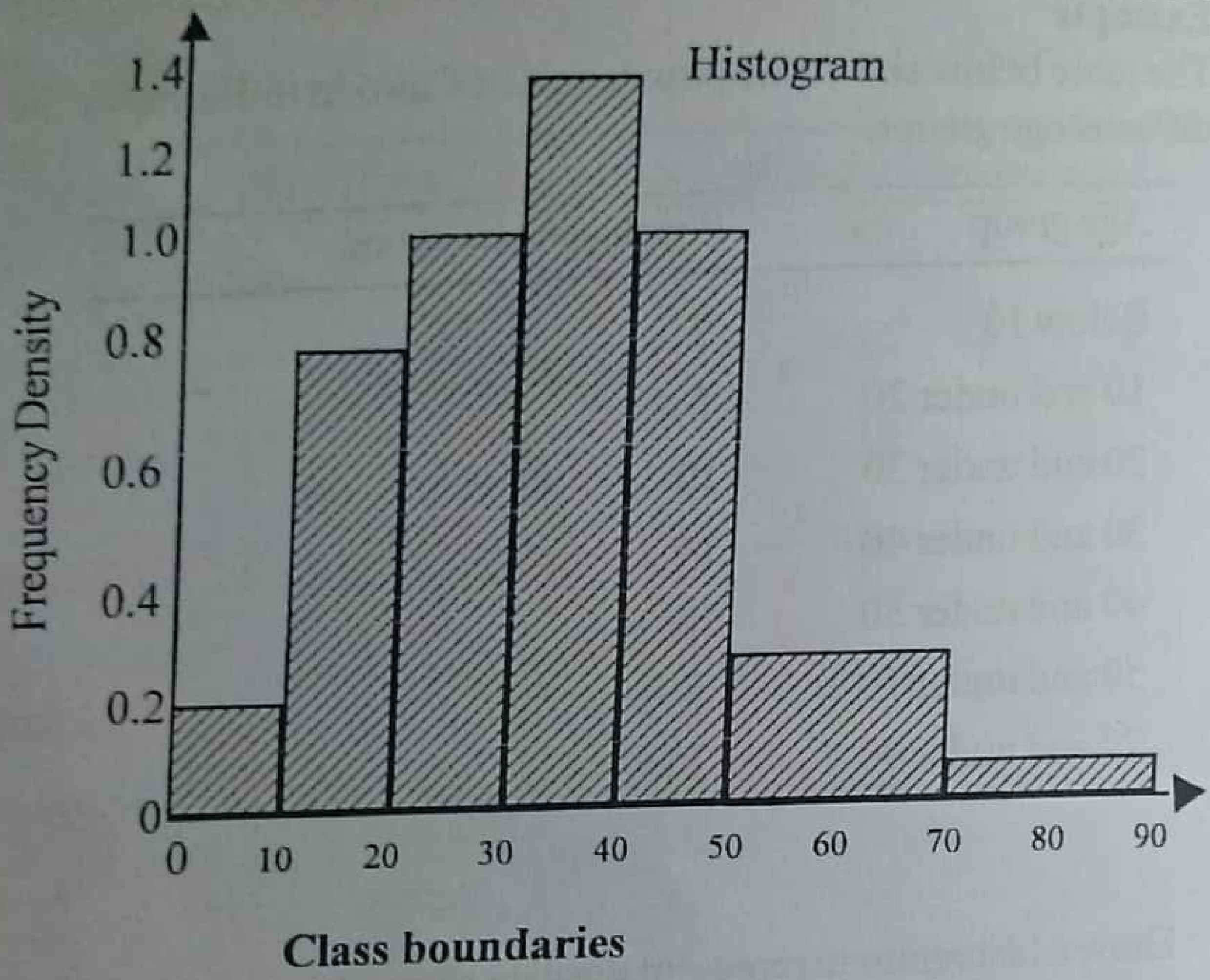
Age group	Population in millions
Below 10	2
10 and under 20	8
20 and under 30	10
30 and under 40	14
40 and under 50	10
50 and under 70	5
70 and under 90	1

Draw a histogram to represent the data above

Solution

Class	Class width	Frequency	***Frequency density***
0 - < 10	10	2	0.2
10 - < 20	10	8	0.8
20 - < 30	10	10	1.0
30 - < 40	10	14	1.4
40 - < 50	10	10	1.0
50 - < 70	20	5	0.25
70 - < 90	20	1	0. 05

70



Assignment: Repeat the above example using standard frequency instead of frequency density.

Example 1

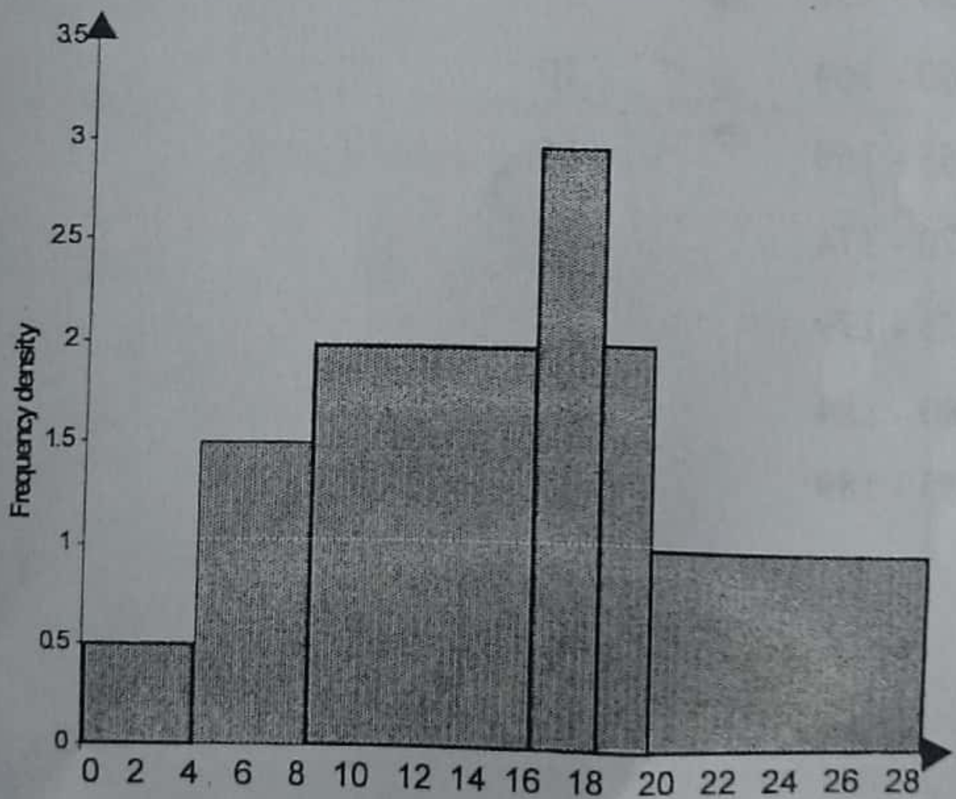
On a particular day the length of stay of each taxi at the New taxi park was recorded as shown in the table below

Length of stay (min)	Frequency
$0 \leq t < 4$	2
$4 \leq t < 8$	6
$8 \leq t < 12$	8
$12 \leq t < 16$	8
$16 \leq t < 18$	6
$18 \leq t < 20$	4
$20 \leq t < 28$	8

Draw a histogram for the data above

Solution

Class	Class width	Frequency	*Frequency density*
0 - < 4	4	2	0.5
4 - < 8	4	6	1.5
8 - < 12	4	8	2.0
12 - < 16	4	8	2.0
16 - < 18	2	6	3.0
18 - < 20	2	4	2.0
20 - < 28	8	8	1.0



Frequency Polygon

A frequency polygon is obtained by plotting class frequencies versus class marks. Then consecutive points are joined using a straight line. *Note*: Additional classes may be added to either end with zero frequency to make the polygon a closed figure.

Example

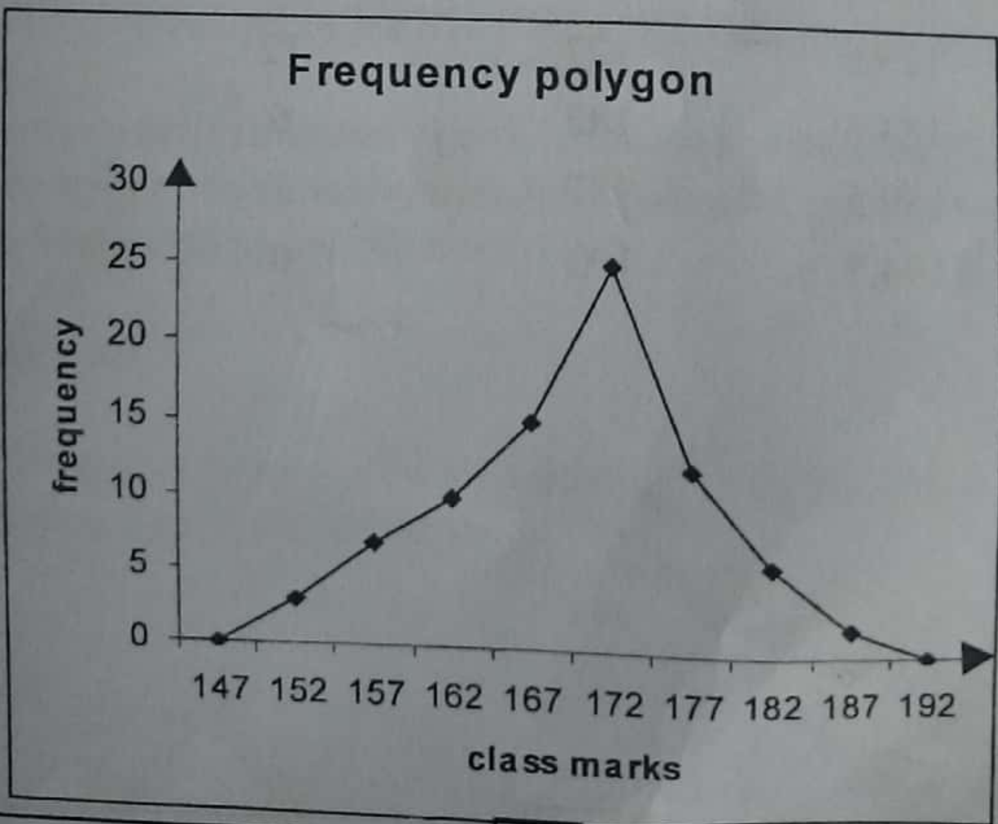
The frequency distribution below shows the heights to the nearest cm, of 80 patients at a Hospital.

Height in cm	frequency
150 - 154	3
155 - 159	7
160 - 164	10
165 - 169	15
170 - 174	25
175 - 179	12
180 - 184	6
185 - 189	2

Construct a frequency polygon for the above data?

Solution

Class limits	Class mark	frequency
Additional class	147	0
150 - 154	152	3
155 - 159	157	7
160 – 164	162	10
165 – 169	167	15
170 – 174	172	25
175 – 179	177	12
180 – 184	182	6
185 – 189	187	2
190 – 194	192	0



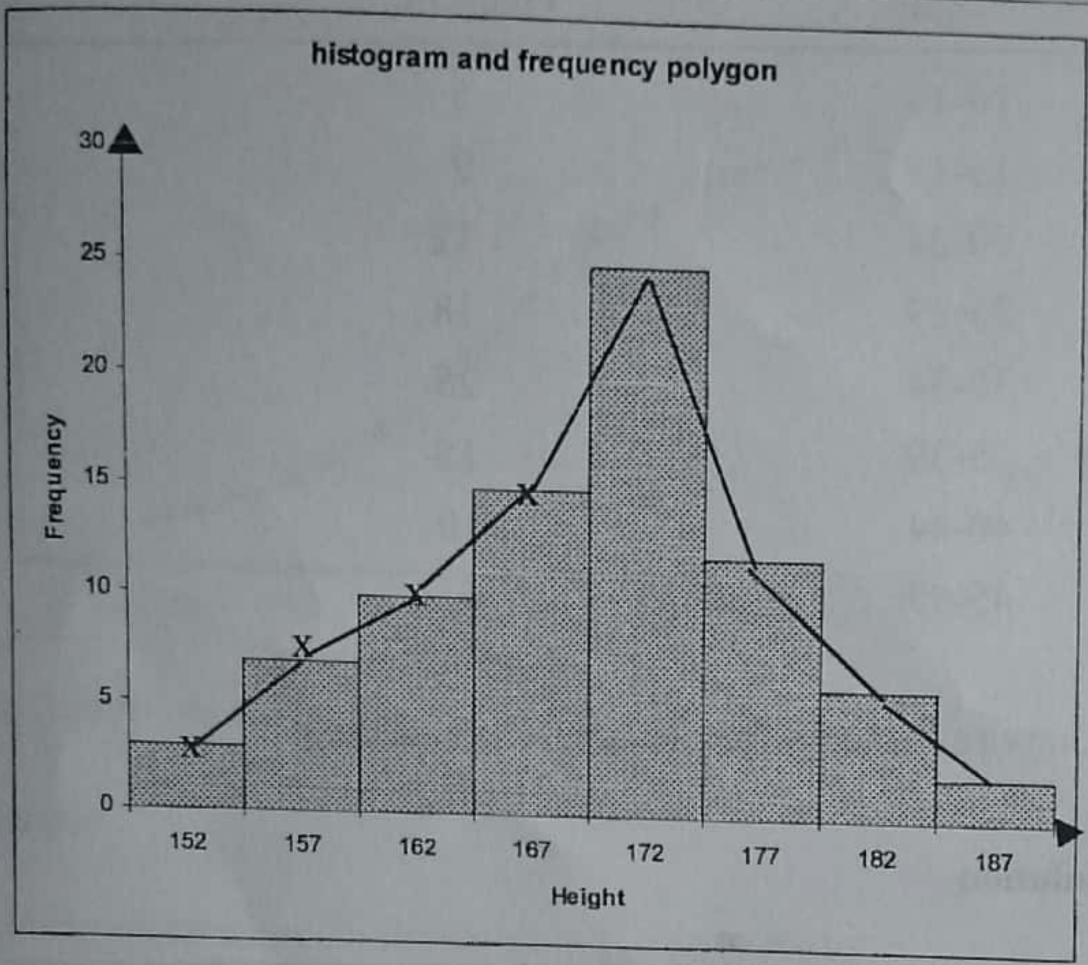
Assignment: Using example 1
Draw a frequency polygon for the data

***Note*:** A frequency polygon can be put on the same graph with the histogram. This is known as superimposing a frequency polygon.

Example

Using example below, draw a histogram and superimpose a frequency polygon.

Class boundaries	*Class mark*	*frequency*
144.5 – 149.5	147	0
149.5 – 154.5	152	3
154.5 – 159.5	157	7
159.5 – 164.5	162	10
164.5 – 169.5	167	15
169.5 – 174.5	172	25
174.5 – 179.5	177	12
179.5 – 184.5	182	6
184.5 – 189.5	187	2
189.5 – 194.5	192	0



THE CUMULATIVE FREQUENCY CURVE (OGIVE)

The cumulative frequency curve is obtained by plotting cumulative frequency versus class boundaries and then the consecutive points are joined using a smooth curve.

Example 2

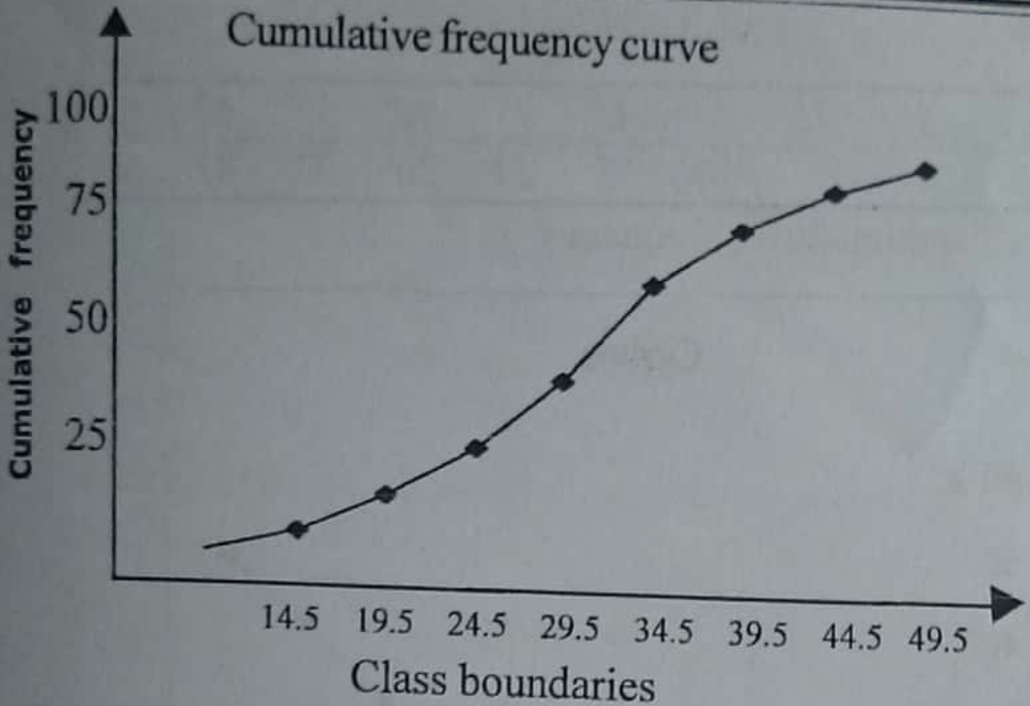
The frequency distribution table shows the weights of 100 children measured to the nearest kg.

Weight	Frequency
10-14	5
15-19	9
20-24	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



Note that points correspond i.e (9.5, 0), (14.5, 5) e.t.c

Example

The table below shows the earnings to nearest dollar of Employees of a certain farm

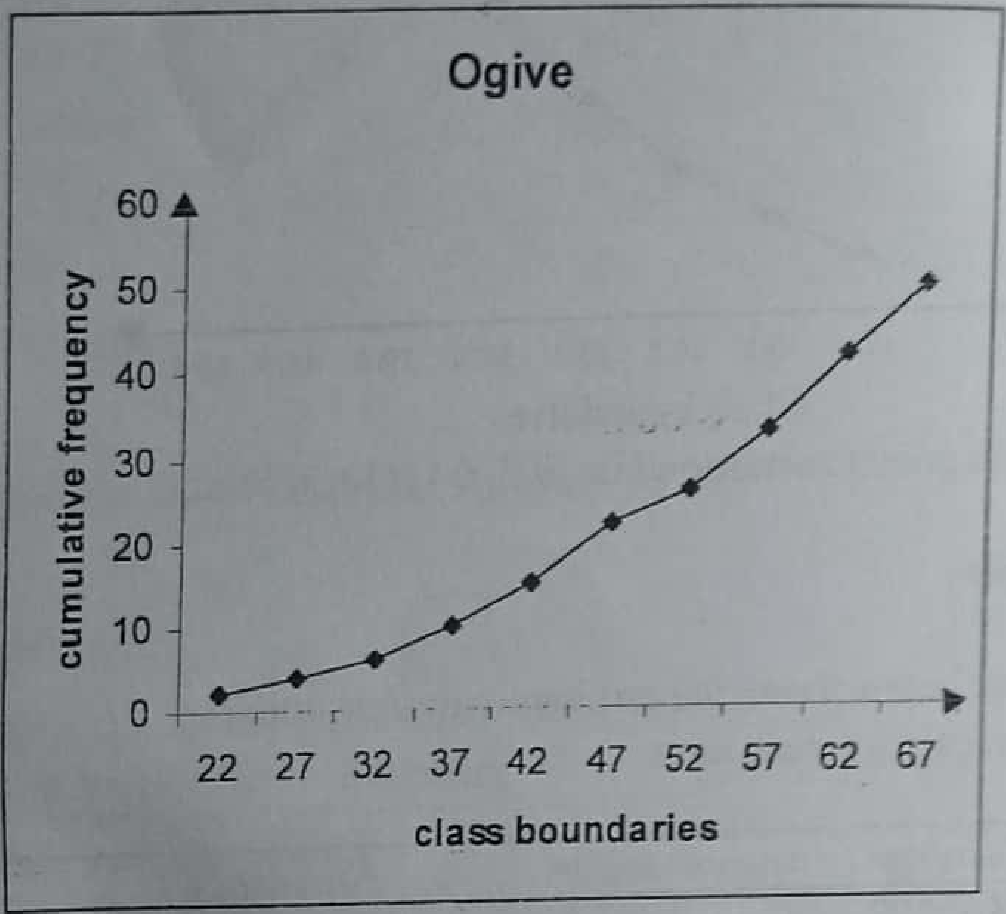
Earnings to nearest dollar	*Frequency*
18 - < 22	2
22 - < 27	2
27 - < 32	2
32 - < 37	4
37 - < 42	5
42 - < 47	7
47 - < 52	4
52 - < 57	7
57 - < 62	9
62 - < 67	8

Draw an Ogive for the data

Solution

	22	27	32	37	42	47	52	57	62	67
Cf	2	4	6	10	15	22	26	33	42	50

Note: Cf is cumulative frequency



Example

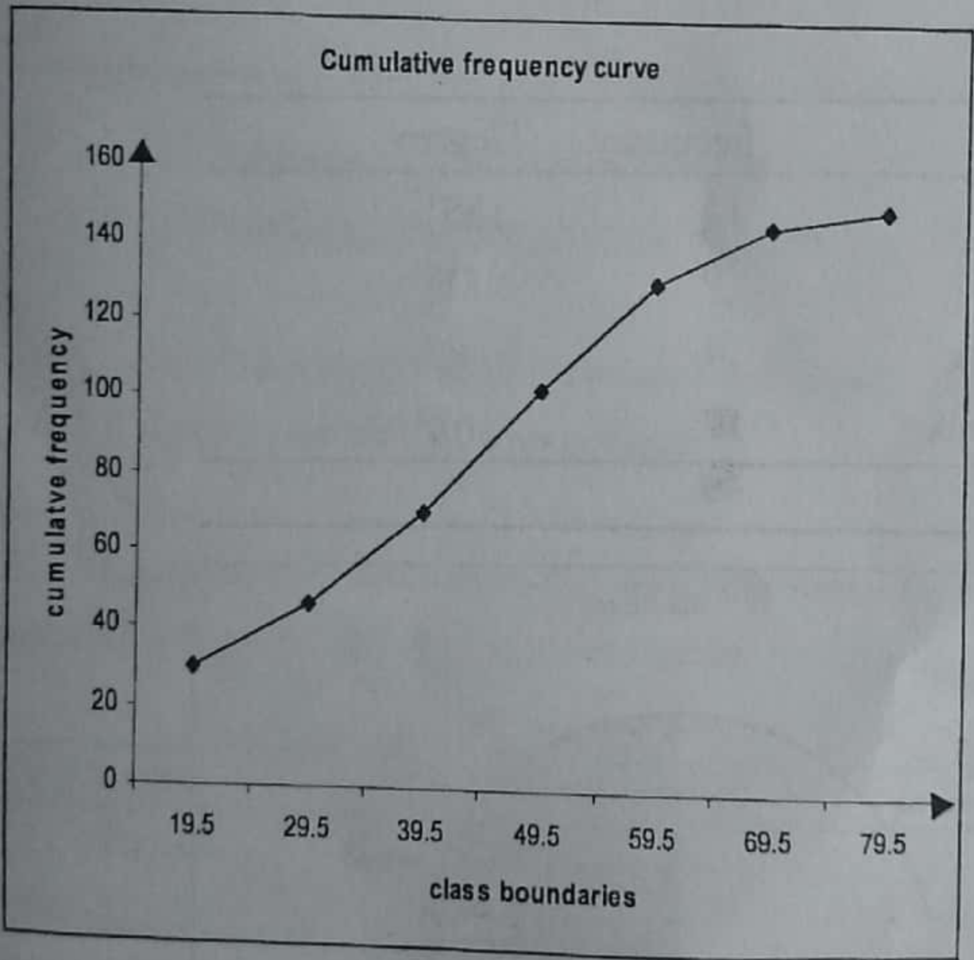
The table below shows the weights to the nearest kg of 150 patients who visited a certain health unit during a certain week

Weight (kg)	*Number of patients (f)*
0 - 19	30
20 - 29	16
30 - 39	24
40 - 49	32
50 - 59	28
60 - 79	20

Draw a cumulative frequency curve for the data

Solution

Class boundaries	*frequency*	*Cumulative frequency*
19.5	30	30
29.5	16	46
39.5	24	70
49.5	32	102
59.5	28	130
79.5	20	150



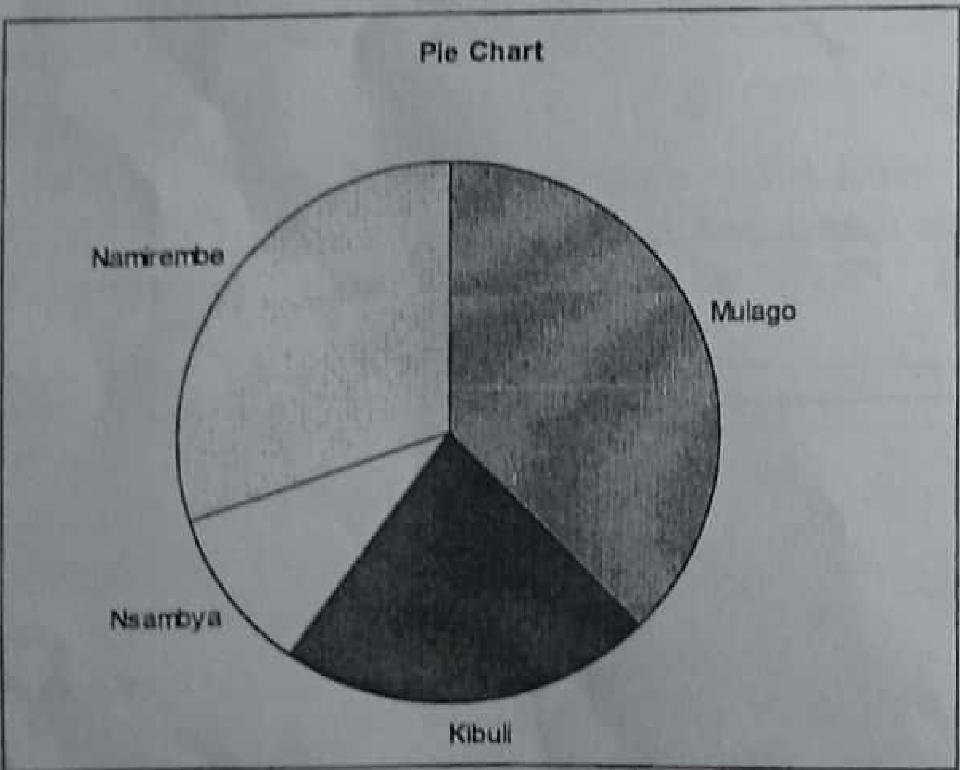
Pie Chart

Draw a pie chart for the information given below

Hospital	Number of children
Mulago	15
Kibuli	9
Namirembe	12
Nsambya	4

Solution

Hospital	frequency	Degree
Mulago	15	135^0
Kibuli	9	81^0
Nsambya	4	36^0
Namirembe	12	108^0
Total	40	



Exercise

1.The table below shows the expenditure of examiners in thousands of shillings during the year 1990 ordinary level marking exercise.

10	11	10	12	14	16	20	25
21	22	13	17	18	24	30	32
27	35	40	44	39	50	54	53
44	37	36	39	52	51	57	15
16	19	34	43	26	38	53	40

i. Form a frequency distribution table with class interval of 5,000 shillings, the lowest class limit being 10,000.

ii. Draw a histogram to represent the above data and superimpose a frequency polygon.

2. The heights of students in S.5 Mathematics class were according to the following frequency table

Height cm	*Frequency (f)*
151 - 153	2
154 - 156	14
157 - 159	13
160 - 162	13
163 – 165	2
166 - 168	1

Plot the cumulative frequency curve for the students' heights. Hence estimate the median, lower and upper quartiles for the heights of the students.

Answer median = 158.3 cm, lower quartile = 155.8 cm,
upper quartile = 160.7 cm

3.The information shows the amount of money in millions (A) given to some districts in Uganda for 'Entandikwa" scheme

A	25-<30	30 -<40	40-<50	50-<60	60 - < 80
f	4	10	4	3	5

i. Use the above information to draw a histogram

ii. Use the histogram to estimate the mode of the distribution.

4. The table below shows the weights in kilograms of 250 boys. Each weight was recorded to nearest 100 grams

weight kgs	*Frequency (f)*
44.0 – 47.9	3
48.0 – 51.9	17
52.0 – 55.9	50
56.0 – 59.9	45
60.0 – 63.9	46
64.0 – 67.9	57
68.0 – 71.9	23
72.0 – 75.9	9

Draw an Ogive and use it to estimate

i. The semi-interquartile range
ii. The second decile
iii. The percentage of boys weighing over 59 kilograms

Answer (i) 5 (ii) 54.55 (iii) 59%

5. At 7.30 am daily a bus leaves Kampala for Jinja. The times (minutes) taken to cover the journey were recorded over a certain period of time and were grouped as shown in the table below.

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 Jinja by bus.

b) Draw a cumulative frequency curve for the data. Use your curve to estimate the:

i. Median time for the journey,

i. Number of times the bus arrived in Jinja between 9.00 – 9.25 am.

ii. Semi-interquartile range of time of travel from Kampala to Jinja.

Answer

(a) mean time = 96.6 minutes

(b) (i) median (96 – 97) minutes

(ii) (118 - 122) times

(iii) (4.5 – 5.5) minutes.

6. The table below is the distribution of weights of a group of animals.

Mass (kg)	*Frequency*
21- 25	10
26 – 30	20
31 - 35	15
36 - 40	10
41 -50	30
51 - 65	45
66 - 75	5

(a) Draw a cumulative frequency curve and use it to estimate the semi-interquartile range.

(b) Find the mode.

1.3 MEASURES OF CENTRAL LOCATION

These are average values that locate values of a variable in a particular part of the number line.
They include: Mean, Median and Mode

Mean or Average

The mean is the sum of the observations divided by the number of items

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

Where $\sum$ sigma, is used for sum.

Example
Determine the average height of the class using the data below
171, 180, 154, 160, 155 and 164

Solution

$$\text{Mean} = \frac{171+180+154+160+155+164}{6} = 164$$

If the class had many students then, some might have the same height. So in this case frequency table is needed

Example

Height	154	155	160	164	171	180
Frequency	4	6	8	5	4	3

Solution

$$\text{Mean} = \bar{X} = \frac{\sum fx}{\sum f}$$

Height (x)	*Frequency (f)*	*fx*
154	4	616
155	6	930
160	8	1280
164	5	820
171	4	684
180	3	540
Total	**30**	**4870**

$$\text{Mean} = \frac{4870}{30} = 162.3333$$

Note: depending on the nature of data, a suitable formula should be used.

Example

The table below shows the weight of some exotic cows in kilograms.

Example

The table below shows the weight of some exotic cows in kilograms.

weight	990	1000	1010	1020	970	980
frequency	5	6	3	2	20	10

Calculate the average weight for the cows

Solution

$$\bar{X} = A + \frac{\sum fd}{\sum f}$$

Let A the working mean be 1000 and d = x - A

Weight (x)	*frequency*	*d*	*fd*
970	20	970 - 1000 = -30	-600
980	10	-20	-200
990	5	-10	-50
1000	6	0	0
1010	3	10	30
1020	2	20	40
Total	**46**		**-780**

$$\bar{X} = 1000 + \frac{-780}{46} = 983.04kgs$$

Mean = 983.04 kgs.

Mean for grouped data

Example

The table shows the frequency distribution of the rate of unemployment in Kampala in 1990.

Unemployment rate	***Frequency***
7.0 – 7.4	2
7.5 – 7.9	4
8.0 – 8.4	5
8.5 – 8.9	4
9.0 – 9.4	3
9.5 – 9.9	2

Calculate the mean of unemployment rate

Solution

Unemployment (x)	*Frequency (f)*	*fx*
7.2	2	14.4
7.7	4	30.8
8.2	5	41.0
8.7	4	34.8
9.2	3	27.6
9.7	2	19.4
Total	**20**	**168.0**

$$\overline{X} = \frac{\sum fx}{\sum f}$$

$$\text{Mean} = \frac{168}{20} = 8.4$$

So mean unemployment rate is 8.4

Use of working mean

Example 3

The table below show the weight of 250 patients at kibúli Hospital

Weight (kgs)	Frequency
44.0 – 47.9	3
48.0 – 51.9	17
52.0 – 55.9	50
56.0 – 59.9	45
60.0 – 63.9	46
64.0 – 67.9	57
68.0 – 71.9	23
72.0 – 75. 9	9

Find the average weight of Patients

Solution

A suitable working mean is selected from any of the middle values of the class marks.

Let A = 57.95

Weight	x	f	d	fd
44.0 – 47.9	45.95	3	-12	-36
48.0 – 51.9	49.95	17	-8	-136
52.0 – 55.9	53.95	50	-4	-200
56.0 – 59.9	57.95	45	0	0
60.0 – 63.9	61.95	46	4	184
64.0 – 67.9	65.95	57	8	456
68.0 – 71. 9	69.95	23	12	276
72.0 – 75. 9	73.95	9	16	144
Total		**250**		**688**

$$\overline{X} = 57.95 + \frac{688}{250} = 60.702kgs$$

Advantages of using the mean include

1) All values are used in the calculation

2) It is easy to understand and calculate

Disadvantages of using mean

One or two very low or high values can distort the information

(II) MEDIAN

The median of a set of numbers is the middle one when they are arranged in order of magnitude. If the number of observations is even, then the median is the average of the two middle values.

Example

Determine the median for given set of data?

(i) 5, 10, 10, 6, 6, 9, 10, 7, 7

(ii) 5,5,5, 9,9,8,9,6,6,8

Solution

Arrange in either ascending or descending order

(i) 5, 6, 6, 7, 7, 9, 10, 10, 10.

Therefore middle value is 7

Median = 7

(ii) 5, 5, 5, 6, 6, 8, 8, 9, 9, 9

The middle values are 6 and 8

$$\text{Median} = \frac{6+8}{2} = 7$$

Therefore the median is 7.

For the grouped data the formula below is used

$$\text{Median} = L_1 + \frac{\left(\frac{N}{2} - F_b\right)C}{f_m}$$

Where

L_1 = Lower class boundary of median class

N = Total number of observations

Fb = Cumulative frequency before the median class

f_m = Frequency of the median class

C = Class width.

Example

Using example 3

Find the median weight of Patients

Solution

$$\text{Median} = L_1 + \frac{(\frac{N}{2} - F_b)C}{f_m}$$

Weight (kgs)	*f*	*Cumulative frequency*
44.0 – 47.9	3	3
48.0 – 51.9	17	20
52.0 – 55.9	50	70
56.0 – 59.9	45	115
60.0 – 63.9	46	161
64.0 – 67.9	57	218
68.0 – 71.9	23	241
72.0 – 75. 9	9	250

Median class is obtained from cumulative frequency just above or

equal to $\frac{250}{2} = 125$. Note: $\frac{N}{2}$

Median class = 60.0 – 63.9

$$\text{Median} = 59.95 + \frac{(125-115)}{46} x4$$

$$= 60.82 \text{ kgs}$$

Advantages of median include

1. It is not affected by extreme values
2. It is easy to understand and calculate

Disadvantages of using median

Is that it is only one or two values that decide the median

Note: that you can use the Ogive to determine the median

(III) Mode

The mode is the value that occurs most frequently in given set of data

Find the mode
3, 3, 3, 4, 4, 6, 6, 7, 7, 7, 7
Since 7 appears four times
Therefore the mode = 7

For grouped data the mode is given by

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

Where

L_1 = Lower class boundary of modal class

Δ_1 = Difference between highest frequency and value before it

Δ_2 = Difference between highest frequency and value after it

C = Class width.

Example

Using example 3

Calculate the mode for data

Solution

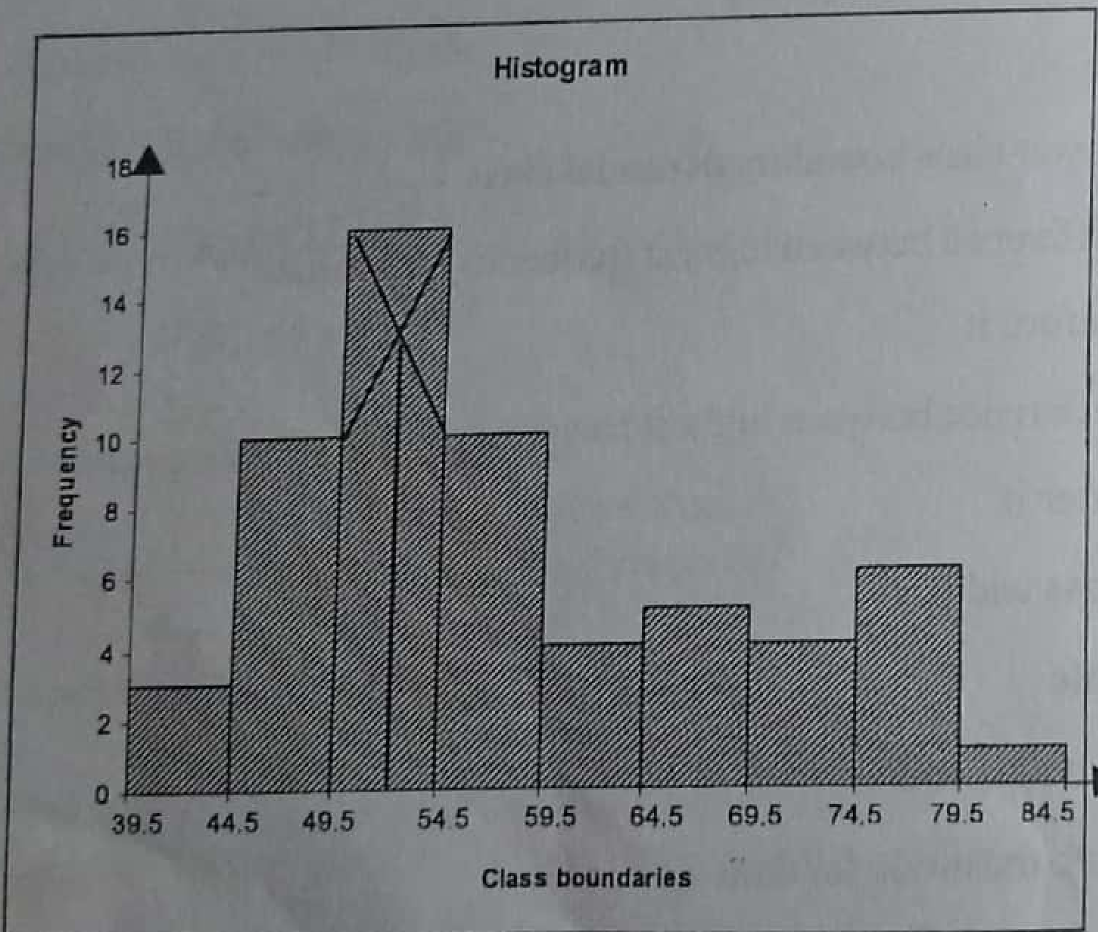
Modal class = 64.0 – 67.9 since 57 is the highest frequency.

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

$$\text{Mode} = 63.95 + \frac{11x4}{11+34}$$

Mode = 64.93 kgs

Note: The mode can be obtained from the histogram. The modal class is identified, the class with highest frequency or frequency density. The position of the mode within the modal class can be estimated from the histogram as shown in the figure below.



Advantage of mode

Very easy to understand and calculate

Example

The age of people in town were as follows

Age	Number in thousands
0 - < 5	4.4
5 - < 15	8.1
15 - < 30	10.5
30 - < 50	14.6
50 - < 70	9.8
70 - < 90	4.7

Determine the median and mean

Solution

age	x	f	fx	Cf
0 - < 5	2.5	4.4	11	4.4
5 - < 15	10	8.1	81	12.5
15 - < 30	22.5	10.5	236.25	23
30 - < 50	40	14.6	584	37.6
50 - < 70	60	9.8	588	47.4
70 - < 90	80	4.7	376	52.1
Total		**52.1**	**1876.25**	

$$\text{Median} = 30 + \frac{(26.05 - 23)}{14.6} x20$$

$$= 34.178 \text{ years}$$

$$\text{Mean} = \frac{\sum fx}{\sum f} = \frac{1876.25}{52.1}$$

Mean = 36. 012 years

1.4 MEASURES OF SPREAD (DISPERSION)

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

(a) RANGE

The range of a set of numbers is the difference between the largest and the smallest value of the set.
For example $S = \{ 45, 54, 64, 76, 86\}$
The range of S is $= 86 - 45 = 41$

(b) THE MEAN DEVIATION

The mean deviation of a set of numbers is given by

$$\sum_{1}^{n} \frac{|x_i - M|}{n} \quad \text{Where M is the Mean}$$

Example

Find the mean deviation of the set of numbers
45, 54, 64, 76, 86

Solution

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

$$M = \frac{45 + 54 + 64 + 76 + 86}{5}$$

$$= 65$$

Mean deviation

$$= \frac{|45-65|+|54-65|+|64-65|+|76-65|+|86-65|}{5}$$

$$= \frac{20+11+1+11+21}{5}$$

Mean deviation $= 12.8$

(c) VARIANCE FOR POPULATION

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

$$\text{Variance} = \frac{\sum_{1}^{n}(x_i - M)^2}{n} \quad \text{Note:}$$

Where M is the mean

For given set of numbers, $x_1, x_2, x_3, \ldots, x_n$

Example

Find the variance of 45, 54, 64, 76, 86

A simplified form of the formula normally used for computations is

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

Solution

x	x^2
45	2025
54	2916
64	4096
76	5776
86	7396
325	**22209**

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

$$= \frac{22209}{5} - (65)^2$$

$$\text{Variance} = 216.8$$

$\text{Variance} = \frac{\sum fx^2}{\sum f} - \left(\frac{\sum fx}{\sum f}\right)^2$ is used for repeated values or grouped data

Note: Standard deviation is the positive square root of the variance

Example

The frequency distribution table shows the marks of some students from 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

Solution

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
Total	**20**		**1411**	**100987**

$\sum fx = 1411$ *and* $\sum fx^2 = 100987$

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

$$= \frac{100987}{20} - \left(\frac{1411}{20}\right)^2$$

$$= 72.05$$

Therefore Standard deviation = $\sqrt{72.05}$

Standard deviation = 8.488 marks.

Example

The frequency distribution table shows the heights of some students from a certain school

Height	154	155	160	164	171	180
Frequency	4	6	8	5	4	3

Determine the variance of the above data

Solution

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

$D = x - A$, let A be 160

x	f	d	fd	fd^2
154	4	-6	-24	144
155	6	-5	-30	150
160	8	0	0	0
164	5	4	20	80
171	4	11	44	484
180	3	20	60	1200
Total	**30**		**70**	**2058**

$$\text{Variance} = \frac{2058}{30} - \left(\frac{70}{30}\right)^2$$

$$= 68.6 - 5.4444$$

$$\text{Variance} = 63.1556$$

Example

Using example 2

Determine the variance and standard deviation

Solution

x	f	x^2	fx	fx^2
12	5	144	60	720
17	9	289	153	2601
22	12	484	264	5808
27	18	729	486	13122
32	25	1024	800	25600
37	15	1369	555	20535
42	10	1764	420	17640
47	6	2209	282	13254
Total	**100**		**3020**	**99280**

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

$$= \frac{99280}{100} - \left(\frac{3020}{100}\right)^2$$

$$= 992.8 - 912.04$$

$$= 80.76$$

Therefore variance = 80.76

Standard deviation = $\sqrt{80.76}$

= 8.987 kgs.

Working mean can be used to obtain variance.

The formula used is

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

Where $d = x - A$ and A is working mean

Example

Using **example 3** and working mean of 57.95 find the variance for given data.

Solution

Weight	*x*	*f*	*d*	*fd*	*fd²*
44.0 – 47.9	45.95	3	-12	-36	432
48.0 – 51.9	49.95	17	-8	-136	1088
52.0 – 55.9	53.95	50	-4	-200	800
56.0 – 59.9	57.95	45	0	0	0
60.0 – 63.9	61.95	46	4	184	736
64.0 – 67.9	65.95	57	8	456	3648
68.0 – 71. 9	69.95	23	12	276	3312
72.0 – 75. 9	73.95	9	16	144	2304
Total		**250**		**688**	**12320**

Variance = $\frac{\sum fd^2}{\sum f} - (\frac{\sum fd}{\sum f})^2$

$$= \frac{12320}{250} - \left(\frac{688}{250}\right)^2$$

$= 49.28 - 7.573504$

Variance $= 41.7065$

Example

The frequency table shows the height (H) of 80 students to nearest cm.

Height (cm)	***Frequency***
150 - 154	3
155 - 159	7
160 – 164	10
165 – 169	15
170 – 174	25
175 – 179	12
180 – 184	6
185 - 189	2

Calculate the mean and standard deviation

Solution

Let A = 167, $t = \frac{d}{c}$ *where* $d = x - A$ and c = 5

x	f	d	t	ft	ft^2
152	3	-15	-3	-9	27
157	7	-10	-2	-14	28
162	10	-5	-1	-10	10
167	15	0	0	0	0
172	25	5	1	25	25
177	12	10	2	24	48
182	6	15	3	18	54
187	2	20	4	8	32
Total	**80**			**42**	**224**

$$Mean = A + \frac{\sum ft}{\sum f}c$$

$$Mean = 167 + \frac{42x5}{80}$$

$$= 169.625\,cm$$

$$\text{Standard deviation} = c\sqrt{\left(\frac{\sum ft^2}{\sum f} - \left(\frac{\sum ft}{\sum f}\right)^2\right)}$$

$$= 5\sqrt{\left(\frac{224}{80} - \left(\frac{42}{80}\right)^2\right)}$$

$$= \quad 7.944\text{ cm}$$

Note: The formula above is used only when the classes are of the same class width.

Example

The information below shows the amount of money in millions (S) given to some districts in Uganda for "Entandikwa" scheme.

S	25 - < 30	30 - < 40	40-<50	50-< 60	60-<80
f	4	10	4	3	5

Find the mean and standard deviation of the distribution

Solution

Note: The classes above are of different class width.

class	x	f	fx	x^2	fx^2
25 - < 30	27.5	4	110	756.25	3025
30 - < 40	35	10	350	1225	12250
40 - < 50	45	4	180	2025	8100
50 - < 60	55	3	165	3025	9075
60 - < 80	70	5	350	4900	24500
Total		**26**	**1155**		**56950**

$$Mean = \frac{\sum fx}{\sum f}$$

$$= \frac{1155}{26}$$

$$Mean = 44.423\ millions$$

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

$\sqrt{\left(\frac{56950}{26}-(\frac{1155}{26})^2\right)}$

$\sqrt{(2190.384615-1973.409763)}$

$\sqrt{216.974852}$

$=14.73$ millions

Variance for Sample

Note: Variance for a population is donated by

Variance of sample (s^2)

$$=\left(\frac{n}{n-1}\right)\quad \frac{\sum x^2}{n}-(\frac{\sum x}{n})^2$$

Example

A sample of 8 students gave the following as the estimates for the length of certain plot in meters as, 44, 48, 57, 53, 50, 52, 49 and 51. Calculate the standard deviation of the plot?

Solution

Sample variance $= \left(\frac{n}{n-1}\right) \quad \frac{\sum x^2}{n} - (\frac{\sum x}{n})^2$

x	x^2
44	1936
48	2304
49	2401
50	2500
51	2601
52	2704
53	2809
57	3249
404	20504

Sample variance $= \frac{8}{7}\left(\frac{20504}{8} - (\frac{404}{8})^2\right)$

$= 14.\ 5714$

Standard deviation $= \sqrt{14.5614}$

$= 3.817$

Example

A certain factory produces ball bearings. A sample of the bearing from the factory produced the following results

Diameter of bearing in mm	frequency
91 – 93	4
94 – 96	6
97 – 99	34
100 - 102	40
103 – 105	13
106 - 108	3

Determine the mean and variance of the diameter of the sample bearings.

Solution

$$\text{Mean} = \frac{\sum fx}{\sum f}$$

$$\text{Sample variance} = \left(\frac{\sum f}{\sum f - 1}\right)\left(\frac{\sum fx^2}{\sum f} - \left(\frac{\sum fx}{\sum f}\right)^2\right)$$

class	*f*	*x*	x^2	*fx*	fx^2
91-93	4	92	8464	368	33856
94 – 96	6	95	9025	570	54150
97 – 99	34	98	9604	3332	326536
100 - 102	40	101	10201	4040	408040
103 – 105	13	104	10816	1352	140608
106 - 108	3	107	11449	321	34347
Total	**100**			**9983**	**997537**

$$Mean = \frac{9983}{100}$$
$$= 99.83\,mm$$

$$\text{Sample Variance} = \left(\frac{100}{99}\right)\left(\frac{997537}{100} - \left(\frac{9983}{100}\right)^2\right)$$

$$= 9.4355 \text{ mm}$$

1.5 PERCENTILES, DECILES, AND QUARTILES

a) PERCENTILES

These are values that divide a set of observations into 100 equal parts. These values, denoted by $P_1, P_2, P_3, \ldots, P_{99}$, are such that 1% of the data falls below P_1, 20% of the data below P_{20}, etc.

Example

Suppose the number of observation n is 200, then

$$P_{10} = \left(\frac{10}{100} x\, 200\right) = 20$$

This implies that 20 observations are below it.

b) DECILES

The values divide a set of observation in 10 equal parts. These are denoted by $D_1, D_2, D_3 \ldots D_9$. Such that 20% of the total frequency fall below D_2

Suppose n = 500, determine D_6

$$D_6 = \frac{6}{10} x 500 = 300$$

c) QUARTILES

Quartiles are values that divide a set of observation into 4 equal parts. These values are denoted by Q_1, Q_2, and Q_3, are such that 25% of the data falls below Q_1, 50% below Q_2 and 75% below Q_3.

NOTE: Percentiles, deciles and quartiles can be obtained from the Ogive. Q_3 is known as the upper quartile while Q_1 is known as the lower quartile.

THE INTERQUARTILE RANGE (QUARTILE RANGE)

= Upper quartile – lower quartile = $Q_3 - Q_1$

$$Upper\ quartile = L_3 + \left(\frac{\frac{3N}{4} - F_b}{f_q}\right) x C$$

Where:

L_3 = lower class boundary of upper quartile class

N = Total number of observations

F_b = Cumulative frequency before the upper quartile class

f_q = frequency of the upper quartile class.

C = class width

$$Lower\ quartile = L_1 + \left(\frac{\frac{N}{4} - F_b}{f_q}\right) x\, C$$

Where:

L_1 = lower class boundary of lower quartile class
N = Total number of observations
F_b = Cumulative frequency before the lower quartile class
f_q = frequency of the lower quartile class.
C = class width

THE SEMI-INTERQUARTILE RANGE

$$= \frac{1}{2}(Q_3 - Q_1)$$

Example

The data below gives the weight of some students from a certain school.

Mass	*Frequency*
40 – 44	3
45 – 49	30
50 – 54	29
55 – 59	33
60 – 64	13
65 – 69	1
70 - 74	1

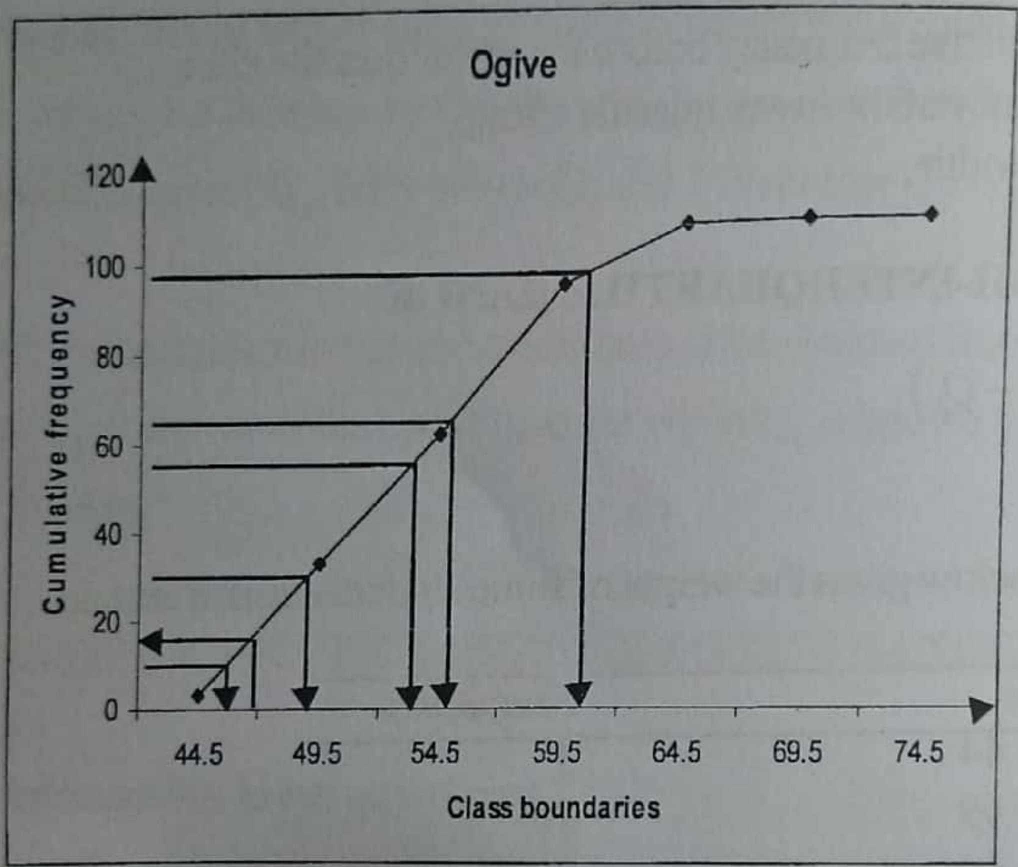
Draw an Ogive and use it to estimate

i. The median mass
ii. 10 -90 percentile range
iii. Interquartile range

(iv) The number of students who weigh above 47 kgs.

Solution

	44.5	49.5	54.5	59.5	64.5	69.5	74.5
Cf	3	33	62	95	108	109	110



i. Median = 53.5 kgs from the graph

ii. $P_{90} - P_{10} = 60.5 - 46.5 = 14$ kgs (graph)

iii. $Q_3 - Q_1 = 57.5 - 48.5 = 9$ kgs (Graph)

iv. Number of people weighing below 47 kgs = 17

Therefore the number above 47 kgs =

$$110 - 17 = 93 \text{ students}$$

Exercise

1. The table below shows the marks obtained by students of mathematics in a certain school

Marks	*Number of students*
30 - < 40	2
40 - < 50	15
50 - < 55	10
55 - < 60	11
60 - < 70	30
70 - < 90	29
90 - < 100	3

a) Draw a histogram for the data and hence determine the modal mark

b) Draw an ogive for the data

c) Calculate the mean, median and standard deviation

Answer (c) mean = 64.575, median = 64
Standard deviation = 13.852

2. The data below shows the amount of cotton (in 100's of bales) produced by Growers' Unions over a certain period of time.

70	41	34	55	45	66	73	77	80	30
50	45	72	50	27	70	55	70	85	70
30	50	60	53	40	45	35	55	20	81
25	51	35	62	60	30	45	35	50	89
53	23	28	65	68	50	65	34	35	76

i. Beginning with the 20 – 29 class and using intervals of equal width, construct a frequency table for the data. Using the frequency table

ii. Draw a cumulative frequency curve for the data and hence estimate the median production

iii. Calculate the mean and standard deviation of production.

Answer (ii) median = 5450 bales (iii) mean = 5370 bales, standard deviation = 1798 bales

3. The amount of money collected by children in a charity walk in thousands of shillings is recorded in the table below:

Amount('000 Ushs)	1	2- 4	5 -1	11 - 15	16 -25
Frequency	20	93	90	58	39

i. Calculate the mean and standard deviation of the children's collection

ii. Draw a histogram and use it to estimate the modal collection

iii. Construct an ogive and use it to estimate the number of children who collected between 3,000/= and 7,000/=.

Answer (i) mean = 8425 /=

standard deviation = 6004.7 /=

From the graph modal collection = 2900 (iii) 68

4. The table below shows the weights to the nearest kg of 150 patients who visited a certain health unit during acertain week.

weight (kg)	*Number of patients (f)*
0 - 19	30
20 -29	16
30 - 39	24
40 - 49	32
50 - 59	28
60 – 69	12
70 - 79	8

a) Calculate the approximate mean and modal weights of the patients.

b) Plot an ogive for the data. Use the ogive to estimate the

i. Median and semi- interquartile range for the weights of patients,

Answer (a) mean = 38.8333 kgs,

Modal weight = 46. 167 kgs, Median = (41.0 – 42.0) kgs

Semi-interquartile range = (14.5 – 15.5) kgs.

5. The ages of people in Kampala were as follows

Age (years)	*Number in thousands*
0 - < 5	4.4
5 - <15	8.1
15 - < 30	10.5
30 - < 50	14.6
50 - < 70	9.8
70 - < 90	4.7

a) Draw a histogram for the data

b) State the modal age interval

c) Estimate

i. Average age of Kampala

ii. Number of people under 18 years

iii. Median age

Answer (b) modal age interal = (0 - < 5)

(c) mean = 36.01 yrs (ii) 14600 (iii) 34. 178yrs

6. Find the mean and standard deviation of the first n integers

Answer: mean $= \frac{1}{2}(n+1)$

Standard deviation $= \sqrt{\frac{(n-1)(n+1)}{12}}$

7. The table below gives the weekly earnings of a random sample of workers in Mukwano soap factory

Weekly earnings (Shs)	Number of workers
Under 1500	1
1500 and under 2000	4
2000 and under 2500	28
2500 and under 3000	42
3000 and under 3500	33
3500 and under 4000	18
4000 and under 4500	13
4500 and under 5000	9
5000 and over	2

Calculate

i. The percentage number of workers earning Shs 3050 and above.

ii. The mean weekly earning and standard deviation of the distribution

iii. If the wages are increased by 10%, determine the new mean weekly wage and new standard deviation

Answer (i) 45% (ii) mean = 3133.333,

standard deviation = 799.364 (iii) mean = 3446.67,

standard deviation = 879.30

8. The heights of a senior five class in a certain school were recorded as 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 students' height

(b) Plot an Ogive

(c) Use your Ogive to estimate the:
 (i) median and inter- quartile range for the data

1.6 INDEX NUMBERS

An index number is a statistical measure, which represents the change in a variable or group of variables with respect to time, environment or other characteristics

Base year

This is the year or period against which all other years or periods are compared.

Current year

This the year (period) *for* which the index is to be calculated

SIMPLE INDEX NUMBER

Price index, quantity index, wage index e.t.c.. are examples of the simple index numbers. Simple price index is often known as a price relative and it is given by $\frac{p_1}{p_0} \times 100$ or $\frac{p_1}{p_0}$

Where p_1 is the price of unit of one commodity in the current period while p_0 is the price of unit of one commodity in the base period.

Example

One litre of petrol costs shs. 1200 in 1998 and shs. 1800 in 2004.Taking 1998 as the base year, find the price relative in 2004

Solution

$$\text{Price relative} = \frac{1800}{1200} x100 = 150$$

This indicates that the price of petrol has gone up by 50%

NOTE: The quantity index can be obtained by using

$$Quantity\ index = \frac{q_1}{q_0} x100$$

Example

The wage of Nurses in Uganda in 1994 was Shs 30,000.The wage of the same Nurse in 1996 was increased by 15,000/=. Using 1994 as the base year calculate the nurses wage index for 1996.

Solution

W_1=30,000+15,000=45,000

W_o=30,000

$$\text{Wage index} = \frac{45000}{30000} x100 = 150$$

Therefore the nurses wage increased by 50 % in 1994.
Note: The percentage sign is always omitted in the answer

Example

In 1990 the price index of a commodity, using 1986 as base year was 112. In 1996, the index using 1990 as base year was 85. What would have been the index in 1996, using 1986 as base year?

Solution

$$\frac{p_{1990}}{p_{1986}} = 1.12$$

$$\frac{p_{1996}}{p_{1990}} = 0.85$$

$\frac{p_{1996}}{p_{1986}} = 1.12 x 0.85$

$= 0.952$

$= 95.2$

Sometimes you can calculate the total price of group of items as ratio of the total price of the same group of items in the base year This index is known as the simple aggregate price index

$$\text{Simple aggregate price index} = \frac{\sum p_1}{\sum p_0}$$

Example

The table shows the price in dollars of sugar and cassava in 1990 and 1995

Item	1990	1995
Sugar (bag)	80	120
Cassava (bag)	10	15

Determine simple aggregate price index for the total cost of one bag of cassava and sugar.

Solution

Item	1990 (P_0)	1995 (P_1)
Sugar	80	120
Cassava	10	15
Total	90	135

$$\text{Simple aggregate price index} = \frac{135}{90} = 1.5$$

Example

An average family in Kampala spent the following amounts per month on the items shown, in the years 1999 and 2000

ITEM	*1999 Amount in Uganda sh.*	*2000 Amount in Uganda Sh.*
Housing	80,000	100,000
Clothing	20,000	20,000
Electricity	40,000	50,000
Water	10,000	12,000
Food	140,000	160,000
Transport	50,000	60,000
Medical	30,000	35,000
Miscellaneous	30,000	40,000

(They save Shs 50,000 per month)

Using 1999 as the base year, calculate the simple aggregate expenditure index for 2000.

Solution

ITEM	*1999 Amount in Uganda sh.*	*2000 Amount in Uganda Sh.*
Housing	80,000	100,000
Clothing	20,000	20,000
Electricity	40,000	50,000
Water	10,000	12,000
Food	140,000	160,000
Transport	50,000	60,000
Medical	30,000	35,000
Miscellaneous	30,000	40,000
Total	400,000	477,000

Simple aggregate expenditure index =

$$= \frac{477,000}{400,00} = 1.1925$$

Therefore the expenditure went by 19.25 % in 2000.

COMPOSITE INDEX NUMBER

Index numbers are at times needed where there is more than just one item, for example an index number that compares the cost of living depends on food, clothing, housing,entertainment,etc...

Example

Find the cost of living based on the following data.

Item	*Price index*	*Weight*
Housing	125	170
Clothing	124	160
Electricity	150	120
Water	110	180
Food	120	172
Transport	135	210
Others	104	140

Solution

ITEM	*Price Index (A)*	*Weight (W)*	*WA*
Housing	125	170	21,250
Clothing	124	160	19,840
Electricity	150	120	18,000
Water	110	180	19,800
Food	120	172	20,640
Transport	135	210	28,350
Others	104	140	14,560
Total		**1152**	**142,440**

$$\text{Cost of living index} = \frac{142{,}440}{1152} = 123.65$$

Example

The following items are used in the assembly of a TV set. 8 transistors,22 resistors, 9 capacitors, 2 diodes and a circuit board. Due to inflation the price of each component has increased as shown in the table:

Year	*Transistor*	*Resistor*	*Capacitor*	*Diodes*	*Circuit*
2000	1200/=	1650/=	1500/=	1600/=	2000/=
2002	1800/=	2100/=	1700/=	1800/=	2500/=

Calculate the composite index number of the assembled TV set in 2002 using 2000 as the base year

Solution

Item	*Weight (W)*		*Price (P_0)*		*WP_0*
	Price (P_1)		*WP_1*		
Transistor	8	1200	9600	1800	14400
Resistor	22	1650	36300	2100	46200
Capacitor	9	1500	13500	1700	15300
Diode	2	1600	3200	1800	3600
Circuit	1	2000	2000	2500	2500
Total			**64600**		**82000**

$$\text{Composite index} = \frac{82000}{64600} = 1.27$$

Note: Composite index is the same as weighted aggregate index.

$$\text{Composite index} = \frac{\sum p_1 w}{\sum p_0 w} x100 = \frac{\sum q_1 w}{\sum q_0 w} x100$$

This is weighted aggregate method

Sometimes the prices of the commodities can be quoted in different units. In such situations the price relatives are used.

Example

Wavamuno imported the following commodities A, B, C from Kenya, U.S.A and England respectively.

Commodity	*Price*	
	1992 = 100	1996
A	K sh 20,000	K sh 25,000
B	$ 4,000	$ 4,800
C	£8,000	£10,,000

Calculate the simple relative price index for 1996

Note: 1992 = 100 means that 1992 is taken as the base year.

Solution

Commodity	*Price*		$\frac{p_1}{p_0}$
	1992 = 100	1996	
A	K sh 20,000	K sh 25,000	1.25
B	$ 4,000	$ 4,800	1.2
C	£8,000	£10,000	1.25
Total			**3.7**

Index for 1996 = $\frac{3.7}{3} x 100 = 123.3$

Therefore the price increased by 23 % in 1996 as compared to 1992.

Note: Price relatives can be used for commodities in the same units too.

Example

Nabisunsa Girls' School bought three types of chicken feed in 1998 and 2000. The weights and the corresponding prices are given in the table. Using 1998 as the base year calculate the weighted average relative price index.

Commodity	*Weight*	*Price 1998*	*Price 2000*
A	120	500	600
B	60	300	360
C	50	250	400

Solution

Commodity	*weight*	*Price 1998*	*Price 2000*	$\frac{p_1}{p_0}$	$\frac{p_1}{p_0} x w$
A	120	500	600	1.2	144
B	60	300	360	1.2	72
C	50	250	400	1.6	80
Total	**230**				**296**

Weighted average relative price index

$$= \frac{296}{230} x 100 = 128.7$$

The commodities increased by 28.7%

Example

The wages and wage bills for a group of workers in a factory for two years are shown in the table.

Let S- skilled, U- unskilled and M- semi-skilled

	Year 1986 Wage (shs)	*Total) wage Bill (shs)*	*Year 1986 Wage (shs)*	*Total wage Bill (shs)*
S	7500	225,000	10,000	320,000
M	4000	160,000	4800	240,000
U	3500	157,000	4200	193,200

Determine

i. The weighted average wage for each year

ii. The wage index in 1989 taking 1986 as the base year.

Solution

Let N be the number of workers, W be wage in shillings.

W_0	*N*	*Total Wage*	W_1	*N*	*Total Wage*
7500	30	225,000	10,000	32	320,000
4000	40	160,000	4800	50	240,000
3500	45	157,500	4200	46	193,200
Total	115	542,500		128	753,200

Weighted average wage for 1986 is

$$\frac{542,500}{115} = 4,717.40/=$$

Weighted average for 1989 is

$$\frac{753,200}{128} = 5,884.40/=$$

The wage index in 1989 taking 1986 as the base year

$$= \frac{5,884.4}{4,717.4} = 124.7$$

Wage index = 124.7

Example

Musa reckons that in 1988, his farm produced 60% of his income, and that farm production in the years 1988, 1990, 1992 were in the ratio 60:55:60. His other income, a salary, in millions shillings was:

1988	1990	1992
8	20	32

However, in real terms this salary must be related to cost of living index, whch was

1988	1990	1992
100	160	240

i) Find his salary in real terms of each of the years given and express it as an index with 1988 as 100.

ii) still keeping 1988 as the base year, find an index of the total real income for 1990 and 1992.

iii) Amend (ii) if the farm production for 1988 formed 75% of his income (the ratio 60:55:50 remaining valid)

Solution.

	Salary Millions	*Farm* Millions	*Cost* of Living	*Total* income (m)
1988	8	12	100	20
1990	20	24.4	160	44.4
1992	32	32	240	64
Total	**60**	**68.4**		**128.4**

Year	*Real salary*	*Salary index (1988=100)*
1988	$8x\frac{100}{100}=8$	$\frac{8}{8}x100=100$
1990	$20x\frac{100}{160}=12.5$	$\frac{12.5}{8}x100=156.3$
1992	$32x\frac{100}{240}=13.33$	$\frac{13.33}{8}x100=166.6$

(ii) Index for total real income

Year	Real income	index for total real income
1988	20	$\frac{20}{20}x100=100$
1990	44.4	$\frac{44.4}{20}x100=222$
1992	64	$\frac{64}{20}x100=320$

Salary was 25% of his total real income

Therefore $\frac{25}{100}x(\textit{total real income})=8)$

Total income = 32 millions

If 60 corresponds to 75%
Then 55 will correspond to y

Therefore y= $\frac{75x55}{60}=68.75$ %

Salary was 31.25 % of total real income in 1990

Year	*Salary*	*Real salary*	*Salary index (1988=100)*
1988	8	$8x\frac{100}{25}=32$	$\frac{32}{32}x100=100$
1990	20	$20x\frac{100}{31.25}=64$	$\frac{64}{32}x100=200$
1992	32	$32x\frac{100}{37.5}=85.3$	$\frac{85.3}{32}x100=266.6$

Example

The cost of making a cake is calculated from the cost of baking flour, sugar, milk and eggs. The table gives the cost of these items in 1990 and 1996

Item	*1990*	*1996*	*Weight*
Flour per kg	600	780	12
Sugar per kg	500	400	5
Milk per litre	250	300	2
Eggs per egg	100	150	1

Using 1990 as the base year

i. Calculate the price relatives for each item. Hence find simple price index for the cost of making a cake

ii. Find the weighted aggregate price index for the cost of making a cake

If the cost of making a cake in 1996 was Sh 300.
Find the cost in 1990 using the two indices in (i) and (ii)

Solution

Item	*1990*	*1996*	*W*	$\frac{p_1}{p_0}$	p_0W	p_1W
Flour/kg	600	780	12	1.3	7200	9360
Sugar/kg	500	400	5	0.8	2500	2000
Milk/l	250	300	2	1.2	500	600
Eggs/egg	100	150	1	1.5	100	150
Total		20	4.8	10300		12110

The fifth column has the price relatives

$$\text{Simple price index} = \frac{4.8}{4} x100 = 120$$

$$\text{Weighted aggregate index} = \frac{12110}{10300} x100 = 117$$

Cost in 1990

Using simple price index = 120

(i) From $\frac{p_1}{p_0} x100 = 120$

$$\frac{300}{P_0} x100 = 120$$

$$P_0 = \frac{300}{120} x100 = 250$$

Cost in 1990 was 250/=

(ii) From $\frac{P_1}{P_0} x100 = 117$

$$\frac{300}{P_0} x100 = 117$$

$$P_0 = \frac{300}{117} x100 = 256.4/=$$

Cost in 1990 was 256.4/=

Sometimes different quantities are quoted for the base year and current year.
In such situations the methods that can be used include

i. Paasche aggregate price index
ii. Laspeyre aggregate price index.

Paasche aggregate price index
The quantities of the current year are used. This implies that you allow the changes in the prices.

Laspeyre aggregate price index
The quantities of the base year are used

Example

Using the figures given below, calculate:

i. Paasche aggregate price index

ii. Laspeyre aggregate price index

	1980		1986	
	Quantity	Price	Quantity	Price
Maize	20	650	25	700
Wheat	10	1500	8	1600
Beans	5	150	8	200

Let Q denote quantity, P denote price

Solution

	1980		1986			
	Q_0	P_0	Q_1	P_1	P_0Q_1	P_1Q_1
Maize	20	650	25	700	16250	17500
Wheat	10	1500	8	1600	12000	12800
Beans	5	150	8	200	1200	1600
Total					**29450**	**31900**

Paasche aggregate price index $= \dfrac{\sum P_1Q_1}{\sum P_0Q_1} x100$

$$= \frac{31900}{29450} x100 = 108.3$$

The prices of 1986 are higher than those of 1980 by 8%

(ii)

	1980		1986			
	Q_0	P_0	Q_1	P_1	P_0Q_0	P_1Q_0
Maize	20	650	25	700	13000	14000
Wheat	10	1500	8	1600	15000	16000
Beans	5	150	8	200	750	1000
Total					28750	31000

Laspeyre aggregate price index = $\frac{\sum P_1Q_0}{\sum P_0Q_0} x100$

$$= \frac{31000}{28750} x100 = 107.8$$

The prices of 1986 are higher than those of 1980 by 8%.

Note: In most cases the two methods give different answers.

Value index number

This is obtained from = $\frac{\sum P_1Q_1}{\sum P_0Q_0} x100$

Using the above example

The value index = $\frac{31900}{28750} x100 = 111$

This shows an increase of 11% in 1986 compared to 1980

Sometimes the value and quantity is given without the price of the commodity.

Example

Commodity	1990		1992	
	Quantity Q_0 (kgs)	Value P_0Q_0	Quantity P_1 (kgs)	Value P_1Q_1
A	25	7500	40	16000
B	30	12000	50	15000
C	10	8000	25	10000
D	20	12000	15	12000

Determine the index for the change in the price and weighted average of price relatives.

Solution

Item	1990 (100)		1992					
	Q_0	P_0Q_0	Q_1	P_1Q_1	P_0	P_1	P_1Q_0	$\frac{P_1}{P_0} x Q_1$
A	25	7500	40	16000	300	400	10000	53.33
B	30	12000	50	15000	400	300	9000	37.5
C	10	8000	25	10000	800	400	4000	12.5
D	20	12000	15	12000	600	800	16000	20.0
		39500	130				39000	123.33

Index for change in price $= \frac{39000}{39500} x 100$

$= 98.7$

Weighted average of price relatives

$$= \frac{123.33}{130} x100 = 94.9$$

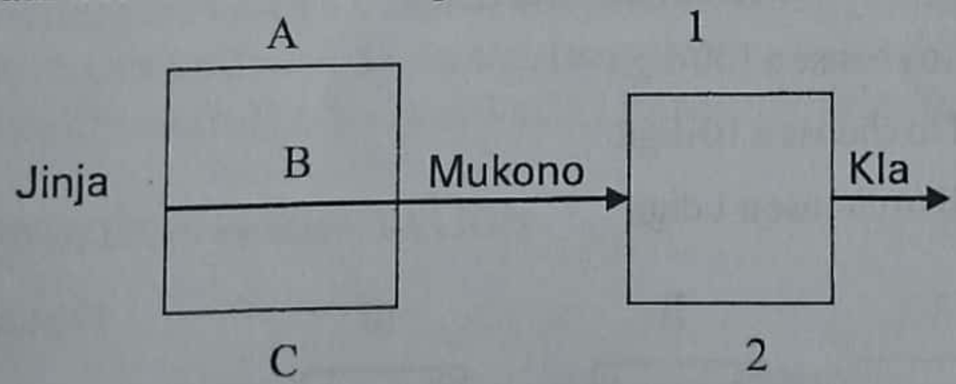
Changing the base year

If you wish to change the base year, multiply the current indices

$$\text{by} \quad \frac{100}{\Pr ice\ index\ of\ new\ base\ year}$$

1.7 PERMUTATION AND COMBINATION

Consider the fundamental counting principle. A bus driver has to drive from Jinja to Kampala via Mukono. There are three different roads that he can take from Jinja to Mukono and two different roads from Mukono to Kampala.



One of the possible route is first take road A and then take 1 or simply A.1 Other possible routes are A2, B1, B2 C1 and C2. Therefore there are six possible routes

NOTE: The number of possible routes can be found without listing any of the routes as follows

Make two boxes for two stages

Jinja I	Mukono II
3	2

Stage I for Jinja to Mukono there are three choices A, B, C.

Stage II for Mukono to Kampala there are two choices 1 and 2

Total number of ways 3 x 2 = 6 ways.

Therefore, permutation is an arrangement of objects in a definite order.

Example

How many three digit numbers can be made from the digits 3, 4, 5, 6, 7 and 8 if no digit is repeated.

There are three stages to make a decision

Stage I to choose a 100 digit

Stage II to choose a 10 digit

Stage III to choose a 1 digit

I	II	III
6	5	4

There are 6 choices for the first position, then 5 choices for the second position and 4 choices for the last position. The total numbers of ways are 6 x 5 x 4 = 120

Example

How many three digits can be formed from the digits 1, 2, 3, 4 if repetitions are allowed?

I	II	III
4	4	4
100s digits	10s digits	1s digits

The total number of ways are 4x4x4 = 64 ways.

Example

Find the number of possible arrangements for five people to sit together o a bench

5	4	3	2	1
Seat 1	Seat 2	Seat 3	Seat 4	Seat 5

Thus there are 5 x 4 x 3 x 2 x 1 = 120 possible arrangements.
Note: 5 x 4 x 3 x 2 x 1 = *5!* Factorial notation.
Hence n! = n x (n-1) x (n - 2)...x 3 x 2 x 1 for positive integers.

CONDITION PERMUTATION

Example

How many permutations of all letters in the word BRIDGE end with the letter I or B

1st	2nd	3rd	4th	5th	6th
5	4	3	2	1	2

Therefore the possibilities are: 5x 4 x 3 x 2 x 1 x 2
= 120 x 2 = 240 ways

THE OR CONDITION

Example

How many even numbers containing one or more digits can be formed from the digits 2, 3, 4, 5, 6 if no digit may be repeated.
Since the numbers required are even, last unit digits must be one of the three digits 2, 4, 6. Note there are five digits so we can form one digit, two digit, three digit, four digit or five digit numbers as follows

One digit numbers	is either 2 0r 4 or 6	3 = 3
Two digit numbers		4x3 = 12
Three digit numbers		4x3x3 = 36
Four digit numbers		4x3x2x3 = 72
Five digit numbers		4x3x2x1x3 = 72
Total		**195**

There are 195 such even numbers

THE AND CONDITION

Example

Four different English books and three different mathematics books are to be arranged on a shelf. In how many ways can these be arranged if the English books must be put together on the left.

English books on left   Math books on the right

4 x 3 x 2 x1   x   3 x 2 x1

4!   x   3!   = 144ways

The symbol $^{n}p_{r}$

The number of r objects chosen from n unlike objects is

$$^{n}p_{r}$$

Example

How many permutations are there of 3 letters chosen from eight unlike letters of the word RELATION

$$^{8}P_{3} = \frac{8!}{(8-3)!} = \frac{8!}{5!}$$

$$= 8 \times 7 \times 6 = 336 \text{ ways}$$

DISTINGUISHABLE PERMUTATION

Example

How many distinguishable six digit numbers can be formed fom the digits of 5,4,8,4,5,4.
The two 5's can be arranged in 2! Ways and the 4's can be arranged in 3! Ways. Hence the number of the six digits are

$$\frac{6!}{3!2!} = \frac{720}{6x2} = 60$$

Example
How many signals can be formed by displaying seven flags if three of them are red, two are green and the other two are blue?

Example $\frac{7!}{3!2!2!} = \frac{5040}{6x2x2} = 210$

How many arrangements can be made of the letters in word TERRITORY?

$$\frac{9!}{2!3!}=\frac{362880}{12}=30240$$

Example

Find the number of ways in which the letters of ISOSCELES can be arranged if the two E's are separated.

Solution

Arrangement without restriction

$$\frac{9!}{3!2!}=\frac{362880}{12}=30240\ ways$$

Let the two E's be together. Then they can be treated as one letter.

Total number of letters is 8

$$\frac{8!}{3!}=\frac{40320}{6}=6720$$

The number of ways when the E's are separated

$= 30240 - 6720 = 23520$ ways

$= 23520$ ways.

CIRCULAR PERMUTATION

The number of circular permutation of n distinct objects is (n-1)!

Example

In how many ways can 5 people sit at a round table?

One person is considered fixed and the other 4 can be arranged in 4! Ways

$n = 5 \Rightarrow (5\text{-}1)! = 4! = 24$ ways

COMBINATIONS

A combination is a selection of an object in which no regard is paid to the order.

Example

List all the three combinations that can be formed from the five students A, B, C, D, E and hence state the number of combinations.

ABC, ABD, ABE, ACD, ACE, ADE, BCD, BCE, BDE and CDE

Thus 10 three-student combinations can be formed.

The number of permutations of three letters from A, B, C, D, E are

$$^5p_3 = \frac{5!}{(5-3)!} = \frac{120}{2} = 60$$

NOTE: The number of permutations

= 3! x (the number of combinations)

$$\therefore \text{The number of combinations} = \frac{5!}{3!2!} = 10$$

In general, the number of combinations of r objects chosen unlike objects is written as ^nC_r or

$$\binom{n}{r} = \frac{n!}{(n-r)!r!} \qquad r \leq n$$

Example

A committee of four people is formed from nine people. In how many ways can this be done

$$^9C_4 = \frac{9!}{(9-4)!4!} = \frac{9!}{5!4!}$$

$$= \frac{9 x 8 x 7 x 6 x 5!}{4 x 3 x 2 x 1 x 5!} = \frac{9 x 8 x 7}{4} = 126 \text{ committee's}$$

Example

A committee of three men and one woman is obtained from five men and three women. In how many ways can the members be chosen?

Three men can be selected from five men in ${}^5C_3$ ways and 1 woman can be selected from three women in ${}^3C_1$ ways. Now for each of the ${}^5C_3$ ways of selecting men, there are ${}^3C_1$ ways of selecting the woman. Hence there are ${}^5C_3 \times {}^3C_1$ ways of selecting the committee.

$${}^5C_3 \times {}^3C_1 =$$

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

Therefore 30 such committees can be formed

$$\frac{120}{4} = 30$$

Example

There are 3 boys and 4 girls at a birthday party. In how many ways can a team of 3 students be formed so as to include at least one boy.

Possible number of boys are 1,2, or all 3. The total number of students should be 3.

Hence number of teams with I boy and 2 girls

${}^3C_1 \times {}^4C_2 = 3 \times 6 = 18$

Number of teams with 2 boys and 1 girl

${}^3C_2 \times {}^4C_1 = 3 \times 4 = 12$

Number of teams with 3 boys and no girl

${}^3C_3 \times {}^4C_0 = 1 \times 1 = 1$

Total is $18 + 12 = 1 = 31$

$\therefore$ 31 such teams can be formed.

Exercise

1. In how many ways can 4 people be seated on a bench?

   Answer 24 ways

2. In how many ways can 9 different books be arranged on a shelf such that

   (a) Three of the books are always together

   (b) Three of the books are all never together.

   Answer (a) 30240 (b) 332640

3. Seven men and two women are to sit on a bench. In how many ways can they arrange themselves so that the women do not sit next to each other?

   Answer 282240

4. How many arrangements can be made of the letters in THIRTIETH?

   Answer 15120

5. Solve for n in ${}^{n}C_{4} = {}^{n}C_{2}$

   Answer $n = 6$

6. A committee of five students to comprise the school council is to be selected from eight male students and five female students. Find how many possible committees can be obtained.

   Answer 1287 committees

7. From 7 capitals, 3 vowels and 5 consonants, how many words of 4 letters each can be formed if each word begins with a capital, contains at least one vowel, all the letters of each word being different.

   Answer 1932 words.

CHAPTER TWO

PROBABILITY THEORY

2.1 INTRODUCTION

The term probability arose from the games of chance. For example tossing a coin, rolling a die, playing cards etc…

Sample Space and generation of the sample space.

Sample space (S) is the set of all possible outcomes of an experiment. Each possible outcome is called a sample point.

Example: rolling a die: S = (1, 2, 3, 4, 5, 6)

Note: 1, 2, 3, 4, 5 and 6 are sample points

Generation of sample space

Ways of generating a sample space include

a) Table of outcomes

b) Permutations and combinations

c) Tree diagram

Review of terms used in set theory

Subset of a sample space S is called an Event.

Example of an Event

(a) Toss a die, then S = {1,2,3,4,5,6}. If the

desired event is that an 'Odd' number shows up.

Odd is the event E and thus E={1,3,5}.

Intersection of *events:*

For any two events A and B, we define a new event called the intersection of A and B and consists of outcomes that are in both A and B or $(A \cap B)$

Union of Events.

Denoted by $A \cup B$ is the set of all sample points in either A or B or both.

Compliment of events

If A is an event of a sample space S, the compliment of A is given by the set containing all sample points in S that are not in A. Denoted by

$\overline{A}$

Mutually Exclusive events

If two events A and B have no sample points in common i.e if $A \cap B = \{\ \}$, we say that A and B are mutually exclusive. These are disjoint as shown in Venn diagram.

Therefore $P(A)=\frac{n(A)}{n(S)}$

Example

Find the probability of choosing a defective computer in a lot of 12 out of which 4 are defective if a single draw is made.

Solution

The number of ways the event can happen is 4. Total number of possibilities is 12

Hence probability $= \frac{4}{12}=\frac{1}{3}$

Example

What is the probability of throwing a number greater than 4 for a die whose faces are numbered from 1 to 6?

$$\{1, 2, 3, 4, 5, 6\} \quad n(S) = 6$$

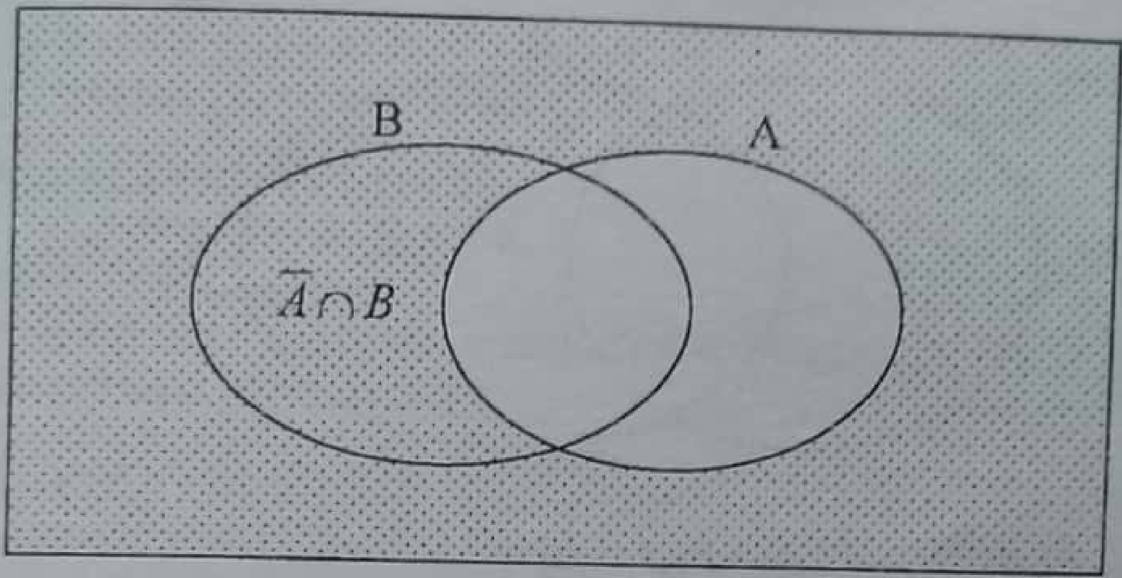
$$A = \{5, 6\} \quad \Rightarrow n(A)=2$$

$$P(A)=\frac{2}{6}=\frac{1}{3}$$

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

2.2 INTERACTION WITH SET THEORY

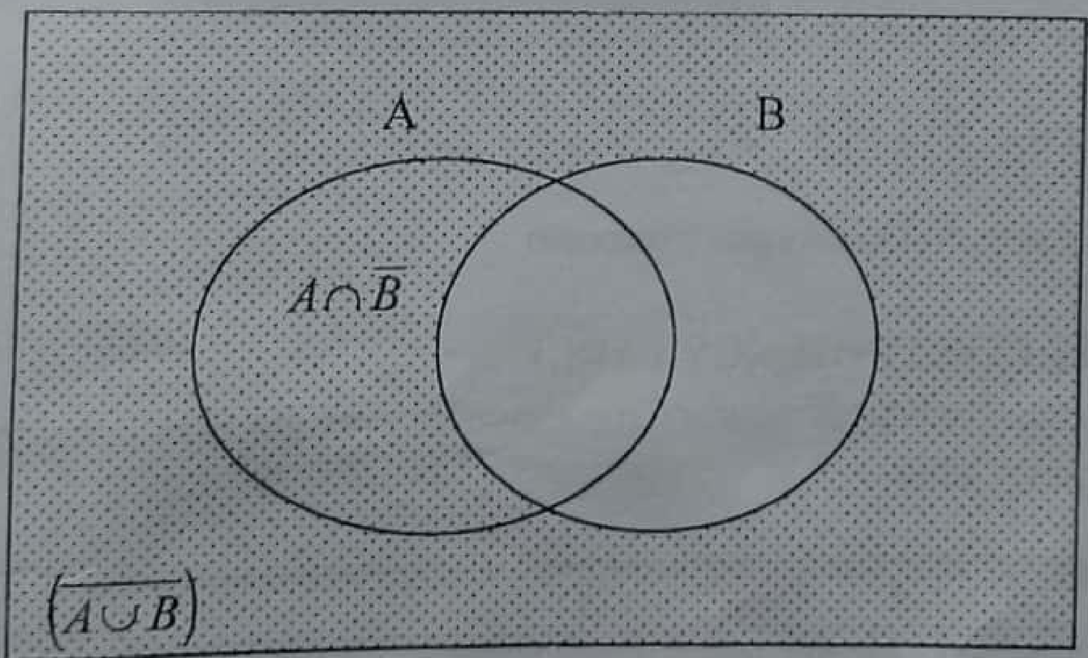
The following results can be deducted from set theory.



Note that $\left(\overline{A \cup B}\right) = \overline{A} \cap \overline{B}$

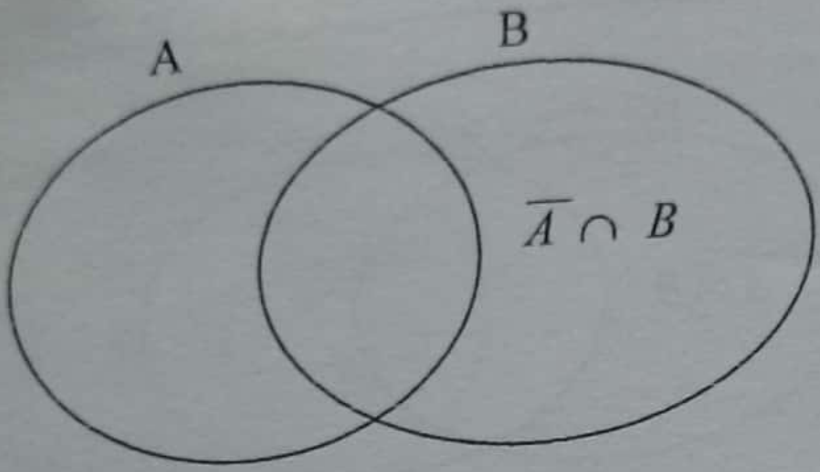
Result 4

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



Result 5

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



$P(A \cup B) = P(A) + P(\overline{A} \cap B)$

$But\ P(\overline{A} \cap B) = P(B) - P(A \cap B)\ from\ result\ 2)$

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

Result 6

For any two events A and B

i. $P(\overline{A} \cap \overline{B}) = P(\overline{A \cup B})$

ii. $P(\overline{A} \cup \overline{B}) = P(\overline{A \cap B})$

This is from De-morgan's theorem

2.3 THE CONTINGENCY TABLE

The alternative way of recalling the above four results, is by using the contingency table.

	A	$\overline{A}$	
B	$P(A \cap B)$	$P(\overline{A} \cap B)$	$P(B)$
$\overline{B}$	$P(A \cap \overline{B})$	$P(\overline{A} \cap \overline{B})$	$P(\overline{B})$
	$P(A)$	$P(\overline{A})$	1

Example

Using the second row implies that

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

Example

Given that A and B are two events such that $P(A)=0.5$, $P(B)=0.7$ and $P(A\cup B)=0.8$

Find (i) $P(A\cap B)$ (ii) $P(A\cap \overline{B})$

Solution

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

$$0.8=0.5+0.7-P(A\cap B)$$

$$\therefore\ P(A\cap B)=1.2-0.8=0.4$$

(ii)

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

$$=0.5-0.4$$

$$\therefore\ P(A\cap \overline{B})=0.1$$

Example

The probability that a student passes Mathematics is $\frac{2}{3}$ probability that he passes Chemistry 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.

Solution

Let M represent mathematics and C chemistry

That implies $P(M)=\frac{2}{3},\ P(C)=\frac{4}{9},\ P(M\cup C)=\frac{4}{5}$

$$P(M \cup C) = P(M) + P(C) - P(M \cap C)$$

$$\Rightarrow \frac{4}{5} = \frac{2}{3} + \frac{4}{9} - P(M \cap C)$$

$$\therefore P(M \cap C) = \frac{2}{3} + \frac{4}{9} - \frac{4}{5} = \frac{14}{45}$$

Example

Given that A and B are mutually exclusive events such that $P(A) = 0.5$, $P(A \cup B) = 0.9$, *find;*

i. $P(\overline{A} \cup B)$

ii. $P(\overline{A} \cap \overline{B})$

Solution

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

$$P(A) = 0.5 \quad \Rightarrow P(\overline{A}) = 1 - 0.5 = 0.5$$

$$P(B) = 0.9 - 0.5 = 0.4 \text{ since } A \text{ and } B \text{ are mutually exclusive}$$

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

$$= 0.4 - 0 = 0.4$$

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

$$= 0.5 + 0.4 - 0.4$$

$$= 0.5$$

$$\therefore P(\overline{A} \cup B) = 0.5$$

(ii) $$P(\overline{A} \cap \overline{B}) = 1 - P(A \cup B)$$

$$= 1 - 0.9 = 0.1$$

$$\therefore P(\overline{A} \cap \overline{B}) = 0.1$$

Example

In a race, the probability that Grace wins is 0.4, the probability that Umar wins is 0.2 and the probability that Denis wins is 0.3. Find the probability that

a) Grace or Denis wins
b) Neither Denis nor Umar wins

Assume that there are no dead heats.

Solution

Let G, U, D represent Grace, Umar and Denis respectively

a)
$$P(G \cup D) = 0.4 + 0.3 = 0.7$$
$$\therefore P(Grace \ \ or \ \ Denis \ \ wins) = 0.7$$

b)
$$P\left(\overline{D \cup U}\right) = 1 - P(D \cup U)$$
$$= 1 - (0.3 + 0.2)$$
$$= 1 - 0.5 = 0.5$$
$$\therefore P(Neither \ \ Denis \ \ nor \ \ Umar) = 0.5$$

2.4 PROBABILITY SITUATIONS

THE OR SITUATION

If A and B are two events, the probability that either event A or B or even both occur is denoted by $P(A \cup B)$,

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

Example

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

Solution

Table of outcome can be used to generate a 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= $\{(3,6)\ (4,5)\ (6,3)\ (5,4)\ (5,5)\ (5,6)\ (6,5)\ (6,4)\ (4,6)\ (6,6)\}$

$n(A) = 6,\quad n(B) = 10$ and $n(A \cap B) = 2$

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

$$= \frac{6}{36} + \frac{10}{36} - \frac{2}{36} = \frac{14}{36}$$

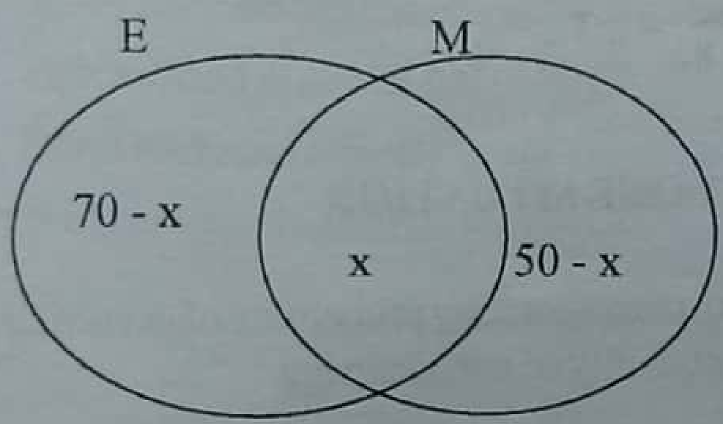
$$\therefore P(\textit{Double or sum greater than } 8) = \frac{7}{18}$$

THE AND SITUATION

a) When there is joint occurrence of events
b) Considering a sequence of events.

Example (a)

In a class of 100 students, 70 offer Economics while 50 students offer Mathematics. Each student offers at least one of the subjects. Determine the probability for the number of students who offer both subjects?



From the venn diagram

$70 - x + x + 50 - x = 100$

$120 - x = 100$

$x = 20$ students

$\therefore P(E \cap M) = 0.2$

Example (b)

Three balls are drawn at random one after the other without replacement from a bag containing 21 white, 9 blue, 40 red and 12 orange balls. Determine the probability that the first ball is blue, the second white and the third red.

Solution

Let B_1 indicate the first ball is blue
W_2 indicate the second ball is white
R_3 indicate the third ball is red

RED	BLUE	WHITE	ORANGE	TOTAL
40	9	21	12	82

$$P(B_1 \cap W_2 \cap R_3) = \frac{9}{82} x \frac{21}{81} x \frac{40}{80}$$
$$= \frac{1}{82} x \frac{7}{3} x \frac{1}{2} = \frac{7}{492}$$

THE ONE AND ONLY ONE SITUATION

If A, B, C, are events with corresponding probability of occurring P(A), P(B), P(C) and probability of not occurring
$P(\bar{A}),\ P(\bar{B})\ \ and\ \ P(\bar{C})$.
Let a = event A occurs, B and C don't occur
b = event B occurs, A and C don't occur
c = event C occurs, A and B don't occur

$$\Rightarrow P(a\, or\, b\, or\, c) = P(a) + P(b) + P(c)$$
$$P(a) = P(A)P(\bar{B})P(\bar{C})$$
$$P(b) = P(\bar{A})P(B)P(\bar{C})$$
$$P(c) = P(\bar{A})P(\bar{B})P(C)$$
$$\therefore P(a \cup b \cup c) = P(A)P(\bar{B})P(\bar{C}) + P(\bar{A})P(B)P(\bar{C}) + P(\bar{A})P(\bar{B})P(C)$$

Example

When three marksmen take part in a shooting contest,

their chances of hitting a target are $\frac{1}{2}, \frac{1}{3}$, and $\frac{1}{4}$

a) Calculate the chance that one and only one bullet will hit the target if all men shoot simultaneously.

b) Determine the probability that the target will be hit (Assume independence)

Solution

Let E_1 = First man hits the target

E_2 = Second man hits the target

E_3 = Third man hits the target

Let N be only one bullet hits the target

$$P(E_1)=\frac{1}{2} \qquad P(\bar{E}_1)=\frac{1}{2}$$

$$P(E_2)=\frac{1}{3} \qquad P(\bar{E}_2)=\frac{2}{3}$$

$$P(E_3)=0.25 \qquad P(\bar{E}_3)=0.75$$

$$P(N)=P(E_1)P(\bar{E}_2)P(\bar{E}_3)+P(\bar{E}_1)P(E_2)P(\bar{E}_3)$$
$$+P(\bar{E}_1)P(\bar{E}_2)P(E_3)$$

$$P(N)=\frac{1}{2}x\frac{2}{3}x\frac{3}{4}+\frac{1}{2}x\frac{1}{3}x\frac{3}{4}+\frac{1}{2}x\frac{2}{3}x\frac{1}{4}$$

$$=\frac{1}{4}+\frac{1}{8}+\frac{1}{12}$$

$$=\frac{11}{24}$$

(b) $P(H) = P(E_1 \cup E_2 \cup E_3)$

The target can be hit by any three or two of them or all the three.

$P(E_1 \cup E_2 \cup E_3) =$

$$P(E_1)+P(E_2)+P(E_3)-P(E_1 \cap E_2)$$
$$-P(E_1 \cap E_3)-P(E_2 \cap E_3)+ P(E_1 \cap E_2 \cap E_3)$$
$$=\frac{1}{2}+\frac{1}{3}+\frac{1}{4}-\frac{1}{2}x\frac{1}{3}-\frac{1}{2}x\frac{1}{4}-\frac{1}{3}x\frac{1}{4}+\frac{1}{2}x\frac{1}{3}x\frac{1}{4}$$
$$=\frac{3}{4}$$

2.5 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 non occurrence of one event does not influence the non-occurrence of the other event.

Therefore $P(A \text{ and } B) = P(A \cap B) = P(A)P(B)$

Example

The probability that a man and his wife pass their driving test at first attempt is $\frac{1}{3}$ and $\frac{2}{5}$ respectively. Assuming that these events are independent, determine the probability that both pass the test at first attempt

Solution

Let M man passes and W woman passes

$P(M \cap W) = P(M)\,P(W)$

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

The probability that both pass $= \frac{2}{15}$

Example

The probability that a student S can solve a certain problem is 0.4 and that of student L can solve it is 0.5. Find the probability that the problem will be solved if both S and L try to solve it independently.

Solution

$P(S \cup L) = P(S) + P(L) - P(S \cap L)$

But $P(S \cap L) = P(S)P(L)$ *for independent events*

$P(S \cup L) = 0.4 + 0.5 - 0.4 x 0.5$

$= 0.9 - 0.2$

$= 0.7$

$\therefore$ *the probability that the probability will be solved* $= 0$

Example

A and B are two independent events, show that A and $\overline{B}$ are independent.

Solution

$P(A \cap B) = P(A)P(B)$ *for independent events*

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

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

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

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

$$=P(A)P(\overline{B})$$

$$Since\ \ P(\overline{B})=1-P(B)$$

$$Hence\ \ P(A\cap\overline{B})=P(A)P(\overline{B})$$

Therefore A and $\overline{B}$ are indpendent events

Example

Events A and C are independent. Probabilities relating to events A, B and C are as follows

$$P(A)=\frac{1}{5}, P(B)=\frac{1}{6}, P(A\cap C)=\frac{1}{20}, P(B\cup C)=\frac{3}{8}$$

Calculate P(C) and show that B and C are independent

Solution

A and C are independent events then

$P(A\cap C)=P(A)P(C)$

$$\frac{1}{20}=\frac{1}{5}P(C)\Rightarrow P(C)=\frac{1}{4}$$

$$But\ P(B\cap C)=P(B)+P(C)-P(B\cup C)$$

$$\frac{1}{6}+\frac{1}{4}-\frac{3}{8}=\frac{4+6-9}{24}=\frac{1}{24}$$

$$P(B)P(C)=\frac{1}{6}x\frac{1}{4}=\frac{1}{24}$$

$$P(B\cap C)=P(B)P(C)=\frac{1}{24}$$

Hence B and C are independent events

2.6 CONDITIONAL PROBABILITY

Suppose we have the information about 20 students given in table below

	Female	Male
Below 18 years old	5	6
Above 18 years old	1	8
Total	**6**	**14**

If a person is selected at random what is probability that is above 18 years

$$=\frac{9}{20}$$

Note if above 18 years is denoted by A and below by B.

Then M to denote male while F to represent female.

The expression P(B| M) is conditional probability that means

The conditional probability of below 18 years given a male.

Definition

If A and B are events, then the conditional probability of A given B is

$$P(A|B) = \frac{P(A \cap B)}{P(B)} \quad \text{provided } P(B) \neq 0$$

The example above

$$P(B|M) = \frac{P(B \cap M)}{P(M)}$$

$$P(B \cap M) = = 0.3$$

$$P(M) = \frac{14}{20}$$

$$P(B|M) = \frac{0.3}{0.7} = \frac{3}{7}$$

Example

If A and B are events and $P(B) = \frac{1}{6}$, $P(A \text{ and } B) = , \frac{1}{12}$

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

Calculate P(A), P(A|B) and $P(A|\overline{B})$

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

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

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

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

$$= \left(\frac{1/12}{\frac{1}{6}}\right)$$

$$P(A|B) = \frac{1}{2}$$

$$P(A \mid \bar{B}) = \frac{P(A \cap \bar{B})}{P(\bar{B})}$$

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

$$= \frac{\frac{1}{4} - \frac{1}{12}}{1 - \frac{1}{6}} = \frac{1}{6} x \frac{6}{5}$$

$$\therefore P\left(A \mid \bar{B}\right) = 0.2$$

Solution

Example

In a certain University

75% of the students are full-time students, 45% of the students are female, 40% of the students are male full-time students. Find the probability that

a) A student chosen at random from the students in the University is a part-time student,

b) A student chosen at random from all students in the University is female and a part-time student,

c) A student chosen at random from all the female students in the University is a part-time student.

Solution

$P(Female) = P(F) = 0.45,\ P(M) = 0.55$

$(a)\ P(S) = 0.75, P(\overline{S}) = 0.25$

$(b)\ P(F \cap S) = 0.75 - 0.4 = 0.35$

$But\ P(F \cap \overline{S}) = P(F) - P(F \cap S)$

$= 0.45 - 0.35$

$\therefore\ P(F \cap \overline{S}) = 0.10$

$$P(\overline{S} \mid F) = \frac{0.1}{0.45}$$

$$\Rightarrow P(\overline{S} \mid F) = \frac{2}{9}$$

Exercise

1. When a die is thrown the score was an even number. What is the probability that it was a prime number.

Answer: P(Prime, given Even) = $\frac{1}{3}$

2. A and B are two events such that $P(A|B) = 0.4$, $P(B) = 0.25$ and $P(A) = 0.2$. Find
   (a) $P(B|A)$ (b) $P(A \cap B)$ (c) $P(A \cup B)$
   Answer (a) 0.5 (b) 0.1 (c) 0.35
3. Events A and B are such that

$$P(A) = \frac{19}{30},\ P(B) = \frac{2}{5} \text{ and } P(A \cup B) = \frac{4}{5}$$

Find $P(A \cap B)$

Answer $P(A \cap B) = \frac{7}{30}$

4. The probability that a regular scheduled flight departs on time is 0.83 and the probability that it arrives on time is 0.92. The probability that it departs on time and arrives on time is 0.78. Find the probability that the plane;
   a) Arrives on time given that it departs on time
   b) Departs on time given that it arrives on time
   Answer (a) 0.94 (b) 0.85

5. Prove that for any two events A and B,

$$P(A|B) + P(\overline{A}|B) = 1$$

6. Three events A, B and C are such that A and B are independent, A and C are mutually exclusive.

Given that

$P(A) = 0.4,\ P(B) = 0.2,\ P(C) = 0.3\ and\ P(B \cap C) = 0.1$

Calculate

(i) $P(A \cup B)$ (ii) $P(C|B)$ (iii) $P(B|A \cup C)$

(iv) the probability that one and only one of the events B, C will occur.

Answer (i) 0.52 (ii) 0.5 (iii) 0.26 (iv) 0.3

7. A box contains 3 red, 2 green and 5 blue crayons. Two crayons are randomly selected from the box without replacement. Find the probability that:

i. The crayons are of the same colour.

ii. At least one red crayon is selected

Answer

i. P(same colour) $= \frac{14}{45}$

ii. P(at least one red crayon) $= \frac{8}{15}$

8. A bag contains 3 black and 5 white balls. 2 balls are drawn at random one at a time without replacement. Find

i. The probability that the second ball is white

ii. The probability that the first ball is white given that the second is white

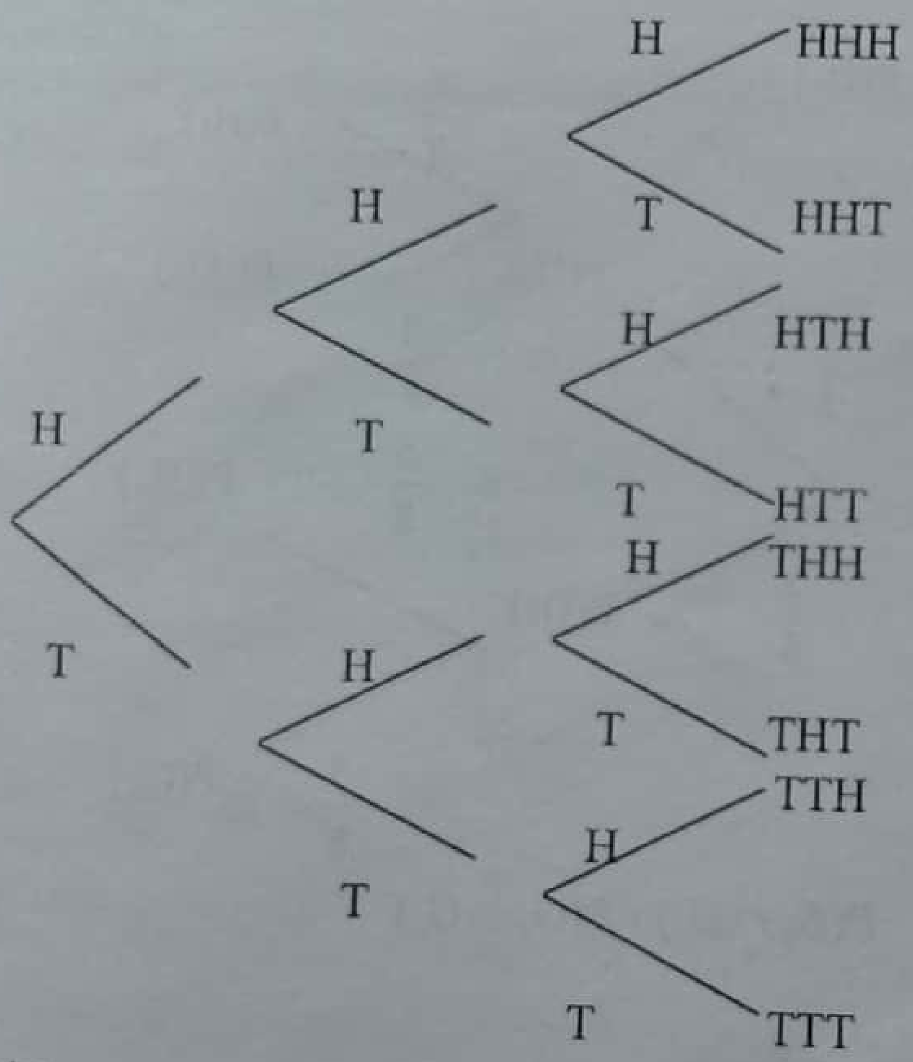
Answer $P(W_2) = \frac{5}{8}$, $(ii)\ P(W_1|W_2) = \frac{4}{7}$

2.7 PROBABILITY TREE DIAGRAMS

Tree diagrams can be used to obtain the possible outcomes of an experiment.

Example

Find the possible outcomes when three coins are tossed.
The possible outcomes are



Note:

The total probability for any one set of branches = 1
The sum of final probabilities is = 1

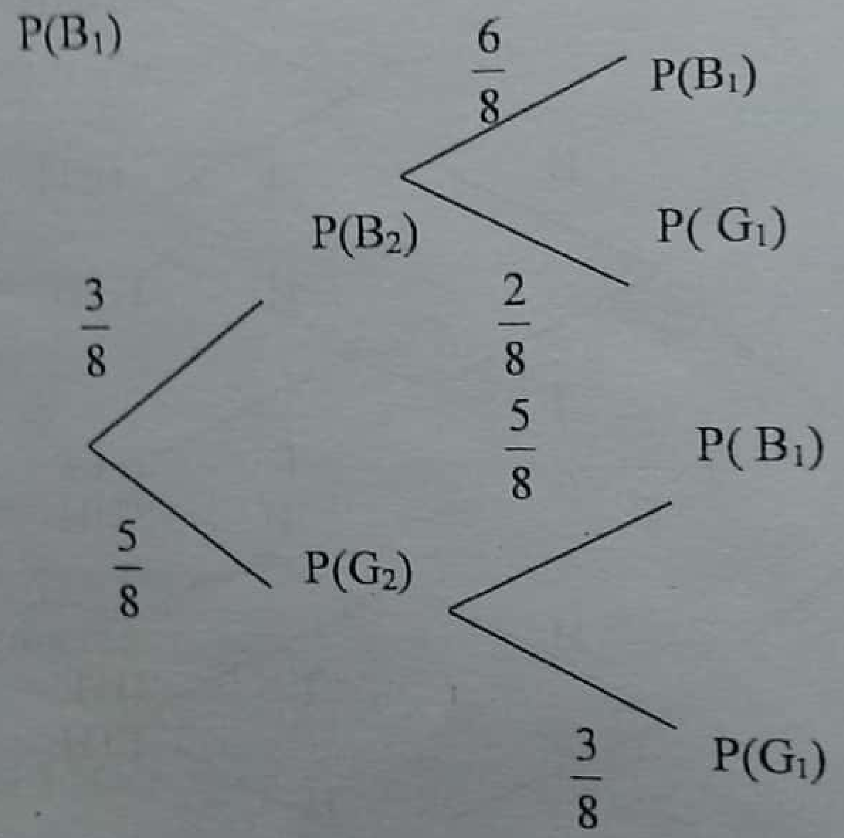
i. The tree shows conditional probabilities

Example

In an experiment, a box contains 2 green and 5 blue balls. A second box contains 5 green and 3 blue balls. One ball is drawn at random from the second box and placed into the first box. What is the probability that a ball drawn now from the first box is green?

Solution

Let B_2 indicate blue ball from second box,
B_1 indicates blue ball from first box.



$$P(G) = P(B_2 \cap G_1) + P(G_2 \cap G_1)$$

$$= \frac{3}{8} x \frac{2}{8} + \frac{5}{8} x \frac{3}{8}$$

$$= \frac{6}{64} + \frac{15}{64}$$

$$\therefore P(G) = \frac{21}{64}$$

Example

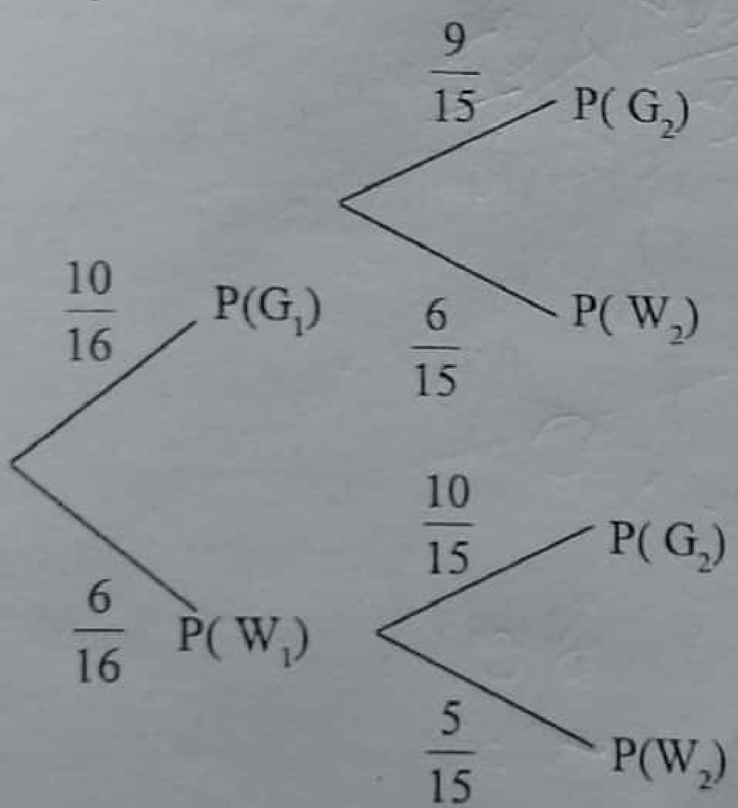
A box contains ten green and six white marbles. A marble is chosen at random, its colour noted and it is not replaced. This is repeated once more.

What is the probability that the marble chosen are of the same colour?

Solution

Let W_1 and W_2 indicate white marble is picked first and second respectively.

Let G_1 and G_2 indicate green marble is picked first and second respectively.



P(marble of the same colour) $= P(W_1 \cap W_2) + P(G_1 \cap G_2)$

$= \frac{6}{16} x \frac{5}{15} + \frac{10}{16} x \frac{9}{15}$

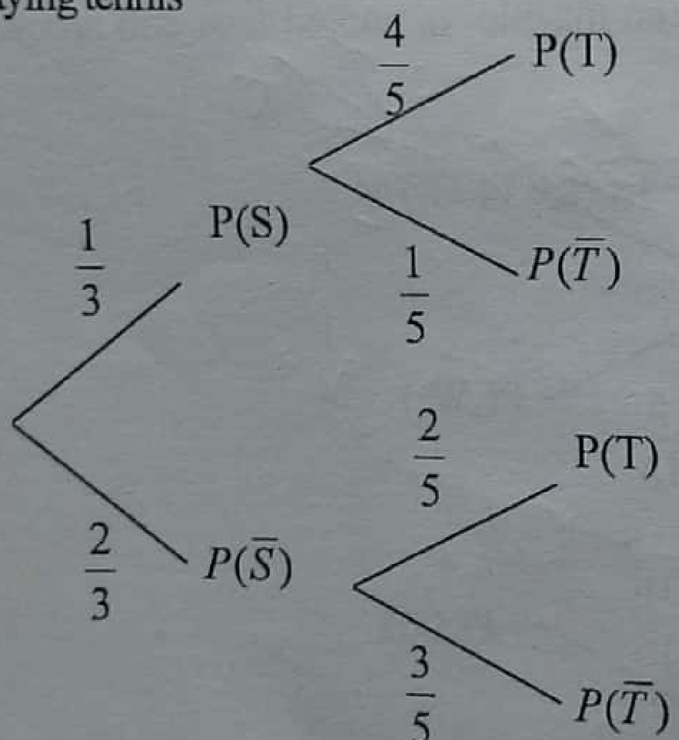
$= \frac{1}{2}$

Example

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

Solution

Let S represent sunny, denote not sunny,
T playing tennis



$$P(T) = P(S \cap T) + P(\overline{S} \cap T)$$

$$= \frac{1}{3} x \frac{4}{5} + \frac{2}{3} x \frac{2}{5}$$

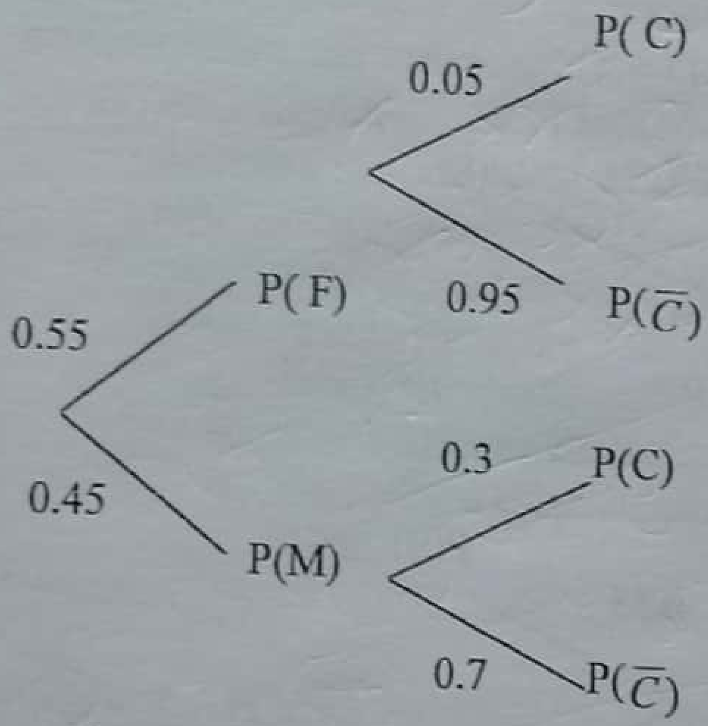
$$\therefore P(T) = \frac{8}{15}$$

Example

The proportion of female students at Kyambogo college is 55%. If 30% of the male students and 5% of female students study Chemistry. What is the probability that a chemistry student chosen at random is a girl?

Solution

Let F be for female, M for male while C is for chemistry



Example

$$P(F|C) = \frac{P(F \cap C)}{P(C)}$$

$$But\ P(C) = P(F \cap C) + P(M \cap C)$$
$$= 0.55x0.05 + 0.45x0.3$$

$$P(F|C) = \frac{0.55x0.05}{0.55x0.05 + 0.45x0.3}$$
$$= 0.1692$$

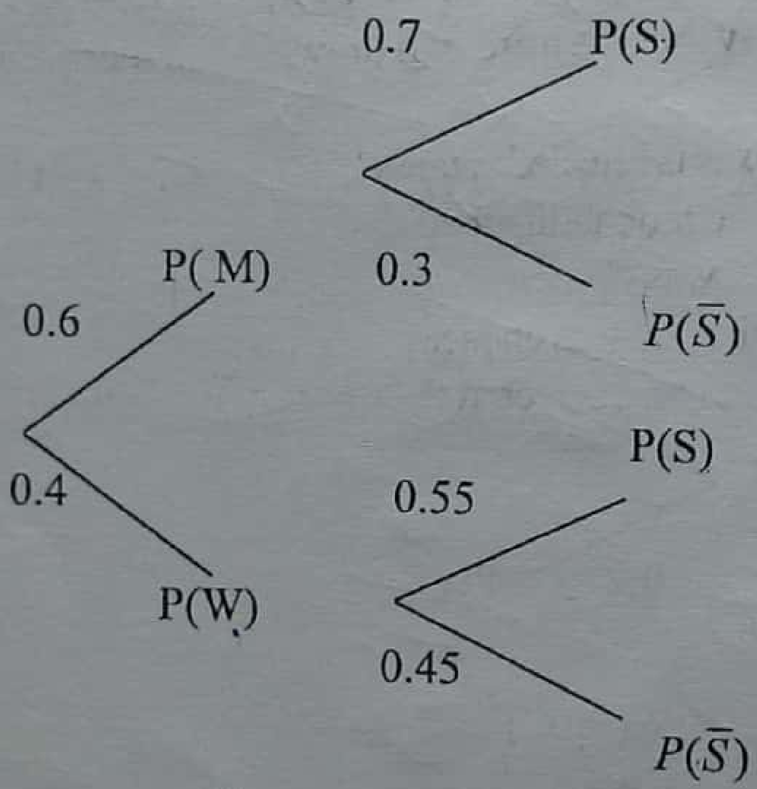
In a survey, 70% of men and 55% of women said that they smoked. If the proportion of men to women is 60 : 40 and a person from the survey was chosen at random and found to be a smoker, what is the probability that this person is a woman?

Solution

Let M and W denote man and woman respectively

Let S indicate a person is a smoker

The required probability is P(W|S)



Note:

$$P(M)=\frac{60}{100},\ P(W)=\frac{40}{100}$$

$$P(W \mid S)=\frac{P(W \cap S)}{P(S)}$$

$$\begin{aligned} But\ \ P(S)&=P(W \cap S)+P(M \cap S) \\ &=0.4x0.55+0.6x0.7 \\ &=0.64 \end{aligned}$$

$$P(W \mid S)=\frac{0.22}{0.64}$$

$$\therefore P(W \mid S)=\frac{11}{32}$$

Example

Three boxes X, Y and Z contain coloured balls. X contains 5 black balls and 4 white balls, Y contains 7 black and 5 white balls and Z contains 3 black and 5 white balls.

a) If the balls are withdrawn from box Z, with replacement, find the probability that the third ball drawn is the second white ball.

b) One of the boxes is selected at random and a ball is withdrawn from it. Find the probability that:

   i. Box X was chosen and the ball was black

   ii. A white ball was chosen

   iii. The ball was selected from box Z, given that it was black.

Solution

For Box Z, $\text{P(black)} = \text{P(B)} = \frac{3}{8}$

$$\text{P(white)} = \text{P(W)} = \frac{5}{8}$$

P(the third ball is the second white one)

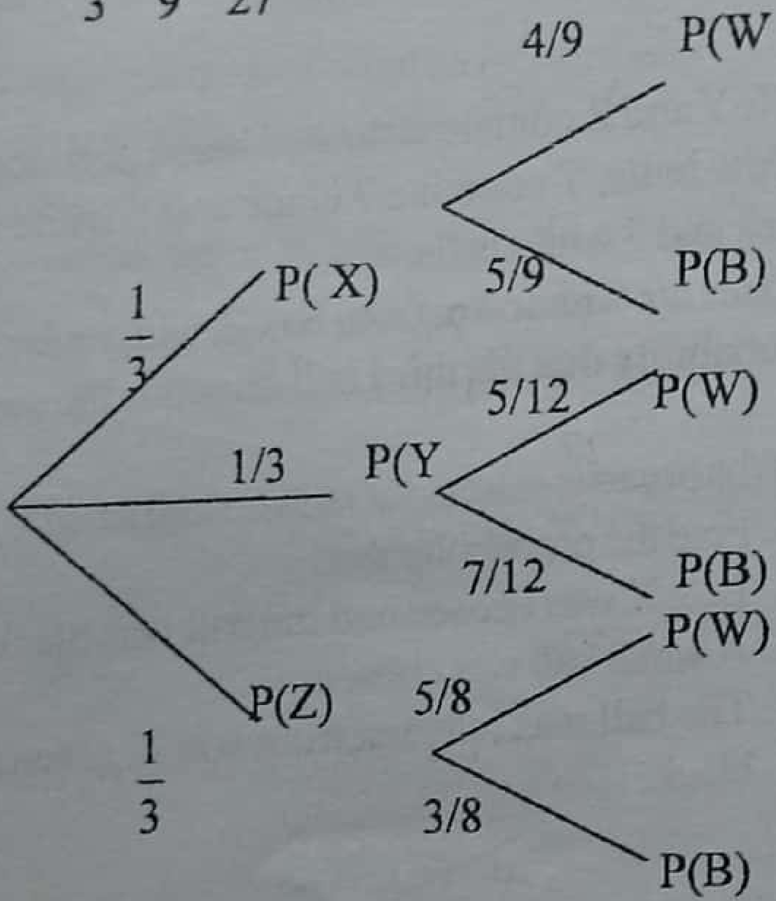
$$\Rightarrow P(BWW)+P(WBW)$$

$$P(BWW)+P(WBW)=\frac{3}{8}x\frac{5}{8}x\frac{5}{8}+\frac{5}{8}x\frac{3}{8}x\frac{5}{8}$$

$$\frac{75}{512}+\frac{75}{512}=\frac{75}{256}$$

b(i) $P(X \cap B) = P(X)\,P(B|X)$

$$= \frac{1}{3} x \frac{5}{9} = \frac{5}{27}$$



b(ii)

$$P(W) = P(X \cap W) + P(Y \cap W) + P(Z \cap W)$$

$$P(W) = \frac{1}{3} x \frac{4}{9} + \frac{1}{3} x \frac{5}{12} + \frac{1}{3} x \frac{5}{8}$$

$$P(W) = \frac{107}{216}$$

(iii) $P(Z|B) = \dfrac{P(Z \cap B)}{P(B)}$

But $P(W) + P(B) = 1$

$$P(B) = 1 - \frac{107}{216} = \frac{109}{216}$$

$$P(Z \cap B) = \frac{1}{3} x \frac{3}{8} = \frac{1}{8}$$

$$P(Z \mid B) = \frac{1}{8} x \frac{216}{109} = \frac{27}{109}$$

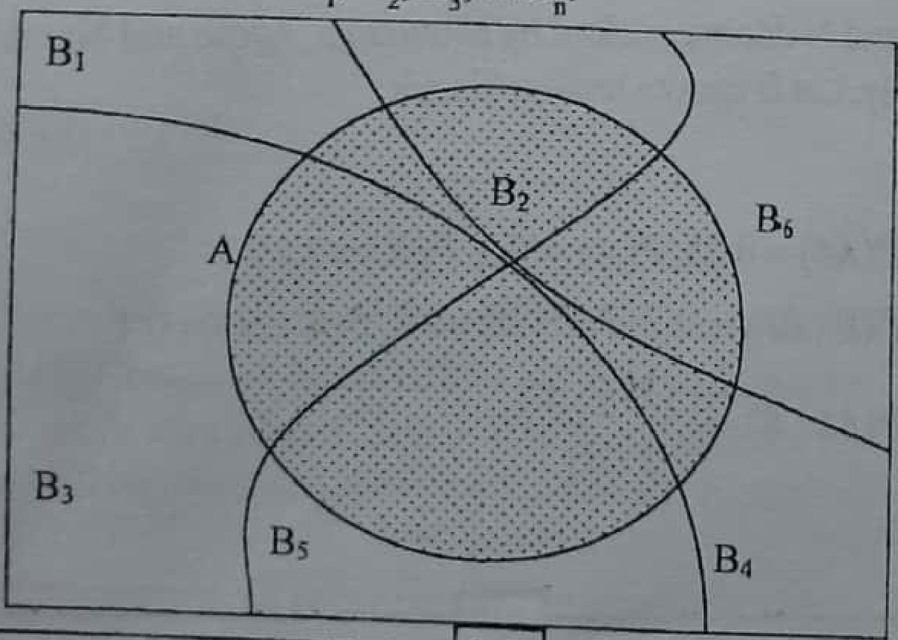
$$\therefore\ P(Z \mid B) = \frac{27}{109}$$

2.8 BAYES' THEOREM

This theorem is obtained from the extension of the conditional probability.

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

Suppose S contains $B_1, B_2, B_3, \ldots B_n$.



Event A can be given by:

$$P(A)=P(B_1\cap A)+P(B_2\cap A)+....+P(B_n\cap A)$$

$$\therefore P(B_i\mid A)=\frac{P(B_i\cap A)}{P(A)}=$$

$$\frac{P(B_i)P(A\mid B_i)}{P(B_1\cap A)+P(B_2\cap A)+...+P(B_n\cap A)}$$

$$\therefore P(B_i\mid A)=\frac{P(B_i)P(A\mid B_i)}{\sum_{k=1}^{n}P(B_k\ P(A\mid B_k)}$$

Where i = 1, 2, 3, …n.

Example

Three girls, Muwanga, Ayolo and Nambi pack biscuits in a factory, from the batch allotted to them, Muwanga packs 55%, Ayolo 30% and Nambi 15%. The probability that Muwanga breaks some biscuits in a packet is 0.7, and the respective probabilities of Ayolo and Nambi are 0.2 and 0.1. What is the probability that a packet with broken biscuits found by the checker was packed by Muwanga

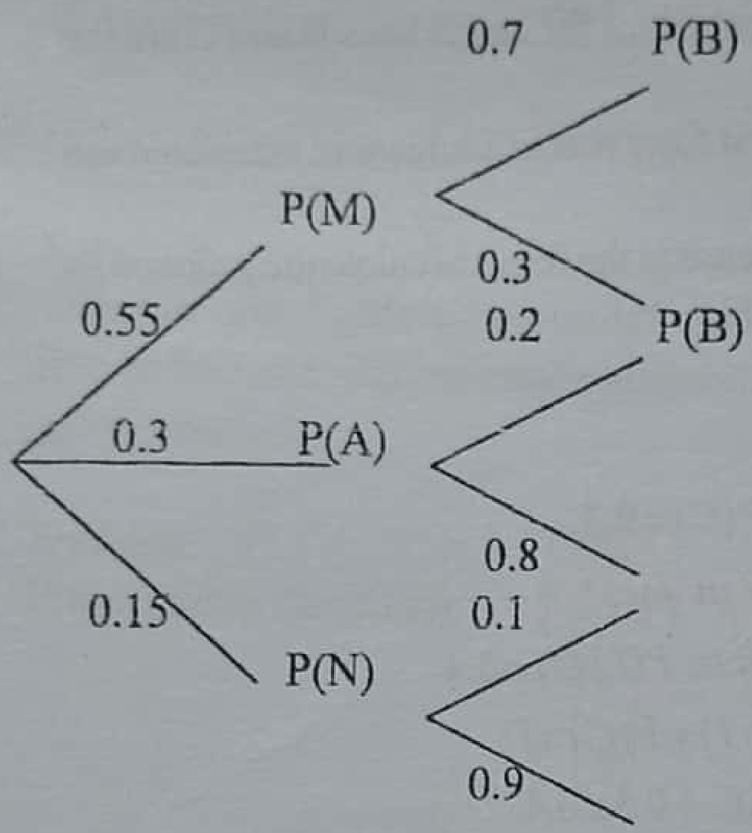
Solution

Let M, A and N denote packed by Muwanga, Ayolo and Nambi respectively. Let B denote broken biscuit.

$$P(M)=0.55, P(A)=0.3, P(N)=0.15$$

$$P(B\mid M)=0.7, P(B\mid A)=0.2, P(B\mid N)=0.1$$

$$P(M\mid B)=\frac{P(M\cap B)}{P(B)}$$



$$P(B) = P(M \cap B) + P(A \cap B) + P(N \cap B)$$

$$= 0.55 x 0.7 + 0.3 x 0.2 + 0.15 x 0.1 = 0.46$$

$$P(M \cap B) = 0.55 x 0.7 = 0.385$$

$$\Rightarrow P(M \mid B) = \frac{P(M \cap B)}{P(B)}$$

$$= \frac{0.385}{0.46}$$

$$\therefore P(M \mid B) = 0.84$$

xample

hree candidates have been nominated for post of welfare prefect in hool. The probability that candidate A will be elected is 0.1, that for ndidate B is 0.2 while candidate C is 0.3. It is expected that, the tertainment fee will be increased if any one of these is elected as prefect. The probability of an increase in the fee when A is elected is 0.5, the corresponding probabilities for candidates B and C are 0.6 and 0.4 respectively.

Determine the probability that there was no increase in entertainment fees?

Given that there was an increase in the fees, calculate the probability that candidate A was elected?

Solution

$P(A)=0.1, P(B)=0.2, P(C)=0.3,$

Let I represent increase in fees

$P(I \mid A)=0.5, P(I \mid B)=0.6,\ P(I \mid C)=0.4$

$P(I)=P(A\cap I)+P(B\cap I)+P(C\cap I)$

$=0.1x0.5+0.2x0.6+0.3x0.4$

$=0.29$

$\Pr obability\ of\ no\ increase\ in\ fees=1-0.29=0.71$

$$P(A \mid I)=\frac{P(A\cap I)}{P(I)}$$

$$=\frac{0.1x0.5}{0.29}$$

$=0.1724$

2.9 Calculations involving combinations

Example

In class with 5 boys and 3 girls. Two students are selected at random from the class. What is the probability that they are all girls.

Solution

The two girls can be selected in ${}^3C_2$ ways from the girls.

But two students can be selected in $^8C_2$ ways from the class

$$P(\text{All girls}) = \frac{^3C_2}{^8C_2} = \frac{3}{28}$$

Example

There are only 3 girls in a group of 8 students. A group of 5 students is to be selected. Determine the probability that two girls are in the group selected.

Solution

$$P(\text{two girls in the group}) = \frac{\binom{3}{2}\binom{5}{3}}{\binom{8}{5}} = \frac{3x10}{56}$$

$$= \frac{15}{28}$$

Example

A bag contains 5 green, 4 yellow and 3 blue balls, from which 4 balls are picked at random. Determine the probability that the 4 balls selected will contain.

i. Exactly 3 yellow balls
ii. At least one green ball

Solution

$$P(Three\ \ yellow\,balls) = \frac{^4C_3\,x\,^8C_1}{^{12}C_4}$$

$$= \frac{4x8}{495} = \frac{32}{495}$$

P(At least one green)

$$= \frac{{}^5C_1 x\, {}^7C_3}{{}^{12}C_4} + \frac{{}^5C_2 x\, {}^7C_2}{{}^{12}C_4} + \frac{{}^5C_3 x\, {}^7C_1}{{}^{12}C_4} + \frac{{}^5C_4}{{}^{12}C_4}$$

$$= \frac{175}{495} + \frac{210}{495} + \frac{70}{495} + \frac{5}{495}$$

$$= \frac{92}{99}$$

2.10 OTHER EXAMPLES

When it is fine day, the probability that Aisha plays basket ball is 0.9 and the probability that Zakia plays is 0.75. If it is not fine, Aisha's probability is 0.5 and Zakia's is 0.25. Their chances are independent. In general it is twice likely to be fine as not.

i. Determine the probability that they both go to play
ii. If they both go to play, what is the probability that it is a fine day?

Solution

Let F, A and Z represent fine, Aisha and Zakia respectively

$$P(A \mid F) = \frac{9}{10}, \quad P(Z \mid F) = \frac{3}{4},$$

$$P(A \mid \bar{F}) = \frac{1}{2}, \quad P(Z \mid \bar{F}) = \frac{1}{4}$$

$$P(F) = \frac{2}{3}, \quad P(\bar{F}) = \frac{1}{3}$$

$P(F)=\frac{2}{3},\ P(\overline{F})=\frac{1}{3}$

P(Both Aisha and Zakia play)

$= P(A\cap Z\cap F)+\ P(A\cap Z\cap \overline{F})$

$= \frac{2}{3}x\frac{9}{10}x\frac{3}{4}+\frac{1}{3}x\frac{1}{4}x\frac{1}{2}$

$= \frac{59}{120}$

$P(F\mid A\cap B)=\frac{P(A\cap Z\cap F)}{P(A\cap B)}$

$=\frac{54}{120}x\frac{120}{59}$

$=\frac{54}{59}$

Example

The table below shows the likely hood where Linda and Sheila spend their Satarday evening

	Linda	*Sheila*
Go dancing	$\frac{1}{2}$	$\frac{2}{3}$
Visit neighbour	$\frac{1}{3}$	$\frac{1}{6}$
Stays at home	$\frac{1}{6}$	$\frac{1}{6}$

i. Find the probability that they both go out?
ii. If we know that they both go out, what is the probability that they went dancing?

Solution

$$P(L_{out})=\frac{1}{2}+\frac{1}{3}=\frac{5}{6}, \quad P(S_{out})=\frac{2}{3}+\frac{1}{6}=\frac{5}{6}$$

$$P(L_{out}\cap S_{out})=\frac{5}{6}x\frac{5}{6}=\frac{25}{36}$$

P(Going to dance | Both go out)

$$=\frac{P(Going\ to\ dance)}{P(Both\ Going\ out)}$$

$$=\frac{1}{2}x\frac{2}{3}x\frac{36}{25}=\frac{12}{25}$$

Example

The probability that two events occur is $\frac{2}{15}$ and the probability that either or both events occur is $\frac{3}{5}$. Determine the individual probabilities of both events.

Solution

$$P(A\cap B)=P(A)P(B)=\frac{2}{15}$$

$$\Rightarrow P(B)=\frac{2}{15P(A)} \quad \text{if let P(A) =Y}$$

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

$$\frac{3}{8} = Y + \frac{2}{15Y} - \frac{2}{15}$$

$$15Y^2 - 11Y + 2 = 0$$

$$Y = \frac{11 \pm 1}{30} = \frac{2}{5} \text{ or } \frac{1}{3}$$

$$\therefore P(A) = \frac{2}{5} \text{ and } P(B) = \frac{1}{3}$$

$$\text{Or } P(A) = \frac{1}{3} \text{ and } P(B) = \frac{2}{5}$$

Example

Bag A contains 2 green and 2 blue balls, while bag B contains 2 green and 3 lue balls. A bag is selected at random and two balls drawn from it without replacement. Find the probability that the balls are of different colours.

Solution

$$P(B_A \cap G_A \cap B_A) + P(B_A \cap B_A \cap G_A) +$$
$$P(B_B \cap G_B \cap B_B) + P(B_B \cap B_B \cap G_B)$$

$$= \frac{1}{2} x \frac{2}{4} x \frac{2}{3} + \frac{1}{2} x \frac{2}{4} x \frac{2}{3} + \frac{1}{2} x \frac{2}{5} x \frac{3}{4} + \frac{1}{2} x \frac{3}{5} x \frac{2}{4}$$

$$= \frac{1}{6} + \frac{1}{6} + \frac{3}{20} + \frac{3}{20}$$

$$P(\text{balls are of different colours}) = \frac{19}{30}$$

Exercise

1. (a) Define the independence of two events A and B. Given that A and B are independent events in a sample space such that

i. $P(A)=\frac{2}{5}$, *and* $P(A\cup B)=0.8$, *find*

ii. $P(B)$ $\quad P(\overline{A}\cup\overline{B})$

(b) In a certain town the probability that a person owns a car is 0.25. Given that the probability that a person who owns a car is a university graduate is 0.2. Find the probability that a person selected at random owns a car and is a university graduate.

Answer (a) (i) P(B) = $\frac{2}{3}$, $\quad$ $(ii)\ P(\overline{A}\cup\overline{B})=\frac{11}{15}$

(b) $P(C\cap G)=0.05$

2. In a certain country, 60% of the cars are privately owned, of the privately owned cars 70% are small, whereas of the cars, which are not privately owned 40% are small. Independent of ownership 20% of the small cars and 30% of the large cars are less than two years old. If a car is choosen at random, calculate the probability that:

i. The car is small
ii. The car is privately owned given that it is large
iii. The car is privately owned, large and more than two years old.
iv. The car is large, given that it is privately owned and more than to years old.

Answer $(i)\ P(S)=0.58,$ $\quad$ $(ii)\ P(P\mid L)=0.43$

$(iii)\ P(P\cap L\cap O)=0.126$ $\quad$ $(iii)\ P(L\mid P\cap O)=0.273$

3. Tom is to travel from Lira to Kampala for an interview. The probabilities that he will be in time for the interview when he travels by bus and taxi are 0.1 and 0.2 respectively. The probabilities that he will travel by bus and taxi are 0.6 and 0.4 repectively.

   i. Find the probability that he will be on time.
   ii. Given that he is not on time what is the probability that he traveled by a taxi?

Answer

(i) $P(T) = 0.14$ (ii) $P(t \mid \bar{T}) = 0.372$

4. When visiting a friend Grace may go by road, air or rail. The proability of using road, air and rail are 0.3, 0.8 and 0.6 respectively. The corresponding probabilities of arriving on agreed time are 0.2, 0.8 and 0.1 respectively. Find the probability of having used the road given that he arrived on time.

Answer $P(R|T) = 0.08$

5. The events A and B are such that $P(A) = 0.5$,

$P(\text{A or B but not both A and B}) = \frac{1}{3}$ $P(B) = 0.25$.

Calculate $P(A \cap B),\ P(\bar{A} \cap B),\ P(A \mid B)$

where $\bar{A}$ is A doesnot occur.

State with reasons whether A and B are:

i. Independent
ii. Mutually exclusive

Answer

$$P(A \cap B) = \frac{5}{24}, \quad P(\bar{A} \cap B) = \frac{1}{24}, \quad P(A \mid B) = \frac{5}{6}$$

6. George Wear was a striker for AC Milan. In any match he may score, 0, 1, 2, 3 and 4 goals for his side with probabilities 0.2, 0.4, 0.2, 0.1 and 0.1 respectively. Assuming that his performance in successive matches are independent, what is the probability that in three consecutive matches:

   i. He scores a total of 10 goals or more

   ii. He does not score exactly three goals in any match

Answer (i) 0.013 (ii) 0.729

7. 95% of car drivers wears seat belts, 64% of cars drivers involved in serious accidents die if not wearing seat belts, where as 12% of those that do wear seat belt die. Calculate to three significant figures, the percentage of drivers involved in serious accidents who died and were not wearing seat belts.

Answer 21.9%

8. A court in Kampala may return any one of these verdicts namely Guilty(G), Not guilty(N) or Not proven(V). Of all the cases tried by this court, 70% of the verdicts were guilty, 20% were not guilty and 10% not proven. Suppose that when the courts verdict is guilty, not guilty and not proven, the respective probabilities of the accused person being innocent are 0.05, 0.95 and 0.25. Find the

i. probability that the person is innocent

ii. an innocent person being tried by this court is found guilty

Answer (i) 0.25 (ii) 0.14

9. Two cards are drawn without replacement from ten cards, which are numbered fom 1 to 10. Find the probability that:

i. the numbers on both cards are even

ii. the number on one card is odd and the number on the other is even.

iii. The sum of the numbers on the two cards exceeds 4.

Answer (i) $\frac{2}{9}$ (ii) $\frac{5}{9}$ (iii) $\frac{43}{45}$

10. Two independent events are such that the probability of both is $\frac{1}{6}$ and that of neither is $\frac{1}{3}$. Determine the probability of each of the events occurring?

Answer $P(A)=\frac{1}{3}$, $P(B)=\frac{1}{2}$, Or $P(A)=\frac{1}{2}$, $P(B)=\frac{1}{3}$

11. Three balls are chosen at random without replacement from a bag containing 3 red, 8 blue and 7 white balls. Determine the probability that the balls chosen will be

(i) all blue (ii) all red (iii) one of each colour

Answer (i) $\frac{7}{102}$, (ii) $\frac{1}{816}$, (iii) $\frac{7}{34}$

12. A bag contains 5 white, 3 red and green counters. 3 counters are drawn without replacement. What is the probability there

(i) is no green counter

(ii) are 2 white counters and a green counter

Answer (i) $\frac{7}{15}$ (ii) $\frac{1}{6}$

13. A football star has injury problems. When he is playing the probability that his team will win is $\frac{3}{4}$ but otherwise it is only $\frac{1}{2}$. The probability that the player will be fit this weekend is $\frac{1}{3}$.

Determine the probability that his team will win the match.

Answer $\frac{7}{12}$

14. Three dice are to be rolled. Determine the probability of scoring a double but not a triple

Answer $\frac{5}{12}$

15. Mariam's chances of passing physics are 0.6, of economics 0.75 and of mathematics 0.80.

i. Determine the chance that she passes one subject only
ii. If it is known that she passed at least two subjects, what is the probability that she failed economics?

Answer (i) 0.17 (ii) $\frac{4}{27}$

16. There are 3 black and 2 white balls in each of the two bags. A ball is taken from the first bag and placed in the second, and then a ball is taken from the second into the first, what is the probability that there is now the same number of black and white balls in each bag as there to begin with?

Answer 0.6

17. A bag contains 5 black beads and 3 white beads. A second bag contains 3 black beads and 5 white beads. A bead is drawn at random from the first bag and placed in the second bag. A bead is now drawn from the second bag and placed in the first. Find the probability that each bag now contains

a) 4 black and 4 white beads

b) The same numbers of each colour as it did initially

Answer (a) $\frac{25}{72}$ $(b)\frac{19}{36}$

18. A box A contains 3 red balls and 4 black balls. A box B contains 3 red balls and 2 black balls. One box is selected at random and then from it, one ball is selected at random.

(a) Find the probability that the ball is red,

(b) the probability that the ball came from box A, given that it is red.

Answer (a) $\frac{18}{35}$, $(b)\frac{5}{12}$

19. A school is divided into two sections: A-level, 400 boys and 200 girls, O-level, 400 girls and 300 boys. A student is selected at random from the school. If this student comes from O-level, a second student is selected from the A- level; if the first students comes from the A- level, the second student is chosen from the O-level. Find the probability that

a) The second student will be a girl

b) If the second student is a boy, he is a member of A-level.

Answer (a) $\frac{121}{273}$ (b) $\frac{49}{76}$

20. A boat hiring company at the local boating lake has two types of crafts 20 blue and 35 herons. The customer has to take the next boat available when hiring. The boats are all distinguishable

by their numbers. A regular customer Amina carefully notes that the numbers of the boats, which she uses, and finds that she has used 15 different blue birds and 20 different herons. Each boat is equally likely to be the next in line. Amina hires two boats at the same time. (one is for a friend). If event X is: Amina has not hired either boat before, and event Y is: both are herons, determine:

(a) P(X) (b) P(Y) (c) P(X|Y)

(d) whether the events are mutually exclusive

Answer

(a) $P(X)=\frac{38}{297}$, $P(Y)=\frac{119}{297}$, $P(X|Y)=\frac{3}{17}$

(d) $P(X\cap Y)\neq 0$, hence not mutually exclusive

21. Two girls Faridah and Monica play a game in which each throws a tetrahedron in turn. The first to get a 3 wins, Faridah tries first. What is the probability that

i. Monica wins on his second trial

ii. Faridah wins

(b) A player can play on either of the two gambling machines A and B. He chooses one of the machines at random, and plays two games. The probability of winning a game on A is $\frac{1}{3}$ and the probability of winning on B is $\frac{1}{4}$. If he loses both of these two games, he plays a third game on the other machine; otherwise he plays the third game on the same machine. Find the probability that he

i. Wins the first game

ii. Changes the machine after the second game

iii. Plays the third game on A

iv. Wins the third game

Answer $(a)\ \frac{27}{256},\ \frac{4}{7}$ $(b)\ (i)\ \frac{7}{24},\ (ii)\ \frac{145}{288}$

$(iii)\ \frac{161}{288}$ $(iv)\ 0.3242$

22. A bag contains 4 white balls and 1 black ball. A second bag contains 1 white ball and 4 black balls. Aball is drawn at random from the first bag and put into the second bag, then a ball is taken from the second bag and put into the first bag. Find the probability that a white ball will be picked when a ball is selected from the first bag.

Answer 0.7

CHAPTER THREE

DISCRETE PROBABILITY DISTRIBUTION

3.1 INTRODUCTION

A probability density function $P(X=x)$ is discrete if it has countable domain. A random variable following such a probability density function (P.D.F) is called a discrete random variable.

Example

Write the probability distribution of the score when an ordinary die is thrown

Solution

(i) $S=(1, 2, 3, 4, 5, 6)$

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

Properties of discrete probability distribution

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

Total probability is 1

(ii) $P(X=x)\geq 0$ for all values of X.

Note: $P(X=x)=f(x)$. f(x) can be used to represent a discrete probability density function.

Example

A random variable X has the following probability distribution

x	1	2	3	4
P(X = x)	$\frac{1}{5}$	$\frac{3}{10}$	$\frac{1}{10}$	$\frac{2}{5}$

Determine

(i) $P(X>2)$, (ii) $P(<2)$ (iii) $P(X\leq 2)$

Solution

(i) $P(X>2) = P(X=3) + P(X=4)$

$$= \frac{1}{10} + \frac{2}{5}$$

$$P(X>2) = 0.5$$

(ii) $P(X<2) = P(X=1)$

$$P(X<2) = 0.2$$

(iii) $P(X\leq 2) = P(X=1) + P(X=2)$

$$= \frac{1}{5} + \frac{3}{10}$$

$$P(X\leq 2) = 0.5$$

Example

A discrete random variable X has P.D.F as shown below

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

Determine

i. Value of a

ii. $P(1<X<3)$

iii. $P(2 \le X \le 4)$

Solution

(i) $\sum P(X=x)=1$

$$0.2+0.25+0.4+a+0.05=1$$

$$0.9+a=1$$

Therefore $a=0.1$

(ii) $P(1<X<3)=P(X=2)$

$P(1<X<3)=0.25$

(iii) $P(2 \le X \le 4) = P(X=2)+P(X=3)+P(X=4)$

$$=0.25+0.4+0.1$$

$P(2 \le X \le 4)=0.75$

Example

Two biased tetrahedrons have each of their faces numbered 1 to 4. The chances of getting anyone face showing uppermost is inversely proportional to the number on it. If the two tetrahedrons are drawn and the number on the uppermost face noted, determine the probability that the faces show the same number.

Solution

$P(X=x) \; \alpha \; \frac{1}{x} \Rightarrow P(X=x)=k\frac{1}{x}$

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

$$k+\frac{k}{2}+\frac{k}{3}+\frac{k}{4}=1$$

$$12k+6k+4k+3k=12$$

$$25k=12$$

$$k=\frac{12}{25}$$

$$P(\text{Same number})=P(1\text{ and }1)+P(2\text{ and }2)$$
$$+P(3\text{ and }3)+P(4\text{ and }4)$$

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

$$=k^2(1+\frac{1}{4}+\frac{1}{9}+\frac{1}{16})$$

$$=(\frac{12}{25})^2(\frac{205}{144})=\frac{205}{625}$$

$$P(\text{same number})=\frac{41}{125}=0.328$$

Example

The probability distribution of a random variable X is given by the function

$$f(x)=\frac{1}{84}\binom{5}{x}\binom{4}{3-x}\quad x=0,1,2,3$$

calculate the numerical probabilities?

Solution $\binom{n}{r}$ *mean combinations*

Note:

$$f(0)=\frac{1}{84}\binom{5}{0}\binom{4}{3-0}$$

$$= \frac{1}{84} x 1 x 4 = \frac{4}{84}$$

$$f(1) = \frac{1}{84}\binom{5}{1}\binom{4}{3-1}$$

$$= \frac{1}{84} x 5 x 6 = \frac{30}{84}$$

$$f(2) = \frac{1}{84}\binom{5}{2}\binom{4}{3-2}$$

$$= \frac{1}{84} x 10 x 4 = \frac{40}{84}$$

$$f(3) = \frac{1}{84}\binom{5}{3}\binom{4}{3-3}$$

$$= \frac{1}{84} x 10 x 1 = \frac{10}{84}$$

x	0	1	2	3
P(X= x)	$\frac{4}{84}$	$\frac{30}{84}$	$\frac{40}{84}$	$\frac{10}{84}$

3.2 MEAN OR EXPECTED VALUE. E(X)

$$E(x) = \sum xP(X = x)$$

The symbol for mean is E(X) or μ

3.3 VARIANCE (VAR(X))

$\text{Var}(X) = EX^2 - (E(X))^2$

Where $EX^2 = \sum x^2 P(X = x)$

and $E(x) = \sum xP(X = x)$

Example

A fair coin is tossed four times and X represents the number of tails that appear. Determine the probability distribution of X and hence expected number of tails.

Solution

$$S = \begin{Bmatrix} TTTT, TTTH, TTHT, THTT, HTTT, \\ TTHH, THTH, HTHT, HHTT, THHT, \\ HTTH, HHHT, HHTH, HTHH, \\ THHH, HHHH \end{Bmatrix}$$

X	$P(X=x)$	$xP(X=x)$
0	$\frac{1}{16}$	0
1	$\frac{4}{16}$	$\frac{4}{16}$
2	$\frac{6}{16}$	$\frac{12}{16}$
3	$\frac{4}{16}$	$\frac{12}{16}$
4	$\frac{1}{16}$	$\frac{4}{16}$
Total	**1**	**2**

Therefore expected value = 2

Example

The random variable X takes on integral values only and had P.D.F given by $P(X=x)=kx \quad x=1,2,3,4,5$

$P(X=x)=k(10-x) \quad x=6,7,8,9$

Find

a) The value of the constant k.

b) Expected value of X

c) Var(X)

Solution

x	x^2	$P(X=x)$	$x\,P(X=x)$	$x^2P(X=x)$
1	1	k	k	k
2	4	2k	4k	8k
3	9	3k	9k	27k
4	16	4k	16k	64k
5	25	5k	25k	125k
6	36	4k	24k	144k
7	49	3k	21k	147k
8	64	2k	16k	128k
9	81	K	9k	81k
Total		25k	125k	725k

(a) $25k=1$

$$\Rightarrow k=\frac{1}{25}$$

(b) $E(X)=\sum xP(X=x)$

$E(X)=125k$

$$= 125x\frac{1}{25} = 5$$

$E(X) = 5$

(c) $Var(X) = EX^2 - (E(X))^2$

$$= (725\ x\frac{1}{25}) - 5^2$$

$$29 - 25$$

$$= 4$$

Variance (x) $= 4$

Example

At school, projector shows follow a discrete random variable. 3 students, Aisha, Brenda and Cathy will watch the show with probability of $\frac{1}{4}$, $\frac{1}{3}$, *and* $\frac{2}{5}$ respectively.

The decision of a student makes to watch the show is independent of the other students.

Find

(i) the probability function of the number of students who watch the projector show

(ii) Hence the expected number of students who watch the Show

Solution

Let X be discrete random variable representing the number of students who watch the show and A, B and C represent the respective events that a student watches the show

$$P(X=0)=P(\bar{A}\cap\bar{B}\cap\bar{C})=P(\bar{A})P(\bar{B})P(\bar{C})$$

$$\frac{3}{4}x\frac{2}{5}x\frac{3}{5}=\frac{3}{10}$$

$$P(X=1)=P(A\cap\bar{B}\cap\bar{C})+$$

$$P(\bar{A}\cap B\cap\bar{C})+P(\bar{A}\cap\bar{B}\cap C)$$

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

$$\frac{1}{10}+\frac{3}{20}+\frac{1}{5}=\frac{9}{20}$$

$$P(X=2)=P(A\cap B\cap\bar{C})+P(A\cap\bar{B}\cap C)+$$

$$P(\bar{A}\cap B\cap C)$$

$$=\frac{1}{4}x\frac{1}{3}x\frac{3}{5}+\frac{1}{4}x\frac{2}{3}x\frac{2}{5}+\frac{3}{4}x\frac{1}{3}x\frac{2}{5}$$

$$\frac{1}{20}+\frac{1}{15}+\frac{2}{20}=\frac{13}{60}$$

$$P(X=3)=P(A\cap B\cap C)$$

$$=\frac{1}{4}+\frac{1}{3}+\frac{2}{3}=\frac{1}{30}$$

Probability function

x	0	1	2	3
P(X=x)	$\frac{3}{10}$	$\frac{9}{20}$	$\frac{13}{60}$	$\frac{1}{30}$

Expected numbers

x	0	1	2	3
P(X=x)	$\frac{3}{10}$	$\frac{9}{20}$	$\frac{13}{60}$	$\frac{1}{30}$
xP(X=x)	0	$\frac{9}{20}$	$\frac{26}{60}$	$\frac{1}{10}$

Means $\frac{59}{60} \cong 1$

Example

A box contains 4 red balls and 2 blue balls. A ball is drawn without replacement until a blue is drawn. If X represents the number of draws required to draw a blue ball, find the probability distribution of X and hence mean and variance.

$$P(X=1) = P(B_1) = \frac{1}{3}$$

$$P(X=2) = P(R_1B_2) = \frac{2}{3} x \frac{2}{5}$$

$$P(X=3) = P(R_1R_2B_3) = \frac{2}{3} x \frac{3}{5} x \frac{2}{4}$$

$$P(X=4) = P(R_1R_2R_3B_4) = \frac{2}{3} x \frac{3}{5} x \frac{2}{4} x \frac{2}{3}$$

$$P(R_1R_2R_3R_4B_5) = \frac{2}{3} x \frac{3}{5} x \frac{2}{4} x \frac{1}{3} x1$$

x	1	2	3	4	5
P(X=x)	$\frac{1}{3}$	$\frac{4}{15}$	$\frac{1}{5}$	$\frac{2}{15}$	$\frac{1}{15}$

x	P(X=x)	xP(X=x)	x^2P(X=x)
1	$\frac{1}{3}$	$\frac{1}{3}$	$\frac{1}{3}$
2	$\frac{4}{15}$	$\frac{8}{15}$	$\frac{16}{15}$
3	$\frac{1}{5}$	$\frac{3}{5}$	$\frac{9}{5}$
4	$\frac{2}{15}$	$\frac{8}{15}$	$\frac{32}{15}$
5	$\frac{1}{15}$	$\frac{5}{15}$	$\frac{25}{15}$
Total	1	$\frac{7}{3}$	7

$$E(X)=\frac{7}{3}=2.3333$$

$$Var(X)=EX^2-(E(X))^2$$

$$Var(X)=7-\left(\frac{7}{3}\right)^2=1.556$$

3.4 PROPERTIES OF MEAN

1. $E(a)=a$ where a is constant
2. $E(aX)=a\,E(X)$
3. $E(aX+b)=aE(X)+b$ where a and b consants.

Example

The discrete random variable X has p.d.f given by

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

If Y is random variable defined by Y = 0.5X + 3. Determine the E(X) and E(Y)

Solution

x	-1	0	1	2
P(X=x)	0.25	0.10	0.45	0.20
xP(X=x)	-0.25	0	0.45	0.40

$\sum xP(X = x) = 0.6$

Therefore E(X) = 0.6

$$\begin{aligned} E(Y) &= E(0.5X + 3) \\ &= 0.5E(X) + 3 \\ &= 0.5 \times 0.6 + 3 \\ &= 3.3 \end{aligned}$$

E(Y) = 3.3

3.5 PROPERTIES OF VARIANCE

1. Var(a) = 0 where a is constant
2. Var(aX) = a^2 Var(X)
3. Var (aX + b) = a^2Var (X)

xample

ie discrete random variable X has p.d.f given by

:	1	2	3	4	5
P(X= x)	0.2	0.25	0.4	0.1	0.05

etermine

) Var (X) (ii) Var(3X-2)

lution

:	x	P(X=x)	xP(X=x)	$x^2P(x=x)$
	1	0.2	0.2	0.2
!	4	0.25	0.5	1
	9	0.4	1.2	3.6
	16	0.1	0.4	1.6
	25	0.05	0.25	1.25
tal		1	2.55	7.65

) $Var(X) = 7.65 - 2.55^2$

$Var(X) = 1.1475$

i) $Var(Y) = Var(3X - 2)$

$= 9\ Var(X) - 0$

$= 9 \times 1.1475$

$= 10.3275$

ierefore Var(3X-2) = 10. 3275.

Example

The discrete random variable X is defined by p.d.f given by

x	0	1	2
P(X=x)	0.4	0.4	0.2

If Y is a random variable defined by Y=2X-1. Determine the mean and variance of X and Y.

Solution

x	y	P(X)	P(Y)	xP(X)	yP(Y)	$x^2P(X)$	$y^2P(Y)$
0	-1	0.4	0.4	0	-0.4	0	0.4
1	1	0.4	0.4	0.4	0.4	0.4	0.4
2	3	0.2	0.2	0.4	0.6	0.8	1.8
Total		1	1	0.8	0.6	1.2	2.6

$E(X)=0.8,$

$Var(X)=1.2-0.8^2$

$Var(X)=0.56$

$E(Y)=0.6$

$Var(Y)=2.6-0.6^2$

$Var(Y)=2.24$

Example

The packets of Omo sold in a shop are of four categories namely, small, medium, large and giant. On a particular day, the stock is such that the ratio of small: medium: large: giant is equal to 4: 2: 1: 1. The costs of the packets are in the ratio: small: medium: large: giant are 350: 500: 800: 1400 respectively.

a) 30 packets are sold randomly on that particular day, the total cost of the sales being S shillings. Calculate
   i. The expected value of S
   ii. The standard deviation of S.

Solution

Let s denote small, m medium, l large and g giant.

$P(s) = \frac{4}{8} = \frac{1}{2}$, $\quad P(m) = \frac{2}{8} = \frac{1}{4}$

$P(l) = \frac{1}{8}$, $\quad P(g) = \frac{1}{8}$

Let X be the cost of a packet

x	$P(X=x)$	$xP(X=x)$	x^2	$x^2P(X=x)$
350	$\frac{1}{2}$	175	122,500	61,250
500	$\frac{1}{4}$	125	250,000	62,500
800	$\frac{1}{8}$	100	640,000	80,000
1400	$\frac{1}{8}$	175	1,960,000	245,000
Total	1	575		448,750

E(X) = 575

E(S) = E(30X)

= 30 E(X)

E(S) = 30 x 575 = 17,250/=

Var(X) = $EX^2 - (E(X))^2$

$= 448{,}750 - 575^2$

Var(X) = 118, 125

Var(S) = 30^2Var(X)

= 900 x 118, 125

Standard deviation = $\sqrt{900 x 118{,}125}$

Standard deviation = 10,310.8/=

3.6 MODE

The mode of a discrete probability density function is the value of X with the highest probability.

Example

A random variable has the following probability distribution

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

Determine the mode

Solution

P(X=2) = 0.4 is the highest probability therefore the mode is 2.

X = 2 is the mode

Example

A random variable has the following distribution

x	1	2	3
P(X=x)	0.35	0.35	0.3

Find the mode

P(X=1)=P(X=2)=0.35. This is highest probability. Therefore the mode is 1 and 2.

Example

A die is loaded such that the chance of throwing a one is $\frac{x}{4}$, the chance of a two is $\frac{1}{4}$, and the chance of a six is $\frac{1-x}{4}$. The chance of a three, four and five is $\frac{1}{6}$. The die is thrown twice. Prove that the chance of throwing total of 7 is $\frac{9x-9x^2+10}{72}$,

Find the value of x, which makes this chance maximum, and find this maximum probability.

Solution

Let r be the number that appears when a die is thrown

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

Possibilities for sum of seven

P(sum of 7) = $P(6 and 1) + P(1 and 6) + P(5 and 2) + P(2 and 5)$

$+P(4 and 3) + P(3 and 4)$

$$\left(\frac{1-x}{4}\right)\left(\frac{x}{4}\right)+\left(\frac{x}{4}\right)\left(\frac{1-x}{4}\right)+\frac{1}{6}x\frac{1}{4}+\frac{1}{4}x\frac{1}{6}+\frac{1}{6}x\frac{1}{6}+\frac{1}{6}x\frac{1}{6}$$

$$=\left(\frac{x-x^2}{16}\right)+\left(\frac{x-x^2}{16}\right)+\frac{1}{24}+\frac{1}{24}+\frac{1}{18}$$

$$=\left(\frac{x-x^2}{8}\right)+\frac{5}{36}$$

$$=\frac{9x-9x^2+10}{72}\text{ as required}$$

For maximum differentiate the expression = 0

$$\frac{d}{dx}\left(\frac{9x-9x^2+10}{72}\right)=0$$

$$9-18x=0$$

Therefore $x=\frac{1}{2}$

$$\text{Maximum probability}=\frac{9x0.5-9x0.5^2+10}{72}$$

$$=\frac{4.5-2.25+10}{72}=\frac{12.25}{72}$$

$$=\frac{49}{288}$$

3.7 CUMULATIVE MASS FUNCTION (DISTRIBUTION FUNCTION).

The cumulative distribution function of a discrete p.d.f is given by:
$F(X) = P(X \leq x)$
Note: Related to how you obtain cumulative frequencies.

Example
Consider the following probability distribution

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

Determine the cumulative mass function

Solution

$$F(1) = P(X \leq 1) = \frac{1}{6}$$

$$F(2) = P(X \leq 2) = P(X = 1) + P(X = 2) = \frac{1}{3}$$

$$F(3) = P(X \leq 3) = P(X = 1) + P(X = 2) + P(X = 3) = \frac{1}{2}$$

$$F(4) = P(X \leq 4) = \frac{4}{6}$$

$$F(5) = P(X \leq 5) = \frac{5}{6}$$

$$F(6) = P(X \leq 6) = 1$$

Normally the last value Of $F(X) = 1$.
Note: $F(X)$ can be given as an formula.

Example

The discrete random variable X has cumulative mass function

$$F(X) = \frac{x}{5} \text{ for } x = 1, 2, 3, 4, 5.$$

Find the probability distribution of X.

Solution

x	1	2	3	4	5
F(X)	$\frac{1}{5}$	$\frac{2}{5}$	$\frac{3}{5}$	$\frac{4}{5}$	1

$P(X=1) = F(1) = 0.2$

$P(X=2) = F(2) - F(1)$

$= 0.4 - 0.2$

$P(X=2) = 0.2$

$P(X=3) = F(3) - F(2)$

$= 0.6 - 0.4$

$P(X=3) = 0.2$

$P(X=4) = F(4) - F(3)$

$P(X=4) = 0.8 - 0.6 = 0.2$

$P(X=5) = 1 - 0.8 = 0.2$

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

3.8 MEDIAN

The median of the probability distribution of random variable X is the smallest value for which F(x) is at least 0.5. If m is the median then

(i) $F(m) \geq 0.5$

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

3.9 GRAPH OF f(x)

This is composed of vertical lines parallel to the y-axis drawn at every specified value of x.

3.10 GRAPH OF F(x)

This is composed of horizontal straight lines drawn parallel to x – axis.

Example

A random variable has the following distribution,
$P(X=1)=0.1$, $P(x=2)=0.4$, $P(X=3)=0.3$, $P(X=4)=0.2$
Find

i. The cumulative distribution function

ii. The median

iii. The graph of f(x) and F(x)

Solution

x	$P(X=x)$	$P(X\leq x) = F(X)$
1	0.1	0.1
2	0.4	0.5
3	0.3	0.8
4	0.2	1.0

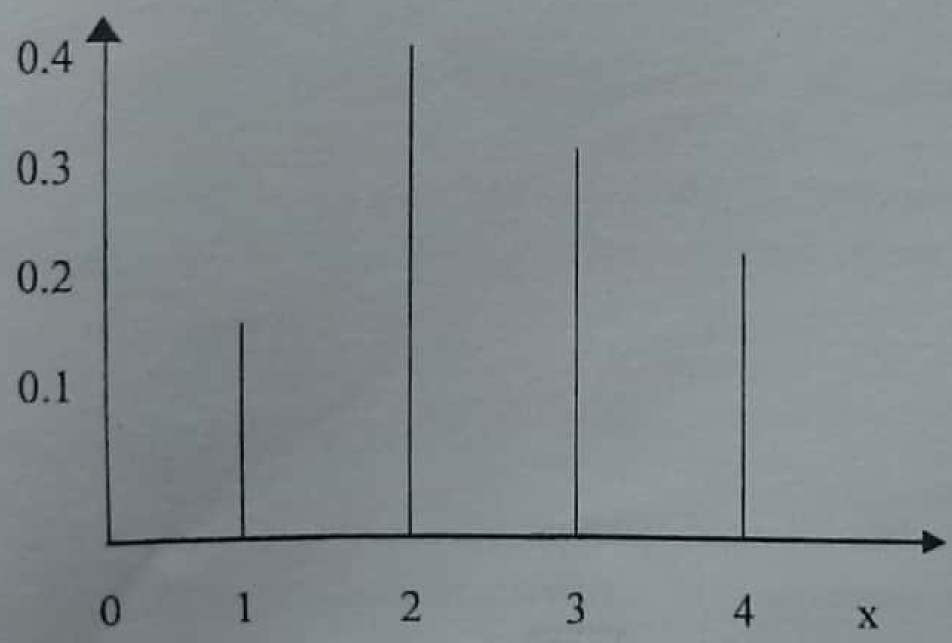
(i) $F(X) = P(X \leq x)$

(ii) $F(m) \geq 0.5$ where m is the median

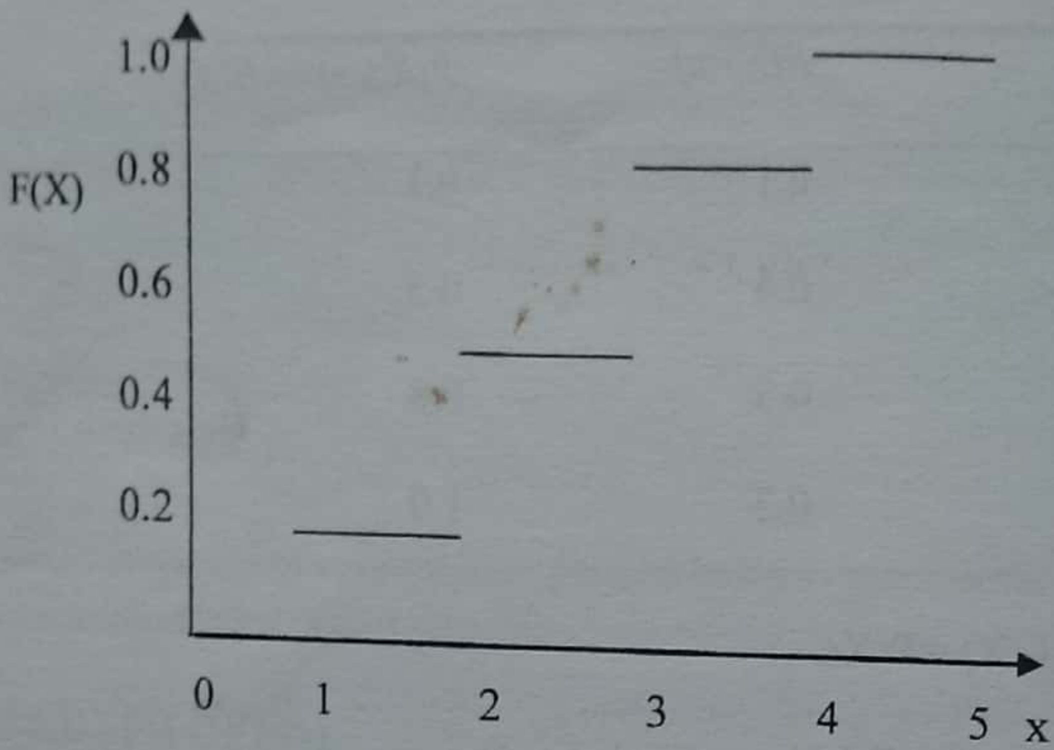
$F(2) = 0.5$

Therefore the median is 2

graph of f(x)



Graph F(X)



Exercise

1. A random variable T has the following probability distribution. $P(T=1)=0.1, P(T=2)=0.2, P(T=3)=0.3, P(T=4)=0.4$, show that the distribution above is a probability distribution

Determine:

i. The expected value and variance of T

ii. $P(T=2 \mid T \geq 2)$

Answer $E(T)=3,\ \ Var(T)=1\ \ P(T=2 \mid T \geq 2)=\frac{2}{9}$

2. For each of the following random vaiable write out the proability distributions

a) The number of heads obtained when two fair coins are tossed
b) The sum of the scores when two ordinary dice are thrown.
c) The number of threes obtained when two tetrahedral are thrown
d) The numerical value of a digit chosen from a set of random number table
e) The number of tails obtained when three fair coins are tossed.
f) The difference between the number when two ordinary dice are thrown

3. A computer is made to produce randomly the number 0, 1, 2, ... 9, what is the expected value and variance of the numbers so produced?

Answer $E(X)=4.5,\ Var(X)=8.25$

4. A discrete random variable X takes integer values between 0 and 5 inclusive with probabilities given by

$$P(X=r)=\begin{cases}\frac{2r+1}{20} & r=0,1,2,3\\ \frac{11-2r}{20} & r=4,5\end{cases}$$

Find the expectation and variance of X

Answer: $E(X)=2.55,\quad Var(X)=1.45$

5. A random variable X can assume values 10 and 20 only. If E(X) = 16, write the P.D.F of X in table form

Answer

x	P(X=x)
10	0.4
20	0.6

6. Two tetrahedral dice are thrown and score is the product of the number on which the dice fall. What is the expected score for a throw?

Answer: 6.25

7. A random variable X has the probability function

$$f(x)=\begin{cases}k2^x & x=0,1,2,\ldots 6\\ 0 & otherwise\end{cases}$$

Determine:

(i) *the value of k*

(ii) $E(X)$

(iii) $P(X<4\mid X>1)$

Answer (i) $=\frac{1}{127}$ (ii) $E(X)=5$ (iii) $\frac{3}{31}$

8. A discrete random variable X has probability function

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

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 variance of X*

(iii) $P(X=2|X\geq 2)$

Answer : (i) n = 4, k = 10

(ii) median = 3, Var(x) = 1

(iv) $P(X=2|X\geq 2)=\frac{2}{9}$

9. The number of times a machine breaks down every month is a discrete random variable X with the probability distribution

$$P(X=x)=\begin{cases}k\ 0.25^x & x=0,1,2,3,...\\ 0 & otherwise\end{cases}$$

where k is a constant

Determine the probability that the machine will break down not more than two times a month.

Answer : 0.984

10. A 5- sided die with sides numbered 1,2,3,4 and 5 is constructed so that 1 and 5 occur twice as often as 2 and 4, which occur three times as often as 3. What is the probability that a prefect square occurs when this die is tossed once?

Answer: $P(4\,Or\ 1)=\frac{9}{19}$

11. A bag contains two red and eight black marbles. A sample of four marbles is to be drawn at random from the bag without replacement:

(a) Show that the probability of obtaining exactly two red marbles in the sample is 2/ 15

(b) Show that the probability of obtaining exactly one red marbles in the sample is 8/ 15

(c) Calculate the expected number of red marbles that will be drawn.

Answer: $E(X) = 0.8$

12. The probability of a number appearing on the top face when an unbalanced die is tossed is proportional to the number. Find the probability that an odd or prme number will appear. If the die is tossed twice, find the probability distribution function for the sum sum of the two numbers that appear on the top face. What is the most likely sum?

x	$P(X=x)$
2	$\frac{1}{441}$
3	$\frac{4}{441}$
4	$\frac{10}{441}$
5	$\frac{20}{441}$
6	$\frac{35}{441}$
7	$\frac{56}{441}$
8	$\frac{70}{441}$
9	$\frac{76}{441}$
10	$\frac{73}{441}$
11	$\frac{60}{441}$
12	$\frac{36}{441}$

Answer: $P(O \cup P) = \frac{11}{21}$

The likely sum is 9

13.A bag contains one 200 Shs note, three 100 Shs notes and n 50 Shs notes. A note is selected at random from the bag, its value noted and then replaced. The process is repeated many times. If the average of the values of the notes after many trials is 110 Shs, determine:

i. The value of n
ii. The expected value of the sum of the two notes selected at random without replacement

Answer: (i) n = 1 (ii) E(X) = 220 /=

14. A basket 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.

Answer : 1

15. A bag Z contains 3 red balls and 6 white balls and the second bag X contains 5 red balls and 4 white balls. A ball is chosen at random from bag Z and placed in bag X. A ball is then chosen at random from bag X and placed in bag Z. If R is the number of red balls in bag Z after these operations, find the probability distribution of R. Hence the expected value of R.

Answer:

$$P(R=2) = \frac{2}{15},\ P(R=3) = \frac{8}{15},\ P(R=4) = \frac{1}{3},\ E(R) = 3.2$$

16. The discrete random variable X has distribution function F(X) where

$F(x)=1-(1-x/4)^x$

$For\ x=1,2,3,4$

(a) Show that $F(3)=63/64$ and $F(2)=3/4$

(b) Obtain the probability distribution of X

(c) Find the $E(X)$ and $Var(X)$

(d) Find $P(X>E(X))$

Answer:

(b)

x	1	2	3	4
P(X=x)	$1/4$	$1/2$	$15/64$	$1/64$

(c) $2\ 1/64$, 0.547 (d) 0.25

17. (a) An examination question has two parts A and B. The probability of a student getting part A correct is $2/3$. If she gets part A correct, the probability that she gets part B correct is $3/4$, otherwise it is $1/6$. There are three marks for a correct solution of part A, two marks for part B and a bonus mark if both parts are correct. Calculate the expecte value and variance of the student's total mark for the question?

(b) In a game where a gambler rolls a dice on a \$ 1 stake, the Casino pays out \$ 10 for a double six and \$ 3 for a score of sum seven (the stake money returned as well) and for other score the gambler loses his stake. Whatis the Casino's expected profit on a \$ 100 stake?

Answer

(a) E(X) = 3.61, Var(X) = 6.683, (b) E(P) = \$ 2.78

CHAPTER FOUR

THE BINOMIAL DISTRIBUTION

4.1 INTRODUCTION

Binomial distribution is an example of discrete probability distribution with a countable domain. A binomial experiment is characterized with repeated trials, and two possible outcomes, one is termed as "success" and other "failure"

Example

A student attempts three questions in a commerce objective section, where each question has four possible alternatives with only one correct answer. Determine the probability distribution of the number of correct answers.

Solution

$P(C)=\frac{1}{4},\ P(\overline{C})=\frac{3}{4}$, *where C represents correct answer.*

Using the theory of Binomial expansion

$(a+b)^3 = a^3 + 3a^2b + 3ab^2 + b^3$ Note: n = 3

If let $a=\frac{3}{4},\ b=\frac{1}{4}$

$$\Rightarrow\left(\frac{3}{4}+\frac{1}{4}\right)^3=\left(\frac{3}{4}\right)^3+3\left(\frac{3}{4}\right)^2\left(\frac{1}{4}\right)+3\left(\frac{3}{4}\right)\left(\frac{1}{4}\right)^2+\left(\frac{1}{4}\right)^3$$

Note: First term has zero correct answers. The second term has one correct answer. Third term has two correct answers. The fourth term has all correct answers.

x		$P(X=x)$
0	$\left(\frac{3}{4}\right)^3$	$\frac{27}{64}$
1	$3\left(\frac{3}{4}\right)^2\left(\frac{1}{4}\right)$	$\frac{27}{64}$
2	$3\left(\frac{3}{4}\right)\left(\frac{1}{4}\right)^2$	$\frac{9}{64}$
3	$\left(\frac{1}{4}\right)^3$	$\frac{1}{64}$

x	0	1	2	3
$P(X=x)$	$\frac{27}{64}$	$\frac{27}{64}$	$\frac{9}{64}$	$\frac{1}{64}$

4.2 PROPERTIES OF BINOMIAL DISTRIBUTION

1. The experiment has a number of repeated trials (n)
2. Each trial results in only two possible outcomes "Success" and "failure"
3. The trials are independent
4. The probability of success p is constant.

4.3 FORMULA FOR BINOMIAL EXPERIMENT

A random variable X is said to have a binomial distribution B(n, p),

If $P(X=r) = \binom{n}{r} q^{n-r} p^r$, $r = 0, 1, 2 \ldots n$

where $0 \leq p \leq 1$ and $q = 1 - p$

Where p is probability of success

where $\begin{pmatrix} n \\ r \end{pmatrix} = {}^nC_r$

EXAMPLE

Find the probability of obtaining exactly three 4's if an ordinary die is tossed 5 times.

Solution:

n = 5, p = 1/6, r = 3

$$P(X = r) = \begin{pmatrix} n \\ r \end{pmatrix} q^{n-r} p^r$$

$$P(X = 3) = \begin{pmatrix} 5 \\ 3 \end{pmatrix} \left(\frac{5}{6}\right)^2 \left(\frac{1}{6}\right)^3$$

$$= 10 \text{ x } \frac{5^2}{6^5} = 0.03215$$

P(X=3) = 0.03215

Example

The probability that a pen drawn at random from a box of pens is defective is 0.1. If a sample of 6 pens is taken, find the probability that it will contain

(a) No defective pens

(b) 5 or 6 defective pens.

(c) Less than 3 defective pens.

Solution

n=6, p=0.1, let X be number of defective pens.

(a) $P(X=0)=\begin{pmatrix}6\\0\end{pmatrix}(0.9)^6(0.1)^0$

$P(X=0) = (0.9)^6 = 0.5314$

$P(X=0) = 0.5314$

(b) P (X=5) + (X=6).

$P(X=r) = \begin{pmatrix}n\\r\end{pmatrix}q^{n-r}P^r$).

$P(X=5) = \begin{pmatrix}6\\5\end{pmatrix}(0.9)^1(0.1)^5$

$= 6(0.9) \text{ x} (0.1)^5$

$= 0.000054$

$P(X=6) = \begin{pmatrix}6\\6\end{pmatrix}(0.9)^0(0.1)^6$

$= (0.1)^6 = 0.000001$

P(5 or 6) = 0.000054 + 0.000001

= 0.000055.

(c) $P(X<3) = P(X \le 2)$

Note: this is a discrete function.

$P(X \le 2) = P(X=0) + P(X=1) + P(X=2)$

$P(X=0) = 0.5314$ from part (a)

$P(X=1) = \begin{pmatrix}6\\1\end{pmatrix}(0.9)^5(0.1)^1$

$0.6 \text{ x } (0.9)^5$

$P(X=1) = 0.3543$

$$P(X=2) = \binom{6}{2}(0.9)^4(0.1)^2$$

$$15 \times 0.01 \times (0.9)^4$$

$P(X = 2) = 0.0984$

$\therefore P(X \leq 2) = 0.5314 + 0.3543 + 0.0984$

$P(X \leq 2) = 0.9841$

Note $\binom{n}{r}$ can be obtained from tables.

Some of the above probabilities can be obtained from the table. The value of n is located from the table and its corresponding p (probability of success noted). The probability can be read off from the table for r = 0, 1, ….n.

Example

Using the example above obtain $P(X = 0)$

Solution:

$n = 6, p = 0.1$

Table

B(n, p), Individual terms.

n	r	Probability of success(p) 0.01	0.05	0.1	0.15	0.2	0.25…
6	0			0.5314			
	1			0.3543			
	2			0.0984			
	3			0.0146			
	4			0.0012			
	5			0.0001			
	6			-			

From the table $P(X=0)=0.5314$. Obtained by reading $n=6, r=0$ and $p=0.1$.

Then $P(X=1)=0.3543$ (Tab)
And $P(X=2)=0.0984$ (Tab).
Check part c. for $P(X=1)$ and $P(X=2)$.

Example

A multiple –choice quiz has 15 questions, each with 4 – possible answers of which only one is the correct answer. Determine the probability that sheer guess work yields.

(a) Exactly five correct answers

(b) Five incorrect answers

Solution

$n=15, p=¼=0.25$

Let X be number of correct answers.

(a) $P(X=5)$

Use $P(X=r)=\binom{n}{r}q^{n-r}p^{r}$.

From the table when $n=15, r=5, p=0.25$

$$\Rightarrow P(X=5)=\binom{15}{5}(0.25)^5(0.75)^{10}=0.1651 \text{ (Tab)}$$

(b) P(five incorrect answers) =
P(ten correct answers)

$P(X=10)$
$n=15, r=10, P=0.25$

$$P(X=10)=\binom{15}{10}(0.25)^{10}(0.75)^5=0.0007 \text{ (Tab)}$$

P(five incorrect answers) = 0.0007.

Example

In certain clan the probability of a family having a girl is 0.6. If there are 5 children in a family, determine the probability that they are all girls.

Solution

$n=5, r=5 \text{ and } p=0.6$

$But\ P(all\ girls) = P(zero\ boys)$

$Therefore\ using\ the\ boys$

$\Rightarrow n=5, r=0, p=0.4$

$$P(X=0) = \binom{5}{0}(0.4)^0(0.6)^5$$

$$=0.0778\,(Tab)$$

The table above gives individual probabilities. Sometimes sums of binomial probability are needed. These are obtained from the cumulative binomial probability ie Σ P. It gives the sum of probabilities from that value and above ($i \geq r$).

Probabilities below a certain value are obtained by using
$P(X \leq r) = 1 - P(X \geq r+1)$

EXAMPLE

The probability that a patient recovers from a disease is 0.4. If 15 people are known to have contracted the disease, what is the probability that

(a) at least 10 will survive

(b) between 3 and 8 inclusive will survive

(c) atmost 5 will survive
(d) between 6 to 10 will survive

Solution:
n = 15, p = 0.4, and q = 0.6
Let X be the number of patients who recover

(a) $P(X \geq 10) = \sum_{10}^{15} \binom{15}{x} (0.4)^x (0.6)^{15-x} = 0.0338(Tab)$.

Read n = 15, p = 0.4 and x = r = 10

(b) $P(3 \leq X \leq 8)$

$$\sum_{3}^{15} \binom{15}{x} (0.4)^x (0.6)^{15-x} - \sum_{9}^{15} \binom{15}{x} (0.4)^x (0.6)^{15-x}$$

$= 0.9729 - 0.0950 = 0.8777$ (Tab)

(c) $P(X \leq 5) = 1 - P(X \geq 6)$.

$P(X \geq 6) = \sum_{6}^{15} \binom{15}{x} (0.4)^x (0.6)^{15-x} = 0.5968(Tab)$

$P(X \leq 5) = 1 - 0.5968 = 0.4032$

(d) $P(6 < X \leq 10) = P(7 \leq X \leq 10)$

$$\sum_{7}^{15} \binom{15}{x} (0.4)^x (0.6)^{15-x} - \sum_{11}^{15} \binom{15}{x} (0.4)^x (0.6)^{15-x}$$

$= 0.3902 - 0.0093 = 0.3809$ (Tab)

4.4 FORMAT FOR TABLE FOR SUMS

		Probability of success (p)								*P(x ≥ r)*
n	r	0.01	0.05	0.1	0.15	0.2	0.25	0.3	.35	0.40
15	1									0.9995
	2									0.9948
	3									0.9729
	4									0.9095
	5									0.7827
	6									0.5968
	7									0.3902
	8									0.2131
	9									0.0950
	10									0.0338
	11									0.0093
	12									0.0019
	13									0.0003
	14									

All the above sums of probability were obtained from the table.

Example

In a certain manufacturing process, it is know that approximately 10% of the items produced are defective. A quality control scheme is set up by selecting twenty items out of a large batch, and rejecting the whole batch if three or more are defective. Find the probability that the batch is rejected.

Solution:

$n = 20, p = 0.1, P(X \geq 3)$

$X \sim B(n, p)$

$$P(X \geq 3) = \sum_{3}^{20} \binom{20}{x} (0.1)^x (0.9)^{20-x} = 0.3231 (Tab)$$

Example

The probability of winning a game is $\frac{4}{5}$. Ten games are played. What is the probability of at least 8 successes in the ten games?

Solution

$n = 10, p = 0.8, \ q = 0.2$

$$P(X \geq 8) = \sum_{8}^{10} \binom{10}{x} (0.8)^x (0.2)^{10-x}$$

$$= \sum_{0}^{2} \binom{10}{x} (0.2)^x (0.8)^{10-x}$$

$$P(X \geq 8) = 1 - \sum_{3}^{10} \binom{10}{x} (0.2)^x (0.8)^{10-x}$$

$$= 1 - 0.3222 = 0.6778 (Tab)$$

Sometimes the probability of success in not contained in the table.

Example

A coin is tossed such that is twice as likely to show heads as tails. Find the probability that in five tosses of the coin

(a) exactly three heads are obtained

(b) more than three heads are obtained

Solution:

(a) Let H denote head and T denote tail

$P(H) = 2(T)$

$P(H) + P(T) = 1$

$$2P(T) + P(T) = 1$$

$P(T) = 1/3$

$P(H) = 2\,P(T) = 2/3$

Probability of success $p = 2/3$

Let X be the number of heads that appear.

$n = 5$

$$P(X = r) = \binom{n}{r} P^r q^{n-r}$$

$$P(X = 3) = \binom{5}{3}\left(\frac{2}{3}\right)^3\left(\frac{1}{3}\right)^2$$

$$= \frac{10 \times 8}{243} = \frac{80}{243} = 0.3292$$

$P(X = 3) = 0.3292$

(b) $P(X > 3) = P(X = 4) + P(X = 5)$

$$P(X = 4) = \binom{5}{4}\left(\frac{2}{3}\right)^4\left(\frac{1}{3}\right) = \frac{5 x 16}{243} = \frac{80}{243}$$

$P(X = 4) = 0.32922$

$$P(X = 5) = \binom{5}{5}\left(\frac{2}{3}\right)^5 = \frac{2^5}{3^5} = \frac{32}{243}$$

$P(X = 5) = 0.131687$

$P(X > 3) = 0.131687 + 0.32922 = 0.460907$

$P(X > 3) = 0.4609$

4.5 EXPECTATION AND VARIANCE

MEAN (EXPECTATION)

If the random variable X is such that

$X \sim B(n, p)$ then

$$E(X) = \sum_{all\,x} x\,P(X = x) = np$$

$\therefore E(X) = \mu = np$

VARIANCE

$Var(X) = npq$

$$\sigma = \sqrt{Vax(x)} = \sqrt{(npq)}$$

Example

In Kampala city, it is known that 1/3 of the voters support the U.P.C party. In a sample of 12 voters, what is the expected value and standard deviation of the number of U.P.C's.

Solution

X □ B(n,p), Let X be the number of U.P.C voters.

$n = 12,\ p = 1/3,\ q = 1 - p = 2/3$

$$E(X) = np = 12 \times \frac{1}{3} = 4$$

$$\sigma = \sqrt{Vax(x)} = \sqrt{(npq)}$$

$$= \sqrt{12x\frac{1}{3}x\frac{2}{3}}$$

$$= \sqrt{\frac{8}{3}} \cong 1.63$$

Example

Given that the binomial distribution B (n, p) has mean 9.6 and standard deviation 2.4, find n and p.

Solution:

$$E(x) = np$$

$$\sqrt{Var(x)} = \sqrt{npq}$$

$$\sigma = \sqrt{npq}$$

$$9.6 = np \ldots\ldots\ldots\ldots\ldots\ldots\ldots(1)$$

$$2.4 = \sqrt{npq\ldots\ldots} \qquad (1)$$

From (2) $npq = (2.4)^2$

$npq = 5.76$

but $np = 9.6$

$\Rightarrow 9.6q = 5.76$

$$q = \frac{5.76}{9.6}$$

$q = 0.6$

$\Rightarrow p = 1 - q = 0.4 \Rightarrow p = 0.4$

From (1) $np = 9.6$

$0.4n = 9.6$

$$n = \frac{9.6}{0.4} = 24$$

$n = 24, p = 0.4$

Hence $n = 24, p = 0.4$

Example

The probability of a student arriving at the school late on any given day is 1/10. What is the probability of his being punctual for a whole week (i.e 5 school days). Calculate the mean and variance of the number of days he will be late in school term consisting of 14 weeks (i.e 70 days). Also calculate the expected number of completely punctual weeks in the term.

Solution:

Let X be number of days he is late.

$n = 5, p = 0.1, P(X = 0)$.

P(punctual for a whole week)

$$P(X = 0) = \binom{5}{0}(0.1)^0(0.9)^5 = 0.5905 \text{ (Tab)}.$$

$E(X) = np$

14 weeks = 14 x 5 = 70 days

$E(X) = 70 \times 0.1 = 7$ days.

$Var(X) = npq$

$= 70 \times 0.1 \times 0.9$

$7 \times 0.9 = 6.3$

$Var(X) = 6.3$

P(punctual for all whole week) = 0.5905.

$n = 14$ weeks

$E(X) = 14 \times 0.5905$

$= 8.267$

$E(X) = 8.267$

Example

If a random variable X is binomially distributed with B(10,p) with p < ½ an variance 15/8. Determine p, the mean and probability of at least one.

Solution:

np = ? n = 10

$$npq = \frac{15}{8}$$

$$pq = \frac{15/8}{10} = \frac{3}{16}$$

$$p(1-p) = \frac{3}{16}$$

$$p - p^2 = \frac{3}{16}$$

$$16p - 16p^2 - 3 = 0$$

$$16p^2 - 16p + 3 = 0$$

$p = ¼$ or $¾$ (Cal)

since $p < ½$ then $p = ¼$

mean = np

= 10 x ¼

= 2.5

$$P(x \geq 1) = \sum_{1}^{10} \binom{10}{x} (0.25)^x (0.75)^{10-x}$$

= 0.9437 (Tab)

Example

A box contains a large number of red and yellow bulbs in the ratio 1:3. Bulbs are picked at random from the box. How many bulbs must be picked so that the probability that there is at least one red bulb among them is greater than 0.95?

Solution

$P = \frac{1}{4}, q = \frac{3}{4}$

$X \sim B(n, p)$

$P(X \geq 1) > 0.95$

$= 1 - P(X = 0)$

$$1 - \binom{n}{0}(0.25)^0(0.75)^n > 0.95$$

$$1 - 0.75^n > 0.95$$

$$0.05 > 0.75^n$$

$$\log 0.05 > n \log 0.75$$

$$-1.301 > -0.1249n$$

$$n > 10.42$$

$n \approx 11$ *therefore the least value of n is* 11

4.6 THE MOST LIKELY VALUE OF X TO OCCUR

The value of X that is most likely to occur is the one with the highest probability. However the working could be tedious to calculate all probabilities.
Therefore the mean is obtained and normally the values close to mean tend to have the highest probabilities.

Example

Of the inhabitants of Katwe village, 80% are known to have Malaria. If 12 people are waiting to see the nurse, what is the most likely number of them to have malaria?

Solution

Mean = np
= 12 x 0.8
= 9.6

Therefore the most likely number is either 9 or 10

P(X = 9) = P(three people without Malaria)
n = 12, p = 0.2, q = 0.8, r = 3

$$P(X=9)=\binom{12}{9}(0.8)^{9}(0.2)^{3}$$

$$Or\ P(X=3)=\binom{12}{3}(0.2)^{3}(0.8)^{9}$$

$$=0.2362\,(Tab)$$

$$P(X=10)=\binom{12}{10}(0.8)^{10}(0.2)^{2}$$

$$Or\ P(X=2)=\binom{12}{2}(0.2)^{2}(0.8)^{10}$$

$$=0.2835\,(Tab)$$

Therefore the most likely number of people is 10 since it has the highest probability.

Exercise.

1. In a certain clan the probability of a family having a boy 0.6. If there are 5 children in a family, determine.
   (i) The expected number of girls.
   (ii) Probability that there are at least three girls.

Answer (i) $E(X)=2$. (ii) $P(X \geq 3)=0.3174$

2. Three people play a game in which each person tosses a coin. The game is a success if one of the players gets an outcome different from the others. Determine the probability that.
   (i) a success will occur at the first trial.
   (ii) in two trials at least one success will occur.

Answer (i) P (a successes at first trial) $= \frac{3}{4}$

(ii) P(at least one success in two trials) $= \frac{15}{16}$

3. Muwanga makes 5 practice runs in the 100m sprint. A run is successful if he runs it in less than 11 seconds. There are 8 chances out of 10 that he is successful. Find the probability that

(a) (i) he records no success at all

(ii) he records at least 2 successes.

(b) If he is successful in the 5 practice runs he makes two additional runs, the probability of success in either of the additional runs is 0.7. Determine the probability that Muwanga will make 7 successful runs in total.

Answer: (a) (i) 0.0003 (ii) 0.9933 (b) 0.16

4. The packets of Omo sold in a shop are of four categories namely, small, medium, large and giant. On a particular day, the stock is such that the ratio of small: medium: large: gaint is equal to 4:2:1:1 respectively. Ten packets are picked at random. Determine the probability that six are medium size packets.

   Answer: $P(X=6)=0.0162$

5. Assuming that a couple is equal likely to produce a girl or a boy. Find the probability that in a family of 5 children there will be more boys than girls.

Answer: 0.5

6. Of the students in a school, 30% travel to school by bus. From a sample of 10 students chosen at random, find the probability that:

   (a) Only 3 travel by bus
   (b) more than 8 travel by bus
   (c) at most 8 travel by bus

   Answer: (a) 0.2668 (b) 0.0001 (c) 0.9999

7. In a group of people the expected number who wear glasses is 2 and the variance is 1.6. Find the probability that:-

   (a) A person, chosen at random from the group wears glasses
   (b) 6 people in the group wear glasses

   Answer a) 0.2, b) 0.0055

8. Seedlings are planted in 10 rows of six each. The probability of a seedling dying before it flowers is 1/8. Calculate the mean and variance of the number of rows in which all the seedlings flower.

   Answer: 4.488, 2. 4738

9. A certain electronic system contains ten components. Suppose that the probability that any individual component will fail is 0.2, and that the components fail independently of each other. Given that at least one of the components has failed, what is the probability that at least two of the components have failed.

Answer: 0.6993

10. A cross world puzzle is published in the times each day of the week, except Sunday. A man is able to complete on average 8 out of 10 of the cross word puzzles.

    (a) Find the excepted value and the standard deviation of the number of completed cross words in a given week.

    (b) Show that the probability that he will complete at least 5 in a given week is 0.655 (to 3 significant figure)

    Answer: (a) 4.8, 0.98

11. At a certain hatchery, records taken over several months have shown that the ratio of male to female chicks hatched has been 4 : 6. Basing on this experience, what is the probability that there will be more female chicks in a random collection of a dozen chicks from this hatchery?

    Answer 0.6652

12. A machine producing electrical components occasionally makes faulty ones. In a batch of 20 such components, 5 are found to be faulty. A sample of 3 of this batch is taken and examined. What is the probability that:-

    (i) it contains faulty components

    (ii) it contains exactly one faulty component?

    Answer (i) 0.5781 (ii) 0.4219

13. A fair die is rolled 6 times. Calculate the probability that

    (i) a 2 or 4 appears on the first throw,

    (ii) four 5s will appear in the six throws

    Answer (i) $\frac{1}{3}$ (ii) 0.0080

CHAPTER FIVE

CONTINUOUS PROBABILITY DENSITY FUNCTION

5.1 INTRODUCTION

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

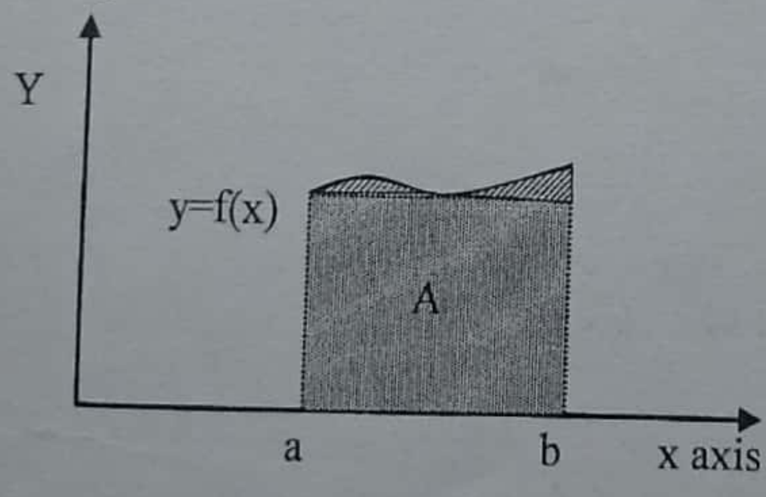
Properties of f (x)

(i) $f(x) > 0$ for all values of x.

(ii) $\int_{\text{all}x} f(x)\,dx = 1$

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

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



5.2 HOW TO OBTAIN PROBABILITIES

The probability that a random variable attains values between x_1 and x_2 given by $P(x_1 < X < x_2)$ is obtained from the area under the curve between x_1 and x_2.

$$\therefore P(x_1 < X < x_2) = \int_{x_1}^{x_2} f(x)dx$$

Example

A continuous random variable has p.d.f given by

$$f(x) = \begin{cases} kx & 0 < x < 4 \\ 0 & \text{otherwise} \end{cases}$$

Determine (a) The value of the constant k.

(b) $P(1 \leq X \leq 2)$

(c) Sketch the graph of f(x)

Solution:

(a) $\int f(x)\,dx = 1 \Rightarrow \int_0^4 kx\,dx = 1$

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

$8k = 1 \Rightarrow \quad k = 1/8$

$$\therefore f(x) = \begin{cases} \frac{x}{8} & 0 \leq x \leq 4 \\ 0 & \textit{otherwise} \end{cases}$$

(b) $P(1 \le X \le 2) = \int_1^2 f(x)dx =$

$$\int_1^2 kx\,dx =$$

$$k\left[\frac{x^2}{2}\right]_1^2 = \frac{3}{2}k$$

$$P(1 \le X \le 2) = \frac{3}{2} x \frac{1}{8}$$

$$= \frac{3}{16}$$

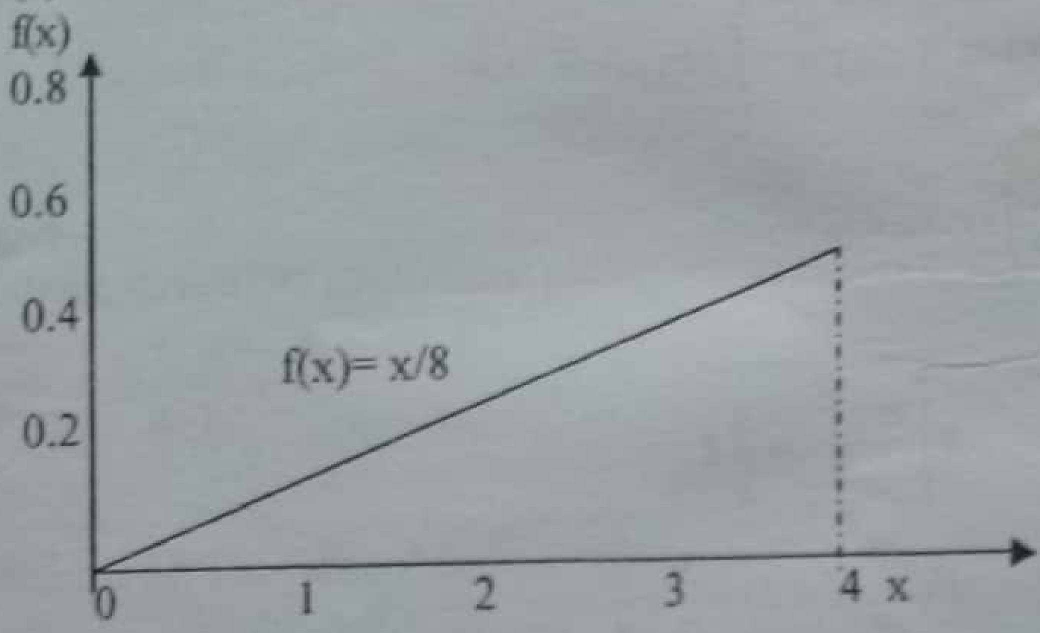
5.3 GRAPH OF f(x)

Look at the function and note whether it is a straight line or simple curve in the interval given. i.e for the above example $f(x) = \frac{x}{8}$ which is a straight line between 0 and 4. With a gradient of $\frac{1}{8}$ and zero intercept

When x = 0, f(0) = 0.

When x = 4, f(4) = 0.5

(c) GRAPH OF f(x)



Example

A continuous random variable has a p.d.f where

$$f(x) = \begin{cases} kx & 0 \le x \le 2 \\ k(4-x) & 2 \le x \le 4 \\ 0 & \text{otherwise} \end{cases}$$

Determine:

(a) Value of the constant k.

(b) $P(1 \le X \le 3)$

(c) Sketch the graph of f(x)

Solution:

(a) $\int_{\text{all } x} f(x)\,dx = 1$

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

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

$$2k + 2k = 4k = 1$$

$$\therefore k = \tfrac{1}{4}$$

(b) $P(1 \le X \le 3) = \int_1^2 kx\,dx + \int_2^3 k(4-x)\,dx$

i.e $P(1 \le X \le 3) = P(1 \le X \le 2) + P(2 \le X \le 3)$

$$k\left[\frac{x^2}{2}\right]_1^2 + k\left[4x - \frac{x^2}{2}\right]_2^3$$

$$= \frac{1}{4}\left(2 - \frac{1}{2}\right) + \frac{1}{4}(7.5 - 6)$$

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

$P(1 \le X \le 3) = 0.75$

(c) GRAPH f(x)

$$f(x) = \begin{cases} x/4 & 0 \le x \le 2 \\ 1 - x/4 & 2 \le x \le 4 \\ 0 & \text{otherwise} \end{cases}$$

for $0 < x < 2$, $y = f(x)$ $y = x/4$

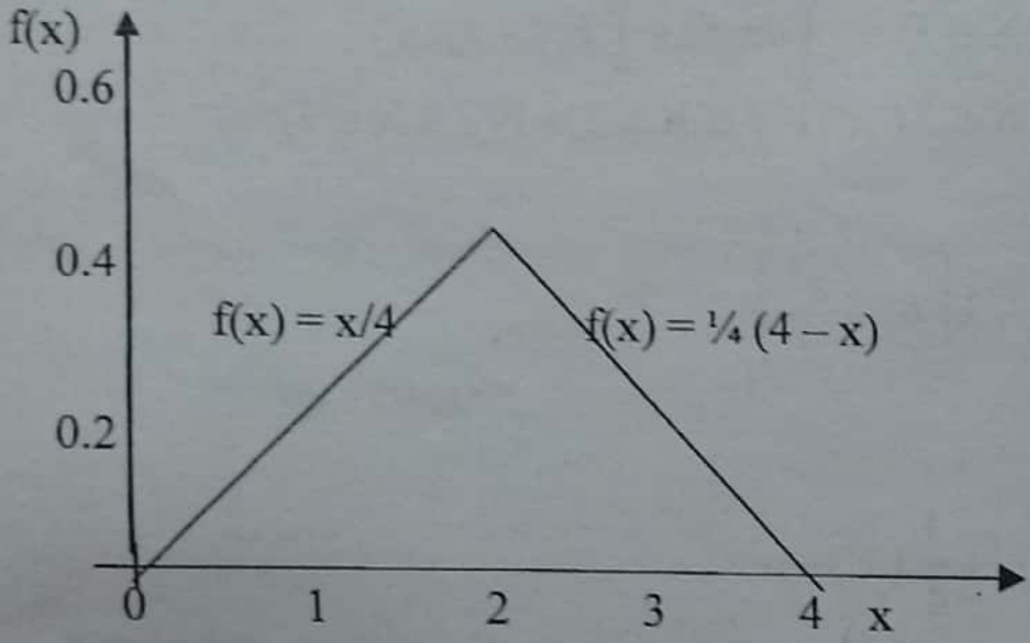
Which is a straight line

With $m = \frac{1}{4}$ and $c = 0$, consider $y = mx + c$

When $x = 0, y = 0, x = 2, \ y = \frac{1}{2}$

For $2 \leq x \leq 4, \ y = 1 - \frac{x}{4}$

When $x = 2, y = 0.5, x = 4, y = 0.$



Example

A continuous random variable X has p.d.f f(x) given by

$$f(x) = \begin{cases} k & 0 \leq x \leq 2 \\ k(2x-3) & 2 \leq x \leq 3 \\ 0 & \text{otherwise} \end{cases}$$

Determine:

(a) value of constant k.

(b) sketch the graph of f(x)

(c) (i) $P(X < 1)$ (ii) $P(X = 1)$

(iii) $P(X > 2.5)$ (iv) $P(0 < X < 2 \mid X > 1)$

Solution:

(a) from $\int_{\text{all}\,x} f(x)dx = 1$

$$k[x]_0^2 + k\left[x^2 - 3x\right]_2^3 = 1$$

$$2k + 2k = 1$$

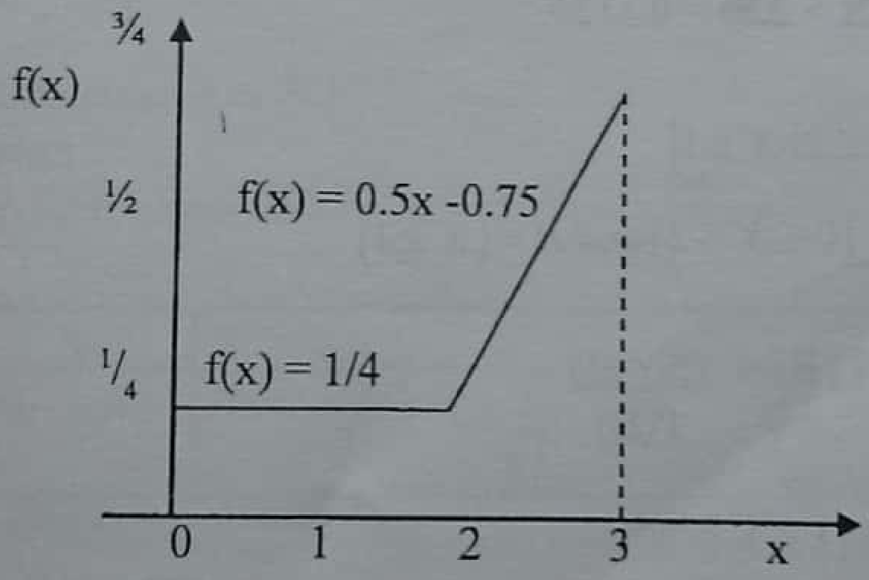
$$k = \tfrac{1}{4}$$

(b) Graph of f (x)

$$f(x) = \begin{cases} 1/4 & 0 \le x \le 2 \\ 0.5x - \tfrac{3}{4} & 2 \le x \le 3 \\ 0 & \text{otherwise} \end{cases}$$

for $0 \le X \le 2$ $\quad y = \tfrac{1}{2}x - \tfrac{3}{4}$

when $x = 3$, $y = 3/2 - \tfrac{3}{4} = \tfrac{3}{4}$, $\quad x = 2$, $y = \tfrac{1}{4}$



(c) (i) P(X <1) =

$$\int_0^1 f(x)dx \ \ Or \ \ Area = \left(\tfrac{1}{4}\right) x 1 = \tfrac{1}{4}$$

$$\int_0^1 \frac{1}{4}dx = \left[\frac{x}{4}\right]_0^1 = \frac{1}{4}$$

$\therefore P(X < 1) = ¼$

(ii) $P(X=1) = 0$ Since X is continuous random variable.

(iii) $P(X > 2.5) = P(2.5 < X < 3) = \int_{2.5}^{3} f(x)dx$

$$= \int_{2.5}^{3} k(2x-3)dx = k\left[x^2 - 3x\right]_{2.5}^{3}$$

$$= ¼\left[x^2 - 3x\right]_{2.5}^{3}$$

$= 0.3125$

⇨ $P(X > 2.5) = 0.3125.$

(iv) $P\left[0 \le X \le 2 \mid X \le 1\right]$

Let A = $(0 \le X \le 2)$ and $B = (X \ge 1)$

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

$A \cap B = (1 \le X \le 2)$

$$P\left[0 \le X \le 2 \mid X \ge 1\right] = \frac{P(1 \le X \le 2)}{P(1 \le X \le 3)}$$

$$\Rightarrow \quad \frac{\int_1^2 1/4dx}{\int_1^2 1/4dx + \int_2^3 1/4(2x-3)dx} = \frac{1/4}{1/4+1/2}$$

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

$$P(0 < X < 2 \mid x \geq 1) = \frac{1}{3}$$

Example

A continuous random variable X has p.d.f given by

$$f(x) = \begin{cases} \frac{3}{32}(x^2) & 0 \leq x \leq 2 \\ \frac{3}{32}(6-x) & 2 \leq x \leq 6 \\ 0 & otherwise \end{cases}$$

Sketch the graph of f(x)

Solution

Note: for the interval $0 \leq x \leq 2$ the function is a curve

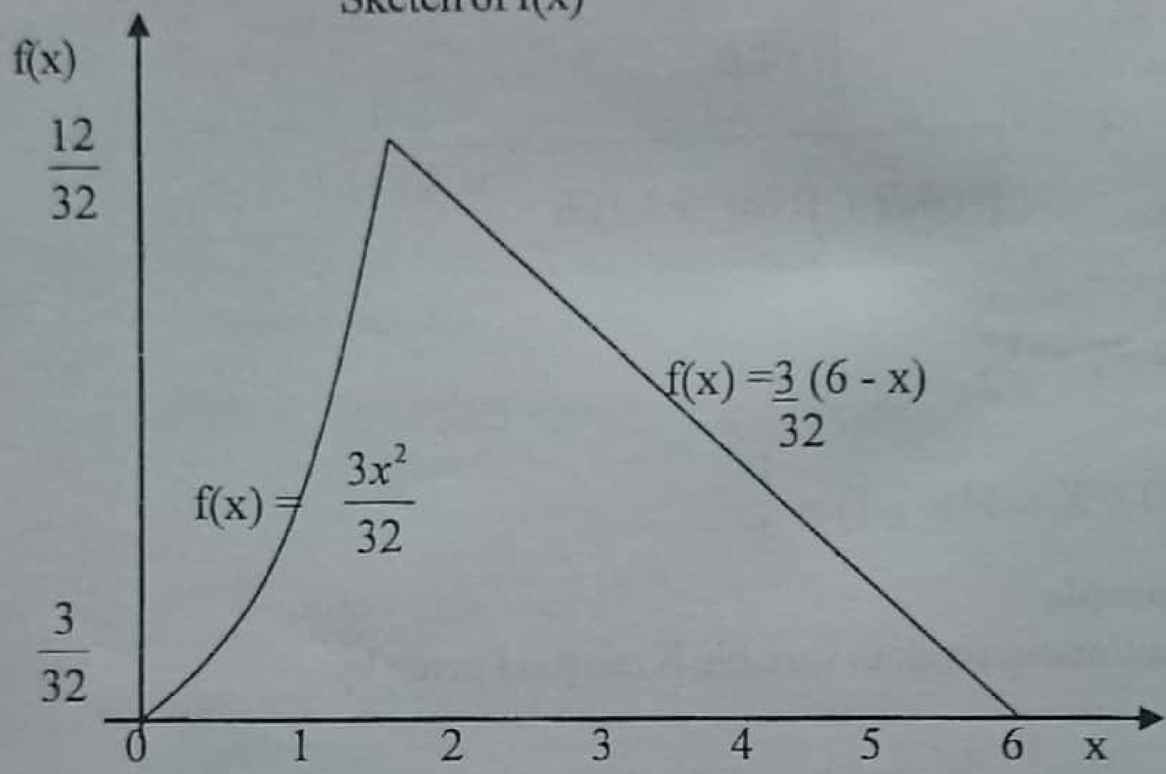
x	*0*	*1*	**2**
f(x)	0	$\frac{3}{32}$	$\frac{12}{32}$

For the interval $2 \leq x \leq 6$. This is a straight line

x	2	6
f(x)	$\frac{12}{32}$	0

Therefore you can use three points .

Sketch of f(x)



Example

A continuous random variable X has a probability density function given by:

$$f(x) = \begin{cases} \alpha, & 2<x<3 \\ \alpha(x-2), & 3<x<4 \\ 0 & otherwise \end{cases}$$

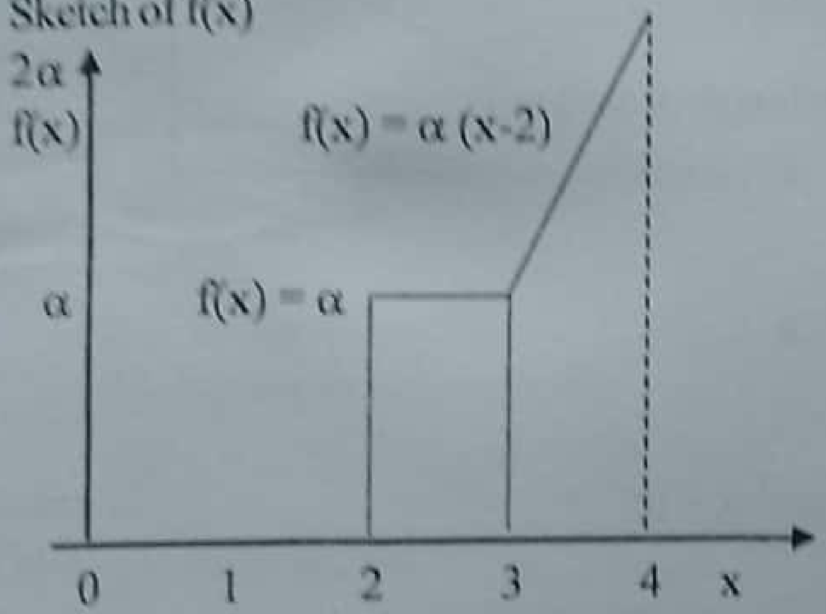
(i) Sketch f(x)

(ii) Find the value of α, hence f(x)

(iii) $P(|x - 2.5| < 0.5)$

Solution

Sketch of f(x)



(ii) From the sketch of f(x)

$\Rightarrow \alpha \times 1 + 0.5 \times 1(\alpha + 2\alpha) = 1$

$So\ 5\alpha = 2$

$\therefore \alpha = \frac{2}{5}$

$$f(x) = \begin{cases} \frac{2}{5}, & 2 < x < 3 \\ \frac{2}{5}(x-2), & 3 < x < 4 \\ 0 & elsewhere \end{cases}$$

(iii) $P(|x - 2.5| < 0.5)$

$P(-0.5 < x - 2.5 < 0.5)$

$P(2.5 - 0.5 < x < 2.5 + 0.5)$

$P(2 < x < 3)$

$= \text{Area of the rectangle from } 2 \text{ to } 3$

$1 x \alpha = \frac{2}{5}$

$P(2 < x < 3) = 0.4$

Note:

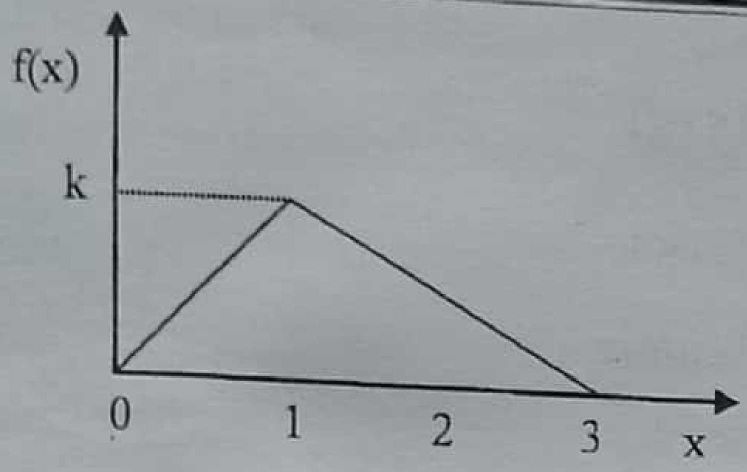
$P(|x| < a) = P(-a < x < a)$

$P(|x| > a) = 1 - P(-a < x < a)$

5.4 How to obtain probability density function from the sketch.

Example

A continuous random variable X has a probability density function f(x) given by $f(x) = 0$ for $x < 0$ and $x > 3$ and between $x = 0$ and $x = 3$ its form is as shown in the graph.



(a) Find the value of k

(b) Express f(x) algebraically

Solution

Total area under the curve = 1

$$\frac{1}{2} x 3 x k = 1$$

$$3k = 2$$

(a)

$$\text{Therefore} \quad k = \frac{2}{3}$$

$$0 \le x \le 1, \ f(x) = \frac{2}{3}x + 0 = \frac{2}{3}x$$

$$1 \le x \le 3, \ (1, \frac{2}{3}) \ \text{and} \ (3, 0)$$

(b)

$$\text{Gradient} = \frac{0 - \frac{2}{3}}{3 - 1} = \frac{-1}{3}$$

$$\text{Equation} = \frac{y - 0}{x - 3} = \frac{-1}{3}$$

$$y = 1 - \frac{x}{3}$$

Otherwise the expression of $f(x) = 0$

5.5 EXPECTATION (MEAN) AND VARIANCE

The mean for continuous p.d.f is given by

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

$$Var(X) = E(X^2) - E(X)^2$$

$$= \int x^2 f(x)dx - [E(X)]^2$$

Not that $\mu = E(X)$

$$\Rightarrow Var(X) = \int_{allx} x^2 f(x)dx - \mu^2$$

and $Var(X) = \sigma^2$

Example

The continuous random variable X has p.d.f f(x) where

$$f(x) = \begin{cases} \frac{x}{8} & 0 \le x \le 4 \\ 0 & \text{otherwise} \end{cases}$$

Determine (i) E(X) (ii) Var (X)
(iii) standard deviation of X.

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

$$\int_{all\,x} xf(x)dx = \int_0^4 x\frac{1}{8}x\,dx$$

$$= \int_0^4 x\frac{x}{8}dx = \int_0^4 \frac{x^2}{8}dx$$

$$\left[\frac{x^3}{24}\right]_0^4 = \frac{64}{24} = \frac{8}{3}$$

Mean = $8/3$

Variance = $E(X^2) - [E(X)]^2$

$E(X^2) = \int_0^4 \frac{x^3}{8}dx$

$$= \left[\frac{x^4}{32}\right]_0^4 = 256/32 = 8$$

$$Var(X) = 8 - \left(\frac{8}{3}\right)^2 = 8 - \frac{64}{9} = \frac{8}{9}$$

$$Var(X) = \frac{8}{9}$$

(iii) Standard deviation (σ) = $\sqrt{8/9}$

$$= \frac{2\sqrt{2}}{3}$$

Properties of E(X)

1. $E(a) = a$
2. $E(ax) = a\,E(x)$
3. $E(aX) + b) = E(X) + b$.

Properties of Var(X)

1. $Var(a) = 0$
2. $Var(aX) = a^2 Var(X)$
3. $Var(aX + b) = a^2 Var(X)$ where a and b are constants.

Example

The outputs of 9 machines in a factory are independent random variables each with p.d.f given by:

$$f(x) = \begin{cases} ax & 0 \le x \le 10 \\ a(20 - x) & 10 \le x \le 20 \\ 0 & \text{otherwise} \end{cases}$$

Find

(i) the value of a.

(ii) the expected value and variance of the output of each machine. Hence or otherwise find the expected value and variance of the total output from all machines.

Solution: (i)

$$\int_{all\,x} f(x)dx = 1$$

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

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

$$50a + a[200 - 150] = 1$$

$$100 = 1 \Rightarrow a = 0.01$$

(ii) $E(X) = \int_{all} xf(x)dx$

$$= \int_0^{10} xf(x)dx + \int_{10}^{20} xf(x)dx$$

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

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

$$= \frac{1000}{300} + \frac{20}{3} = \frac{10}{3} + \frac{20}{3} = \frac{30}{3} = 10$$

For 9 machines $E(9X) = 9E(X)$

$9 \times 10 = 90$, Expected value is 90.

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

but $E(X^2) = \int_{all x} x^2 f(x)dx$

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

$$E(X^2) = \frac{1}{100}\left[\frac{x^4}{4}\right]_0^{10} + \frac{1}{100}\left[\frac{20x^3}{3} - \frac{x4}{4}\right]_{10}^{20}$$

$$E(X^2) = \frac{1}{100}\left[\left(\frac{160,000}{3} - 40,000\right) - \left(\frac{20,000}{3} - \frac{10,000}{4}\right)\right]$$

$$E(X^2) = 25 + \frac{1,400}{3} + 25 - 400$$

$$E(X^2) = 116.667$$

$$Var(X) = 116.6667 - (10)^2 = 16.6667$$

$$Var(9X) = 9^2 Var(X) = 81\,Var(X) =$$

$$81 \times 16.6667 = 1,350.$$

Assignment

The mass X kg of loaves of bread produced per hour in certain bakery is modeled by a continuous random variable with a probability function given by:

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

(a) find the value of k

(b) Given that a 1 kg loaf is sold at Ushs. 1500 and the running costs of baking is 1200/= per hour. Taking Y/= as the profit made in each express Y in terms of X, hence find E(Y).

5.6 THE MEDIAN

The median divides the area under the curve into two halves, consider m to be the median then,

$$\int_{-\infty}^{m} f(x)dx = 0.5 \quad or \quad \int_{m}^{\infty} f(x)dx = 0.5$$

Example

If X is a continuous random variable with p.d.f given by

$$f(x) = \begin{cases} \frac{x}{8} & 0 \le x \le 4 \\ 0 & otherwise \end{cases}$$

Obtain the median m.

Solution

$$\int_{0}^{m} f(x)dx = 0.5 \qquad \Rightarrow \int_{0}^{m} \frac{x}{8} dx = 0.5$$

$$\left[\frac{x^2}{16}\right]_0^m = 0.5$$

$$m^2 = 8 \therefore m = \pm 2.828$$

Note $0 < m < 4$, Median = 2.828 since -2.828 is outside the interval.

Example

A continuous random variable X has p.d.f f(x)

$$f(x) = \begin{cases} 0.4x + 0.8 & -1 \le x \le 0 \\ 0.8 - 0.8x & 0 \le x \le 1 \\ 0 & \text{elsewhere} \end{cases}$$

Obtain the median

We first integrate the first interval to establish if it is at least 0.5. If not then median will be in the second interval.

$$\int_{-1}^{0} f(x)dx = \int_{-1}^{0}(0.4x+0.8)dx$$

$$\left[0.2x^2+0.8x\right]_{-1}^{0} = 0-(0.2-0.8) = 0.6$$

Therefore the median lies in the first interval $-1 \leq x\, 0$

$$\left[0.2x^2+0.8x\right]_{-1}^{m} = .05$$

$0.2m^2 + 0.8m + 0.6 = 0.5$
either m = -0.129 or -3.871 (Cal)
m = -0.129 since -3871 is outside the interval

Example
A random variable X has probability density function given by

$$f(x) = \begin{cases} \frac{2}{3}(x+1); & -1 < x \leq 0 \\ \frac{1}{3}(2-x); & 0 < x \leq 2 \\ 0 & \text{otherwise} \end{cases}$$

where a is constant, determine the median of X

Solution:
Let m be the median
We integrate the first interval to establish whether it is at least 0.5

$$\int_{-1}^{0} f(x)dx = \int_{-1}^{0} \frac{2}{3}(x+1)dx$$

$$= \frac{2}{3}\left[\frac{(x^2+x)}{2}\right]_{-1}^{0}$$

$$\frac{2}{3}x\left(0+1-\frac{1}{2}\right) = \frac{1}{3}$$

$$\Rightarrow \frac{1}{3} < \frac{1}{2}$$

m lies between 0 and 2.

$$\Rightarrow \int_{-1}^{0} f(x)dx + \int_{0}^{m} \tfrac{1}{3}(2-x)dx = 0.5$$

$$\frac{1}{3} + \frac{1}{3}\left(2x - \frac{x^2}{2}\right)_0^m = 0.5$$

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

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

$$4m - m^2 = 1 \Rightarrow m^2 - 4m + 1 = 0$$

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

$$m = \frac{4 + \sqrt{16-4}}{2}$$

$$m = 2 + \frac{\sqrt{12}}{2}$$

$$m = 2 \pm \frac{3.464}{2} \quad or \quad 2 \pm 1.732$$

0.268 lies between 0 and 2 while 3.732 does not.
Hence the median is 0.268.

5.7 MODE

This is the value of x for which f'(x)= $\frac{d}{dx} f(x) = 0$

Example
The continuous random variable x has p.d.f f(x) where

$$f(x) = \frac{3}{80}(2+x)(4-x) \quad 0 \leq x \leq 4.$$

Find the mode.

Solution
For mode use f'(x)=0

$$f(x) = \frac{3}{80}\left(-x^2 + 2x + 8\right) \Rightarrow f'(x) = {}^3\!/_{80}\left(-2x + 2\right) = 0$$

$$\frac{3}{80}(2-2x) = 0 \text{ , when } 2 - 2x = 0 \Rightarrow x = 1$$

Note if you have more than one value check for one which gives a maximum i.e $f''(x) < 0$.

$$f(x)=\begin{cases}\frac{x}{108}(6-x)^2 & 0<x\le 6\\ 0 & \text{otherwise}\end{cases}$$

obtain the mode

$$f(x)=\frac{x}{108}(6-x)^2$$

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

$$f'(x)=\frac{36-24x+3x^2}{108}=0$$

$$\Rightarrow\frac{3}{108}\left(x^2-8x+12\right)=0$$

$$\frac{3}{108}(x-6)(x-2)=0 \ \textit{Either}\ x=6\ \ Or\ \ x=2$$

$$f''(x)<0, f'(x)=\frac{3}{108}\left(x^2-8x+12\right)$$

$$f''(x)=\frac{3}{108}(2x-8), \textit{when}\ x=6\Rightarrow f''(6)>0$$

$$f''(2)=\frac{3}{108}(4-8)=\frac{-12}{108}\Rightarrow f''(2)<0$$

Hence the mode is 2.

5.8 THE CUMULATIVE DISTRIBUTION FUNCTION

The cumulative probability function for any variable X is defined by

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

Properties of F(x)

(i) $F(-\infty) = 0$ (ii) $F(\infty) = 1$

(iii) None decreasing function.

Example

Given the following p.d.f f(x) determine F(x)

$$f(x) = \begin{cases} \frac{x}{8} & 0 \le x \le 4 \\ 0 & otherwise \end{cases}$$

Solution

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

For $x < 0$, $f(t) = 0 \Rightarrow F(x) = 0$ and $F(0) = 0$

For $0 < x < 4$, $f(t) = \frac{t}{8}$

$$F(x) = F(0) + \int_0^x \frac{t}{8} dt$$

$$= 0 + \left[\frac{t^2}{16}\right]_0^x$$

$$\Rightarrow F(x) = \frac{x^2}{16} \qquad \Rightarrow F(4) = \frac{16}{16} = 1$$

for $x > 4$ $f(t) = 0$

$$F(x) = F(4) + \int_0^x f(t)dt$$

$$F(x) = F(4) + 0 = 1$$

$$F(x) = \begin{cases} 0 & x<0 \\ \dfrac{x^2}{16} & 0 \le x \le 4 \\ 1 & x \ge 4 \end{cases}$$

Example

Given the following p.d.f f(x) determine F (x)

$$f(x) = \begin{cases} \dfrac{x}{3} & 0 \le x \le 2 \\ \dfrac{-2x}{3}+2 & 2 \le x \le 3 \\ 0 & otherwise \end{cases}$$

Solution

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

For the interval $x < 0$, $f(t) = 0$, $F(0) = 0$

for $0 \le x \le 2$ $\quad f(t) = \dfrac{t}{3}$

$$\Rightarrow F(x) = F(0) + \int_0^x \frac{t}{3}\, dt,$$

$$F(x) = \frac{x^2}{6}, \Rightarrow F(2) = \frac{4}{6} = \frac{2}{3}$$

For $2 < x < 3 \Rightarrow f(t) = \dfrac{-2t}{3} + 2$

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

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

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

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

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

$$\Rightarrow F(3) = \frac{-9}{3} + 6 - 2 = 1$$

for $x > 3$, $f(t) = 0$, $F(x) = F(3) + 0 = 1$

$$F(x) = \begin{cases} 0 & x < 0 \\ \dfrac{x^2}{6} & 0 \le x \le 2 \\ \dfrac{-x^2}{3} + 2x - 2 & 2 \le x \le 3 \\ 1 & x \ge 3 \end{cases}$$

Note: The cumulative probability function F(x) can be used to obtain.

5.9 Probaliites, median, p.d.f and constants

Example

Given the following distribution function F(x) determine:

(i) $p(1 \leq x \leq 4)$ ii) Median   iii) p.d.f f(x)

$$F(x) = \begin{cases} 0 & x < 0 \\ \frac{x^2}{12} & 0 \leq x \leq 3 \\ 2x - \frac{x^2}{4} - 3 & 3 \leq x \leq 4 \\ 1 & x \geq 4 \end{cases}$$

i) $P(x_1 \leq X \leq x_2) = P(X \leq x_2) - P(X \leq x_1)$

$= F(x_2) - F(x_1)$

$P(1 \leq X \leq 4) = P(X \leq 4) - P(X \leq 1)$

$= F(4) - F(1)$

$F(4) = 1$

$F(1) = 1/12$

$\therefore P(1 \leq X \leq 4) = 1 - \frac{1}{12} = \frac{11}{12}$

ii) $F(m) = 0.5$ ie $\int_{-\infty}^{m} f(x)dx = 0.5$

For $0 < x < 3$, $F(3) = \frac{9}{12} = \frac{3}{4}$

$F(3) > 0.5$

$\Rightarrow$ m lies between $0 < x < 3$

$F(m) = 0.5$

$\Rightarrow \quad F(x) = \frac{x^2}{12}, \qquad \frac{1}{2} = \frac{m^2}{12}$

$6 = m^2 \Rightarrow m = \pm\sqrt{6}$

$m = 2.449$ the median is 2.449 since -2.449 is outside the range

iii) p.d.f(x) $f(x) = F^1(x)$

$$f(x) = \frac{d}{dx}\left(\frac{x^2}{12}\right) = \frac{x}{6} \quad \text{for } 0 < x < 3$$

$$\text{for } 3 < x < 4,\ f(x) = F^1(x) \Rightarrow \frac{d}{dx}\left(2x - \frac{x^2}{4} + 3\right)$$

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

otherwise $f(x) = 0$

$$\therefore f(x) = \begin{cases} \frac{x}{6} & 0 \le x \le 3 \\ 2 - \frac{1}{2}x & 3 \le x \le 4 \\ 0 & \text{otherwise} \end{cases}$$

Example

A continuous random variable X has a probability density function:

$f(x) = kx(3-x)$ for $0 \leq x \leq 2$

$f(x) = k(4-x)$ for $2 \leq x \leq 4$

$f(x) = 0$ otherwise

Determine

i) The value of k
ii) Mean
iii) F(x) the cumulative distribution function
iv) $P(1 \leq X \leq 3)$

Solution

(i) $$\int_{all} f(x)dx = 1$$

$$k\int_0^2 (3x - x^2)dx + K\int_2^4 (4-x)dx = 1$$

$$k\left[\frac{3x^2}{2} - \frac{x^3}{3}\right]_0^2 + k\left[4x - \frac{x^4}{2}\right]_2^4 = 1$$

$$k\left[6 - \frac{8}{3}\right] + k[8-6] = 1$$

$$\frac{16}{3}k = 1 \quad k = \frac{3}{16}$$

(ii) The mean $E(X) = \int_{all x} xf(x)dx$

$$E(X) = \frac{3}{16}\left[\int_0^2 (3x^2 - x^3)dx + \int_2^4 (4x - x^2)dx\right]$$

$$\frac{3}{16}\left(\left[x^3 - \frac{x^4}{4}\right]_0^2 + \left[2x^2 - \frac{x^3}{3}\right]_2^4\right)$$

$$E(X) = \frac{3}{16}\left[(8-4) + (32 - \frac{64}{3}) - (8 - \frac{8}{3})\right]$$

$$= \frac{7}{4} = 1\frac{3}{4} \quad \Rightarrow E(X) = 1\frac{3}{4}$$

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

For $x < 0$, $F(x) =$ since $f(t) = 0$

For $0 < x < 2$, $f(t) = kt(3 - t)$

$$F(x) = F(0) + \frac{3}{16}\int_0^x (3t - t^2)dt$$

$$= \frac{3}{16}\left[\frac{3t^2}{2} - \frac{t^3}{3}\right]_0^x \Rightarrow F(x) = \frac{3}{16}\left(\frac{3x^2}{2} - \frac{x^3}{3}\right)$$

$$F(2) = \frac{3}{16}\left(6 - \frac{8}{3}\right) = \frac{10}{16}$$

$2 < x < 4$

$$F(x) = F(2) + \frac{3}{16}\int_2^x (4-t)\,dt$$

$$= \frac{10}{16} + \frac{3}{16}\int_2^x (4-t)\,dt$$

$$= \frac{10}{16} + \frac{3}{16}\left[4t - \frac{t^2}{2}\right]_2^x$$

$$= \frac{10}{16} + \frac{3}{16}\left[(4x - \frac{x^2}{2}) - (8-2)\right]$$

$$= \frac{10}{16} + \frac{3}{16}\left[4x - \frac{x^2}{2} - 6\right]$$

$$F(x) = \frac{3}{16}(4x - \frac{x^2}{2}) - \frac{1}{2}$$

$F(4) = 1$

For $x > 4$ $\qquad F(x) = 1$

$$F(x) = \begin{cases} 0 & x \le 0 \\ \frac{3}{16}\left(\frac{3x^2}{2} - \frac{x^3}{3}\right) & 0 \le x \le 2 \\ \frac{3}{16}\left(4x - \frac{x^2}{2}\right) - \frac{1}{2} & 2 \le x \le 4 \\ 1 & x \ge 4 \end{cases}$$

(iv) $P(1 \le X \le 3) = P(X \le 3) - P(X \le 1)$

$= F(3) - F(1)$

$$F(3) = \frac{3}{16}(12 - \frac{9}{2})\frac{1}{2} = \frac{3}{16}\left(\frac{15}{2}\right) - \frac{1}{2}$$

$$= \frac{45}{32} - \frac{1}{2} = \frac{29}{32}$$

$$F(1) = \frac{3}{16}\left(\frac{3}{2} - \frac{1}{3}\right) = \frac{3}{16} x \frac{7}{6} = \frac{7}{32}$$

$$P(1 \le X \le 3) = \frac{29}{32} - \frac{7}{32} = \frac{22}{32} = \frac{11}{16}$$

$$\therefore P(1 \le X \le 3) = \frac{11}{16}$$

Assignment

A random variable X has F(x) given by

$$F(x) = \begin{cases} 0 & x < 0 \\ \dfrac{x^2}{4} & 0 \le x \le 1 \\ \alpha x + k & 1 \le x \le 2 \\ \dfrac{1}{4}(5-x)(x-1) & 2 \le x \le 3 \\ 1 & x \ge 3 \end{cases}$$

Determine

(a) The value of the constants α and k

(b) $P(3 < 2X < 5)$

(c) $P(|X-2| < 1)$

Note: Graph F(x) is plotted the same way as that of f(x) however its sketch should be given from $-\infty$ to $+\infty$

Example

Sketch the graph of F(x) for the function given below

$$F(x) = \begin{cases} 0 & x \le 0 \\ \dfrac{x^2}{6} & 0 \le x < 2 \\ \dfrac{-x^2}{3} + 2x - 2 & 2 \le x < 3 \\ 1 & x \ge 3 \end{cases}$$

Solution

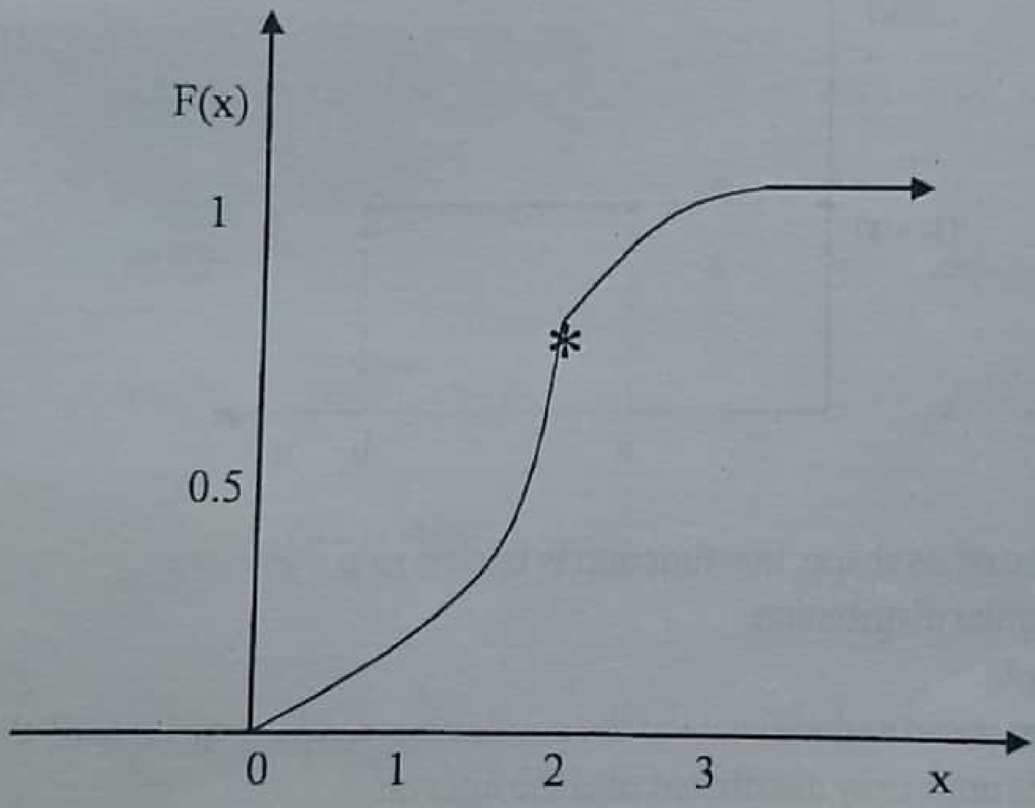
When $x < 0$, $F(0) = 0$

For $0 \leq x \leq 2$, $F(x) = x^2/6$ and $F(2) = 2/3$

For $2 \leq x \leq 3$, $F(x) = \dfrac{-x^2}{3} + 2x - 2$ and

$$F(3) = \frac{-3^2}{3} + 2x3 - 2 = 1$$

For $x > 3$, $F(x) = 1$



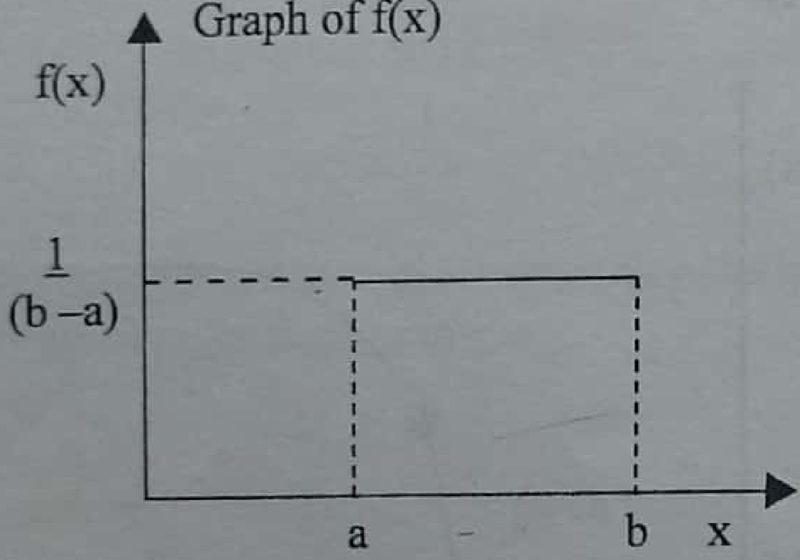
5.10 UNIFORM DISTRIBUTION

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

$$f(x) = \begin{cases} \frac{1}{b-a} & a \leq x \leq b \\ 0 & \text{otherwise}. \end{cases}$$

Where $a \leq b$

Graph of f(x)



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

Example

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

(a) 0 to 1

(b) 2 to k

$$\text{(a) } f(x) = \begin{cases} \dfrac{1}{b-a} & a \le x \le b \\ 0 & otherwise \end{cases}$$

$$f(x) = \begin{cases} 1 & 0 \le x \le 1 \\ 0 & otherwise \end{cases}$$

where a = 0 and b = 1

$$E(X) = \int_{all} x f(x) dx \quad = \int_0^1 x dx = \left[\frac{x^2}{2}\right]_0^1 = \frac{1}{2}$$

$$E(X) = \frac{1}{2}$$

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

$$E(X^2) = \int_{all} x^2 f(x) dx$$

$$\int_0^1 x^2 dx = \left[\frac{x^3}{3}\right]_0^1 = \frac{1}{3}$$

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

$$\text{Var}(X) = \frac{1}{3} - \frac{1}{4} = \frac{1}{12}$$

(b) $$E(X) = \int_2^k \frac{x}{k-2}\,dx$$

$$= \frac{1}{k-2}\left[\frac{x^2}{2}\right]_2^k$$

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

$$E(X) = \frac{k+2}{2}$$

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

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

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

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

$$E(X^2) = = \frac{k-2}{k-2}\left(\frac{(k-2)^2 + 6k}{3}\right)$$

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

$$\text{Var}(X) = = \left(\frac{k^2+2k+4}{3}\right) - (\frac{k+2}{2})^2$$

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

$$= \left(\frac{k^2-4k+4}{12}\right) = \frac{(k-2)^2}{12}$$

Note for uniform distribution,

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

$$\therefore \text{Var}(X) = \frac{(k-2)^2}{12}$$

where k = b and a = 2

Assignment

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 vechiles 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 vechiles leave the stage

Exercise

1.A continuous random variable X has the distribution function

$$F(x) = \begin{cases} 3kx\left(1-\frac{x^2}{3}\right) & 0 \le x \le 1 \\ 1 & x > 1 \end{cases}$$

Determine

i) the value of k
ii) the probability density function of X
iii) the mean of X
iv) $P(X > 0.5 \mid 0.25 < X < 1)$

Answer: i) k = ½

$$\text{(ii) } f(x) = \begin{cases} \frac{3}{2}\left(1-x^2\right) & 0 \le x \le 1 \\ 0 & \textit{otherwise} \end{cases}$$

(iii) $E(x) = \frac{3}{8}$ iv) 0.494

2. The probability density function of a random variable X is

$$f(x) = \begin{cases} k\sin x \text{ for } & 0 \le x \le \pi \\ 0 & \text{otherwise} \end{cases}$$

Determine

i) the value of k

ii) $P(X > \frac{\pi}{3})$ (iii) the median of X

Answer

(i) $k = \frac{1}{2}$, (ii) 0.75 (iii) $\frac{\pi}{2}$

3. A continuous random variable X has a probability density function given by

$$f(x) = \begin{cases} \frac{1}{4} & 0 < x < 1 \\ \frac{1}{8} & 2 < x < 8 \\ 0 & \text{elsewhere} \end{cases}$$

Find

(i) the distribution function and expectation of X

(ii) $P[0.5 < X < 3]$

Answer

(i) $$F(x) = \begin{cases} 0 & x < 0 \\ \frac{x}{4} & 0 \leq x \leq 1 \\ \frac{1}{4} & 1 \leq x \leq 2. \\ \frac{x}{8} & 2 \leq x \leq 8 \\ 1 & x \geq 8 \end{cases}$$

$E(X) = 3.875$ ii) 0.25

4. A random variable X has the probability density function given by

$$f(x) = \begin{cases} k(1-x^2) & 0 \le x \le 1 \\ 0 & \text{otherwise} \end{cases}$$

where k is a constant, find

i) the value of the constant k
ii) the mean of X
iii) the variance of X

Answer

i) $k = 1.5$ ii) $E(X) = \frac{3}{8}$ (iii) $\frac{19}{320}$

5. The number X of cars crossing Owen falls Dam daily is uniformly distributed between 1,026 to 3,025 cars.

(i) Find the probability that at least 1,625 cars cross the bridge.
(ii) What is the expected number of cars that will cross the bridge on any given day?

Answer

i) 0.7 ii) E(X) = 2,026 cars.

6. A random variable X has the probability density function

$$f(x) = \begin{cases} \frac{2}{3a}(x+a) & -a \le x \le 0 \\ \frac{1}{3a}(2a-x) & 0 \le x \le 2a \\ 0 & \textit{otherwise} \end{cases}$$

where a is constant

Determine

i. the value of a
ii. the median of X

(iii) $P(X \leq 1.5 \mid X > 0)$

(iv) The cumulative distribution function F(x)

(v) Sketch the graph of F(x)

Answer

i) $a = 1$ ii) median $= 0.268$,

iii) $P[(x \leq 1.5) \mid (x \leq)] = 0.9375$

$$F(x) = \begin{cases} 0 & x \leq -1 \\ \frac{2}{3}\left(\frac{x^2}{2} + x + \frac{1}{2}\right) & -1 \leq x \leq 0 \\ \frac{1}{3} + \frac{2x}{3} - \frac{-x^2}{6} & 0 \leq x \leq 2 \\ 1 & x \geq 2 \end{cases}$$

7. A.p.d.f is given by

$$f(x) = \begin{cases} \frac{1}{6}x & for \quad 0 \leq x \leq 3 \\ \frac{1}{2}(4-x) & for \quad 3 \leq x \leq 4 \\ 0 & for \quad x < 0 \ and \ x > 4 \end{cases}$$

sketch the graph of f(x)
calculate

(a) the probability that X occurs in the interval (1,2)

(b) the probability that $x > 2$. Obtain the cumulative probability function and hence or otherwise, find the median of the distribution.

Answer (a) 0.25 (b) $\frac{2}{3}$

$$F(x) = \begin{cases} 0 & x < 0 \\ \frac{x^2}{12} & 0 \le x \le 3 \\ 2x - \frac{x^2}{4} - 3 & 3 \le x \le 4 \\ 1 & x > 4 \end{cases}$$

$m = \sqrt{6} = 2.449$

8. The probability that a transistor in a radio lasts less than t hours is $1 - e^{-t/2000}$. Find the p.d.f for the lifetime of a transistor?

a) What is the probability that the transistor lasts more than 4,000hours?

b) What is the probability that a transistor ceases to function after 2,000 hours, of use but before 3,000 hours?

c) If a radio contains 8 transistors, what is the probability that none of them fails before 1,000 hours of use?

Answer a)0.135 b)0.145 c)0.18

$$f(t) = \frac{1}{2000} e^{\frac{-t}{2000}}$$

9. A random variable X has cumulative probability function.

$$F(x) = \begin{cases} 0 & x \le a \\ \dfrac{x-a}{b-a} & a \le x \le b \\ 1 & x \ge b \end{cases}$$

Find the p.d.f f(x) and sketch the graph of f(x). Obtain the mean and variance of X.

Answer

$$f(x) = \begin{cases} \dfrac{1}{b-a} & a \le x \le b \\ 0 & \text{otherwise} \end{cases}$$

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

10. The continuous random variable X has p.d.f f(x) given by

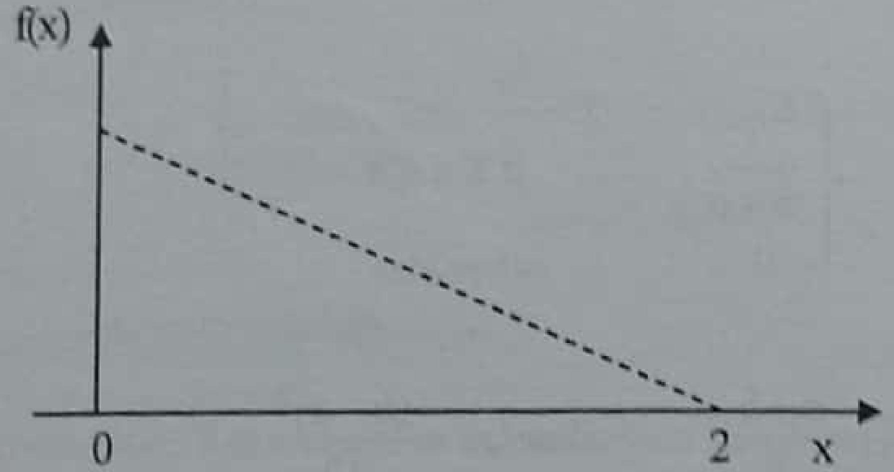
$$f(x) = \begin{cases} k & 0 \le x \le 2 \\ k(3-x) & 2 \le x \le 3 \\ 0 & \text{otherwise} \end{cases}$$

a) Find the value of k
(b) Hence evaluate E(X) and the median of X
(c) Prove that σ the standard deviation of X is 0.75, correct to 2 decimal places.
(d) Denoting E(X) by μ, find the $P[(X < \mu - \sigma)]$

Answer

a) $k = 0.4$ b) $E(X) = \frac{19}{15}, \frac{5}{4}$ d) 0.207

11. The probability density function f(x) of a random variable X takes on the form shown in the diagram below.



Determine the expression for f(x). Hence obtain the expression for the cumulative probability density function of X.

Answer $f(x) = \begin{cases} 1 - x/2 & 0 \leq x \leq 2 \\ 0 & \text{otherwise} \end{cases}$

$$F(x) = \begin{cases} 0 & x < 0 \\ x - \frac{x^2}{4} & 0 \leq x \leq 2 \\ 1 & x > 2 \end{cases}$$

12. The quality of an animal feedstuff depends both upon the raw materials used and the production process. One measure of this quality is the nutritional index, which varies between 0 and 1. A particular mill produces an animal feedstuff in batches of constant size, and the nutritional index of any batch may be considered to be having probability density function.

$$f(x) = \begin{cases} kx(1-x)^2 & 0 \le x \le 1 \\ 0 & \text{otherwise} \end{cases}$$

(i) show that $k = 12$ and sketch $f(x)$

ii. calculate $P(x < 0.25)$

iii. batches of this mill's feedstuff may be sold for 500/= each if the index is 0.8 or more, and 350/= otherwise. The cost of producing a batch is 300/=. What is the expected profit per batch?

Answer

(ii)0.262 (iii) $E(x) = 0.4$ (iv) 54.08/=

13. The continuous random variable X has cumulative distribution function F(x) given by

$$F(x) = \begin{cases} 0 & x < -1 \\ \dfrac{1+x}{8} & -1 \le x \le 0 \\ \dfrac{1+3x}{8} & 0 \le x \le 2 \\ \dfrac{5+x}{8} & 2 \le x \le 3 \\ 1 & x \ge 3 \end{cases}$$

i) Sketch the graph of the probability density function f(x)

ii) Determine the expectation of X and the variance of X

iii) Determine $P(3 \le 2X \le 5)$

Answer

ii) $E(X) = 1$ $\quad Var(X) = \frac{5}{6}$ iii) $P(3 \leq 2X \leq 5) = 0.25$

14. Petrol is delivered to a garage every Monday morning. At this garage the weekly demand of petrol, in thousands of units is a continuous random variable X distributed with a probability function of the form.

$$f(x) = \begin{cases} ax^2(b-x) & 0 \leq x \leq 1 \\ 0 & \text{otherwise} \end{cases}$$

I. Given that the mean weekly demand is 600 units, determine the value of a and b.

II. If the storage tanks at this garage are filled to their capacity of 900 units every Monday morning, what is the probability that in any given week the garage will be unable to meet the demand of petrol.

III. What is the mode of X.

Answer

I) $a = 12$ $\quad b = 1$ $\quad$ (II) 0.0523 $\quad$ (III) $\frac{2}{3}$

15. A random variable X has F(x) given by

$$F(x) = \begin{cases} 0 & x < 0 \\ \frac{x^2}{4} & 0 \leq x \leq 1 \\ ax + k & 1 \leq x \leq 2 \\ \frac{1}{4}(5-x)(x-1) & 2 \leq x \leq 3 \\ 1 & x \geq 3 \end{cases}$$

Determine

(a) The value of the constants a and k

(b) $P(3 \le 2X \le 5)$

(c) The probability density function f(x) sketch graph of f(x),and hence or otherwise, deduce the mean of X.

Answer

(a) a = 0.5, k = -0.25 b) $\frac{3}{2}$, Var (X) $= \frac{5}{12}$

$$f(x) = \begin{cases} \frac{1}{2}x & 0 \le x \le 1 \\ \frac{1}{2} & 1 \le x \le 2 \\ -\frac{1}{2}x + \frac{3}{2}. & 2 \le x \le 3 \\ 0 & \text{otherwise} \end{cases}$$

16. A continuous random variable has probability function f(x)

f(x) = 0	x < 0
f(x) = x	$0 \le x \le 1$
f(x) = ½	$1 \le x \le 2$
f(x) = 0	$x \ge 2$

(a) Sketch f(x) and show that it satisfies the conditions to be a probability density function.

(b) Find the cumulative distribution function F(x).

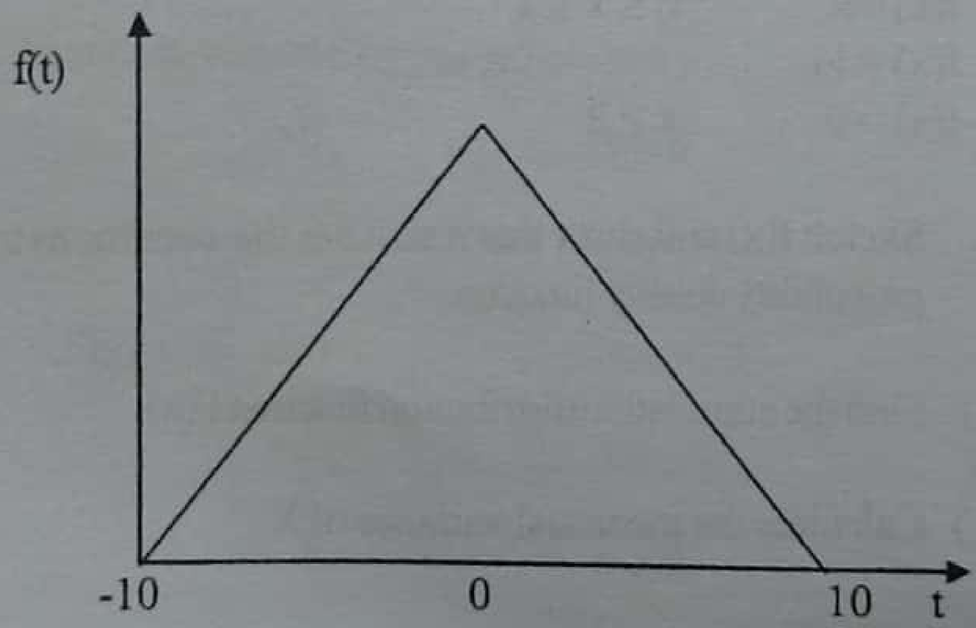
(c) Calculate the mean and variance of X.

Answer

$$F(x) = \begin{cases} 0 & x < 0 \\ \frac{1}{2}x^2 & 0 \le x \le 1 \\ \frac{1}{2}x & 1 \le x \le 2 \\ 1 & x > 2 \end{cases}$$

(c) $E(X) = \frac{13}{12}$ $Var(X) = 0.243$

17. A student cycles to school each day. He is supposed to be at school by 8:50 am, but aims at getting there by 8:45. Over a full year his arrival times are noted and recorded. The student never arrives earlier than 8.35am and never later than 8.55am. An examination of this data suggests that an appropriate model of the distribution of his arrival times is given by the probability density function f(t) shown in the figure below, where the students arrives t minutes after 8.45am.



State the value of f(t) when $t=0$

Find for a day chosen at random

a) Probability that he arrives by 8.42am

b) The probability in terms of t that he arrives before the time given by t where $-10<t<0$.

c) The probability in terms of t that he arrives before the time given by t where $0<t<10$, hence specify fully the distribution function F(t), sketch the graph of F(t).

Answer 0.1 a)0.245 b) $\frac{1}{200}(10+t)^2$

c) 1- $\frac{1}{200}(10+t)^2$

18. A continuous random variable X has p.d.f given by

$f(x)=ax-bx^2$ for $0\le x\le 2$

$f(x)=0$ elsewhere

observation on X indicates that the mean is 1

a) Obtain two simultaneous equations for a and b, show that $a=1.5$ and find the value of b.

b) Find the variance of X.

c) If F(x) is the probability that $X\le x$, find F(x) and verify that $F(2)=1$

d) If two independent observations are made on X what is the probability that at least one of them is less than 0.5?

Answer a) $b=0.75$ b) $Var(x)=0.2$ d) 0.288

19. The time taken to perform a particular task, t hours, has the p.d.f given by

$$f(t)=\begin{cases} 10Ct^2 & 0\ t\ 0.6 \\ 9C(1-t) & 0.6\ t\ 1.0 \\ 0 & \text{otherwise} \end{cases}$$

Where C is a constant.

a) Find the value of C and sketch the graph of this distribution.
b) Write down the most likely time.
c) Find the expected time.
d) Determine the probability that the time will be-

i. More than 48 minutes

(ii) Between 24 and 48 minutes

Answer

a) $C = \frac{25}{36}$ b) most likely time is 36 minutes

c) $E(T) = 35.5$ minutes d) (i) 0.125 (ii) 0.727

20. The continuous random variable X has cumulative distribution function.

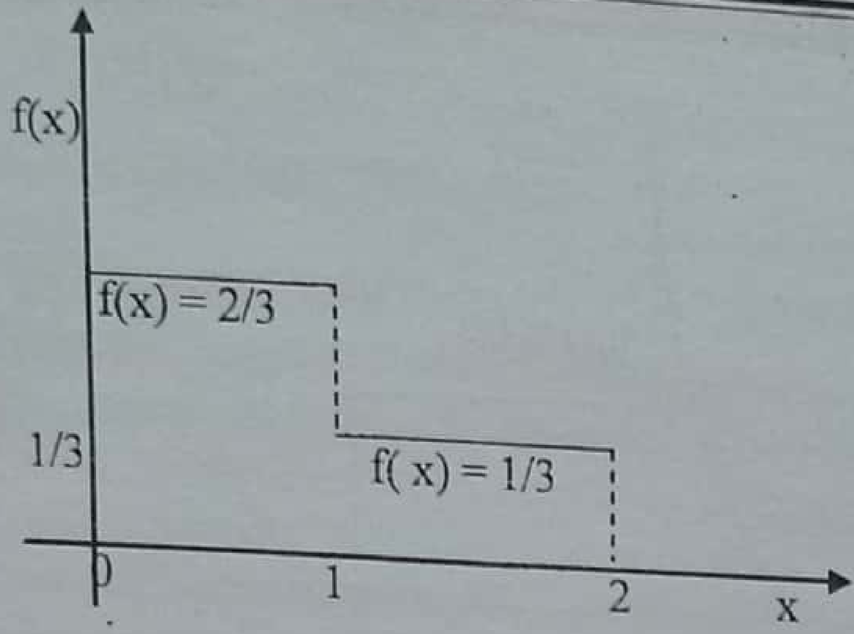
$$F(x) = \begin{cases} 0; & x \le 0 \\ k_1 x; & 0 \le x \le 1 \\ x/3 + k_2 & 1 \le x \le 2 \\ 1; & x \ge 2 \end{cases}$$

Find:

(i) the values of k_1 and k_2
(ii) p.d.f f(x) and sketch it
(iii) mean and variance of X
(iv) $P(X < 1.5 \mid X > 1)$

Answer (i) $k_1 = 2/3$, $k_2 = 1/3$

$$\text{(ii) } f(x) = \begin{cases} 2/3; & 0 \le x \le 1 \\ 1/3 & 1 \le x \le 2 \\ 0 & \text{otherwise} \end{cases}$$



Mean = 5/6 variance = 19/36 (iv) 0.4998

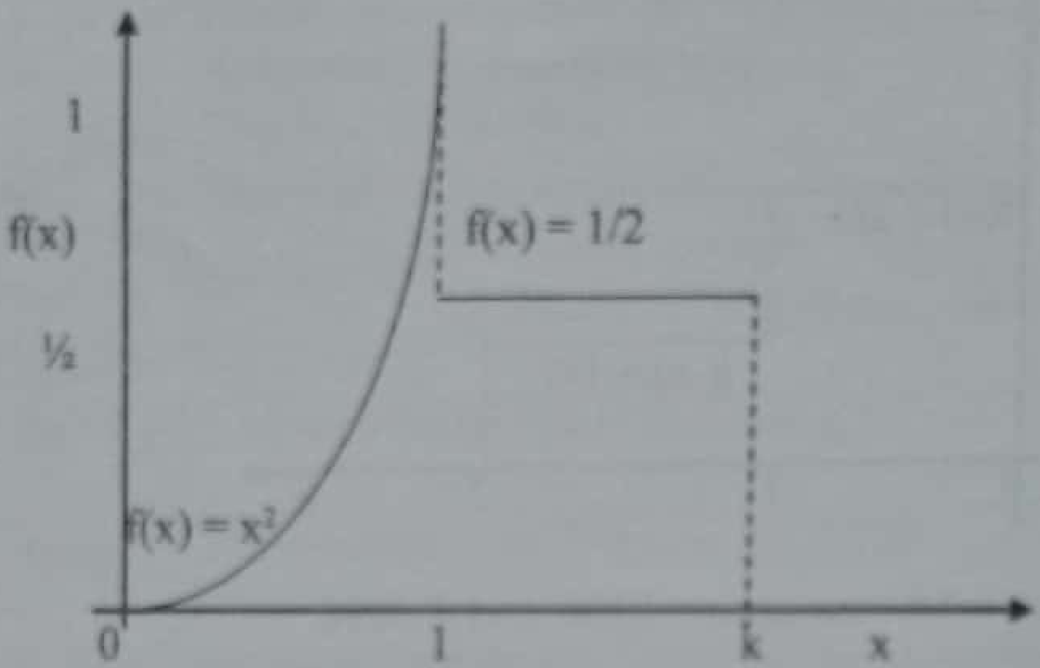
21. A probability density function of a random variable X is defined by

$$f(x) = \begin{cases} 0 & x < 0 \\ x^2; & 0 \leq x \leq 1 \\ \frac{1}{2} & 1 < x < k \\ 0 & x > k \end{cases}$$

Where k is constant

i. sketch the graph of f(x)
ii. find the value of k and hence the mean of the distribution.
iii. Calculate the median of the distribution

Answer k = 7/3, mean = 49/36, median = 4/3

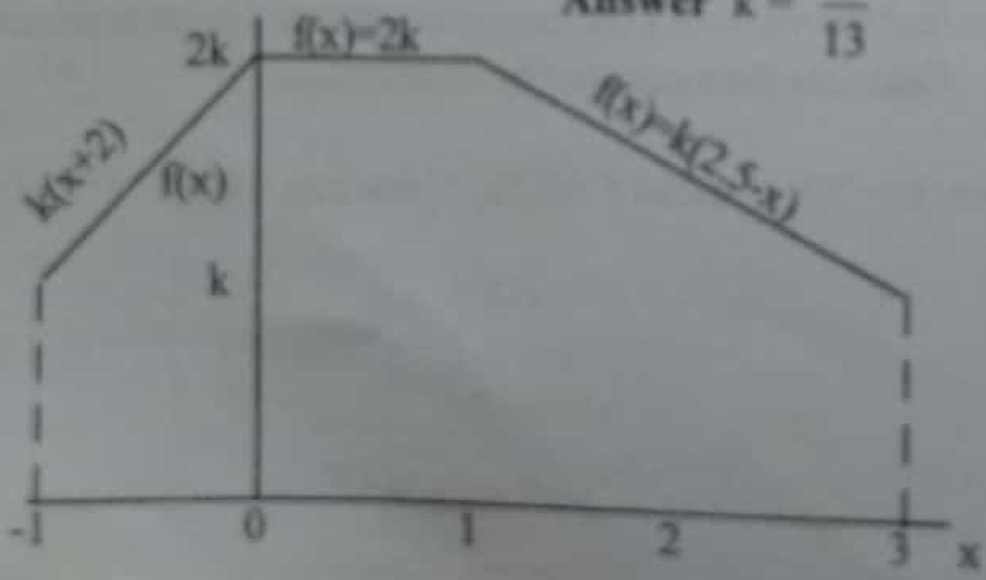


22. The probability density function of a random variable X is given by

$$f(x)=\begin{cases} k(x+2); & -1<x<0,\\ 2k & 0\le x\le 1,\\ k(2.5-x); & 1<x\le 3\\ 0 & elsewhere\end{cases}$$

Sketch the graph of f(x) and find the value of k

Answer $k = \dfrac{2}{13}$



CHAPTER SIX

THE NORMAL DISTRIBUTION

6.1 INTRODUCTION

The normal distribution is one of the most important distributions in statistics. Many quantities in natural science follow a normal distribution and under certain circumstances it is also a useful approximation to binomial distribution.

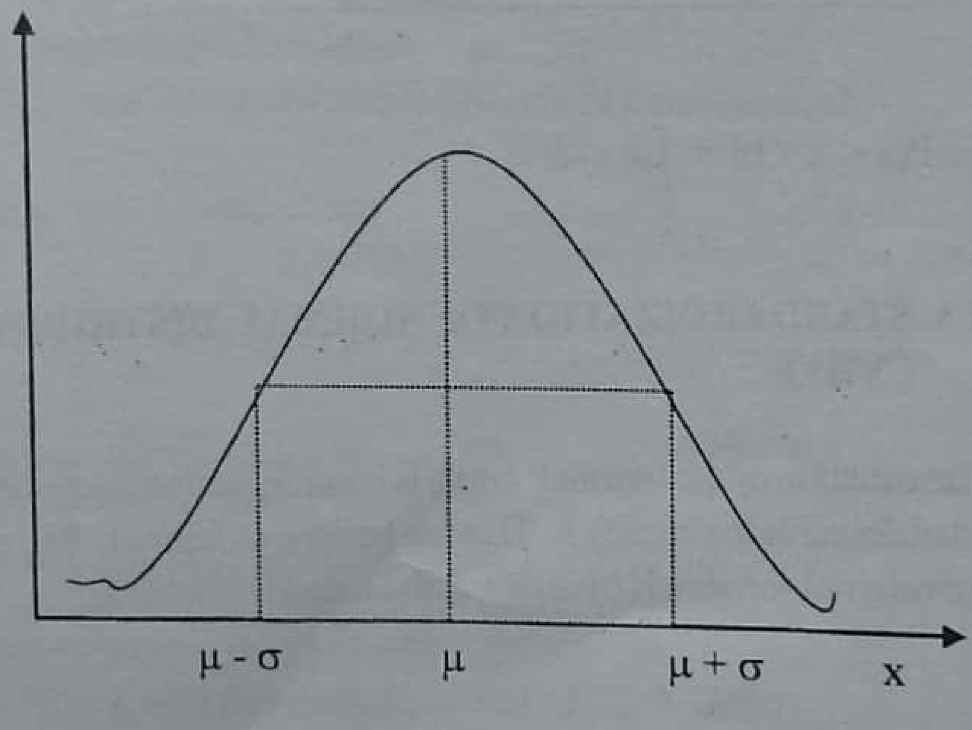
If X follows a normal distribution, then we can represent it by

$X \sim N(\mu, \delta^2)$

Where μ is the mean and δ^2 is the variance

The sketch of ϕ (x) is given below

$$\phi(x) = \frac{1}{\sqrt{2\pi}\,\delta} e - (x-\mu)^2 / 2\sigma^2 \quad -\infty < x < \infty$$



The curve for ϕ (x) known as normal curve

6.2 PROPERTIES OF THE NORMAL CURVE

(i) The curves is symmetrical about the mean μ.

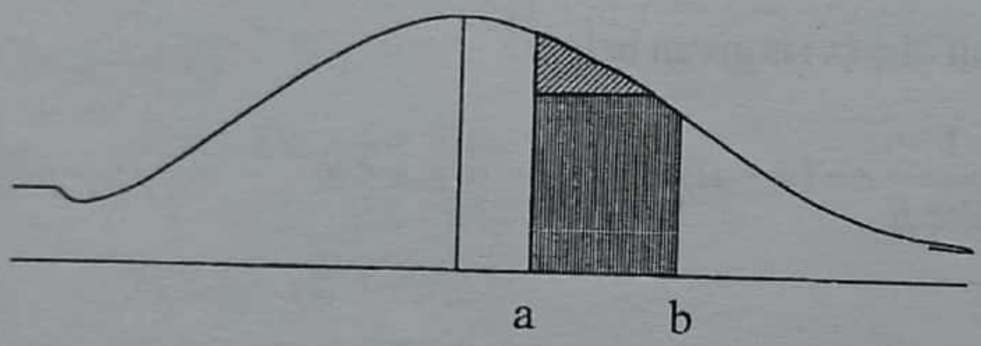
(ii) The curve never touches the x-axis i.e it is asymptotic to the x-axis.

(iii) The total area under the curve and above the x-axis is 1.

(iv) The mean, median and mode coincide at the maximum value of the function i.e the mode occurs at a point on the horizontal axis where $x = \mu$.

Area under the normal curve.

The area under the normal curve is used to obtain probabilities $P(a < x < b)$



$$\Rightarrow P(a < x < b) = \int_a^b \phi(x)dx = A$$

6.3 STANDARDIZATION OF NORMAL DISTRIBUTION CURVE

The equation of the 'normal' curve is complicated to integrate. It is often desirable to simplify it. This can be done as follows. We replace our original variable X by a new standardized variable Z,

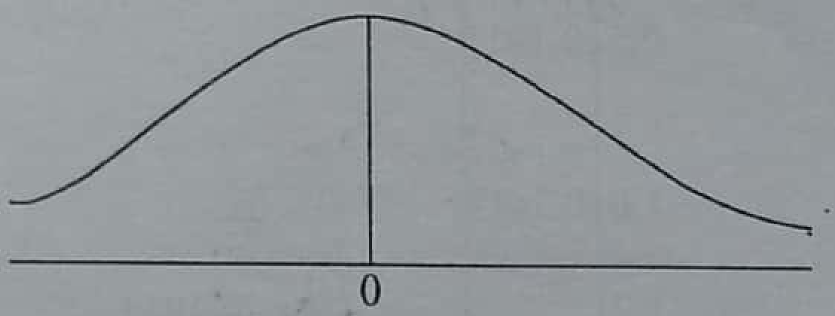
where $Z = \dfrac{x - \mu}{\sigma}$

The equation $\phi(x) = \frac{1}{\sigma\sqrt{2\pi}} e^{-(x-\mu)^2/2\sigma^2}$

Then reduces to $y = \frac{1}{\sqrt{2\pi}} e^{-z^2/2}$

This curve still has the same basic shape but is now symmetrical about the y-axis, so that its mean is 0 and its standard deviation is 1.

Standard normal curve



How to obtain probabilities.

The values of the normal distribution are first standardized

$$U\sin g\ \ Z = \frac{X-\mu}{\sigma}$$

$$\Rightarrow Z_1 = \frac{X_1-\mu}{\sigma}\ \ and\ \ Z_2 = \frac{X_2-\mu}{\delta}$$

$$P(X_1 < X < X_2) = P(Z_1 < Z < Z_2)$$

6.4 HOW TO READ TABLES

Probabilities are given in the table for all positive values of Z.
The $P(0 < Z < Z_1)$ can easily be obtained from the table.
Using the property of symmetry $P(Z < 0)$ can be obtained.

Example

(i) $P(0 < Z < 2) = 0.4772$ (Tab)

From table $P(0<Z<2)=0.4772$. The first column gives Z values to one decimal place. The other nine columns give Z to two decimal places. Therefore $Z=2.00$ implies read $Z=2.0$ and its probability is obtained by reading the second column with zero in the first row.

Z	0	1	2	3	4	5	6	7	8	9
0.0			0.0080							
0.5	0.1915									
1.0	0.3413									
1.2						0.3944				
1.8										
1.9										
2.0	0.4772									
2.1										
2.2						0.4878				

Probabilities of $P(0<Z<z_p)$ can easily be obtained from the table

$P(0 < Z < 1) = 0.3413$ (Tab)

$P(0 < Z < 2.25) = 0.4878$ (Tab)

$P(0 < Z\ 0.024) = 0.0080 + 0.0016 = 0.0096$

Note that the table above gives $Z = 0.02$.

$P(0<Z<0.02)=0.0080$. For Z values with more than two decimal places, the third decimal place is obtained from additional columns not shown.

(i) $P(0<z<1.252)=0.3944+0.0004=0.3948$

(ii) $(P(-0.5<Z<0)=P(0<Z<0.5)$ by symmetry.

$$P(0<Z<0.5)=0.1915$$
$$\Rightarrow p(-0.5<Z<0)=0.1915$$

(iii) $P(Z<-2)=P(Z>2)$ by symmetry.

$$P(Z>2)=0.5-P(0<Z<2)$$
$$P(Z>2)=0.5-0.4772$$
$$P(Z>2)=0.0228 \text{ (Tab)}$$
$$\therefore P(Z<-2)=0.0228$$

(iv) $P(0<Z<0.42)=0.1628$ (Tab)

(v) $P(-1.25<Z<0)=0.3944$ (Tab)

(vi) Note: $P(-3.5<Z<0)=0.5000$ (Tab)

(vii) $P(Z>1.36)=0.0869$ (Tab)

(viii) $P(Z>0.25)=0.4013$ (Tab)

(ix) $P(Z<-0.25)=0.4013$ (Tab)

(x) $P(Z<1)=0.5+0.3413=0.8413$

(xi) $P(Z>-2)=0.5+0.4772=0.9772$

(xii) $P(-1<Z<-2)=0.4772-0.3413$ (Tab)
$=0.1359$

(xiii) $P(-1<Z<2)=0.3413+0.4772$ (Tab)
$=0.8185$

(xiv) $P(0.5<Z<2.5)=0.4938-0.1915$ (Tab)
$=0.3023$

(xv) $P(|Z|<1.78)=P(-1.78<Z<1.78)$
$=2 \times 0.4625$ (Tab)
$=0.9250$

(xvi) $P(|Z|>2.326)=1-P(-2.326<Z<2.326)$
$=1-2 \times 0.4900$ (Tab)
$=0.0200$

Note. All above probabilities can be verified using a calculator. PRESS MODE, MODE, 1 SHIFT, 3 then area needed. Either P Or , Q Or R.

Example

The heights of students at a particular University follow a normal distribution with mean 150.3 cm and standard deviation 5 cm. Find the probability that a student picked at random from the university has height

a) Less than 153 cm,
b) More than 158 cm
c) Between 150 cm and 158 cm

solution

$X \sim N(150.3, 25)$

$P(X < 153) = P\left(Z < \frac{153 - 150.3}{5}\right)$

$P(Z < 0.54) = 0.5 + 0.2054(Tab)$

$\therefore P(X < 153) = 0.7054$

(b) $P(X > 158) = P\left(Z > \frac{158 - 150.3}{5}\right)$

$P(Z > 1.54) = 0.5 - 0.4382(Tab)$

$\therefore P(X > 158) = 0.0618$

(c) $P(150 < X < 158) =$

$P\left(\frac{150 - 150.3}{5} < Z < \frac{158 - 150.3}{5}\right)$

$P(-0.06 < Z < 1.54) = 0.4382 + 0.0239(Tab)$

$\therefore P(150 < X < 158) = 0.4621$

Example

A random variable X is normally distributed with mean 50 and standard deviation 10. Find the probability that X assumes a value between 45 and 62.

Solution: $\sigma = 10, \mu = 50, \quad P(45 < X < 62)$

$$Z = \frac{X-\mu}{\sigma}$$

$\Rightarrow P(45<X<62)=$

$$P\left(\frac{45-50}{10} < Z < \frac{62-50}{10}\right)$$

$$P(-0.5 < Z < 1.2) = 0.1915 + 0.3849(Tab)$$

$$P(45 < X < 62) = 0.5764$$

Example

In an external examination, the scores are normally distributed with mean 120 and standard deviation 15. If a score of 100 is required to pass the examination, what is the probability that a candidate chosen at random has failed the examination?

Solution:

$\mu = 120, \sigma = 15$, from $Z = \frac{x-\mu}{\sigma}$

$$Z = \frac{100-120}{15} = -1.333$$

P(He scores less than 100) = $P(Z < -1.333)$

$P(Z < -1.333) = (Z > 1.333)$ by symmetry

$P(Z > 1.333) = 0.5 - P(0 < Z < 1.333)$

$= 0.5 - 0.4087$ (Tab)

$P(Z < -1333) = 0.0913.$

P(A candidate failed) = 0.0913.

Example

A certain firm sells flour bags of mean weight 40 kg and standard deviation 2kg. Given that the weight is normally distributed find:

(i) the probability that the weight of any bag taken at random will lie between 41.0 and 42.5kg.

(ii) the percentage of bags whose weight exceeds 43kg.

(iii) the number of bags rejected out of a 500 bag purchase by a retailer whose consumers cannot accept a bag whose weight is below 38.5kg.

Solution:

(i) $\mu = 40, \sigma = 2$

$P(41.0 < X < 42.5) = P(Z_1 < Z < Z_2)$.

$$\Rightarrow Z_1 = \frac{41-40}{2} = 0.5 \quad \text{and}$$

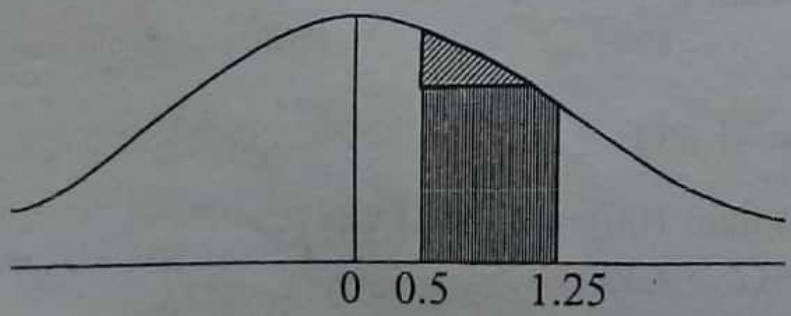
$$Z_2 = \frac{42.5-40}{2} = 1.25$$

$$\Rightarrow \quad P(41.0 < X < 42.5) = P(0.5 < Z < 1.25)$$

$$P(0.5 < Z < 1.25) = P(0 < Z < 1.25) - P(0 < Z < 0.5)$$

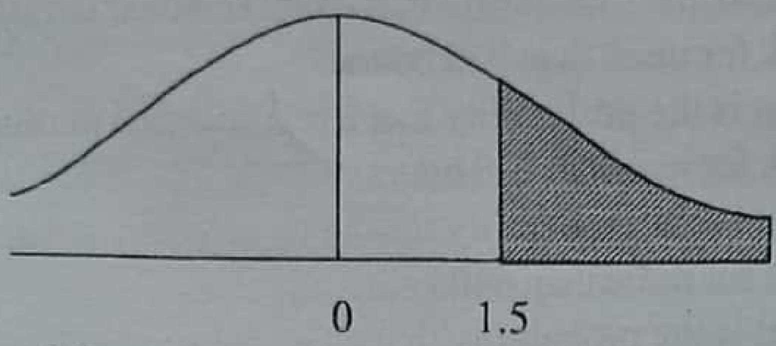
$$= 0.3944 - 0.1915 \text{ (Tab)} = 0.2029$$

$$P(41.0 < X < 42.5) = 0.2029$$



(ii) $P(X > 43) = P(Z > \frac{43-40}{2} = 1.5)$

$P(X > 43) = P(Z > 1.5)$



$P(Z > 1.5) = 0.5 - P(0 < Z < 1.5)$

$0.5 - 0.4332 = 0.0668$

Percentage $= 0.0668 \times 100 = 6.68\%$

(iii) Total number of bags is 500

$P(X < 38.5) = P(Z <) \dfrac{38.5 - 40}{2}$

$Z_1 = -0.75$

$P(Z < -0.75) = P(Z > 0.75)$ by symmetry.

$P(Z > 0.75) = 0.5 - P(0 < Z < 0.75)$

$= 0.5 - 0.2734(\text{Tab}) = 02266$

$P(Z < -0.75) = 0.2266$

Bags rejected are $0.2266 \times 500 = 113.3$ bags.

Assignement

1.Cartons of milk from a particular supermarket are advertised as containing 1 litre, but in fact the volume of the contents is normally distributed with a mean of 1012 ml and a standard deviaion of 5 ml.

a) Find the probability that a randomly chosen carton contains more than 1010 mls
b) In a batch of 1000 cartons, estimate the number of cartons that contain less than the advertised volume of milk.

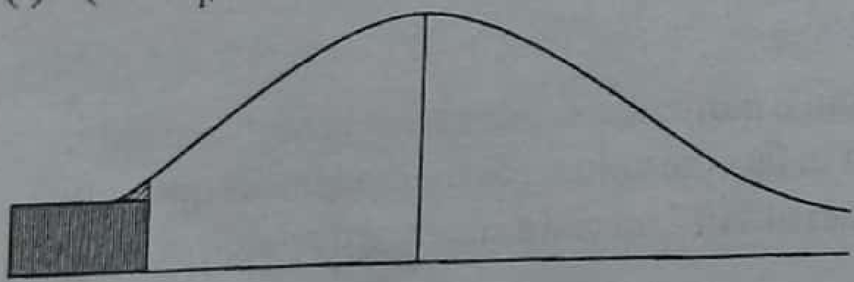
2. The life time of pens produced by a certain factory is normally distributed with mean of 800 hours and standard deviation 80 hours.

i. What is the probability that a pen selected at random will work for more than 920 hours.
ii. What is the probability that a pen selected at random will work for less than 720 hours.
iii. What is the probability that a pen selected at random will work for more than 640 hours.
iv. What is the probability that a pen selected at random will work for less than 880 hours.
v. What is the probability that a pen selected at random will work for between 840 hours and 920 hours.

6.5 HOW TO OBTAIN Z-VALUES FROM A GIVEN PROBABILITY

Sometimes, you are given information leading to the probability of a certain value of Z. This probability can be used to determine the value of Z using the tables. e.g

(i) $P(Z < Z_1) = 0.0968$. Find Z_1



$P(Z < Z_1) < 0.5$

Means Z_1 is less than mean $\mu = 0$.

$P(Z < Z_1) = P(Z > -Z_1) = 0.0968$

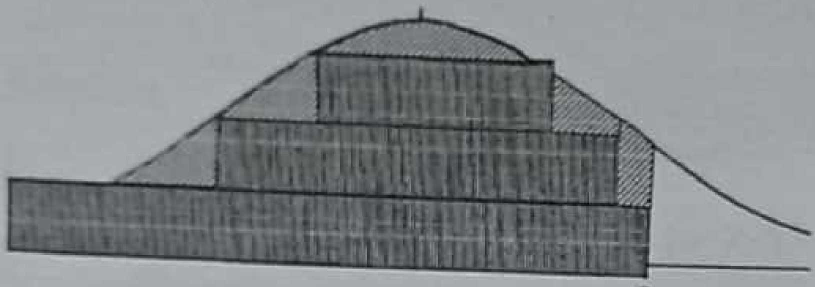
$P(Z > -Z_1) = 0.0968$.

$P(0 < Z < -Z_1) = 0.5 - 0.0968 = 0.4032$

$P(0 < Z < -Z_1) = -1.3$, (Tab)

$\therefore Z_1 = -1.3$

(ii) $P(Z<a)=0.787$



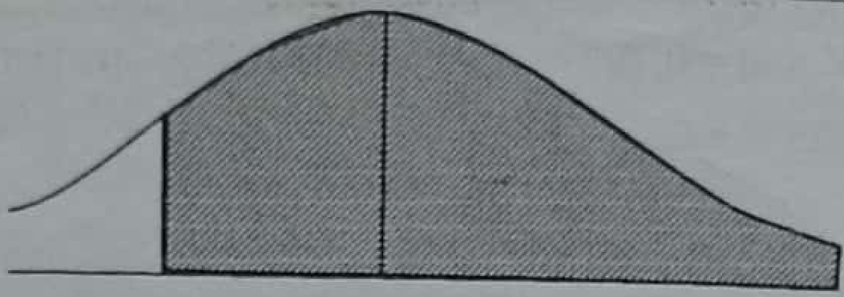
$P(Z<a)=P(Z<0)+P(0<Z<a)$
$P(Z<0)=0.5$
$P(0<Z<a)=0.787-0.5$
$P(0<Z<a)=0.287$
From the table $a=0.796$

Note: From the table the value indicated is 0.2852; the difference is obtained from the last part where 6 corresponds to 18, to make the value 0.2870.

(iii) $P(Z_1>b)=0.0100$
$P(0<Z<b)=0.5-0.0100$
$P(0<Z<b)=0.49$
$b=2.326$

Note: the value of b can be obtained by using the critical points of the normal distribution. 0.01 corresponds to Q and this gives the value of b as 2.326 (Tab)

(iv) $P(Z>a)=0.812.$
$P(Z>a)=P(a<Z<0)+P(Z>0)$
$P(a<Z<0)=0.812-0.5$
$P(a<Z<0)=0.312$



$P(a < Z < 0) = P(0 < Z < -a) = 0.312$

$-a = 0.885$

$a = -0885$

(v) $P(Z > b) = 0.030$

P	*Q*	*z*
0.470	0.030	1.881

Therefore z = 1.881.

(vi) $P(Z < b) = 0.040$

P	*Q*	*z*
0.460	0.040	1.751

Therefore z = -1.751. Since z is on the left of zero.

(vii) $P(Z < b) = 0.650$

P	*Q*	*z*
0.150	0.350	0.385

Therefore z = 0.385.

Example

The period of a certain machine approximately follows a normal distribution with mean of 5 years and standard deviation of 1 year. Given that the manufacturer of this machine replaces the machines that fail under guarantee, determine the length of the guarantee required so that not more than 2% of the machines that fail are replaced.

Determine the proportion of the machines that would be replaced if the guarantee period was 4 years.

Solution:

$\mu = 5, \sigma = 1$

let x_0 be guarantee period

$X \sim N(\mu, \sigma^2)$

$P(X < x_0) = 0.02$

$-Z_0$ corresponds to 2.054 (Tab) using the critical points

$Z_0 = -2.054$

but $Z = \dfrac{X-\mu}{\sigma} \Rightarrow X = Z\sigma + \mu$

$Xo = -2.054 \times 1 + 5$

$Xo = 2.946$ years,

the guarantee period is 2.946 years.

When $X_0 = 4$, $P(X < Xo) = P(X < 4)$

$Z = \dfrac{X-\mu}{\sigma} = \dfrac{4-5}{1} = -1$

$P(X < 4) = P(Z < -1)$

$P(Z < -1) = P(Z > 1)$ by symmetry

but $P(0 < Z < 1) = 0.5 - P(Z > 1)$

$P(Z > 1) = 0.5 - P(0 < Z < 1)$

$P(Z - 1) = 0.5 - 0.3413$

$P(Z > 1) = P(Z < -1) = 0.1587$

$P(X < 4) = 0.1587$

The proportion is $0.1587 \times 100\%$

15.87%

Example

A total population of 700 students sat an examination for which the pass mark was 50. Their marks were normally distributed. 28 students scored below 40 marks while 35 scored above 60.

(a) Find the mean mark and standard deviation of the students.

(b) What is the probability that a student chosen at random passed the examination?

(c) Suppose the pass mark is lowered by 2 marks, how many more students will pass?

Solution:

(a) Let X be the marks obtained by a student.

$X \sim N(\mu, \sigma^2)$.

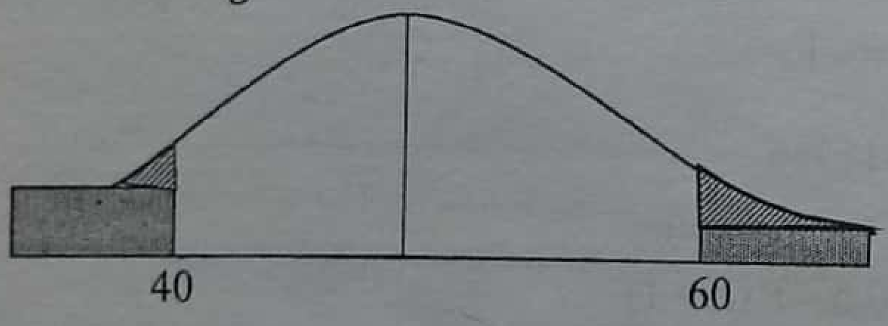
Let the mean be μ and standard deviation σ.

$P(X < 40) = {}^{28}/_{700} = 0.04$

$P(X < 40) = 0.04$

$P(X > 60) = {}^{35}/_{700} = 0.05.$

Note $Z = \dfrac{X - \mu}{\sigma}$



$P(X < 40) = P(Z < Z_1) = 0.04$

$P(Z < Z_1) = 0.04$

$-Z_1 = 1.751$ from the table of critical values

$\therefore Z1 = -1.751$

$\Rightarrow -1.751 = \dfrac{40 - \mu}{\sigma}$

$\Rightarrow \mu - 1.75\sigma = 40 \ldots(1)$

$P(X > 60) = 0.05$

$P(X < 60) = P(Z > Z_2) = 0.05$

$Z_2 = 1.645$ from table of critical values

$$\Rightarrow 1.645 = \frac{60 - \mu}{\sigma}$$

$\Rightarrow \mu - 1.645\sigma = 60 \ldots(2)$

$\Rightarrow \mu - 1.75\sigma = 40 \ldots\ldots(1)$

Equation $(2)-(1)$

$1.645\sigma + 1.751\sigma = 20 \ldots(1)$

$3.396\sigma = 20$

$\sigma = 5.889$

From (1) $\mu = 1.751\sigma + 40$

$\mu = 10.3 + 40$

$\mu = 50.3$

Standard deviation $\sigma = 5.889$, $\mu = 50.3$

(b) The pass mark is 50.

$P(X < 50)$

$$Z_1 = \frac{X - \mu}{\sigma} = \frac{50 - 50.3}{5.889}$$

$Z_1 = -0.051$

$P(Z > -0.051) = 0.5 + P(-0.0051 < Z < 0)$

$P(-0.051 < Z < 0) = P(0 < Z < 0.051)$ By symmetry

$P(0 < Z\ 0.051) = 0.203$ from the table

$P(0 < Z < 0.051) = 0.5203$

$P(Z > -0051) = 0.5203$

(c) The pass mark

$P(X > 48) = P(Z > Z_1)$

$$Z_1 = \frac{48 - 50.3}{5.889}$$

$Z_1 = -0.391$

$P(Z > -0.391) = P(-0.391 < Z < 0) + P(Z > 0)$

$P(Z > 0) = 0.5$

$P(-0.391 < Z < 0) = P(0 < Z\ 0.391)$ By symmetry

$P(0 < Z < 0.391) = 0.1521$

$P(Z > -0.391) = 0.5 + 0.1521$

$P(X > 48) = 0.6521$

The probability of more students is

$0.6521 - 0.5203 = 0.1318$

number of students

0.1318×700

$= 92$ students

6.6 DISTRIBUTION OF SAMPLE MEAN FROM A NORMAL POPULATION

If $x_1, x_2 \ldots x_n$ is a random sample of size n taken from a normal distribution with mean μ and variance σ^2 such that $X \sim N(\mu, \sigma^2)$ then the distribution of $\bar{x}$ is also normally distributed with mean μ and variance σ^2/n.

Example

A random sample of size 16 is taken from a normal population with mean 60 and standard deviation 4. Find the probability that the mean of the sample is less than 58.

Solution

$$P(\overline{X} < 58) = P(Z < \frac{58-60}{4/4})$$

$$P(Z < -2) = 0.5 - 0.4772$$

$$P(\overline{X} < 58) = 0.0228$$

Example

The height of students are normally distributed with mean 164cm and standard deviation 7.2cm. Calculate the probability that the mean height of a sample of 36 students will be between 162 and 166cm

Solution

$$P(162 < x < 166) = P\left(\frac{162-164}{(7.2/6)} < Z < \frac{166-164}{(7.2/6)}\right)$$

$$= P(-1.667 < Z < 1.667)$$

$$= 2P(0 < Z < 1.667)$$

$$= 2 \times 0.4522$$

$$P(162 < x < 166) = 0.9044$$

Assignment

The days production of milk sold in packets of certain firm is normally distributed with mean of 0.5 litres and standard deviation of 0.08 litres. If a sample of 16 packets is drawn from days production, find the probability that the mean is between 0.46 and 0.47 litres.

Exercise

1.The length of a type A rod is normally distributed with mean of 15cm and a standard deviation of 0.1cm. The length of another type B rod is also normally distributed with a mean of 20cm and standard deviation 0.16cm. For a type A rod to be acceptable, its length must be between 14.8cm and 15.2cm and for a type B rod the length must be between 19.8cm and 20.2cm.

i. What proportion of type A rod is of acceptable length?
ii. What is the probability that one of them is of acceptable length?

Answer

i) Acceptable length for type A is 95.44%.
Acceptable length for type B is 78.88% (ii) 0.7528 (iii) 0.2375

2.The lifetime of a bulb is normally distributed with a mean of 800 hours and standard deviation of 80 hours. The manufacturer guarantees to replace bulbs, which blow before 660 hours.

i. What percentage of the bulbs will he have to replace under the guarantee?
ii. The manufacturer is only willing to replace a maximum of 1% of the bulbs. What should be the guaranteed lifetime of the bulbs?
iii. Instead of reducing the guaranteed life time as in (ii). the mean lifetime was increased by superior technology. What should be the new mean so that only 1% are replaced if the guaranteed life time remains at 660 hours but the standard deviation is reduced to 70 hours.

Answer: i) 4.01% ii) 613.92 hours
iii) $\mu = 822.82$ hours

3.In an examination 30% of the candidates fail.and 10% achieve distinctions. Last year the pass mark (out of 200) was 84 and the minimum mark required for a distinction was 154. Assuming that the marks of the candidates were normally distributed, estimate the mean mark and standard deviation.

Answer: $\mu = 104.31$ $\sigma = 38.76$

4. The marks of 500 candidates in an examination are normally distributed with a mean of 45 marks and standard deviation of 20 marks.

a) Given that the pass mark is 41, estimate the number of candidates who passed the examination.
b) If 5% of the candidates obtained a distinction by scoring x marks, or more, estimate the value of x.
c) Estimate the interquartile range of the distribution.

Answer a) 290 b) 78 c) 27

5. Given a normal distribution with mean 40 and standard deviation 6, find

a) The area below $x = 32$.
b) The area above $x = 34$
c) The area between $x = 42$ and $x = 51$
d) The $x -$value that has 45% of the area below it
e) The x-value that has 13% of the area above it

Answer: a) 0.0913 b) 0.8413, c) 0.3362 d) 39.244 e) 46.756

6.The marks obtained by 1000 candidates in an examination were normally distributed with a mean of 55 and standard deviation 8,

a) If a mark of 71 or more is required for A – pass, estimate the number of A passes awarded.
b) If 15% of the candidates failed the examination, estimate the minimum mark required for a pass.

c) Calculate the probability that two candidates chosen at random both passed the examination.

Answer: a) 23, b) 47 c) 0.7225

7. A certain type of electricity light bulb has a burning life of H hours, where it has a normal distribution with mean 1300 hours and standard deviation 125 hours.

a) What is the probability that a bulb selected at random will burn for more than 1500 hours?
b) If a manufacturer guarantees to replace any bulb which burns for less than 1050 hours, what percentage of the bulbs will have to be replaced?
c) If two bulbs are installed at the same time, what is the probability that both will burn less than 1400 hours but more than 1200 hours?

Answer: a) 0.0548 b) 2.28 c) 0.3320

8. The heights of boys in a certain school are normally distributed. 10% are over 1.8 meters and 20% are below 1.6 meters. Determine the mean height μ meters and standard deviation σ meters. Hence, find the interquartile range

Answer: $\mu = 1.68$ $\sigma = 0.09, 0.13$

9. Tests made on two types of electric bulbs show the following. Type A, lifetime was distributed normally with an average of 1150 hours and standard deviation of 30 hours. Type B, long life bulbs, average lifetime of 1900 hours, with standard deviation of 50 hours.

a) What percentage of bulbs of type A could be expected to have a life time more than 1200 hours.
b) What percentage of type B would you expect to last longer than 1800 hours?
c) What lifetime limit would you estimate would contain the central 80% of production of type A.?

Answer: a) 4.78% b) 97.72%
c) 1111.54 hours, 1188.46 hours

10. A factory produces two types of bars of soap A and B. Their lengths are normally distributed with type A having mean 115cm and standard deviation 3 cm. Type B having mean 190 cm and standard deviation 5 cm

Determine the percentage of type.

(i) A bars that have a length of more than 120cm

(ii) B bars that have a length of more than 180 cm

Answer (i) 4.78% (ii) 97.72%

11. The weight of the contents of a "2Kg" packet of frozen food is normally distributed with a mean of 2.05 kg and a standard deviation of 0.015kg. Estimate 99% confidence limits for the weight of contents in a single 2kg packet.

Answer 2.011 kg, 2.089kg

12. The r. v X is such that $X \sim N(\mu, 4)$. A random sample size n, is taken from the population. Find the least value of n such that

$$P(|X - \mu| < 0.5) > 0.95.$$

Answer 62

13. Boxes made in a factory have weights which are normally distributed with a mean of 4.5 kg and a standard deviation of 2.0 kg. Find the probability of there being a box with a weight of more than 5.4 kg when a box is chosen at random. If a sample of 16 boxes is drawn, find the probability that the mean is between

(i) 4.6 and 4.7 kg

(ii) 4.3 and 4.7 kg

Answer: 0.3264 (i) 0.0761 (ii) 0.3108

14. Biscuits are produced with Weight (Wg) where W is N(10, 4) and are packed at random into boxes consisting of 25 biscuits. Find the probability that;

(a) a biscuit chosen at random weighs between 9.25g and 10.75g
(b) the contents of a box weighs between 245g and 255g.
(c) the average weight of the biscuits in the box lies between 9.7g and 10.3g.

Answer (a) 0.2924 (b) 0.0796 (c) 0.5468

15. Observation of a very large number of cars at a certain point on a motorway establishes that the speeds are normally distributed. 90% of cars have speeds less than 77.7 km/h and only 5% of cars have speeds less than 63.1 km/h.
Determine the mean speed μ and standard deviation σ

Answer $\mu = 71.305$, $\sigma = 4.988$.

16. The marks in an examination were normally distributed with mean μ and standard deviation σ. 10% of the candidates scored more than 75 marks and 20% scored less than 40 marks.
   i. 25 candidates were chosen at random from those who sat for the examination.
   Find the probability that their average mark exceeds 60.
   ii. If a sample of 8 candidates were chosen, find the probability that not more than 3 scored between 45 and 65 marks

Answer $\mu = 53.87$, $\sigma = 16.473$,
(i)0.0313 (ii) 0.5419

17. The life time of batteries produced by a certain factory is normally distributed. Out of 10,000 batteries selected at random, 668 have life time less than 130 hours and 228 have life time more than 200 hours.
   i. find the mean and standard deviation of the battery life time.

ii. find the percentage of the batteries with life time between 150 and 180 hours.
iii. If the sample of 25 batteries is selected at random, find the the probability that the mean of the life time exceeds 165 hours.

Answer (i) $\mu = 160,\ \sigma = 20$
(ii) 53.28% (iii) 0.1056

18. The volume of a soft drink bottled by a certain company is approximately normally distributed with mean 300 mls and standard deviation 2 mls. Determine the probability that in a sample of 10 bottles at least two contain less than 297.4 mls.

Answer 0.2515

6.7 THE NORMAL APPROXIMATION TO THE BINOMIAL DISTRIBUTION

Under certain circumstances, the normal distribution can be used as an approximation to the binomial distribution. One practical advantage is that calculations are much less tedious to perform.

Condition

For a large value of $n > 20$ and the probability of success not too small or too large, i.e p is very close to 0.5

If $X \sim B(n, p)$, then

$E(X) = np$

$Var(X) = npq$ where $q = 1 - p$.

Then $X \sim N(np, npq)$. Approximately

The value of Z can be obtained from

$$\therefore Z = \frac{X \pm 0.5 - np}{\sqrt{npq}}$$

i.e substituting for $\mu = E(X) = np$ and

$\sigma = \sqrt{npq}$ in $\frac{x-\mu}{\sigma}$

The $\pm$ 0.5 is used to make the binomial distribution continuous.

Note: $P(X \leq x_1) = P(Z \leq Z_1)$ and corresponding value of Z_1 is

$$Z_1 = \frac{X + 0.5 - np}{\sqrt{npq}}$$

$$P(X \geq X_2) = P(Z \geq Z_2) \Rightarrow Z_2 = \frac{X - 0.5 - np}{\sqrt{npq}}$$

Example

If an unbiased coin is tossed 100 times, what is the probability that

(a) There will be more than 60 heads?

(b) There will be at least 45 and at most 55 heads?

(c) There will be fewer than 43 heads?

Solution

a) $n = 100, p = ½$

$X \sim B(n, p)$

$X \sim N(np, npq)$ since n is large

$\mu = np = 100 \times 0.5 = 50$

$Var(X) = npq = 50 \times ½ = 25$

$P(X > 60) = P(X \geq 61)$

$P(X \geq 61) = P(Z > Z_1) =$

$$\Rightarrow Z_1 = \frac{61 - 0.5 - 50}{\sqrt{25}} = \frac{60.5 - 50}{5}$$

$Z1 = 2.1$

$P(Z > 2.1) = 0.5 - P(0 < Z < 2.1)$

$= 0.5 - 0.4821$

$P(Z > 2.1) = 0.0179$

$P(X > 60) = 0.0179$

b) $P(45 \le X \le 55)$

$\mu = np, \sigma = \sqrt{npq} = 5$

$P(45 \le X \le 55) = P(44.5 \le X \le 55.5)$

$$\Rightarrow P\left(\frac{44.5-50}{5} < Z < \frac{55.5-50}{5}\right)$$

$P(-1.1 < Z < 1.1) =$

$2\ x P(0 < Z < 1.1) = 2x0.3643(Tab)$

$P(45 \le X \le 55) = 0.7286$

P(at least 45 and at most 55 heads)

$= 0.7286$

c) $P(X < 43)$

$P(X \le 42) = P(Z \le Z_1)$

$$P(Z < \frac{42.5-50}{5}) =$$

$P(Z < -1.5) = 0.0668(Tab)$

P(will be fewer than 43 heads) $= 0.0668$

Example

Among the spectators watching a football match 80% were the home team's supporters while 20% were the visiting team's supporters. If 2,500 of the spectators are selected randomly, what is the probability that there are more than 540 visitors in the sample?

Solution

$n = 2500,\ p = 0.2$ and $q = 0.8$

n is large, hence $X \sim N(np, npq)$

$\mu = np = 2500 \times 0.2 = 500$

$\delta = \sqrt{2500x0.8x0.2} = 20$

$P(X > 540) = P(X \geq 541) =$

$$P(Z > \frac{540.5 - 500}{20})$$

$P(Z > 2.025) = 0.5 - 0.4785(Tab)$

$P(X > 540) = 0.0215$

Example

It is known that 72% of UBC viewers watch a particular rogramme known as Cuando Seas Mia. What is the probability that in a sample of 500 viewers chosen at random

(a) More than 350 watch the programme.

(b) Fewer than 340 watch the programme.

(c) Exactly 350 watch the program

Solution

$n = 500, p = 0.72$

$\mu = np = 500 \times 0.72 = 360$

$\sigma^2 = npq = 360 \times 0.28 = 100.8$

$\sigma = 10.04$

$P(X > 350) = P(x \geq 351)$

$P(Z > \frac{350.5-360}{10.04}) =$

$P(Z > -0.946) = 0.5 + 0.3280(Tab)$

$P(X > 350) = 0.8280$

b) $P(X < 340) = P(X \leq 339)$

$P(Z < \frac{339.5-360}{10.04})$

$P(Z < -2.042) = 0.5 - 0.4794(Tab)$

$P(X < 340) = 0.0206$

(c)

$$P(X = 350) = P(349.5 < X < 350.5)$$

$$= P(\frac{349.5-360}{10.04} < Z < \frac{350.5-360}{10.04})$$

$$= P(-1.046 < Z < -0.946)$$

$$= 0.3522 - 0.3280\,(Tab)$$

$$= 0.0242$$

Assignment

1. Twenty percent of the eggs supplied by a poultry farm have racks on them. Determine the probability that a sample of 900 eggs supplied by the farm will have more than 200 eggs with cracks.

2. A question paper contained 100 objective questions with 4 alternative answers but only one correct answer. A candidate who sat for the paper attempted 80 questions purely by guessing and left out the remaining 20. Determine the probability that;

(i) he got 35%

(ii) he scored between 20% and 30%

Exercise

1. Statistics records from Uganda Police traffic department shows that on weekend nights, one out of every ten drivers on the road is drunk. A random sample of four hundred drivers is checked on a weekend.
Find the probability that the number of drunk drivers is at least 35 but less than 47.

Answer: 0.6807

2. A pair of dice is tossed 144 times and the sum of the out comes recorded. Find the probability that a sum of 7 occurs at least 26 times.

Answer: 0.3688

3. In an examination which consists of 100 questions, a student has a probability of 0.6 of getting each question correct. The student fails the examination if he obtains a mark less than 55, and obtains a distinction for a mark of 68, or more. Calculate:

(a) the probability that he fails the examination.

(b) The probability that he obtains a distinction.

Answer: a) 0.1308 b) 0.0629

4. If a fair die is thrown 300 times, what is the probability that:
(a) there will be more than 60 sixes
(b) There will be fewer than 45 sixes.

Answer: a) 0.0519 b) 0.1971

5. A coin is biased such that head is twice as likely to occur as a tail. The coin is tossed 120 times. Find the probability that there will be:
a) between 42 and 51 tails inclusive
b) 48 tails or less
c) less than 34 tails
d) between 72 and 90 heads inclusive.

Answer: a) 0.3729 b) 0.9501
c) 0.1039 d) 0.9290

6. Four hundred students sat a test which consists of 80 true / false questions. None of the candidates knows any of the answers and so guesses.

(a) If the pass mark is 38, how many of the candidates would be expected to pass?

(b) What should the new pass mark, if it is decided that only 115 candidates pass?

Answer: a) 285 b) 43

7. A lorry of potatoes has on average one rotten potato in 6. A green grocer tests a random sample of 100 potatoes and decides to turn away the lorry if he finds more than 18 rotten potatoes in the sample. Find the probability that he accepts the consignment.

Answer: 0.6886

CHAPTER SEVEN

ESTIMATION

7.1 INTRODUCTION

Statistical estimation is a statistical procedure used to describe the unknown characteristics of the population by using sample characteristics.A sample is a representation of a population. Parameters are population constants, e.g population mcan μ, population variance σ^2

There are two types of estimation

a) Point estimation
b) Interval estimation.

7.2 Point Estimation

Point estimation means single value estimation of a parameter. The value got is the *estimate* while the method used is called the *estimator*. An estimator is a statistical technique used to obtain a sample characteristic. It is a function of the sample observations,

Example If x = 3, 4, 6, 7

$$\bar{x} = \frac{\sum x}{n} = \frac{3+4+6+7}{4} = 5$$

$\bar{x}$

$\bar{x} = 5$ which is an estimate.

and the formula $\frac{\sum x}{n}$ *is* the estimator

7.3 Interval Estimation

Interval estimation deals with a range between which the parameter lies e.g $5 < \mu < 10$.

7.4 Confidence Interval

If a and b are limits of a parameter, i.e $a \leq \mu \leq b$, a and b are called confidence limits, a being the lower confidence limit while b is the upper confidence limit.

7.5 Confidence Coefficient or Degree of Confidence

The probability that the true population parameter will lie within the stated interval is denoted by $(1 - \alpha)$ and this is referred to as confidence coefficient.

7.6 Estimation of the Mean

Consider the population to be normally distributed. The point estimator for μ is $\bar{x}$ (the sample mean)

Note: The mean of a random sample drawn from a population which is normally distributed with mean μ and variance σ^2 is also normally distributed with mean μ and variance σ^2/n

$$\text{Var}(\bar{x}) = \frac{\sigma^2}{n} \quad \text{Remember } Z = \frac{x-\mu}{\sigma}. \text{Implies } Z = \frac{\bar{x} - E(\bar{x})}{\sqrt{\text{var}(\bar{x})}}$$

Note

If $\bar{x}$ is the mean of a random sample of size n from a normal population with unknown mean μ and known variance σ^2, then

$(1 - \alpha)\%$ confidence limit for μ is

Note: For a population the (confidence limit for

μ is $\quad \bar{x} \pm Z_{\alpha/2}\, \delta$

Example

A factory produces two types of bars of soap A and B. Their lengths are normally distributed with type A having mean 115cm and standard deviation 3 cm. Find the 95% confidence limits for the mean of lengths of type A bars of soap.

Solution

$95\% C.I \Rightarrow Z_{0.025} = 1.96 (Tab)$

$Upper\ \lim it = 115 + 1.96 x 3 = 120.88$

$Lower\ \lim it = 115 - 1.96 x 3 = 109.12$

Example

It is known that an examination paper is marked in such a way that the standard deviation of the marks is 15.1. In a certain school, 80 candidates take the examination and they have an average mark of 57.4. Find

(a) 95% and

(b) 99% confidence limits for the mean mark in the examination.

Solution

$\sigma = 15.1n, n = 80,$

(1 - α)% confidence interval is given by $\bar{x} \pm Z_{\alpha/2} \frac{\sigma}{\sqrt{n}}$

(a) $95\% \Rightarrow 1 - \alpha = 0.95$

$0.05 = \alpha$, implies that $\alpha/2 = 0.025$

$P(-Z_{0.025} \leq Z \leq Z_{0.025}) = 1 - 0.05 = 0.95$

But $P(0 < Z < Z_{0.025}) = 0.475$

$Z_{0.025} = 1.96(Tab)$

$$Upper\ limit = 57.4 + 1.96x\frac{15.1}{\sqrt{80}} = 60.709$$

$$Lower\ limit = 57.4 - 1.96x\frac{15.1}{\sqrt{80}}$$

$$= 54.091$$

$Z_{0.005} = 2.575(Tab)\ from\ critical\ points$

$$Upper\ limit = 57.4 + 2.575x\frac{15.1}{\sqrt{80}}$$

$$= 61.746$$

$$Lower\ limit = 57.4 - 2.575x\frac{15.1}{\sqrt{80}}$$

$$= 53.053$$

Example

The mean and standard deviation of a random sample of size 100 is 900 and 60 respectively. Given that the population is normally distributed, find a 96% confidence interval of the population mean.

Solution

$n = 100,\ \bar{x} = 900,\ \delta = 60\ and\ Z_{0.02} = 2.054(Tab)$

$$Upper\ limit = 900 + 2.054x\frac{60}{\sqrt{100}} = 912.324$$

$$Lower\ limit = 900 - 2.054x\frac{60}{\sqrt{100}} = 887.676$$

Example

The weights of a sample of 36 chicken from a poultry farm were recorded as follows

Weight (kg)	*3.30*	*3.60*	*3.90*	*4.20*	*4.50*	*4.80*
Frequency	3	6	9	11	5	2

Calculate (i) the mean sample weight and standard deviation of the chicken

(ii) 97.5% confidence interval of the mean weight of the poultry on the farm

Solution

x	f	fx	fx^2
3.30	3	9.90	32.67
3.60	6	21.60	77.76
3.90	9	35.10	136.89
4.20	11	46.20	194.04
4.50	5	22.50	101.25
4.80	2	9.60	46.08
Total	**36**	**144.9**	**588.69**

$$Mean = \frac{\sum fx}{\sum f} = \frac{144.9}{36}$$

$$\therefore Sample\ mean = 4.025$$

$$Sample\ Variance = \frac{n}{n-1}\left(\frac{\sum fx^2}{\sum f} - \left(\frac{\sum fx}{\sum f}\right)^2\right)$$

$$Sample\,Variance = \frac{36}{35}\left(\frac{588.69}{36} - \left(\frac{144.9}{36}\right)^2\right)$$

$$= 0.156$$

$$\therefore standard\ deviaton = \sqrt{0.156} = 0.395$$

$$97.5\% \Rightarrow 1 - 0.975 = 0.025$$

$$\Rightarrow Z_{0.0125} = 2.24\,(Tab)\ Used\ P = 0.4875(Tab)$$

$$Upper\ \ limit = 4.025 + 2.24\,x\frac{0.395}{6}$$

$$= 4.1725$$

$$Lower\ \ limit = 4.025 - 2.24\,x\frac{0.395}{6}$$

$$= 3.8775$$

$$\therefore Therefore\ confidence\ interval\ is\ [3.8775, 4.1725]$$

Note:

$$\text{Standard error} = \frac{\delta}{\sqrt{n}}$$

$$\text{Maximum error} = Z_{\alpha/2}\frac{\delta}{\sqrt{n}}$$

Class width = Upper limit − lower limit

Confidence interval for μ (when $n < 30$) and σ^2 is unknown.
Here the t- distribution is used

$$T = \frac{x-\mu}{s/\sqrt{n}} \text{ from which the } (1-\alpha)\%$$

C.I for μ *is given by*

Where $\bar{x} \pm t_{\alpha/2}, (n-1)\frac{s}{\sqrt{n}}$, *where*

$n-1$ *indicates the degree of freedom*

Assignment

The table below shows the distribution of weights of a random sample of 16 tins taken from a large consignment.

Weight (gm)	97	98	99	100	101	102
Frequency	2	1	2	3	6	2

Assuming the weghts are normally distributed, determine a 95% confidence interval for the mean weight of all the tins.

CHAPTER EIGHT

SCATTER GRAPHS AND CORRELATION

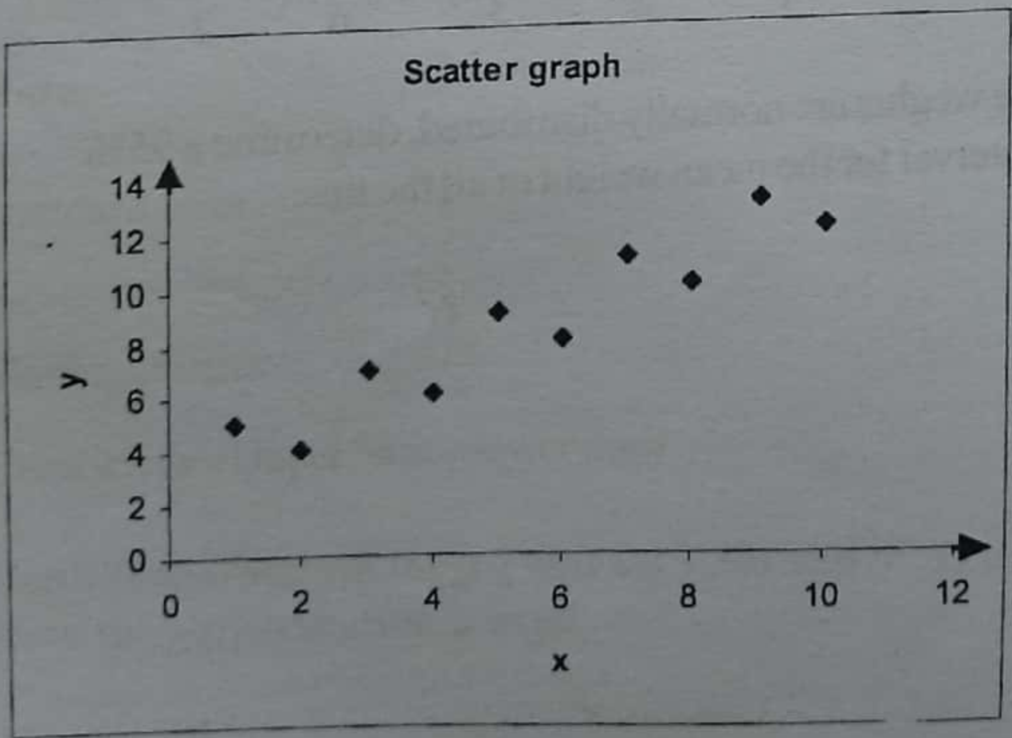
8.1 INTRODUCTION – SCATTER GRAPHS

Relation between two variables can be shown on diagrams or graphs known as scatter diagrams or graphs. A scatter graph is obtained by representing scores of one variable on the vertical axis and the other scores on the horizontal axis.

Example

Draw a scatter diagram for the following data.

x	1	2	3	4	5	6	7	8	9	10
y	5	4	7	6	9	8	11	10	13	12



8.2 REGRESSION LINE

When a scatter graph is plotted, a line of best fit can be drawn through the points. This line is called the regression line.

Note: The regression line should pass through $(\bar{x}\ \bar{y},)$

where

$$\bar{x} = \frac{\sum x}{n} \text{ and } \bar{y} = \frac{\sum y}{n}$$

Drawing a regression line "by eye"

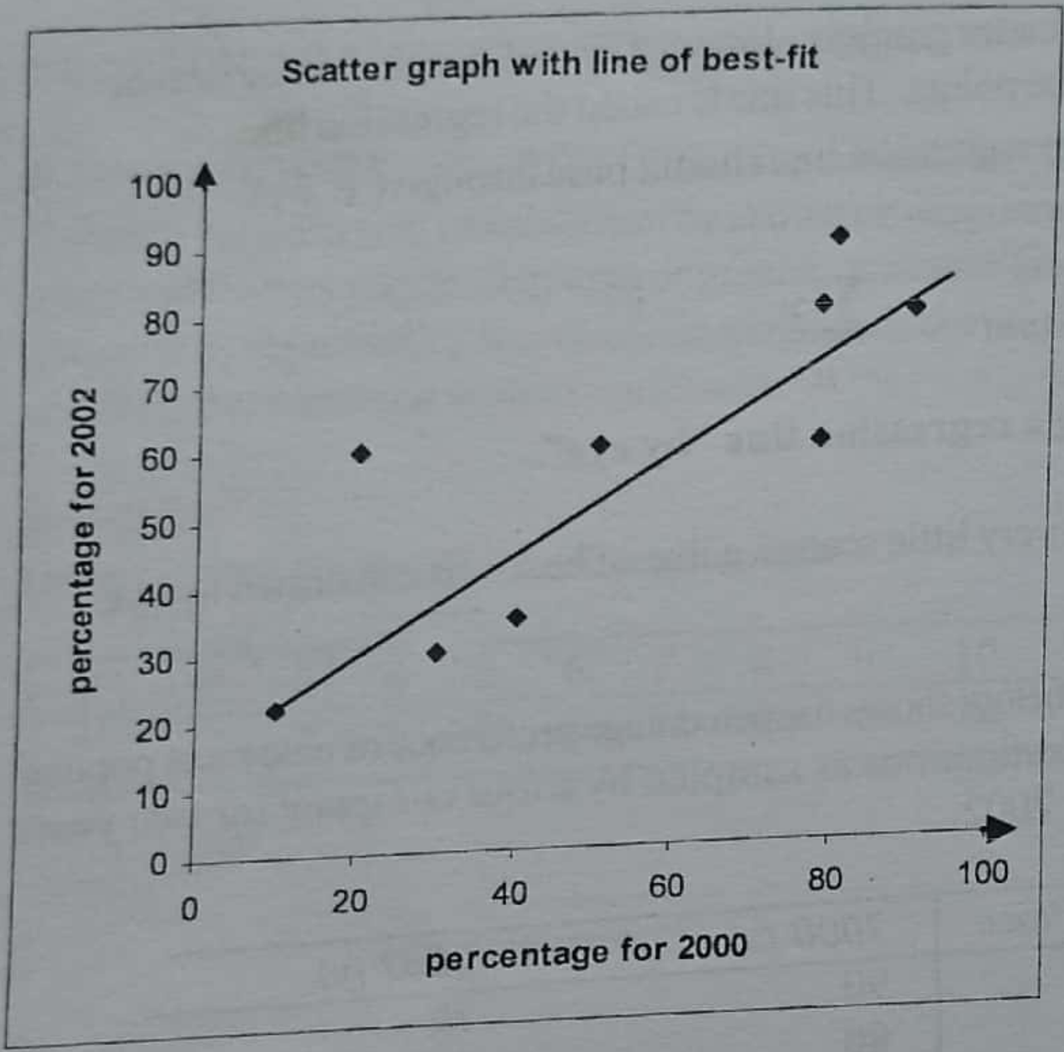
If there is very little scatter, a line of best – fit can drawn by eye.

Example

The table below shows the percentage preference of nine most popular holiday destinations as sampled by a tour company for two years 2000 and 2002.

Holiday Place	*2000 (x)*	*2002 (y)*
F	90	79
G	80	90
S	78	80
I	78	60
A	50	60
Y	40	35
C	30	30
H	20	60
B	10	22

i. Plot a scatter diagram for the data, and include the line of best fit.



The line above is used to predict y from x and hence known as y on x.

If the points are widely scattered two different lines can be drawn. One is y on x and the other x on y.

Steps for drawing the line y on x by eye.

Assume the values of x are accurate.

(i) Find the mean M $(\bar{x}, \bar{y})$ of the distribution.

(ii) Through M draw a line parallel to y – axis.

(iii) Find the (mean) M_L of points on the left

(iv) Find the (mean) M_R of the points on the right.

(v) Then draw the line of best fit through M, M_L and M_R

Steps for drawing the line x on y by eye.

Assume the values of y are accurate.

i) Find the mean $M(\bar{x}, \bar{y})$ of the distribution.
ii) Through M draw a line parallel to x-axis.
iii) Find the M_A of the points above
iv) Find the M_B of the points below.
v) Then draw the line of best fit through M, M_A and M_B

Example:

X	2	3	4	5	6	7	8	10
Y	13	7	11	16	8	15	12	16

Draw on a scatter diagram and the line of best fit for predicting y from x.

Solution

$$M(\bar{x}, \bar{y}) = \left(\frac{\sum x}{n}, \frac{\sum y}{n}\right) = (5.6, 12.3)$$

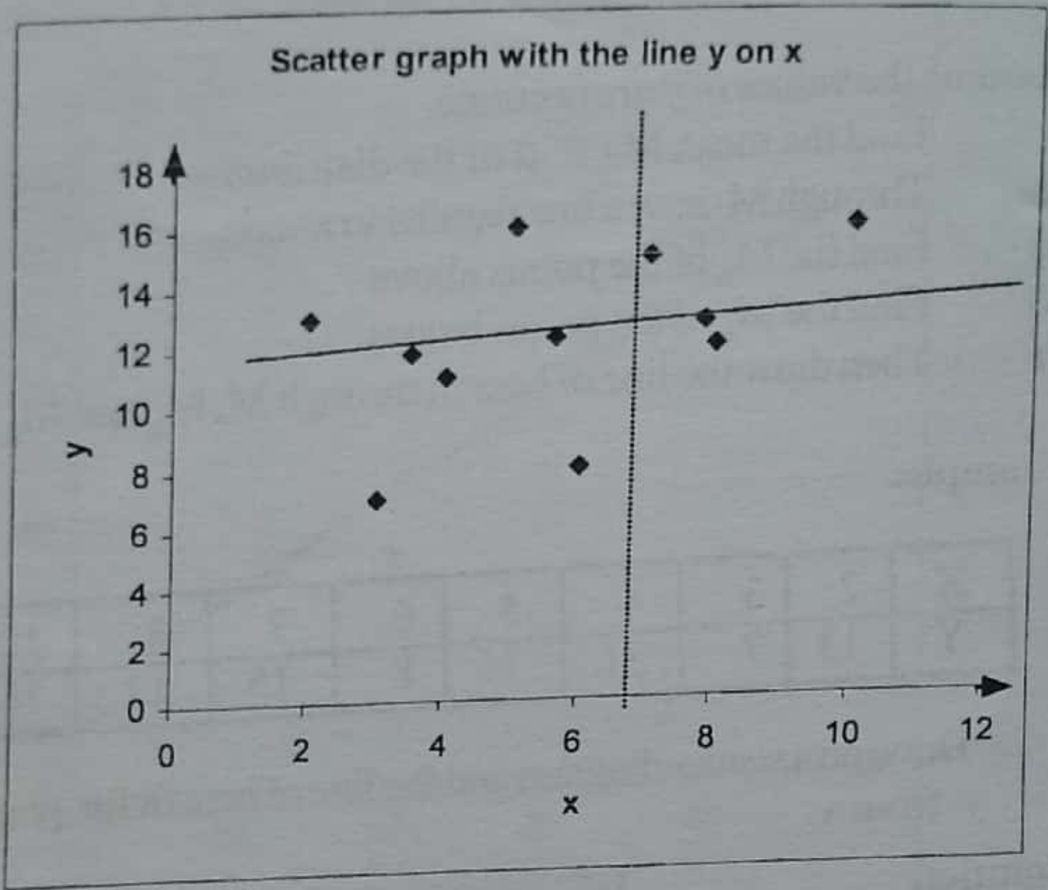
$$M_L = \left(\frac{2+3+4+5}{4}\right), \left(\frac{13+7+11+16}{4}\right) = (3.5, 11.8)$$

$$M_R = \left(\frac{6+7+8+10}{4}\right), \left(\frac{8+15+12+16}{4}\right) = (7.8, 12.8)$$

The dotted line is $x = 5.6$ which is parallel to y – axis

Note: In order to draw x on y two other points are needed and the line $y = 12.3$. Therefore line x on y has been left as an exercise.

Note: The slope of a regression line is also known as the regression coefficient.



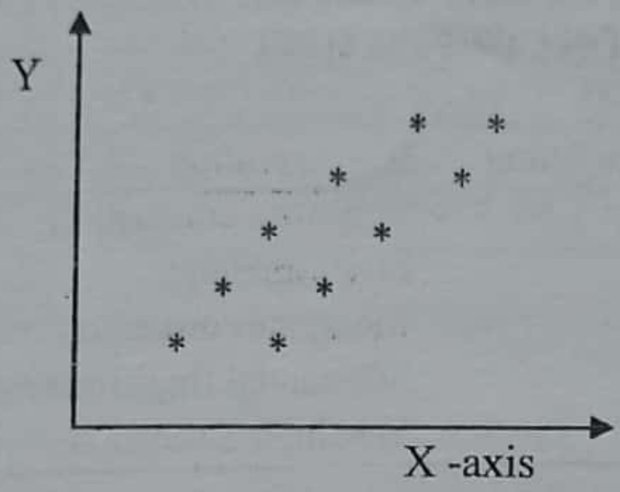
8.3 CORRELATION

Correlation is method used to determine relationships between two or more variables. Correlation coefficient is the index used to measure the degree of correlation

Types of Correlation

Positive correlation

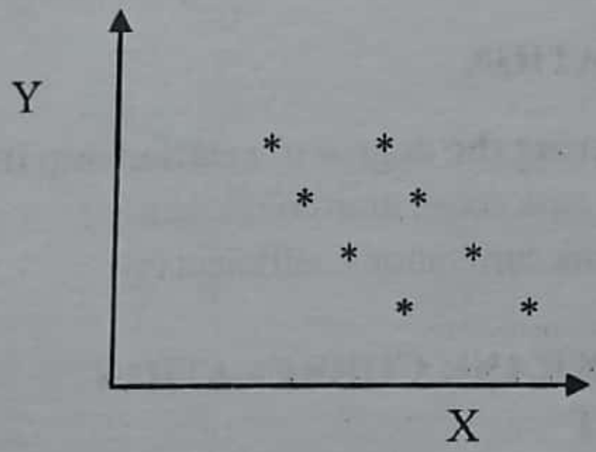
If y tends to increase as x increases, then there is positive correlation. The correlation coefficient is between 0 and 1



Negative correlation

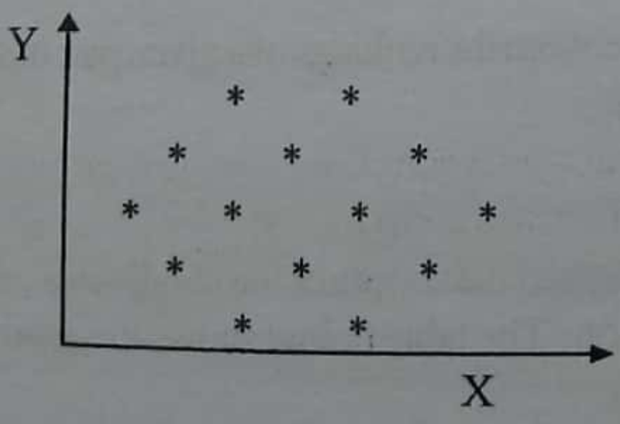
If y tends to decrease as x increases, then there is negative correlation.

The correlation coefficient is between -1 and 0.



Zero or No correlation

If there is no relationship between x and y, then there is no correlation.



8.4 INTERPRETATION OF THE MAGNITUDE OF CORRELATION COEFFICIENT

Correlation coefficient	*interpretation*
0 – 0.19	Very low correlation
0.2 – 0.39	Low correlation
0.4 – 0.59	Moderate correlation
0.6 – 0.79	Substantial (high) correlation
0.8 – 1.0	Very high correlation

Note:
The sign associated with correlation coefficient will be the one responsible for the type of correlation.
Example –0.5 would indicate a moderate negative linear correlation.

RANK CORRELATION

Method for measuring the degree of relationship include:-

1. Spearman's rank correlation coefficient (ρ)
2. Kendall's rank correlation coefficient (τ)

8.5 SPEARMAN'S RANK CORRELATION COEFFICIENT

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

where d is the difference between the rankings of a given pair of scores and n is number of pairs.

Example
Two examiners x and y marked the scripts of ten candidates who sat a mathematics examination. The table below shows the examiners rankings of the candidates.

Examiner	A	B	C	D	E	F	G	H	I	J
x	8	5	9	2	10	1	7	6	3	4
y	5	3	6	1	4	7	2	10	8	9

Calculate spearman's correlation coefficient (ρ)

Solution:

	Rx	*Ry*	*D=Rx-Ry*	D^2
A	8	5	3	9
B	5	3	2	4
C	9	6	3	9
D	2	1	1	1
E	10	4	6	36
F	1	7	-6	36
G	7	2	5	25
H	6	10	-4	16
I	3	8	-5	25
J	4	9	-5	25
				186

But $\rho = 1 - \frac{6\sum D^2}{n(n^2-1)}$

$\rho = 1 - \frac{6x186}{990} = 1 - \frac{1116}{990}$

$\rho = -0.13$

Example

Two examiners Y and Z each marked the scripts of eight candidates who sat a mathematics contest. The table below shows the examiner's marks of the candidates.

Calculate spearman's correlation coefficient and comment about your results.

Examiners	A	B	C	D	E	F	G	H
y	72	60	56	76	68	52	80	64
z	56	44	60	74	66	38	68	52

Solution:

	Rx	*Ry*	*D=Rx-Ry*	D^2
A	3	5	-2	4
B	6	7	-1	1
C	7	4	3	9
D	2	1	1	1
E	4	3	1	1
F	8	8	0	0
G	1	2	-1	1
H	5	6	-1	1

$$\sum D^2 = 18$$

But $\rho = 1 - \dfrac{6\sum D^2}{n(n^2-1)}$

$$\rho = 1 - \frac{6x18}{8(8^2-1)} = \frac{11}{14}$$

$\therefore \rho = 0.7857$

Since ρ= 0.7857 shows that there is a high positive linear correlation between the two examiners x and y.

Example

The following table shows the marks of eight students in biology and chemistry. Rank the results and find the value of spearman's coefficient of rank correlation. Test for the significance of this at 5% level.

Biology x	65	65	70	75	75	80	85	85
Chemistry y	50	55	58	55	65	58	61	65

Solution:

x	y	Rx	Ry	$D=Rx-Ry$	D^2
65	50	7.5	8	-0.5	0.25
65	55	7.5	6.5	1	1
70	58	6	4.5	-1.5	2.25
75	55	4.5	6.5	-2	4
75	65	4.5	1.5	3	9
80	58	3	4.5	-1.5	2.25
85	61	1.5	3	-1.5	2.25
85	65	1.5	1.5	0	0
					$\sum D^2 = 21$

But $\rho = 1 - \frac{6\sum D^2}{n(n^2-1)}$

$$\rho = 1 - \frac{6x21}{8(8^2-1)}$$

$$\rho = 1 - \frac{126}{504}$$

$$\therefore \rho = 0.75$$

From the table, for eight pairs at 5% level of significance $\rho = 0.71$
since from above $\rho = 0.75$ which is greater than 0.71
$\Rightarrow$ it is significant at 5% level.

8.7 KENDALL'S RANK CORRELATION COEFFICIENT (τ)

Kendall's rank correlation coefficient is given by

$$\tau = \frac{agreements - disagreements}{total\ number\ of\ pairs} = \frac{S}{\frac{1}{2}n(n-1)}$$

Where S is the total sum of all scores. For set of n objects the total number of pairs is $\frac{1}{2}n(n-1)$

Obtaining scores.

1. The data is arranged in two rows, the first row should be in descending order.
2. Arrange the data in the second row in accordance with that of the first row.
3 Obtain ranks for each of the rows.
4. Compare each of the score in the second row with the rest of the members starting from the first member. A score of +1 is allotted to a pair of objects in the right order and – 1 for a pair not in the right order.

The total sum of scores given (S).

Example

Two examiners Y and Z each marked the scripts of ten candidates who sat a mathematics examination. The table below shows the examiners ranking of the candidates.

Examiner	A	B	C	D	E	F	G	H	I	J
Y	5	3	6	1	4	7	2	10	8	9
Z	6	3	7	2	5	4	1	10	9	8

Calculate Kendell's correlation coefficient for the two examiners.

Solution

Examiner	D	G	B	E	A	C	F	I	J	H
Ry	1	2	3	4	5	6	7	8	9	10
RZ	2	1	3	5	6	7	4	9	8	10

DG	DB	DE	DA	DC	DF	DI	DJ	DH	Score
-1	1	1	1	1	1	1	1	1	7
	GB	GE	GA	GC	GF	GI	GJ	GH	
	1	1	1	1	1	1	1	1	8
		BE	BA	BC	BF	B1	BJ	BH	
		1	1	1	1	1	1	1	7
			EA	EC	EF	EI	EJ	EH	
			1	1	-1	1	1	1	4
				AC	AF	A1	AJ	AH	
				1	-1	1	1	1	3
					CF	C1	CJ	CH	
					-1	1	1	1	2
						F1	FJ	FH	
						1	1	1	3
							IJ	IH	
							-1	1	10
								JH	
								1	1

Total score is 35 $\therefore S = 35$

$$\tau = \frac{2S}{n(n-1)} = \frac{35\times 2}{10\times 9} = \frac{35}{45}$$

$\tau = 0.78$

Example

In many Government Institutions, officers complain about typing errors. A test was designed to investigate the relationship between typing speed and errors made.

Twelve typists A, B, C, D …L. were picked at random to type the same text. The table below shows the rankings of the typists according to speed and errors made. [N.B lowest ranking in errors implies least errors made]

Typist	A	B	C	D	E	F	G	H	I	J	K	L
Speed	3	4	2	1	8	11	10	6	7	12	5	9
Error	2	6	5	1	10	9	8	3	4	12	7	11

Calculate the rank correlation coefficient (Kendall's). Test the assertion made by the officers and comment on your result.

	D	C	A	B	K	H	I	E	L	G	F	J
Rx	1	2	3	4	5	6	7	8	9	10	11	12
Ry	1	5	2	6	7	3	4	10	11	8	9	12
P	11	7	9	6	5	6	5	2	1	2	1	=55
Q	0	3	0	2	2	0	0	2	2	0	0	=11

S = Agreements (P) – Disagreements (Q)
= 55 – 11 = 44

$$\tau = \frac{2S}{n(n-1)} = \frac{88}{12\times 11} = \frac{88}{132}$$

$\tau = 0.67$

From the table for 12 pairs of observation at 1% level of significance, $\tau = 0.58$

But $0.67 > 0.58$

$\Rightarrow$ Typing speed causes errors at 1% level of significance based on twelve observations.

Example

The following are the final examination scores, which the 12 students obtained in Mathematics (X) and Economics (Y)

X	35	56	65	78	49	82	22	90	77	53	52	93
Y	57	72	63	76	53	100	38	82	82	19	43	79

Solution

Rx	Ry	P	Q
1	4	8	3
2	2.5	8	1
3	1	9	0
4	5	7	1
5	2.5	7	0
6	7	5	1
7	6	5	0
8	12	0	4
9	10	1	2
10	9	1	1
11	8	1	0
12	11		
	Total	52	13

$S = 52 - 13 = 39$

$$\tau = \frac{39x2}{12x11}$$

$$\tau = 0.591$$

Note: In case of ties in the first row, the best alternative would be to use Spearman's method since the questions are normally open.

Otherwise if you have the ties in first row, the student with a better mark in the second paper is ranked first. Also when you are comparing the second row make sure you put into consideration the firs row because of the ties.

Exercise

1.Three persons P, Q and R were asked in order of importance nine features of a house (A, B, C …I). Calculate spearman's rank correlation coefficients between the pairs of preferences as shown in the following table.

	A	B	C	D	E	F	G	H	I
P	1	2	4	8	9	7	6	3	5
Q	1	4	5	8	7	9	2	3	6
R	1	9	6	8	7	4	2	3	5

How far does this help to decide which pair from the three would be most likely to be able to compromise on a suitable house.

Answer: P, Q: 0.75 Q, R: 0.567 P, R: 0.317

2. The price (p) of matooke is found to depend on the distance (d) the market is away from the nearest town. The table below gives the average price of matooke for markets around Kampala City.

d(km)	40	8	17	20	24	30	10	28	16	28
p(sh)	120	160	140	130	135	125	150	130	145	125

(i) Plot this data on a scatter diagram.

(i) Draw the line of best fit on your diagram.

(iii) Find the equation of your line in form of

$$P = \alpha + \beta d$$

where α and β are constants.

Hence estimate the price of matooke when $d = 5$

Answer: (iii) $p = 161.4 - 1.111d$,

When $d = 5$ Implies $p = 156$ shs.

3. The following table gives the order in which six candidates were ranked in two tests x and y

x	E	C	B	F	D	A
y:	F	A	D	E	C	C

Calculate the coefficient of rank correlation(Kendall's) and comment on your results.

Answer: $\tau = 0.4$

4. In a certain commercial institution, a speed and error typing examination was administered to 12 randomly selected candidates A, B, C, ... L of the institution. The table below shows their speeds (y) in seconds and the number of errors in their typed script (x).

	A	B	C	D	E	F	G	H	I	J	K	L
x	12	24	20	10	32	30	28	15	18	40	27	35
y	130	136	124	120	153	160	155	142	145	172	140	157

(i) Plot the data on a scatter diagram.

(ii) Draw the line of best fit on your diagram and comment on the likely association between speed and the errors made.

(iii) determine the equation of your line in the form $y = xk + b$ where k and b are constants.

(iv) By giving rank 1 to the fastest student and the student with the fewest errors, rank the above data and use it to calculate the rank correlation coefficient. Comment on your results.

Answer: (iii) $y = 1.42x + 110.1$

(iv) $\rho = 0.84$, $\tau_0 = 0.7$

5. The following table shows the marks of 8 students in two different subjects

Maths	65	60	70	75	75	80	85	50
Geography	50	55	54	60	65	58	65	61

Calculate the rank correlation coefficient

Answer $\tau = \frac{2}{7}$, $\rho = 0.435$

6. 8 students took an exam in Math and Physics an their grades were as follows:

Math	A	B	A	D	O	D	B	F
Physics	C	B	C	E	E	C	A	F

Calculate Spearman's rank correlation coefficient

Answer: $\rho = 0.7083$

CHAPTER NINE

ERRORS

9.1 INTRODUCTION

Errors arise from the many calculations and approximations used when working out numbers.
Suppose a calculation involved the fraction $\frac{1}{3}$, then using 0.33 (to two decimal places) causes a round-off error.

9.2 TYPES OF ERRORS

Random errors
These type of errors that occur due machine or human failure and therefore they can not be treated numerically.
Suppose a student is given 54 and uses 45 instead.

Rounding errors
Some numbers are normally corrected to a number of decimal places or significant figures. An error usually arises because most values used in the computations are normally approximations of the true values and such are referred to as rounding errors.

Example
Round off 3.896234 to four decimal places

3.896234 = 3.8962 (to four decimal places)

Round off $\frac{2}{3}$ to three decimal places

$\frac{2}{3} = 0.6666666666666 = 0.667$ (3 decimal places)

Round off 0.00652673 to three significant figures

0.00652673 = 0.00653 (3 significant figures)

Round off 7.00214 to four significant figures

7.00214 = 7. 002 (to 4 significant figures)

Round 5,415,000 to 3 significant figures

5,415,000 = 5,420,000 (3 significant figures)

Note that rounding is done in a single step.

Truncation errors

Occurs when an infinite process is terminated at a some point.

Example

Truncate $\frac{2}{3}$ to three significant figures

$$\frac{2}{3} = 0.666666666666 = 0.666 \text{ (3 s.f's)}$$

Sometimes infinite series are terminated at some point such that higher terms are neglected since these values that are likely to be too small to affect the result.

Example

Find the approximate value of $e^{0.6}$ to 2 decimal places

$$e^x = 1 + x + \frac{x^2}{2!} + \frac{x^3}{3!} + \ldots$$

$$e^{0.6} = 1 + 0.6 + \frac{0.6^2}{2!} + \frac{0.6^3}{3!} + .\frac{0.6^4}{4!} + \ldots \quad = 1.822$$

$$= 1.82 \text{(2.d.p's)}$$

COMMON TREMS USED

Include error, absolute error, relative error and percentage error.

Error. If x represents an approximation to the value of X

and ∂x is the error in this approximation.

Then ∂x = X- x = True value - Approimate value

Absolute error is the actual size of error disregarding the sign.

Therefore absolute error $|\partial x| = |X - x| = e$ which is at times used. So absolute error only takes positive values only.

Therefore

$$|x| = \begin{cases} x, & x \geq 0 \\ -x, & x \leq 0 \end{cases}$$

Example

Round off 32.52632 to 2 decimal places and determine the bsolute error.

Solution

32.52632 = 32.53 (to 2 decimal places)

So X = 32.52632 and x = 32.53

Error = X- x = 32.52632 – 32.53 = -0.00368

Therefore the absolute error

$= |-0.00368| = 0.00368$

Relative error = $\dfrac{\text{Absolute error}}{\text{Exact value or Approximate value}}$

The relative error gives an indication of whether the error is relatively large or small.

For example if in measuring a distance of 50 metres an error of 1 cm is introduced, then relative error is

$$\frac{1}{5000} = 0.0002$$

However if an error of 1 cm is made in 1metre

Then $\frac{1}{100} = 0.01$

Relative error = $\left|\frac{\partial x}{X}\right|$ or $\left|\frac{\partial x}{x}\right|$

Percentage error or percentage relative error

The relative error is expressed as a percentage i.e multiplied by 100.

So percentage error = Relative error x 100

$$\left|\frac{\partial x}{X}\right| \text{ x } 100$$

Example

Using your calculator, state as accurately as possible the value of √3. Find the percentage error involved in approximating this value to two decimal places.

Solution

$\sqrt{3} = 1.732050808$ which is 1.73 to two decimal places

Absolute error = $|1.732050808 - 1.73|$
$= 0.002050808$

Percentage error = $\left|\frac{0.002050808}{1.732050808}\right| \times 100\% = 0.118$

Note that the percentage (%) sign can be ignored.

Maximum possible error(s)

This depend on the number of decimal places or significant figures used in the computation. If a number is corrected to n decimal places, then the maximum possible error in it becomes 0.5×10^{-n}
If the maximum and minimum possible errors are known then the

$$\text{absolute error} = \frac{\max \textit{value} - \min \textit{value}}{2}$$

9.3 Limits of accuracy

We can state the upper and lower limit of given figures
Example If a student weighs 70 kgs.
Then this value can range from 69.5 kg to 70.5 kg

Example
If X is given to a stated level of accuracy. State the lower and upper bounds of X.
(a) 7.47
(b) 0.184
(c) 7.0

Solution
(a) error = $0.5 \times 10^{-2} = 0.005$
Lower bound = 7.47 – 0.005 = 7.465
Upper bound = 7.47 + 0.005 = 7. 475

(b) error = $0.5 \times 10^{-3} = 0.0005$
Lower bound = 0.184 – 0.0005 = 0.1835
Upper bound = 0.184 + 0.0005 = 0.1845

(c) error = $0.5 \times 10^{-1} = 0.05$

Lower bound = 7.0 – 0.05 = 6.95
Upper bound = 7.0 + 0.05 = 7.05

Example

The sides of a rectangle are measured as 5.24 m and 6.38 m, to the nearest centimeter.
(a) Calculate minimum value of the perimeter.
(b) Calculate the maximum value of the area
(c) Calculate the minimum value of the area
hence determine the absolute error

Solution

W = 5.24, then it ranges from 5.235 to 5.245
L = 6.38, then it ranges from 6.375 to 6.385

$$\text{Perimeter} = 2(L + W)$$

$$\text{Minimum perimeter} = 2(5.235 + 6.375) = 23.22 \text{ m}$$

$$\text{Maximum area} = 5.245 \times 6.385 = 33.489325 \text{ m}^2$$

(c) $\text{Minimum area} = 5.235 \times 6.375 = 33.373125$

$$\text{absolute error} = \frac{\max value - \min value}{2}$$

$$= \frac{33.489325 - 33.373125}{2}$$

$$= 0.0581$$

Example

If $x = 5.356$ and $y = 6.81$, both numbers are rounded. Find the maximum and minimum value of:

(a) $y - x$

(b) $\frac{y}{x}$

Solution

y ranges from 6.805 to 6.815

Then x ranges from 5.3555 to 5.3565

Maximum $(y - x) = 6.815 - 5.3555$
$= 1.4595$

Minimum $(y - x) = 6.805 - 5.3565$
$= 1.4485$

(b) Maximum $\frac{y}{x} = \frac{\max y}{\min x} = \frac{6.815}{5.3555}$

$= 1.272524$

Minimum $\frac{y}{x} = \frac{\min y}{\max x} = \frac{6.805}{5.3565}$

$= 1.270419$

Example

Given that $N = \frac{12.4}{4.20} - \frac{10.80}{6.124}$

(a) Write the possible error in each of the values given

(b) Estimate the range of values within which N Lies. Hence determine the absolute error in N

Solution

Error in 12.4 is 0.05

Error in 4.20 is 0.005

Error in 10.80 is 0.005

Error in 6.124 is 0.0005

(b) Maximum value of N= $\frac{12.45}{4.195}-\frac{10.795}{6.1245}=1.20523$

Minimum value of N= $\frac{12.35}{4.205}-\frac{10.805}{6.1235}=1.17247$

Range of value is [1.17247, 1.20523]

absolute error = $\frac{\text{max } value - \text{min } value}{2}$

$= \frac{1.20523-1.17247}{2}$

$= 0.01638$

Example

Obtain the range of values within which the exact value of

$2.7654+\frac{3.8006x15.178}{0.9876}$ lies

Solution

Max value = $2.76545+\frac{3.80065}{0.98755}x\,15.1785=61.18089$

Min value = $2.7635+\frac{3.80055}{0.98765}x15.1775=61.16949$

Range of values is [61.16949, 61.18089]

Example

A trader in tea and coffee makes an annual profit in tea of Shs 1080 million with a margin error of $\pm 10\%$ and an annual loss in coffee of Shs 560 million with a margin of error of $\pm 5\%$.

(i) Find the range of values corresponding to his gross income.

(ii) Given that his annual income tax is Shs 75 million, express this as a percentage of his gross income giving your answer as a range of values.

Solution

Max profit from tea = 1080 x 110% = 1188 m
Min profit from tea = 1080 x 90% = 972 m

Max loss from coffee = 560 x 105% = 588 m
Min loss from coffee = 560 x 95% = 532 m

Range of values for the gross income
Max income = max profit – min loss
= 1188 – 532 = 656 millions
Min income = min profit – max loss
= 972 – 588 = 384 millions

Range [384, 656]

(ii) lower limit $= \frac{75x100}{656} = 11.43\,\%$

Upper limit $= \frac{75x100}{384} = 19.53\%$

Example

A value of L = 200.27m was obtained when measuring the length of football pitch. Given that the relative error in this value was 0.08%, find the limits within which the value of L lie.

Solution

$L = 200.27$

$\frac{\partial L x100}{L} = 0.08$ so $|\partial L| = \frac{200.27x0.08}{100} = 0.160$

Upper limit $= 200.27 + 0.160 = 200.43$
Lower limit $= 200.27 - 0.160 = 200.11$

Triangular inequality

It states that $|a \pm b| \leq |a| + |b|$. It is important when deducing maximum possible error.

Example

If a = 3 and b = 7 then

$|3 + 7| \leq |3| + |7|$

If a =7 and b = -4

$|7 - 4| \leq |7| + |-4|$

$3 \leq 11$

9.4 Deriving formular for error propagation

ADDITION

If x and y are approximation to X and Y,

and ∂x and ∂y be error respectively.

then true value $= X + Y$

Working value $x + y = z$

So $X + Y = (x + \partial x) + (y + \partial y)$

$z + \partial z = (x + \partial x) + (y + \partial y)$

$\partial z = \partial x + \partial y$

Maximum possible absolute error in addition

$$|\partial x + \partial y| \leq |\partial x| + |\partial y|$$

from triangular inequality

Maximum possible error in addition $= |\partial x| + |\partial y|$

Maximum relative error in addition

$$\frac{|\partial x| + |\partial y|}{x + y}$$

Percentage error

$$\left(\frac{|\partial x| + |\partial y|}{x + y}\right) \times 100$$

Example

If x and y are numbers rounded to 2 and 3 decimal places respectively, Calculate the max possible error in $x + y$ if $x = 2.84$ and $y = 6.364$

Solution

$\partial x = 0.005$

$\partial y = 0.0005$

$e_{x+y} = 0.005 + 0.0005 = 0.0055$

Maximum possible error $= 0.0055$

SUBTRACTION

If x and y are approximation to X and Y,

And ∂x and ∂y be error respectively.

Then true value $= X - Y$

Working value $x - y = z$

So $X - Y = (x + \partial x) - (y + \partial y) = \partial x - \partial y$

Max possible error $|\partial x-\partial y|\leq|\partial x|+|-\partial y|=|\partial x|+|\partial y|$

Example

If x = 6.364 and y = 2.84

Determine lower and upper limit in (x-y)

Max possible error

$$\partial x=0.005$$

$$\partial y=0.0005$$

$$e_{x-y}=0.005+0.0005=0.0055$$

Lower limit = (6.364-2.84)-0.0055 = 3.5185

Upper limit = (6.364-2.84) +0.0055 = 3.5295

MULTIPLICATION

If x and y are approximation to X and Y,

And ∂x and ∂y be errors respectively.

True value = XY

z = xy

$$z+\partial z=(x+\partial x)(y+\partial y)$$

$$z+\partial z=xy+x\partial y+y\partial x+\partial x\partial y$$

dx is too small and dy is too small

Implies $\partial x\partial y\cong 0$

$$z+\partial z=xy+x\partial y+y\partial x$$

$$dz=x\partial y+y\partial x$$

$$|dz|=|x\partial y+y\partial x|\leq|x\partial y|+|y\partial x|$$

From the triangular inquality

Therefore maximum absolute error in the product is $|x\partial y|+|y\partial x|$

Relative error

$$\frac{\partial z}{z} = \frac{x\partial y + y\partial x}{xy}$$

$$= \frac{\partial x}{x} + \frac{\partial y}{y}$$

Max possible relative error

$$\left|\frac{\partial x}{x} + \frac{\partial y}{y}\right| \leq \left|\frac{\partial x}{x}\right| + \left|\frac{\partial y}{y}\right|$$

Percentage error

$$\left(\left|\frac{\partial x}{x}\right| + \left|\frac{\partial y}{y}\right| \right) \times 100$$

Example

The sides of a rectangle are measured as 5.24 m and 6.38 m, to the nearest centimeter.

(i) Determine the maximum possible area of the rectangle

(ii) Calculate the percentage error in measuring of the area

Solution

$$\partial w = 0.005$$

$$\partial l = 0.005$$

$$\partial lw = l|\partial w| + w|\partial l|$$

$= 6.38 \times 0.005 + 5.24 \times 0.005$

$= 0.0581$

Maximum possible area $= 5.24 \times 6.38 + 0.0581$

$= 33.4893$

$$\text{Percentage error} = \left(\left|\frac{\partial w}{w}\right| + \left|\frac{\partial l}{l}\right| \right) x\ 100$$

$$\left(\left|\frac{0.005}{5.24}\right| + \left|\frac{0.005}{6.38}\right| \right) x\ 100 = 0.174$$

Example

Two positive numbers y_1 and y_2 are each rounded off to three decimal places to give x_1 and x_2 respectively. Find in terms of x_1 and x_2 the maximum relative error in using x_1x_2 as an approximation for y_1y_2.

Solution

$\partial x_1 = 0.0005 \qquad y_1 = x_1 + 0.0005$

$\partial x_2 = 0.0005 \qquad y_2 = x_2 + 0.0005$

$$y_1y_2 = x_1x_2 + e$$

$$= (x_1 + 0.0005)(x_2 + 0.0005)$$

$$x_1x_2 + 0.0005x_2 + 0.0005x_1 + 0.0005x0.0005$$

$0.0005\ is\ too\ small$

$0.0005x0.0005 \cong 0$

$$e = 0.0005x_1 + 0.0005x_2$$

$$\text{Relative error} = \left|\frac{0.0005x_1 + 0.0005x_2}{x_1x_2}\right|$$

Max relative error $\leq \left|\frac{0.0005}{x_1}\right| + \left|\frac{0.0005}{x_2}\right|$

Max relative error $= \left|\frac{0.0005}{x_1}\right| + \left|\frac{0.0005}{x_2}\right|$

Note if m = xyz

Then the absolute error in the product

$$|\partial m| = yz|\partial x| + xz|\partial y| + yx|\partial z|$$

The max relative error

$$\left|\frac{\partial m}{m}\right| \leq \left|\frac{\partial x}{x}\right| + \left|\frac{\partial y}{y}\right| + \left|\frac{\partial z}{z}\right|$$

$$= \left|\frac{\partial x}{x}\right| + \left|\frac{\partial y}{y}\right| + \left|\frac{\partial z}{z}\right|$$

DIVISION

If x and y are approximation to X and Y,

And ∂x and ∂y be errors respectively

True value $= \frac{X}{Y}$ but when $z = \frac{x}{y}$ is used

$$z + \partial z = \frac{x + \partial x}{y + \partial y}$$

$$= \frac{x + \partial x}{y + \partial y} \quad \frac{(y - \partial y)}{(y - \partial y)}$$

$$z + \partial z = \frac{xy - x\partial y + y\partial x - \partial x\partial y}{y^2 - y\partial y + y\partial y - (\partial y)^2}$$

$$z + \partial z = \frac{xy - x\partial y + y\partial x - \partial x\partial y}{y^2 - (\partial y)^2}$$

dx is too small and dy is too small

Implies $\partial x\partial y \cong 0, (\partial y)^2 \cong 0$

$$z + \partial z = \frac{xy - x\partial y + y\partial x}{y^2}$$

$$\partial z = \frac{\partial x}{y} - \frac{x\partial y}{y^2}$$

Max absolute error $= |\partial z| = \frac{y|\partial x| + x|\partial y|}{y^2}$

Relative error in division

$$\frac{\partial z}{z} = \frac{\partial x}{x} - \frac{\partial y}{y}$$

Max possible relative error in division

$$\left|\frac{\partial x}{x} - \frac{\partial y}{y}\right| \leq \left|\frac{\partial x}{x}\right| + \left|\frac{-\partial y}{y}\right|$$

$$= \left|\frac{\partial x}{x}\right| + \left|\frac{\partial y}{y}\right|$$

Example

Given the numbers a = 23.037 and b = 8.4658, measured to their nearest number of decimal places indicated,

(i) State the maximum possible errors in a and b.

(ii) determine the absolute error in $\frac{a}{b}$

(iii) find the limits within which $\frac{a}{b}$ lies, correct to 4 decimal places

Solution

(i) $\partial a = 0.0005$

$\partial b = 0.00005$

(ii) $\partial(\frac{a}{b}) = \frac{b|\partial a| + a|\partial b|}{b^2}$

$= \frac{8.4658|0.0005| + 23.037|0.00005|}{8.4658^2}$

$= 0.0000751$

(iii) lower limit $= \frac{23.037}{8.4658} - 0.0000751 = 2.7211$

Upper limit $= \frac{23.037}{8.4658} + 0.0000751 = 2.7213$

Example

Given that Y_1 and Y_2 are approximations to X_1 and X_2 with errors E_1 and E_2 respectively, show that the maximum possible relative error in $\frac{X_1}{X_2}$ is $\left|\frac{E_1}{Y_1}\right| + \left|\frac{E_2}{Y_2}\right|$

Solution

$X_1 = Y_1 + E_1, \quad X_2 = Y_2 + E_2$

$$= \frac{X_1}{X_2} = \frac{Y_1 + E_1}{Y_2 + E_2}$$

$$= \frac{(Y_1 + E_1)(Y_2 - E_2)}{(Y_2 + E_2)(Y_2 - E_2)}$$

$$= \frac{Y_1Y_2 + Y_2E_1 - Y_1E_2 - E_1E_2}{Y_2^{\,2} - E_2^{\,2}}$$

Assumptions

E_1 is too small, E_2 is too small

$E_1E_2 \cong 0, (E_2)^2 \cong 0$

$$\frac{Y_1}{Y_2} + E = \frac{Y_1}{Y_2} + \frac{E_1}{Y_2} - \frac{Y_1E_2}{Y_2^2}$$

$$E = \frac{E_1}{Y_2} - \frac{Y_1E_2}{Y_2^2}$$

$$\text{Relative error } = \left(\frac{E_1}{Y_2} - \frac{Y_1E_2}{Y_2^2}\right) x \frac{Y_2}{Y_1}$$

$$= \frac{E_1}{Y_1} - \frac{E_2}{Y_2}$$

$$= \left|\frac{E_1}{Y_1} - \frac{E_2}{Y_2}\right| \le \left|\frac{E_1}{Y_1}\right| + \left|\frac{-E_2}{Y_2}\right|$$

$$\text{Max relative error} = \left|\frac{E_1}{Y_1}\right| + \left|\frac{E_2}{Y_2}\right|$$

NOTE: The maximum possible relative error in multiplication and division are the same

ERRORS IN FUNCTIONS

Maclaurin's theorem

$f(x) = f(0) + f'(0) + \frac{f''(0)}{2!}x^2 + -$

Let z = f(x)

$$z + \partial z = f(x + \partial x)$$

$$f(x + \partial x) = f(x) + f'(x)\partial x + \frac{f''(x)(\partial x)^2}{2!} + \ldots$$

dx is too small

$$(\partial x)^2 \cong 0$$

$$\partial z = f'(x)\partial x$$

$$|\partial z| = |f'(x)||\partial x|$$

$$\text{Relative error} = \frac{|\partial z|}{z} = \left|\frac{f'(x)\partial x}{f(x)}\right|$$

Example.

If $z = x^5$. Find expressions for the absolute error and maximum relative error in z.

Solution

$z = f(x) = x^5$

$$f'(x) = 5x^4$$

$$\partial z = 5x^4 \partial x$$

$$|\partial z| = 5x^4 |\partial x|$$

$$R.E = \left|\frac{\partial z}{z}\right| = \left|\frac{5x^4 \partial x}{x^5}\right|$$

$$= 5\left|\frac{\partial x}{x}\right|$$

Example

If z = sinx. Determine the expression for the absolute error and maximum relative error.

Solution

f(x) = sinx

$$f'(x) = \cos x$$

$$|\partial z| = |\cos x||\partial x|$$

$$R.E = \left|\frac{\cos x}{\sin x}\right||\partial x|$$

$$= |\cot x||\partial x|$$

Example

Given that the error in measuring an angle is 0.5^0.
Find the maximum possible error in $\frac{\sin x}{\cos x}$
If $x = 30^0$.

Solution

$$\text{Max relative error} = \left|\frac{E_1}{Y_1}\right| + \left|\frac{E_2}{Y_2}\right|$$

Let $Y_1 = \sin x$ and $Y_2 = \cos x$

$$E_1 = |\cos x \partial x|$$

$$E_2 = |-\sin x \partial x| = |\sin x \partial x|$$

Note: $E = f'(x)\partial x$

$$R.E = \left|\frac{\cos x \partial x}{\sin x}\right| + \left|\frac{\sin x \partial x}{\cos x}\right|$$

$$= |\cot x||\partial x| + |\tan x||\partial x|$$

$$\partial x = 0.5^0 = \frac{\pi}{360}$$

$$R.E = |\cot 30|\left|\frac{\pi}{360}\right| + |\tan 30|\left|\frac{\pi}{360}\right|$$

$$= 0.015115 + 0.00504$$
$$= 0.02015$$

Percentage error $= 0.02015 \times 100$
$= 2.015$

Example

If m, y and z are 3 numbers which have been rounded.
Evaluate $\frac{m^2 + y}{y + z}$ with its error bounds
if $m = 2.12$, $y = 3.8$ and $z = 0.31$

Solution

$$\frac{m^2+y}{y+z}=\frac{2.12^2+3.8}{3.8+0.31}$$

Error in numerator =

$$|f'(m)\partial m|+|\partial y|=2m\partial m+|\partial y|$$

$$=2 \text{ x } 2.12 \text{ x } 0.005+0.05=0.0712$$

Error in denominator = $\dfrac{p}{n}=\dfrac{m^2+y}{y+z}=\dfrac{8.2944}{4.11}=2.0181$

$$|\partial y|+|\partial z|=0.05+0.005=0.055$$

If we let u= $\partial u=\left|\frac{\partial p}{p}\right|+\left|\frac{\partial n}{n}\right|=(\left|\frac{0.0712}{8.2944}\right|+\left|\frac{0.055}{4.11}\right|)x2.0181$

$$=0.0443$$

Range $\quad 2.0181 \pm 0.0443$

$[1.9738, 2.0624]$

Exercise

1. The numbers A and B are approximated by the numbers X and Y respectively such that $A = X - a$, $B = Y - b$, where a, b are small numbers compared to A and B. Given that $Y = f(X)$ and $B = f(A)$. Show that $|b| = |a||f'(A)|$. If $f(A) = A^p$ where p is a constant deduce that $|b| = |a| pA^{p-1}$ and find the expression for the relative error.

Answer $\frac{|ap|}{|A|}$

2. The numbers 2.6754, 4.8006, 15.175 and 0.92 have been rounded off correct to the given number of decimal places. Find the range of values within which the exact value of

$$2.6754(4.8006 - \frac{15.175}{0.92})$$ can be expected to lie.

Answer [-31.529, -31.045]

3. Given the numbers; $x = 2.678$ and $y = 0.8765$ measured to the nearest number of decimal places indicated,
   (a) state the maximum possible errors in x and y
   (b) determine the absolute error in xy
   (c) find the limits within which the product xy lies, correct to 4 decimal places

Answer (a) $\partial x = 0.0005 \quad \partial y = 0.00005$

(b) $|\partial xy| = 0.000572$

(c) Lower limit = 2.3467, Upper limit = 2.3478

4. The radius of a circle is measured as 5.34 m to the nearest cm. Calculate the lower bound of its area, correct to three significant figures.

Answer: 89.4 m^2

5. Two decimal number x and y are rounded to give X and Y with errors E_1 and E_2 respectively. Show that the maximum relative error made in approximating x^2y by X^2Y is given by

$$2\left|\frac{E_1}{X}\right|+\left|\frac{E_2}{Y}\right|$$

6. Determine the maximum absolute error in $\frac{\sqrt{z}}{x^2y^3}$,

given that x = 2.4, y = 5.4 and z = 1.8 all numbers rounded

Answer 0.000123

7. If the error in each of the values of e^x and e^{-x} is ± 0.0005,

find the maximum and minimum value of the quotient, $\frac{e^x}{e^{-x}}$

when x = 0.004, giving your nswer correct to 3 decimal places.

Answer: Max = 1.084 and Min = 1.082

8. The variable v, r and h are related by the formula

$$v=\frac{r^2}{3h}+5$$

In an experiment, the value of r and h were found to be 4.1 and 40.8 respectively. Calculate the lower and upper limits of v, correct 3 decimal places.

Answer

lower limit = 5.134, upper limit = 5.141

9. Given that Y and Z are measured with possible Errors ΔY and ΔZ respectively, show that the relative error in the product YZ is

$$\frac{Z|\Delta Y| + Y|\Delta Z|}{YZ}$$

10. A company had a capital of sh. 500 million. The profit in a certain year was sh. 25.8 million in section A of the company and sh. 14.56 million in section B of the company. There was possible error of 5% in section A and an 8% error in section B. Find the maximum and minimum values of the total profit of the sections as a percentage of the capital.

Answer: Max = 8.56%, Min = 7.58%

11. Given that $A = |x||y|\sin\theta$

(a) Deduce that the maximum possible relative error in A is given by $\left|\frac{\partial x}{x}\right| + \left|\frac{\partial y}{y}\right| + \cot\theta|\partial\theta|$

Where ∂x, ∂y and $\partial\theta$ are small numbers compared to x, y and θ respectively. Find the percentage error made in the area if $|x|$ and $|y|$ are measured with errors of $\pm$ 0.05 and angle with an error $\pm 0.5^0$, given that $|x|$ = 2.5 cm, $|y|$ = 3.4 cm and $\theta = 30^0$

Answer 4.98

12. Two numbers A and B are approximated by a and b with errors x and y respectively. Show that maximum percentage error in approximating $\frac{A}{B}$ is given by $\left(\left|\frac{x}{a}\right|+\left|\frac{y}{b}\right|\right)x100\%$

Hence find the maximum percentage error in $\frac{3.25^2}{4.562}$ to 2 decimal places.

Answer 0.32

13. Given that a = 42.326, b = 27.26 and c = -12.93 are rounded off to the given decimal places, find the range within which the exact value of the expression, $\frac{A}{B+C}$ lies.

Note, the numbers A and B are rounded off to a and b.

Answer

14. A cylindrical pipe has a radius of 2.5 cm measured to the nearest unit. If the relative absolute error made in calculating its volume is 0.125, find the absolute relative error made in measuring its height.

Answer 0.085

CHAPTER TEN

LINEAR INTERPOLATION

10.1 Interpolation

Linear interpolation is a process where by the non-tabulated values of the function are estimated on the assumption that the function behaves sufficiently smooth between the tabular points.

Example

x	1	2	3
f(x)	2	8	11

Using interpolation, find f(x) when x = 1.15

Solution

x	1	1.15	2
f(x)	2	f	8

$$\frac{\textit{Change in } Y}{\textit{Change in } X} = \frac{8-2}{2-1} = \frac{f-2}{1.15-1}$$

$$f = 2 + 6x0.15 = 2.9$$

$$\textit{Therefore } f(1.15) = 2.9$$

Example

X^0	50.0^0	50.2^0	50.4^0	50.6^0
Cos X^0	0.6428	0.6401	0.6374	0.6347

Find the cos 50.3^0

Solution

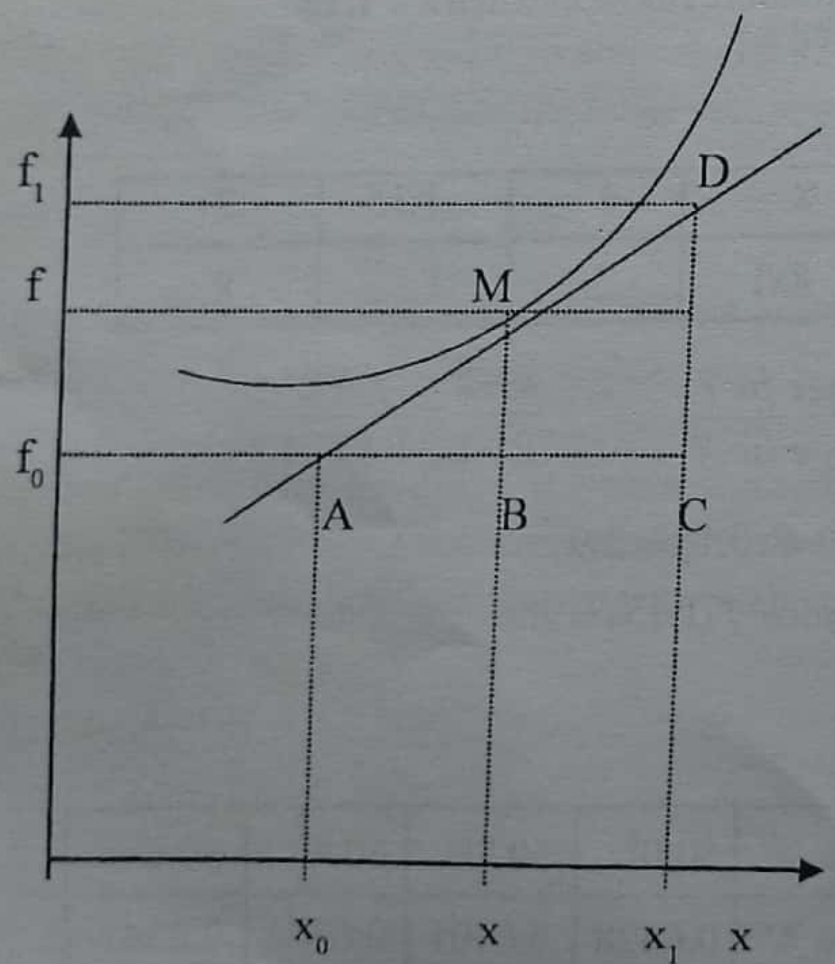
X^0	50.2^0	50.3^0	50.4^0
$\text{Cos } X^0$	0.6401	f	0.6374

$$\frac{0.6374-0.6401}{50.4^0-50.2^0}=\frac{f-0.6401}{50.3^0-50.2^0}$$

$$f=0.6401-0.00135=0.63875$$

$$\therefore \cos 50.3^0=0.6388$$

10.2 Expression for linear interpolation



$f(x)=f \qquad f(x_0)=f_0 \qquad f(x_1)=f_1$

i) Identify the curve.

ii) Draw a straight line for the curve.

iii) Find the gradient of the straight line $=\frac{CD}{AC}$

$$=\frac{f_1-f_0}{x_1-x_0}$$

Gradient AM $=\frac{BM}{AB}=\frac{f-f_0}{x-x_0}$

$$f-f_0=\frac{(x-x_0)}{(x_1-x_0)}(f_1-f_0)$$

let $\frac{x-x_0}{x_1-x_0}=\delta, let(f-f_0)=\Delta f_0$

$$\Rightarrow f=f_0+\delta\Delta f_0$$

Example

Certain curve y = f(x) passes through the points (4, 1.88) and (5, 1.84). Find the value of f (4.2)

Solution

$$\frac{f_1-f_0}{x_1-x_0}=\frac{f-f_0}{x-x_0}$$

X	4	4.2	5
f(x)	1.88	f	1.84

$$\frac{1.84-1.88}{5-4}=\frac{f-1.88}{4.2-4}$$

-0.04 x 0.2 = f − 1.88

1.88 − 0.008 = f

$f(4.2) = 1.872$

Example

Given the following values

X^0	0.85^0	0.86^0	1.85^0	1.86^0
$\sin x^0$	0.7513	0.7578	0.9612	0.9585

Use linear interpolation to estimate;

i) $\sin 0.857^0$

ii) $\sin 1.857^0$

Solution

$x = 0.857^0$

$$\frac{f_1 - f_0}{x_1 - x_0} = \frac{f - f_0}{x - x_0}$$

$$\frac{0.7578 - 0.7513}{0.86 - 0.85} = \frac{f - 0.7513}{0.857 - 0.85}$$

$$\frac{6.5x10^{-3}}{0.01} = \frac{f - 07513}{7 \times 10^{-3}}$$

$\Rightarrow 4.55 \text{ x } 10^{-3} + 7.513 \text{ x } 10^{-3} = 0.01f$

$f = 0.75585$

$\Rightarrow \sin 0.857 = 0.7559$

If $x = 1.857^0$

$$\frac{f_1 - f_0}{x_1 - x_0} = \frac{f - f_0}{x - x_0}$$

$$\frac{0.9585-0.9612}{1.86-1.85}=\frac{f-0.9612}{1.857-1.85}$$

$$\frac{0.0027}{0.01}=\frac{f-0.9612}{7\times10^{-3}}$$

$f=0.95931$

$\Rightarrow \sin 1.857=0.9593$

Example

Use 4 figure tables and linear interpolation to evaluate

(a) $\sqrt{60.535}$

(b) sec 78.38^0

Solution

(a)

$\sqrt{60.535}$

$\sqrt{60.535}=7.78,\quad \sqrt{60.54}=7.781$

x	60.53	60.535	60.54
f(x)	7.780	f	7.781

$$f=f_0+\frac{(x-x_0)}{(x_1-x_0)}(f_1-f_0)$$

$$=7.78+\left(\frac{60.535-60.53}{60.54-60.53}\right)(7.781-7.78)$$

$$= 7.78 + \left(\frac{60.535-60.53}{60.54-60.53}\right)(7.781-7.78)$$

$$= \quad 7.78 + \frac{5\times10^{-3}}{0.01} 1\times10^{-3}$$

$$= 7.7805$$

(b) Sec 78.38^0

x	78.3	78.38	78.4
f(x)	4.9313	f	4.9732

$$f = f_0 + \frac{(x-x_0)}{x_1-x_0}(f_1-f_0)$$

$$\Rightarrow 4.9313 +) \left(\frac{78.38-78.3}{78.4-78.3}\right)(4.9732-4.9313$$

$$4.9313 + \left(\frac{0.08}{0.01}\right)(0.0419)$$

$$4.9313 + 0.03352 = 4.96482$$

$$\sec 78.38^0 = 4.9648$$

10.3 Inverse interpolation

From $(f-f_0) = \frac{(f_1-f_0}{(x_1-x_0}(x-x_0)$

$$(x-x_0) =) \frac{f-f_0}{f_1-f_0}(x_1-x_0$$

$$x = x_0 + \frac{f - f_0}{f_1 - f_0}(x_1 - x_0)$$

$$\left(\frac{x_1 - x_0}{f_1 - f_0}\right) = \left(\frac{x - x_0}{f - f_0}\right)$$

Example

In an experiment the following observations were noted:

T	0	12	20	30
σ	6.6	2.9	-0.1	-2.9

Use interpolation to find T when σ = -1

$$\frac{x_1 - x_0}{f_1 - f_0} = \frac{x - x_0}{f - f_0}$$

$$\frac{30 - 20}{-2.8 - -0.1} = \frac{x - 20}{-1 - -0.1}$$

$$\frac{10}{-2.8} = \frac{x - 20}{-0.9}$$

$-9 = -2.8x + 56$

$-9 - 56 = -2.8x$

$23.21^0 = x = T$

Example

Give the table below

x	-0.11	-0.10	-0.09	-0.08
f(x)	1.1821	1.1800	1.1781	1.1764

Find the value of x such that f(x) = 1.1785

Solution

$$\frac{x_1 - x_0}{f_1 - f_0} = \frac{x - x_0}{f - f_0}$$

$$\frac{-0.10 - 0.09}{1.1800 - 1.1781} = \frac{x + 0.09}{1.1785 - 1.1781}$$

$-4 \times 10^{-6} = 1.09 \times 10^{-3}x + 1.71 \times 10^{-4}$

$$\frac{-0.01}{1.09 \times 10^{-3}} = \frac{x + 0.09}{4 \times 10^{-4}}$$

$-4 \times 10^{-6} = 1.09 \times 10^{-3}x + 1.71 \times 10^{-4}$

x = -0.092

10.4 Extrapolation

Involves approximating the value of a function f(x) for given values of x outside the given tabulated values.

Example

x	x^2
3.35	11.223
3.36	11.280

Find $(3.363)^2$

Using: $\frac{f_0 - f_0}{x_1 - x_0} = \frac{f - f_0}{x - x_0}$

$$\frac{11.290 - 11223}{3.36 - 3.35} = \frac{f - 11.223}{3.363 - 3.35}$$

$$\frac{0.067}{0.01} = \frac{f - 11.223}{0.013}$$

$0.0871 + 11.223 = f$

$11.3101 = f$

$3.363^2 = 11.3101$

ASSIGNMENT

1. In an experiment to measure the rate of cooling of an object, the following temperatures, (θ^0 C) against times, T were recorded:

Temperature, θ^0 C	80	70.2	65.8	61.9	54.2
Time, T (s)	0	10	15	20	30

Use linear interpolation to find

(i) the value of θ when $T = 18$ s

(ii) T when $\theta = 60^0$.

2. In the table below is an extract of part of log x to base 10, $\log_{10} x$:

x	80.00	80.20	80.50	80.80
$\text{Log}_{10} x$	1.9031	1.9042	1.9058	1.9074

Use linear interpolation to estimate:

(a) Log_{10} 80.759,
(b) The number whose logarithm is 1.90388.

3. The bus stages a long Jinja- Kampala road are 10 kms apart. An express bus traveling between the two towns only stops at these stages except in case of emergency when it is permtted to stop at a point in between the two stages. The fares up to the firs, second, third and fourth stages from Jinja are Sh 110, Sh 150, Shs 185 and Sh 200 respectively. On certain day a passenger paid to travel from Jinja in the bus up to the fourth stage but fell sick and had to be left at a health center 33 kms away from Jinja. Given that he was refunded money for part of journey he had not traveled, find the approximate amountt he recieved.
Another person had only Sh 165 was allowed t board the bus but be left at a point worth his money. How far from Jinja was he to be left?

Answer balance Sh 10.5, 24.29 km

4. The table shows the values of function f(x) at a set of points

x	0.9	1.0	1.1	1.2
f(x)	0.266	0.242	0.218	0.192

Using linear interpolation to find
(i) the value of f(1.04)
(ii) the value of x corresponding to f(x) = 0.25

Answer (i) 0.2324 (ii) 0.967

CHAPTER ELEVEN

APPROXIMATE NUMERICAL METHODS

11.1 LOCATION OF ROOTS $f(x) = 0$

Two methods can be used.

i – Graphical
ii – Analytical

Analytical method.

Example

Locate the ranges of root of the equation
$x^3 - x - 1 = 0$

Let $f(x) = x^3 - x - 1 = 0$

x	-3	-2	-1	0	1	2
f(x)	-25	-7	-1	-1	-1	5

The root f(x) of the equation lies between 1 and 2 because
$f(1) < 0$ and $f(2) > 0$

Example

Show that the equation $x^4 - 12x^3 + 12 = 0$ has a root between 1 and 2

$f(x) = x^4 - 12x^3 + 12 = 0$
$f(1) = 1 - 12 \times 1 + 12 = 1$
$f(2) = 16 - 12 \times 8 + 12 = -68$

Since f(1) > 0 and f(2) < 0, implies the root lies between 1 and 2.

Example

Show that the equation $x = \ln(8 - x)$ has a root between 1 and 2

Solution

$f(x) = x - \ln(8 - x)$
$f(1) = 1 - \ln(7) = -0.946$
$f(2) = 2 - \ln(6) = 0.208$

Since $f(1) \times f(2) < 0$, implies $1 < x_r < 2$

GRAPHICAL METHOD

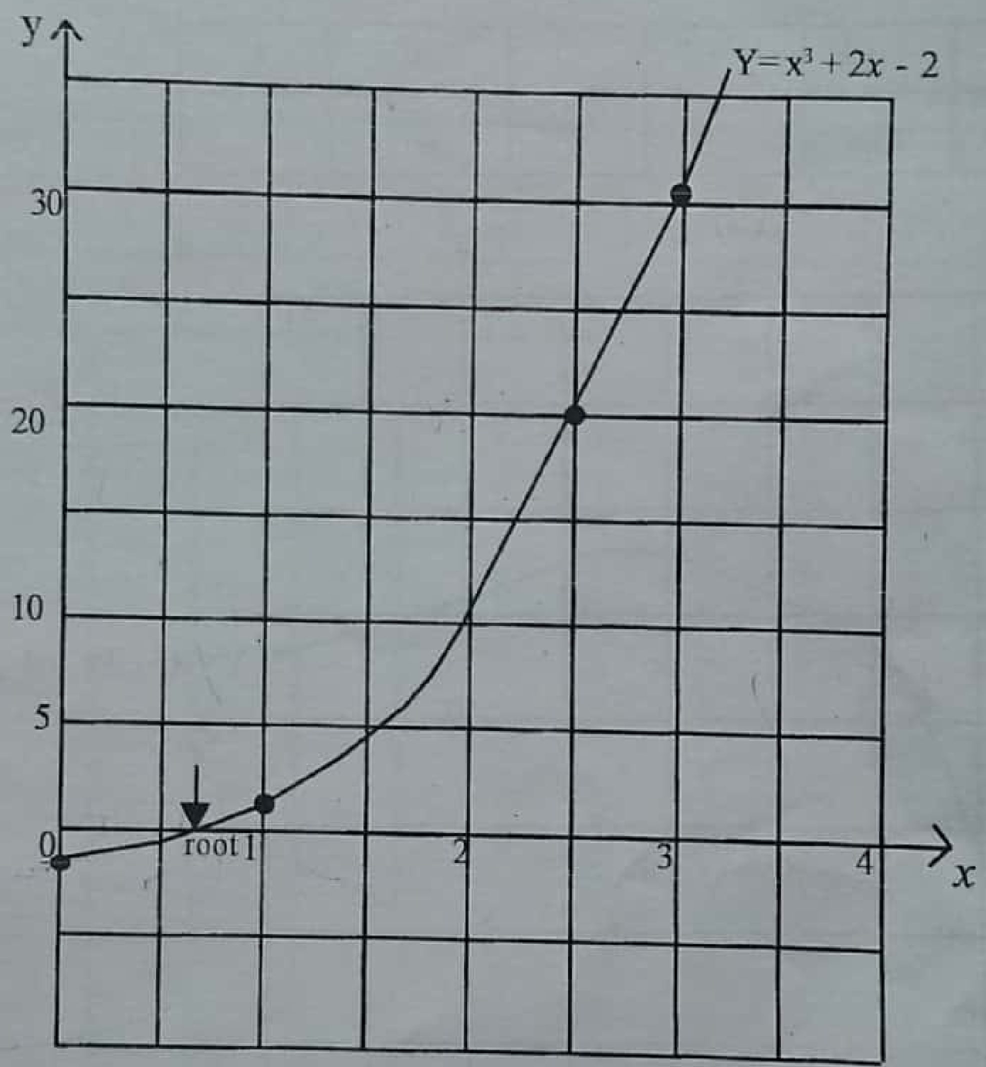
Using the graphical method, single graph could be plotted or two or more graphs could be plotted.

Example

Show graphically that there is **one** positive real root of the equation $x^3 + 2x - 2 = 0$

Solution

x	0	1	2	3
f(x)	-2	1	10	31



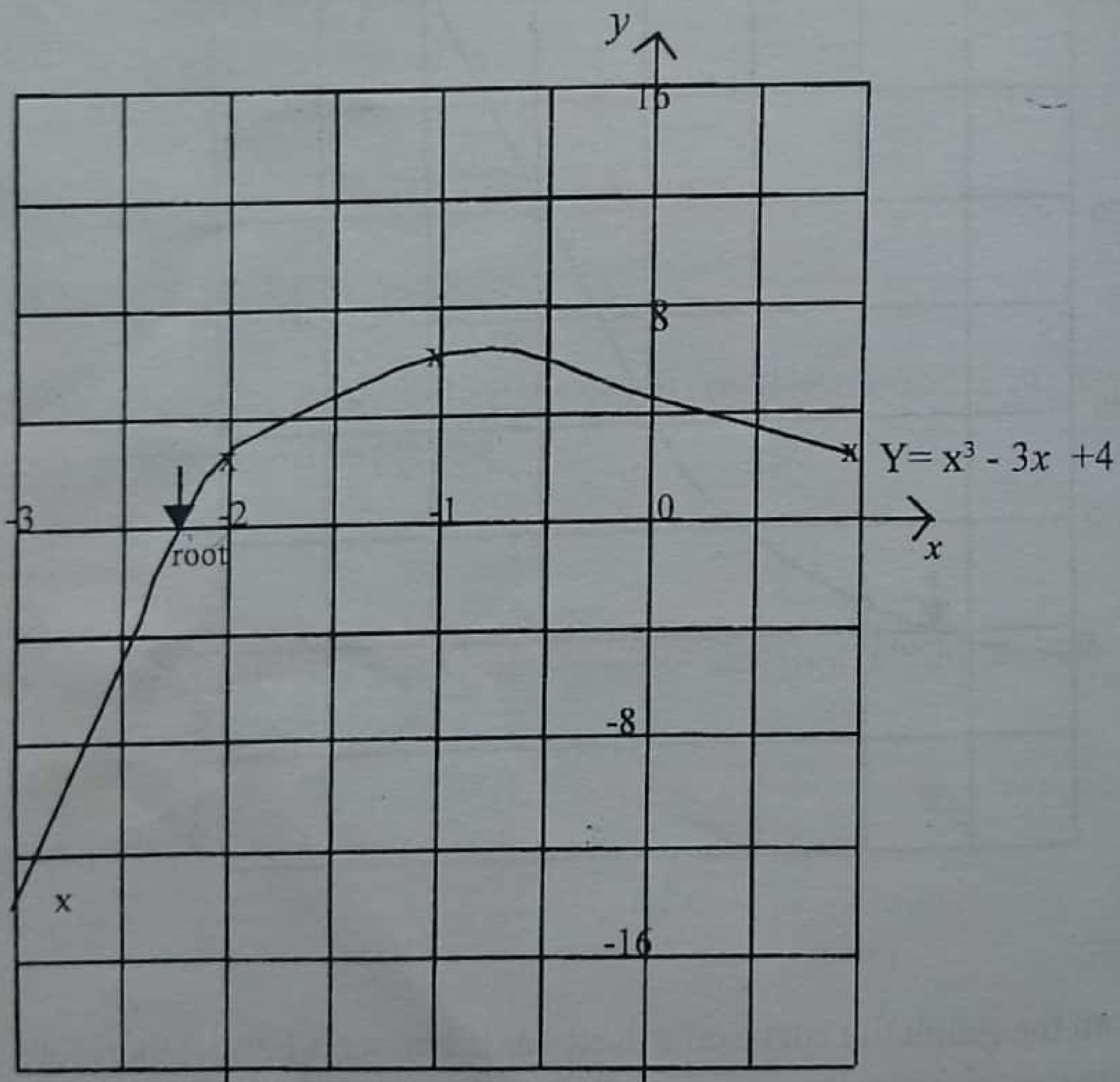
From the graph the curve cuts the x –axis between 1 and 2 therefore there is positive root.

Example

Use a graphical method to find a first approximation to the real root of $x^3 - 3x + 4 = 0$

Solution

x	-3	-2	-1	0	1
f(x)	-14	2	6	4	2



Assignment

Using $f(x) = x^3 - 3x - 3$, show that it's root lies between 2 and 3.

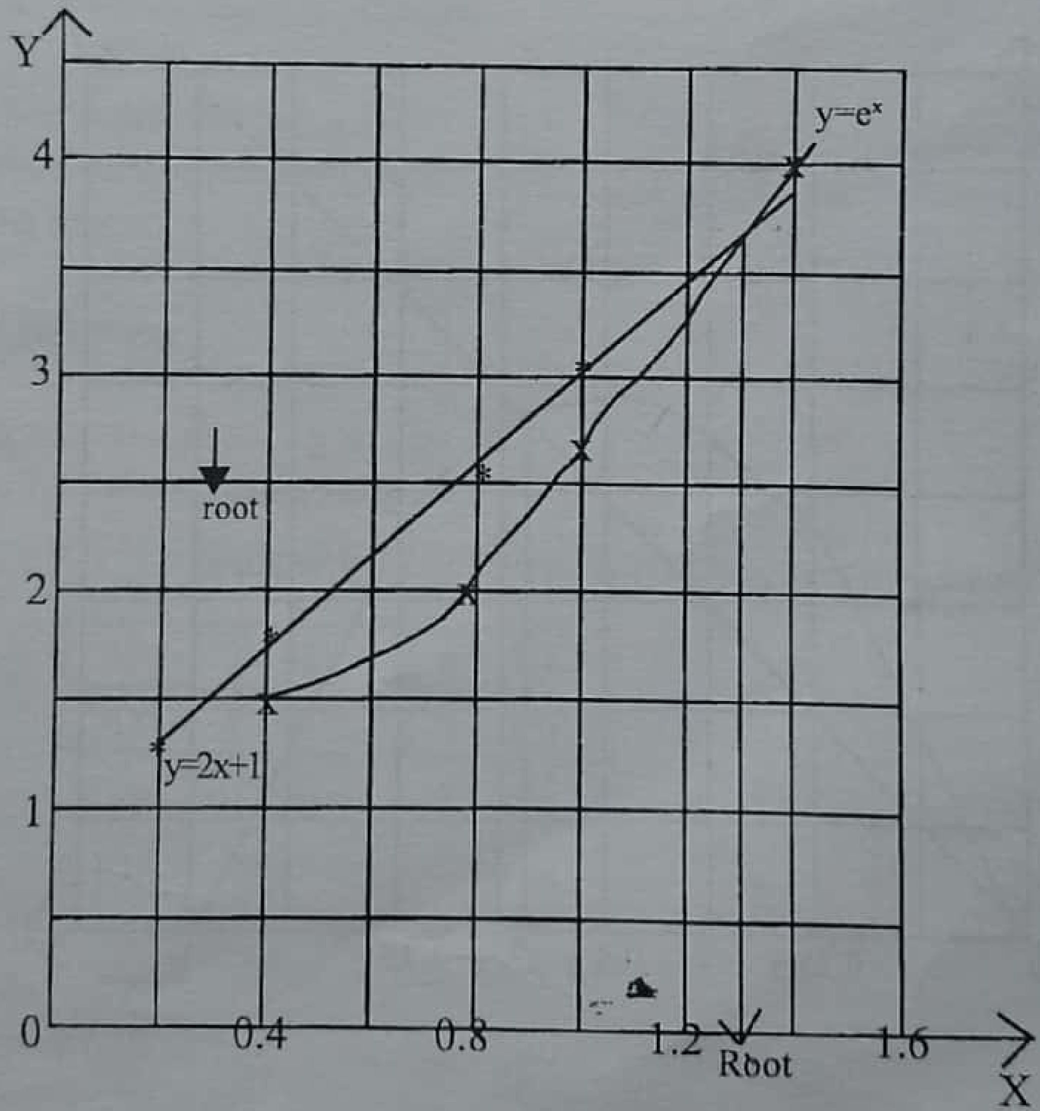
Example

Use a graphical method to show that the equation $e^x + x - 4 = 0$ has only one real root.

$f(x) = e^x - 2x - 1$ $\qquad$ $e^x = 2x + 1$

Let $y_1 = e^x$, $y_2 = 2x + 1$

x	0.2	0.4	0.6	0.8	1.0	1.2	1.4	
y_1	1.2	1.5	1.8	2.2	2.7	3.3	4.1	
y_2	1.4	1.8	2.2	2.6	3.0	3.4	3.8	

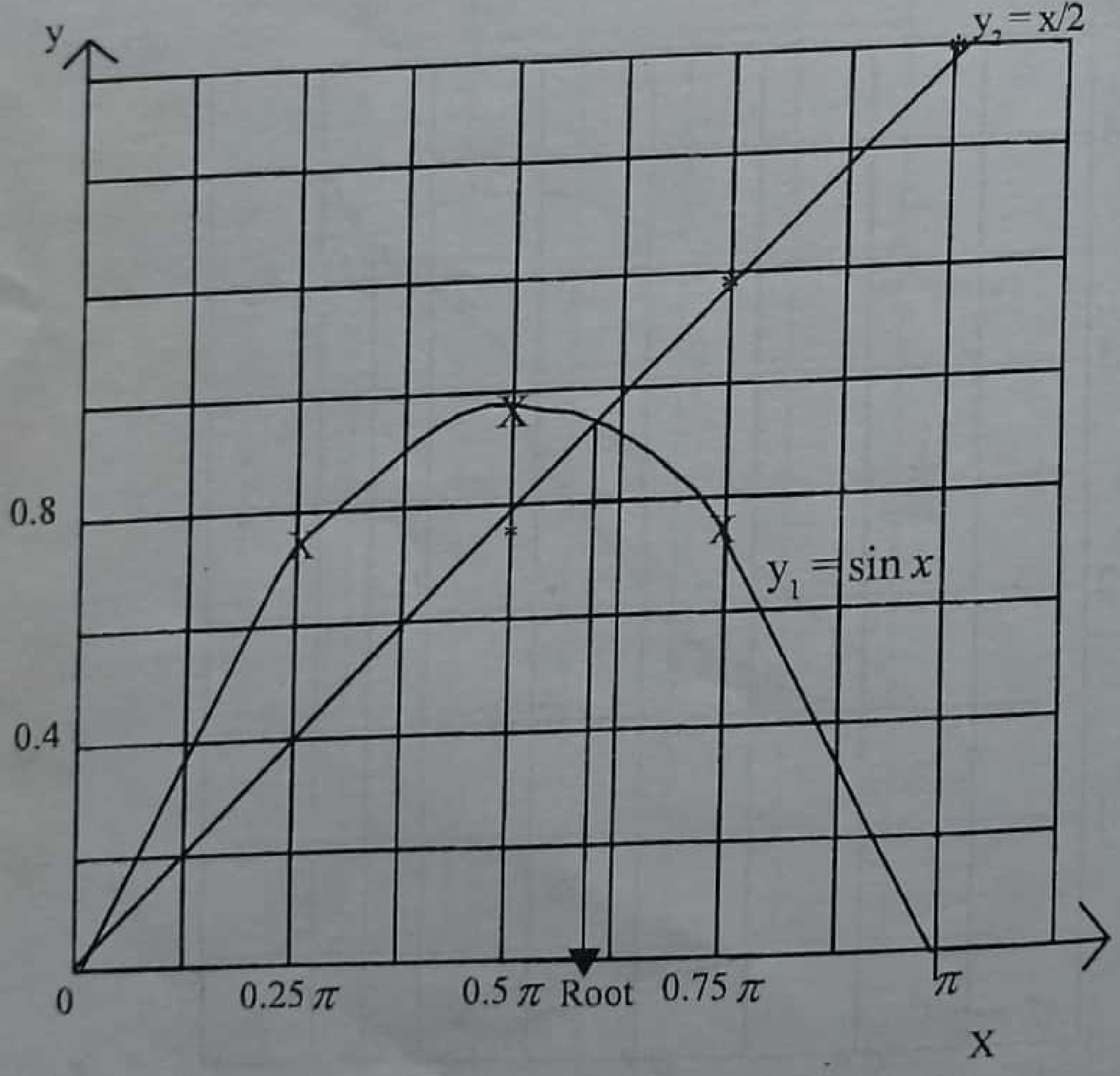


Example

Given the equation $y = \sin x - x/2$, show by plotting suitable graphs on the same axes that the root lies between $\pi/2$ and $3\pi/4$.

$y_1 = \sin x$, $y_2 = x/2$

x	0	$\pi/4$	$\pi/2$	$3\pi/4$	π
y_1	0	0.71	1.00	0.71	0
y_2	0	0.39	0.76	1.18	1.57



Therefore the root lies between $\pi/2$ and $3\pi/4$

11.2 Approximations / Iterations.

The process of finding successive approximations to a quantity is known as an iterative process. Each use of particular formulae is an iteration.

The methods of finding real roots of f(x) = 0 include:-

- Interpolation
- General iteration method
- Newton Raphson method

11.2.1 INTERPOLATION

Show that the equation $x^3 - 3x - 12$ has a root between 2 and 3. Hence use linear interpolation once to get the first approximation to the root.

Solution

$f(2) = 2^3 - 3x2 - 12 = -10$
$f(3) = 3^3 - 3x3 - 12 = 6$
Since $f(2) < 0$ and $f(3) > 0$ implies that root lies between 2 and 3.

X	2	x	3
f(x)	-10	0	6

$$\frac{3-2}{6+10} = \frac{x-2}{10}$$

$16x - 32 = 10$
$16x = 42$
$x_0 = 2.625 \Rightarrow$ first approximation

Example

Show that the equation $3x^2 + x - 5 = 0$ has a real root between $x = 1$ and $x = 2$. Hence use linear interpolation to calculate the root to 2 decimal places.

Solution

$f(1) = 3x\ 1^2 + 1 - 5 = -1$

$f(2) = 3x2^2 + 2 - 5 = 9$

Since $f(1)\,x\,f(2) < 0$, implies $1 < x_r < 2$

Since $x = 2$ gives 9 which is too far from zero, then other values can be used, $x = 1.5$

$f(1.5) = 3x\ 1.5^2 + 1.5 - 5 = 3.25$

x	1	x	1.5
f(x)	-1	0	3.25

$$\frac{1.5-1}{3.25+1} = \frac{x-1}{1}$$

$x = 1.1176$

$f(1.1176) = 3x\ 1.1176^2 + 1.1176 - 5 = -0.1353$

We can use any other x value between 1.1 and 1.5

Using $f(1.2) = f(1.2) = 3x\ 1.2^2 + 1.2 - 5 = 0.52$

x	1.1176	x	1.2
f(x)	-0.1353	0	0.52

$$\frac{1.2-1.1176}{0.52+0.1353}=\frac{x-1.1176}{0.1353}$$

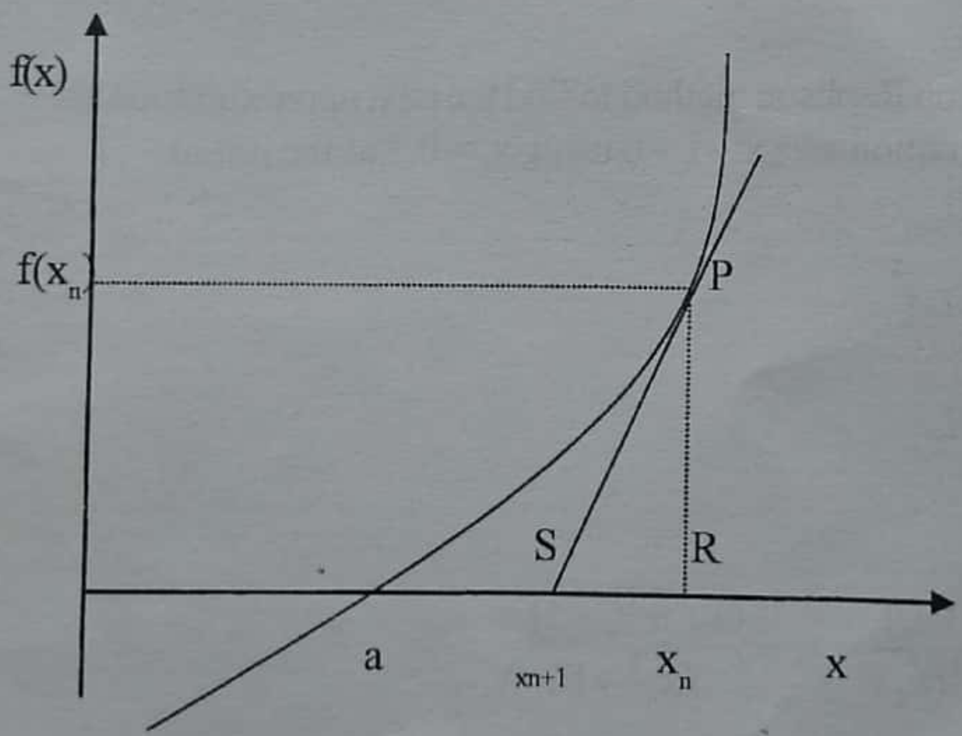
$x = 1.1176 + 0.0170 = 1.1346$

$f(1.1346) = 3x\ 1.1346^2 - 5 = -0.0034$

$x = 1.13$ root correct to 2 d.p

11.2.2 NEWTON RAPHSON METHOD

Expression for Newton Raphson method



From the graph, $f(a) = 0$ if we let x_n be the first approximation

The gradient of point $P = \frac{PR}{SR}$

$$= \frac{f(x_n)}{x_n - (x_{n+1})}$$

From the graph, Grad $P = f^1(x_n)$

$$\Rightarrow \frac{f(x_n)}{x_n - x_{n+1}} = f'(x_n)$$

$$x_n - x_{n+1} = \frac{f(x_n)}{f'(x_n)}$$

$$\therefore x_{n+1} = x_n - \frac{f(x_n)}{f'(x_n)}$$

Example

Use the Newton Raphson method to find the next approximation to root of the equation $x^3 + x - 1 = 0$ using $x_0 = 0.5$ as the initial approximation

$f(x) = x^3 + x - 1$

$f^1(x) = 3x^2 + 1$

$x_0 = 0.5$

$$x_{n+1} = x_n - \frac{f(x_n)}{f^1(x_n)} = x_n - \frac{(x_n^3 + x_n - 1)}{3x_n^2 + 1}$$

$$x_{n+1} = \frac{2x^2_n + 1}{3x^2_n + 1}$$

note $x_0 = 0.5$

$$\frac{2x_0^2+1}{3x_0^2+1} = \frac{2x(0.5)^2+1}{3x(0.5)^2+1}$$

$x_1 = \underline{0.7143}$

Example

Use the N.R.M to show the cube root of a number N is given as

$$\frac{1}{3}(2x_n + \frac{N}{x_n^{\ 2}})$$

Solution

$x = \sqrt[3]{N} \Rightarrow x^3 = N$

$x^3 - N = 0 \Rightarrow f(x) = x^3 - N = 0$

using N.R.M

$$x_{n+1} = x_n - \frac{f(x_n)}{f^1(x_n)}$$

$f(x) = x^3 - N \Rightarrow f^1(x) = 3x^2$

$$x_{n+1} = x_n - \frac{(x_n^3 - N)}{3x_n^2}$$

$$= \frac{3x_n^{\ 3} - x_n^3 + N}{3x_n^{\ 2}} = \frac{2x_n^3 + N}{3x_n^{\ 2}}$$

$$\frac{2x_n^2 + N}{3x_n^2} = \frac{1}{3}\left(2x_n + \frac{N}{x_n^2}\right)$$

Example

Using the interactive formula, show that the fourth root of the number N is $\frac{3}{4}x_n + \frac{N}{4x_n^3}$

(b) and hence show that

$(45.7)^{1/4} = 2.600$ (correct to 3 decimal places)

Solution

$x = \sqrt[4]{N} \Rightarrow x^4 = N$

$x^4 - N = 0 \Rightarrow f(x) = x^4 - N = 0$

Using NRM; $x_{n+1} = x_n - \dfrac{f(x_n)}{f^1(x_n)}$

$f(x) = x^4 - N \Rightarrow f^1(x) = 4x^3$

$$x_{n+1} = x_n - \frac{(x_n^4 - N)}{4x_n^3}$$

$$= \frac{4x_n^4 - x_n^4 + N}{4x_n^3}$$

$$= \frac{3x_n^4 + N}{4x_n^3}$$

$$= \frac{3x_n^4 + N}{4x_n^3} = \frac{3}{4}x_n + \frac{N}{4x_n^3}$$

b) $x = \sqrt[4]{45.7}$

let $x_o = 2.5$, $N = 45.7$

$$x_{n+1} = \frac{3}{4}x_n + \frac{N}{4x_n^3}$$

Substituting

$$\Rightarrow \frac{3}{4}\times 2.5 + \frac{45.7}{4(2.5)^3} = 1.875 + 0.7312$$

$x_1 = 2.6062$

$$x_2 = 2.6062 \text{ x } \frac{3}{4} + \frac{45.7}{4(2.6062)^3} = 1.95465 + 0.64540$$

$$= 2.60006$$

$x_2 \cong 2.6001$

$\Rightarrow \sqrt[4]{45.7} \cong 2.600$ (3 decimal places)

Assignment

Show that the Newton – Raphson formula for for approximating the K^{th} root of a number N is given by:

$$x_{n+1} = \frac{1}{k}\left[(K-1)x_n + \frac{N}{x_n^{K-1}}\right]$$

(ii) Use your formula to find the positive square root of 67 correct to **four** significant figures

Example

Derive the simplest interactive formula for the NRM for the root of the equation $e^{3x} - 3 = 0$. Using your formula with $x_0 = \frac{1}{3}$, find the root correct to 4 decimal places and hence find and log e^3 correct to 4 decimal places.

Solution.

$f(x) = e^{3x} - 3 = 0$

$f^1(x) = 3e^{3x} \Rightarrow$

Using $x_{n+1} = x_n - \frac{f(x_n)}{f^1(x_n)}$

$$= x_n - \frac{(e^{3x_n} - 3)}{3e^{x_n}}$$

$$x_{n+1} = \frac{(3x_n e^{3x_n} - e^{3x_n} + 3)}{3e^{3x_n}}$$

$$= \frac{e^{3x_n}(3x_n - 1) + 3}{3e^{3x_n}}$$

$$x_0 = \frac{1}{3} = x_n$$

$$x_1 = \frac{e^{3x\frac{1}{3}}(3 \times \frac{1}{3} - 1) + 3}{3e^{3x\frac{1}{3}}}$$

$$x_1 = \frac{e(0)+3}{3e} = \frac{3}{3e} = \frac{1}{e} = 0.36787944$$

$$x_2 = \frac{e^{3\times 0.36787944}(3\times 0.36787944-1)+3}{3e^{3\times 0.36787944}}$$

$$= \frac{3.0151(0.103638+3}{9.045345} = 0.36621$$

$$x_3 = \frac{e^{3\times 0.36621}(3\times 0.36621-)+3}{3e^{3\times 0.36621}}$$

$$= \frac{3.0000(0.0986)+3}{9} = \frac{3.2958}{9}$$

$= 0.36620$

$\Rightarrow$ Root is 0.3662

Hence find log e^3

Let log $e^3 = x \quad \Rightarrow e^x = 3$

$x \ln e = \ln 3 \quad \Rightarrow x = \dfrac{\ln 3}{\ln e}$

$e^{3x} - 3 = 0 \quad \Rightarrow 3x \ln e = \ln 3$

$\Rightarrow$ In $3 = 3x$, but $x = 0.3662$

$\Rightarrow$ In $3 = 3 \times 0.3662 = 1.0986$

11.2.3 THE GENERAL ITERATIVE METHOD

If f(x) = 0, then let x = g(x)

Now x can be expressed as the root to f(x)

Example

Given $f(x)=x^3-3x-12=0$, generate equations in the form of $x_{n+1}=g(x_n)$ that can be used to solve the equation f(x) = 0.

Solution

$$f(x)=x^3-3x-12=0$$

$$3x=x^3-12$$

$$\Rightarrow x=\frac{x^3-12}{3}=g(x)$$

$$\therefore x_{n+1}=\frac{x_n^{\ 3}-12}{3} \quad ....(1)$$

$$x^3=3x+12$$

$$x=\sqrt[3]{(3x+12)}$$

$$\therefore x_{n+1}=\sqrt[3]{(3x_n+12)} \quad ....(2)$$

$$x^3=3x+12$$

$$x^3-3x=12$$

$$x\left(x^2-3\right)=12$$

$$x=\frac{12}{x^2-3}$$

$$\therefore x_{n+1}=\frac{12}{x_n^{\ 2}-3} \quad ....(3)$$

$$x^3 = 3x + 12$$

$$x(x^2 - 3) = 12$$

$$(x^2 - 3) = \frac{12}{x}$$

$$x = \sqrt{\left(3 + \frac{12}{x}\right)}$$

$$\therefore x_{n+1} = \sqrt{\left(3 + \frac{12}{x_n}\right)} \quad \text{..........(4)}$$

$$x^3 = 3x + 12$$

$$x = \frac{3}{x} + \frac{12}{x^2}$$

$$\therefore x_{n+1} = \frac{3x_n + 12}{x_n{}^2} \quad \text{.........(5)}$$

Example

Show that the iteratve formula for solving the equation

$$x^3 - x - 1 = 0 \;\; is \;\; x_{n+1} = \sqrt{\left(1 + \frac{1}{x_n}\right)}$$

Hence find the root of the equation correct to 3 s.fs.

Solution

$$f(x) = x^3 - x - 1 = 0$$

$$divide\ by\ x \Rightarrow \frac{x^3}{x} = \frac{1+x}{x}$$

$$x = \sqrt{\left(1 + \frac{1}{x}\right)}$$

$$\therefore x_{n+1} = \sqrt{\left(1 + \frac{1}{x_n}\right)}$$

Using $x_0 = 1.2$

$$x_1 = \sqrt{\left(1 + \frac{1}{1.2}\right)} = 1.3540$$

$$|x_1 - x_o| = |1.354 - 1.2| = 0.154$$

$$x_2 = \sqrt{\left(1 + \frac{1}{1.354}\right)} = 1.3185$$

$$|x_2 - x_1| = |1.3185 - 1.354| = 0.0355$$

$$x_3 = \sqrt{\left(1 + \frac{1}{1.3185}\right)} = 1.3261$$

$$|x_3 - x_2| = |1.3261 - 1.3185| = 0.0076$$

$$x_4 = \sqrt{\left(1 + \frac{1}{1.3261}\right)} = 1.3244$$

$$|x_4 - x_3| = |1.3244 - 1.3261| = 0.0017$$

$$|x_4 - x_3| = 0.0017 < 0.005$$

$$\therefore root = 1.32$$

11.3 Test for convergence

Using the general iterative method a suitable formula can be rearranged in the form $x_{n+1} = g(x_n)$

In most cases at least one equation can be obtained.

The best equation is the one whose $|g'(x_n)| < 1$, otherwise it tends to diverge.

Example

Given the two iterative formulae

(i) $x_{n+1} = \frac{x_n^{\,3} - 1}{5}$ (ii) $x_{n+1} = \sqrt{\left(5 + \frac{1}{x_n}\right)}$

Using $x_0 = 2$, deduce a more suitable formula for solving the equation.

Hence, find the root correct to 2 decimal places.

Solution

(i) $g(x) = \frac{x^3 - 1}{5}, g^1(x) = 3x^2/5$

$$g'(2) = \frac{3x2^2}{5} = 2.4$$

since $|g'(2)| > 1$ then it will not form a convergent sequence

(ii) $g(x) = \sqrt{\left(5 + \frac{1}{x}\right)} \Rightarrow g'(x) = -\frac{1}{2}x^{-2}\left(5 + \frac{1}{x}\right)^{-0.5}$

$$g'(2) = -\frac{1}{2}x2^{-2}\left(5 + \frac{1}{2}\right)^{-0.5} = -0.0533$$

since $|g^1(2)| < 1$, then the second formula will produce a convergent sequence.

Hence using equation (ii)

$$x_{n+1} = \sqrt{\left(5 + \frac{1}{x_n}\right)} \; but \; x_0 = 2$$

$$x_1 = \sqrt{\left(5 + \frac{1}{2}\right)} = 2.3452, |x_1 - x_0| = 0.3452$$

$$x_2 = \sqrt{\left(5 + \frac{1}{2.3452}\right)} = 2.3295, |x_2 - x_1| = 0.0157$$

$$x_3 = \sqrt{\left(5 + \frac{1}{2.3295}\right)} = 2.3301, \; |x_2 - x_1| = 0.0006$$

Since $0.0006 < 0.005$

Therefore the root $= 2.33$

11.4 Trapezium rule

$$\text{Area} = \frac{1}{2}d(y_1 + y_2) + \frac{1}{2}d(y_2 + y_3) + \ldots + \frac{1}{2}d(y_{n-1} + y_n)$$

$$\text{Area} = \frac{1}{2}d\left[y_1 + 2(y_2 + y_3 + \ldots + y_{n-1}) + y_n\right]$$

$$\int_{x_0}^{x_n} f(x)dx \cong \int_{x_o}^{x_n} \frac{1}{2}h[y_o + 2(y_1 + y_2 + \ldots + y_{n-1}) + y_n]$$

Example

Use the trapezium rule to estimate the area under the curve $\frac{1}{x}$ from $x = 1$ to $x = 2$

x	1.0	1.2	1.4	1.6	1.8	2.0
y	1	0.833	0.7143	0.625	0.556	0.50
y_o	y_1 - y_0	y_2 y_1	y_3 y_2	y_5 y_3	y_5 y_4	y_5

$\int f(x).dx = \frac{1}{2} h[y_0 + 2(y_1 + y_2 + y_3 + y_4) + y_5]$

$$\int_{1.0}^{2.0} \frac{1}{x} dx \cong \frac{1}{2} \times 0.2[1 + 2(0.833 + 0.7143 + 0.625 + 0.556) + 0.5]$$

$$= 0.69572 \cong 0.696$$

Example

Use the trapezium rule with 7 ordinates to estimate

$$\int_0^3 \frac{1}{1+x}\,dx, \text{ correct to 3 decimal places}$$

Solution

x	y_0, y_6	$y_1, \dots y_5$
0	1.0000	
0.5		0.6667
1.0		0.5
1.5		0.4
2.0		0.3333
2.5		0.2857
3.0	0.2500	
Sum	1.25	2.1857

$$\int_0^3 \frac{1}{1+x}\,dx = \frac{1}{2} x 0.5(1.25 + 2x2.1857)$$

$$= 1.40535 \cong 1.405$$

Note: $d = \dfrac{x_n - x_0}{\text{number of sub-intervals}} = \dfrac{3-0}{6} = 0.5$

Sub-intervals, sub-divisions and strips are same
However ordinates = (sub-intervals + 1)

Example

i. Use the trapezium rule to estimate the area of $y = 3^x$ between the x-axis, x = 1 and x = 2 using five strips. Give your answer correct to 4 significant figures.

ii. Find the exact value of $\int_1^2 3^x\,dx$.

iii. Find the percentage error in calculations (i) and (ii) above

Solution

x	y_0, y_5	$y_1, \dots y_4$
1.0	3.0000	
1.2		3.7372
1.4		4.6555
1.6		5.7995
1.8		7.2247
2.0	9.0000	
Sum	12.0000	21.4169

$$\int_1^2 3^x\,dx = \frac{1}{2}x0.2(12+2x21.4169)$$

$$=5.48338 \cong 5.483$$

(ii) $\int_1^2 3^x\,dx = \left[\frac{3^x}{\ln 3}\right]_1^2 = 5.461$

(iii) percentage error =

$$\frac{|5.483-5.461|}{5.461}x100 = 0.401$$

Example

Use the trapezium rule with six sub-intervals to estimate

$\int_0^{\pi} x\sin x\,dx$ correct to 2 decimal places. Determine the error in your estimation and suggest how this error may be reduced.

Solution

x	0	π/6	2π/6	3π/6	4π/6	5π/6	π
x sinx	0	0.262	0.907	1.571	1.814	1.309	0

$$=\frac{1}{2}x\frac{\pi}{6}(0+2(0.262+0.907+1.571+1.814+1.309))$$

$$=0.5236 \text{ x } 5.862$$

$$=3.069$$

$$=3.07$$

$$\int_0^{\pi} x\sin x\,dx=[-x\cos x]_0^{\pi}+\int_0^{\pi}\cos x\,dx$$

$$=\pi+\sin x=\cong 3.14$$

$$\text{error}=3.14-3.07=0.07$$

Increasing the number of sub intervals can reduce the error.

Exercise

1. An iterative formula for solving an equation is given by:

$$x_{n+1} = \sqrt[3]{(3x_n + 3)} \quad \text{for n} = 0, 1, 2, ..$$

i) Find the equation whose root is being sought

ii) Show that the equation has one root and using an appropriate starting value, find the root correct to 2 decimal places.

Answer 2.10

2. By drawing a graph, show that the root of the equation 2tanx = 3x lies between $\frac{\pi}{6}$ and $\frac{\pi}{3}$. Hence by using Newton Raphson's method find the root of the equation to 2 decimal places.

Answer 0.97

3. Draw using the same axes, the graphs of $y = x^2$ and $y = \sin 2x$ for $0 \le x \le \frac{\pi}{2}$. From your graph obtain to one decimal place an approximation of the non-zero root of the equation $x^2 - \sin 2x = 0$. Using Newton's method, calculate to 2 decimal places a more suitable approximation.

Answer 0.97

4. Show that the equation $3x^3 + x - 5$ has a real root between x = 1 and x = 2.

(i) Using linear interpolation find the first approximation for this root to two decimal places.

(ii) Using Newton Raphson formula twice, find the value of this root correct to 2 decimal places.

Answer (i) 1.13 (ii) 1.10

5. By sketching the graphs of 2x and tanx show that the equation 2x = tanx has only one root between x = 1.1 and 1.2. Use linear interpolation to find the value of the root correct to 2 decimal places.

Answer 1.17

6. Show graphically that there is only one positive real root of the equation; $xe^{-x} - 2x + 5 = 0$. Using the Newton Raphson method, find this root correct to 1 decimal place.

Answer 2.6

7. (a) Show that the Newton Raphson's formular for finding the smallest positive root of the equation

$3\tan + x = 0$ is $\dfrac{6x_n - 3\sin 2x_n}{6 + 2\cos^2 x_n}$

(b) By sketching the graphs of $y = \tan x$, $y = \dfrac{-x}{3}$ or otherwise, find the first approximation to the required root and use it to find the actual root correct to 3 decimal places.

(hint work in radians)

Answer 2.456

8. Find an approximation value of $\int_0^{0.5} 1 - x^2 dx$ by using the trapezoidal rule with intervals of 0.1. Show by integration that the magnitude of the error in the approximation is less than 0.001.

Answer 0.4778; 0.4783

9. Show that the equation $x = \cos x - 3$ has a root between -4 and -3. Use an iterative method to calculate this root correct to 3 decimal places.

Answer -3.794

10. Show that the equation $f(x) = x^3 + 3x - 9$ has a root between x=1 and x=2. Using the Newton Raphson formula once, estimate the root of the equation, rounded off to **two** significant figures.

Answer 1.6

11. Use the trapezium rule with 7 ordinates to find the value of

$\int_0^{\pi} \sqrt{(1+\sin x)}\, dx$, *correct to* **two** decimal places

Answer 3.98

12. (a) Use the trapezium rule to estimate the area of $y = 5^{2x}$ between the x-axis, x = 0 and x = 1 using **five** sub-intervals. Give your answer correct to 3 decimal places.

(b) Find the exact value of $\int_0^1 5^{2x}\, dx$.

(c) Find the percentage error in the two calculations in (a) and (b) above

Answer (a) 7.712 (b) 7.4560 (c) 3.43

13 Show that the equation $e^x + x - 4 = 0$ has a real root between 1 and 1.2. Use the Newton-Raphson method to find the root of the equation correct to 3 significant figures.

Answer 1.07 (3 sf)

14. Show graphical that equation $e^x + x - 8 = 0$ has only one real root. Use Newton's Raaphson method to find the approximation of $x = \ln(x-8)$ correct to 3 decimal places

Answer 1.821

15.) Use the trapezium rule to estimate the area of

$y = e^{-2x}$ between the x – axis, x = 1 and x =2, using six ordinates. Give your answer correct to 3 significant fiures.

Answer 0.0593

CHAPTER TWELVE

FLOW DIAGRAM/ CHARTS

12.1 INTRODUCTION

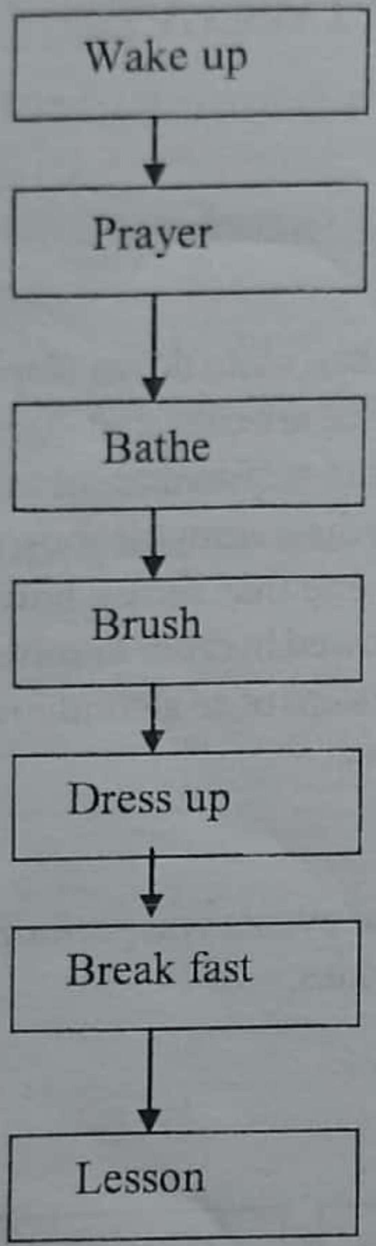
If one has a problem to solve, you can write down steps that he/she needs to solve it. These steps required are called an algorithm. So an algorithm is a step by step procedure required to solve a problem. These steps can be represented in a diagrammatic form using a flow chart. A flow chart or diagram is one that shows how the logical sequence of steps that must be followed in order to solve a problem. Or it is a visual picture that gives the steps of an algorithm and the flow of control between the various steps.

Example

1. Draw a flow chart as a student of events you perform before the 8.00 am lesson using the given activities.

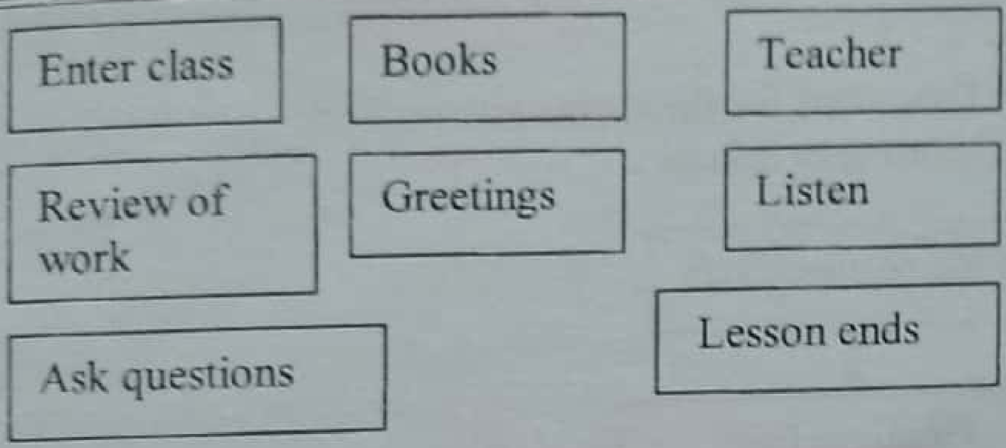
Break fast	Bathe	Dress up
Wake up	Lesson	Prayer
	Brush	

Solution

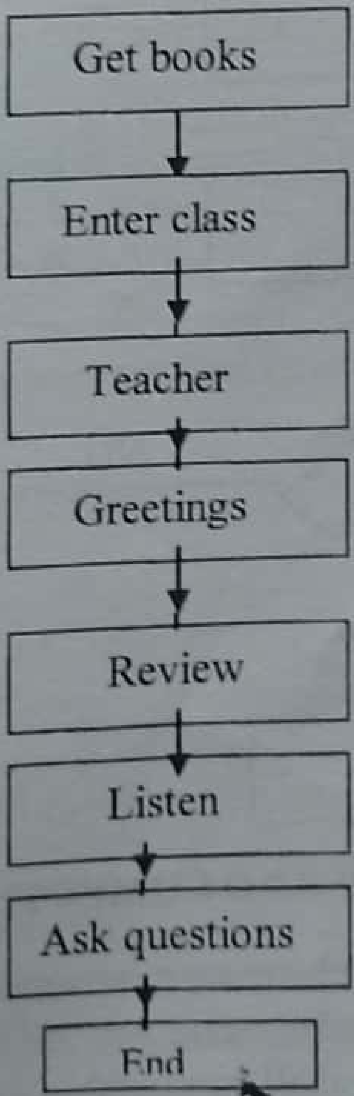


Example

Draw a flow chart as a student of the activities you perform during a Mathematics lesson using the given information.

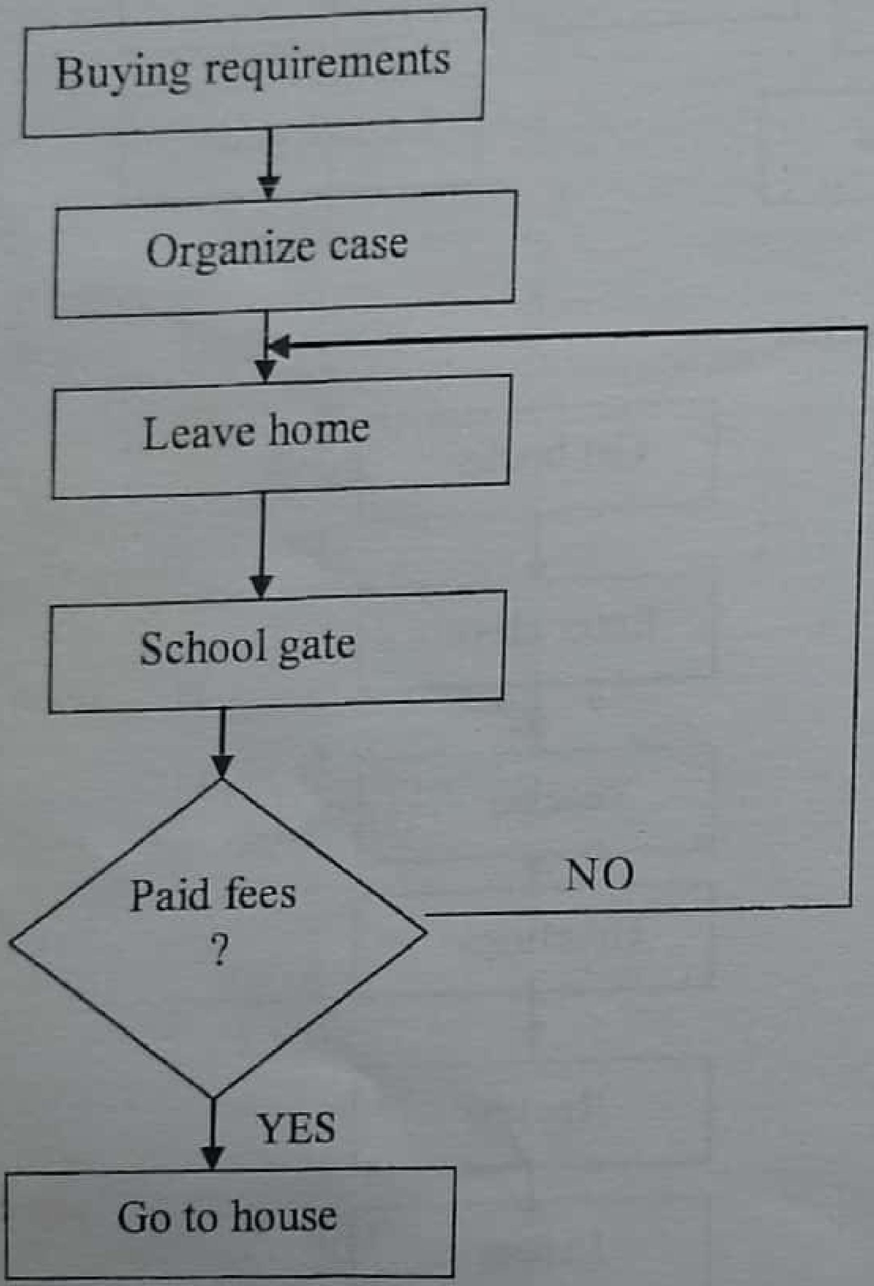

Enter class
Books
Teacher
Review of work
Greetings
Listen
Ask questions
Lesson ends

Solution


Get books
Enter class
Teacher
Greetings
Review
Listen
Ask questions
End

Example

Draw a flow chart for the events you perform when reporting to school

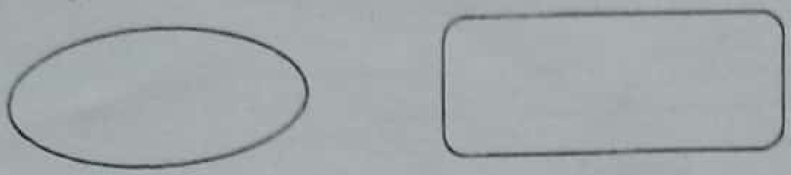


12.2 FLOW CHARTS AND COMPUTER PROGRAMS

Computer based flow charts are more modified for every day life type. They are constructed in such a way that a mathematical problem is broken down into small computable steps. These figures are connected to each other by arrows such that the information flow is followed. In such steps different figures are used.

They include the following

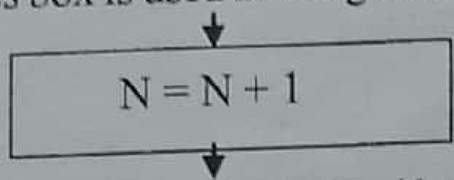
1.



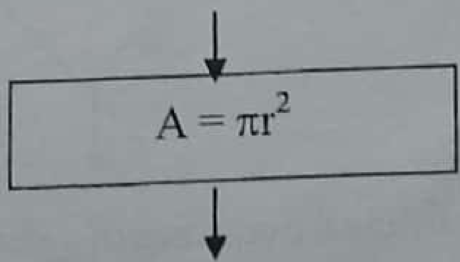
Starts with word **Start/Begin** in first box
Ends with word **Stop/End** in the final box.
Any of the above boxes can be used.

In between the start and end boxes, there are a number of other boxes with different shapes that can be used depending on the nature of the instructions.

2. Rectangles box is used as assignment box



This indicates that the new N is obtained by adding one to the previous N.
So it means let N become N + 1.

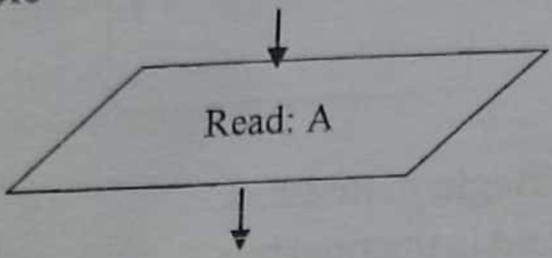


Example

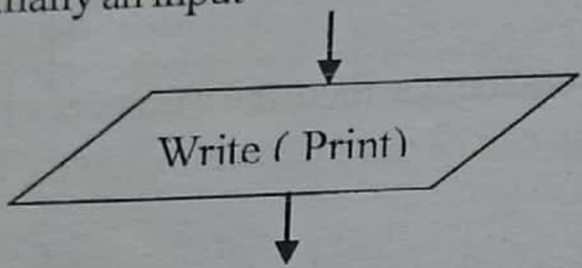
So the box calculates the area of circle. Therefore output will be the area.

3. Other shapes include rhombus, parallelogram
These can be used for information input or output

Example

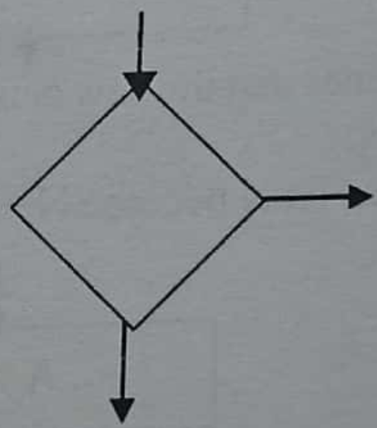


Read is normally an input



Print or Write is an output box

Diamond box is a conditional box

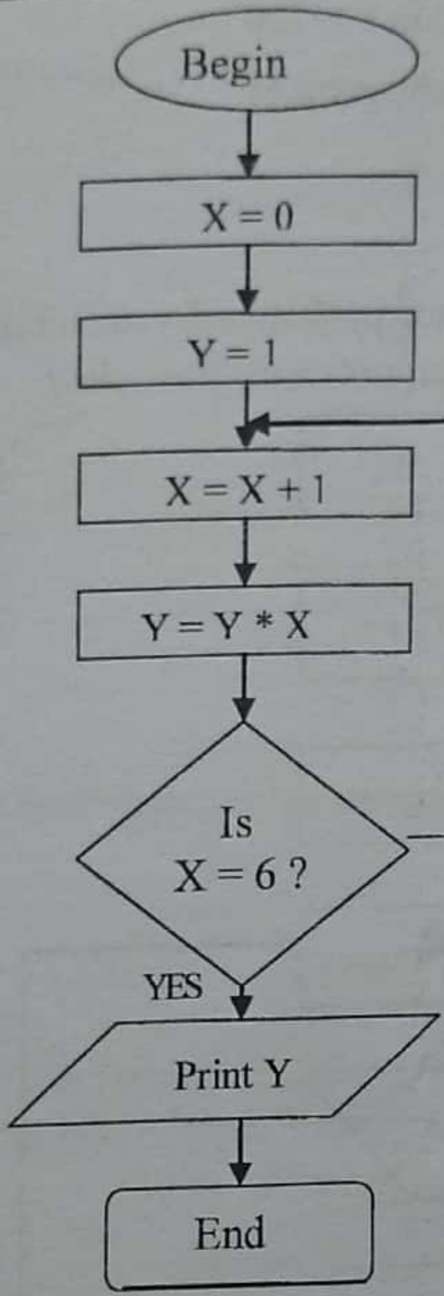


12.3 Dry run

This is the method for predicting the outcome of a given flow chart that produces a table of contents that can be continually updated

Example

Perform a dry run and state the purpose of the flow chart.



Solution

X	Y
0	1
1	1
2	2
3	6
4	24
5	120
6	**720**

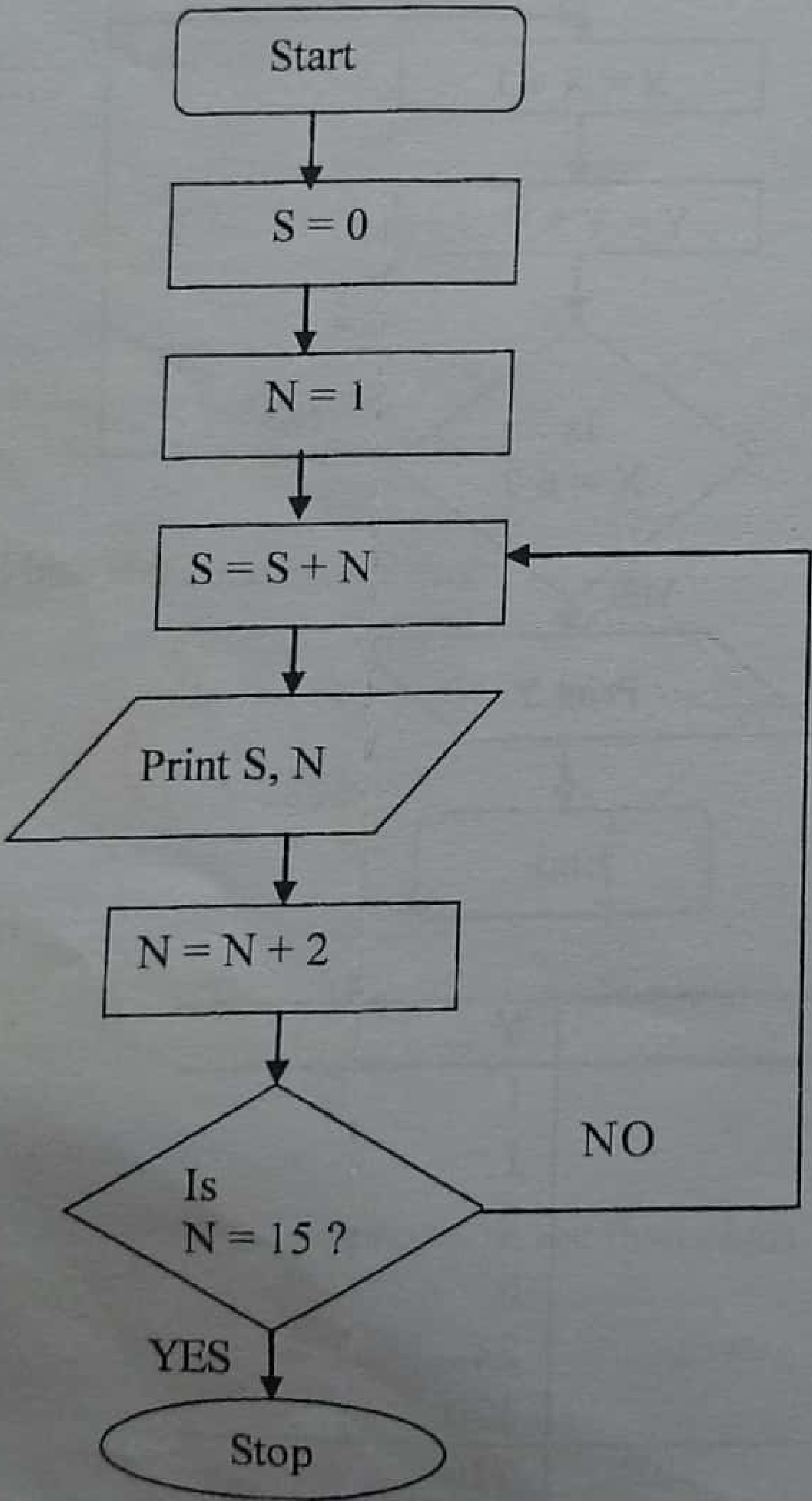
Therefore the print out is Y = 720

Purpose is to compute and print 6!

Relationship is Y = X!

Example

Study the flow chart given and perform a dry run for the flow chart. Hence state the purpose of the flow chart



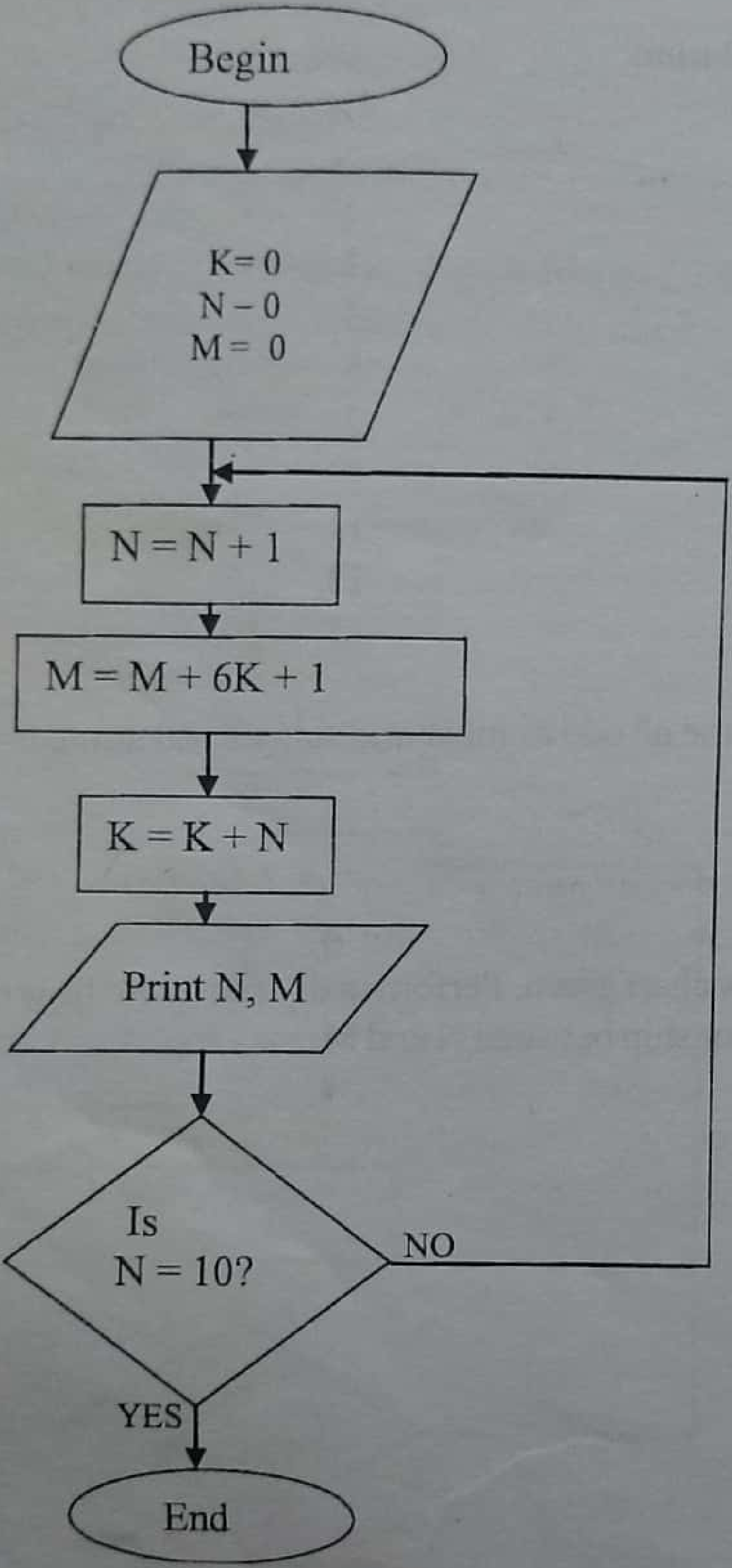
Solution

S	N
0	1
1	1
4	3
9	5
16	7
25	9
36	11
49	13
-	15

Prints the n^{th} odd number and sum of odd numbers up to 7^{th}

Example

Study the flow chart given. Perform a dry run of the flow chart and state the relationship between N and M.

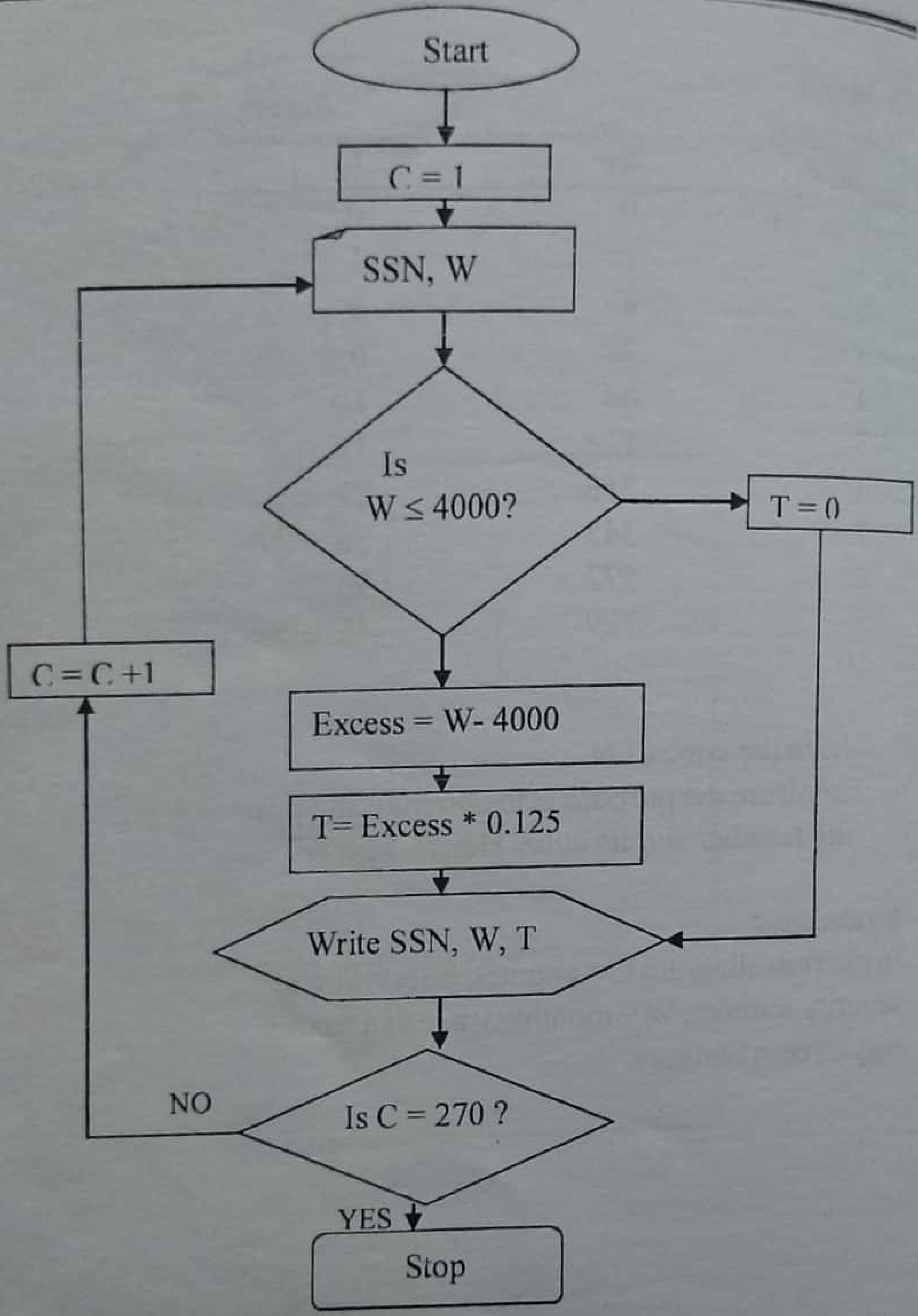

Begin
K= 0
N – 0
M = 0
N = N + 1
M = M + 6K + 1
K = K + N
Print N, M
Is
N = 10?
NO
YES
End

Solution

N	M	K
0	0	0
1	1	1
2	8	3
3	27	6
4	64	10
5	125	15
6	216	21
7	343	28
8	572	36
9	729	45
10		

M is the cube of N
Therefore the purpose is to compute and print the number and its cube. Hence $M = N^3$

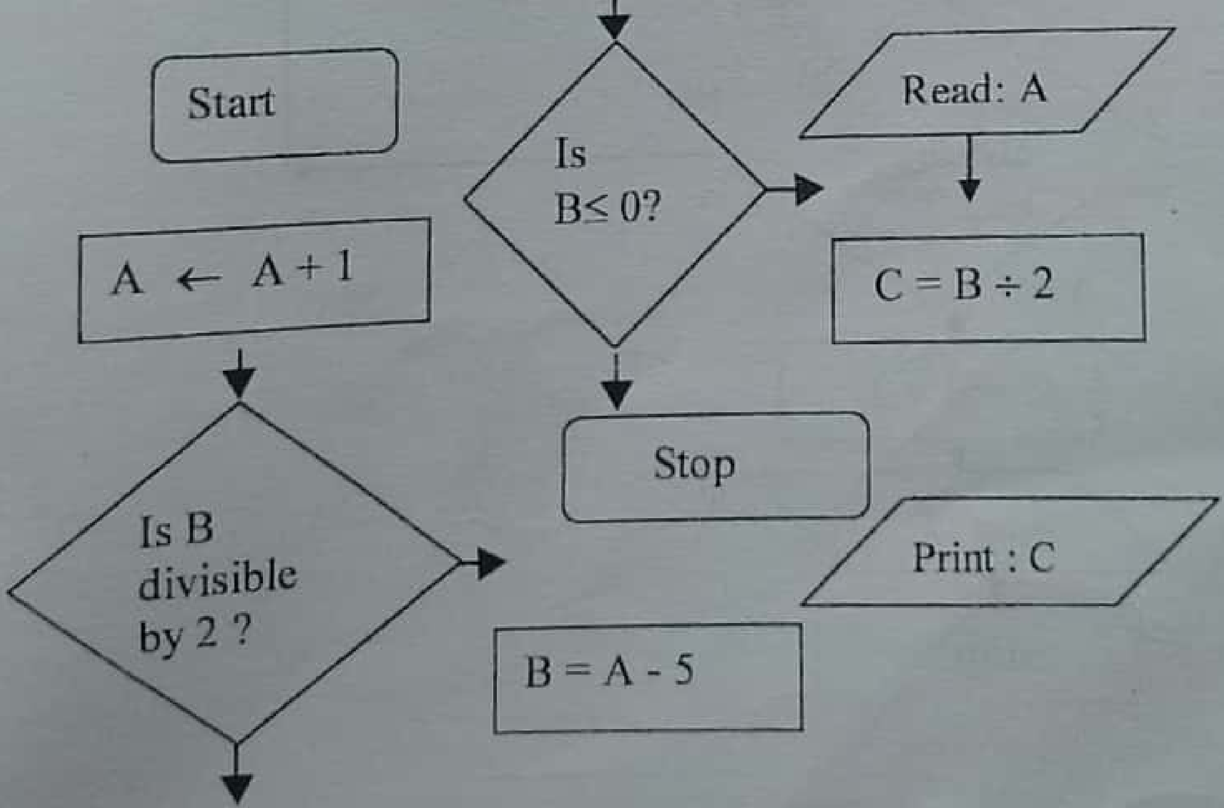
Example 4
In the flow diagram C = counter, SSN = Social security number, W = monthly wage of a factory employee in Shillings.


Start
C = 1
SSN, W
Is
W ≤ 4000?
T = 0
C = C +1
Excess = W- 4000
T= Excess * 0.125
Write SSN, W, T
NO
Is C = 270 ?
YES
Stop

Copy and complete the following table.
For how many employees is this program designed?

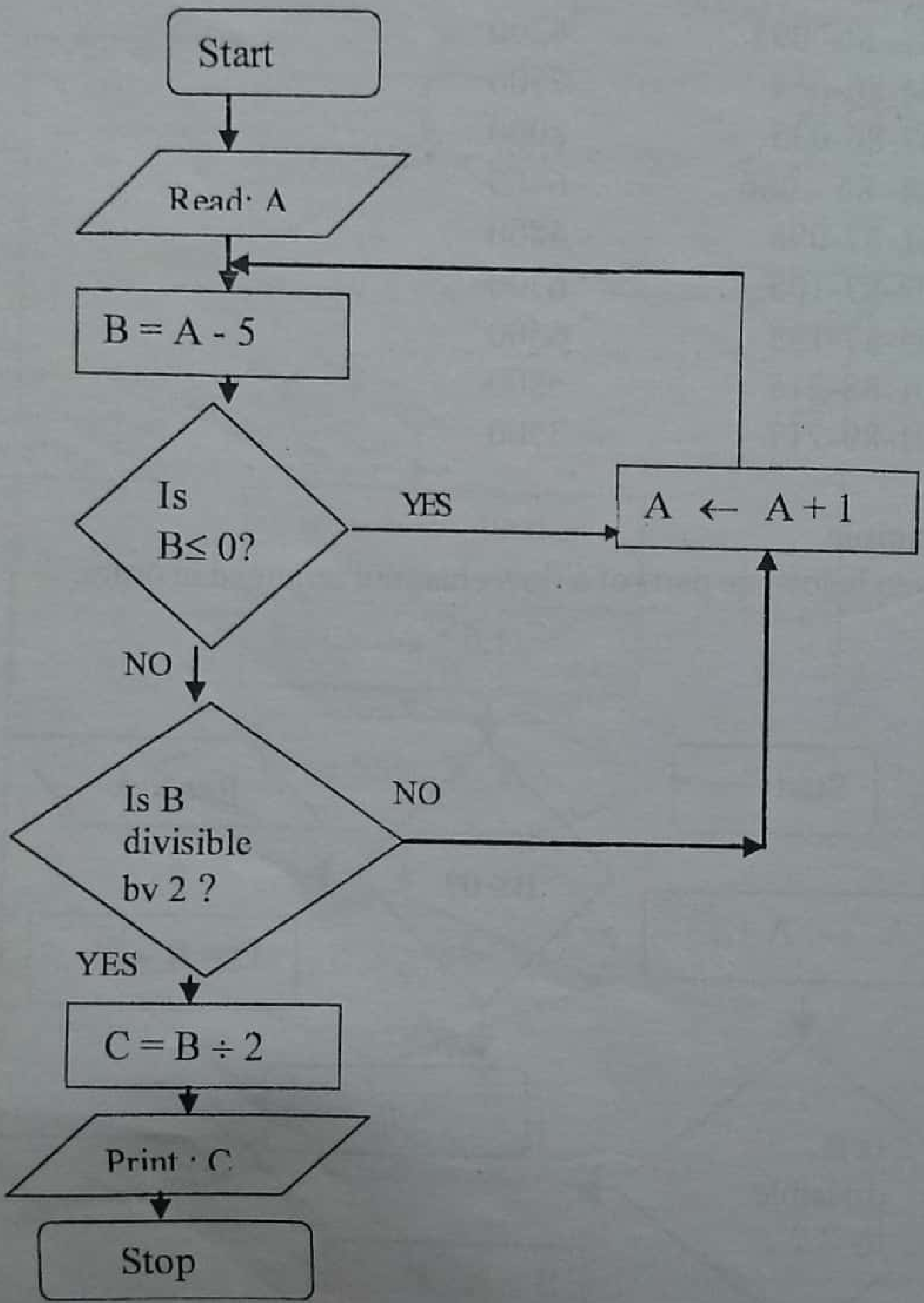
SSN	W	T
01-86-003	8400	
03- 86- 095	8200	
04-86-064	7500	
02-86-035	8000	
04- 86 - 066	6400	
01-87-098	4800	
02-87-105	6300	
03-87-135	5500	
01-88-215	3800	
01-89-217	3500	

Example
Given below are parts of a flow chart not arranged in order.



Rearrange them and draw a complete logical flow chart

Solution



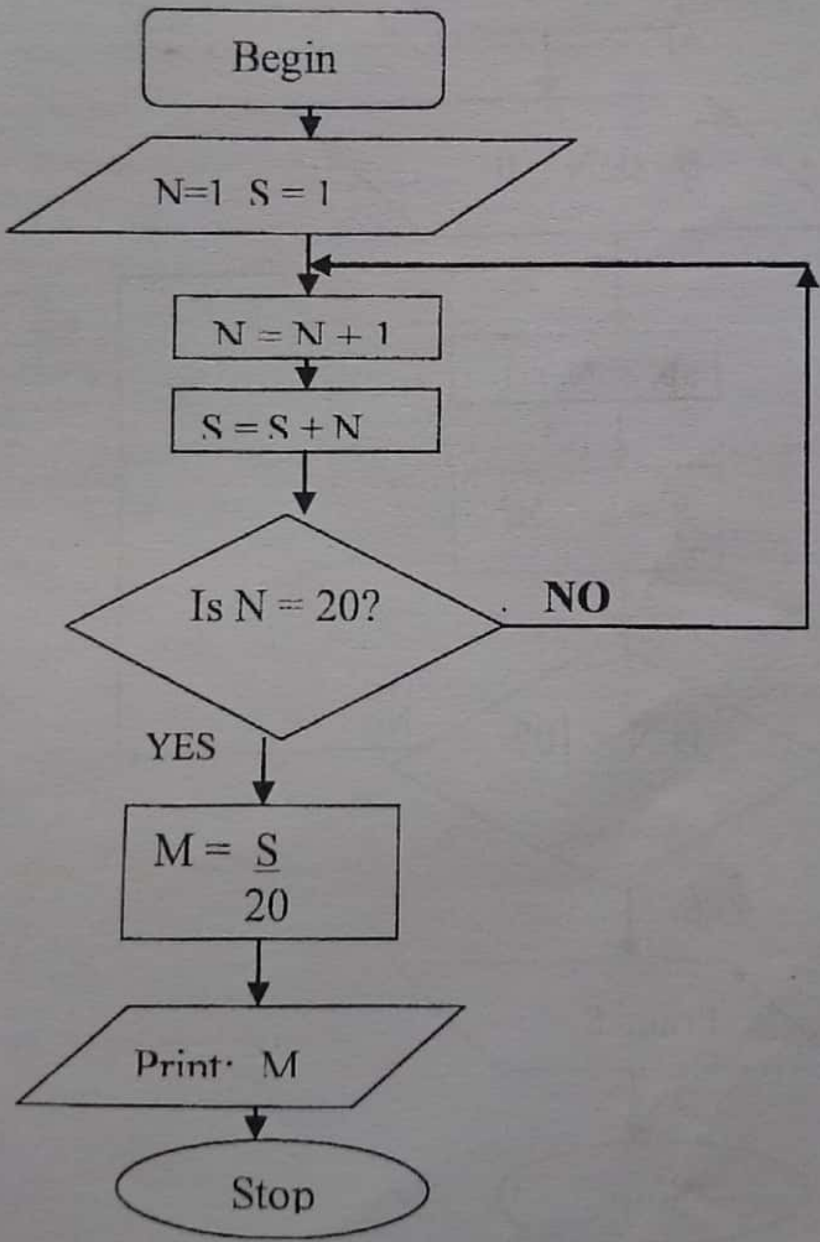
12.4 Constructing flow charts

Example

Draw a flow chart that reads and prints the mean of the first twenty counting numbers.

Solution

Let S represent sum and M the mean

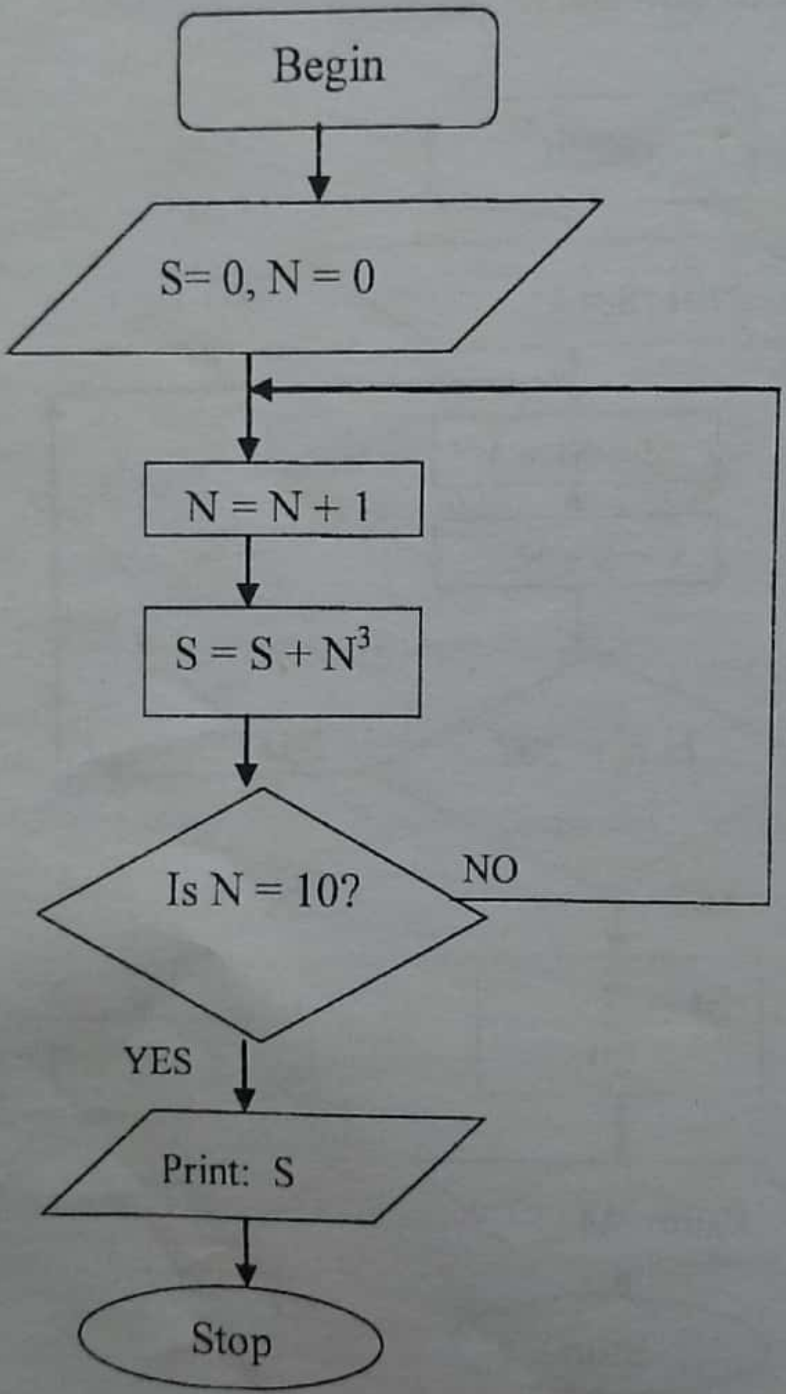


Example

Draw a flow chart that computes and prints the sum of the cubes of the first ten natural numbers.

Solution

Let S represent sum



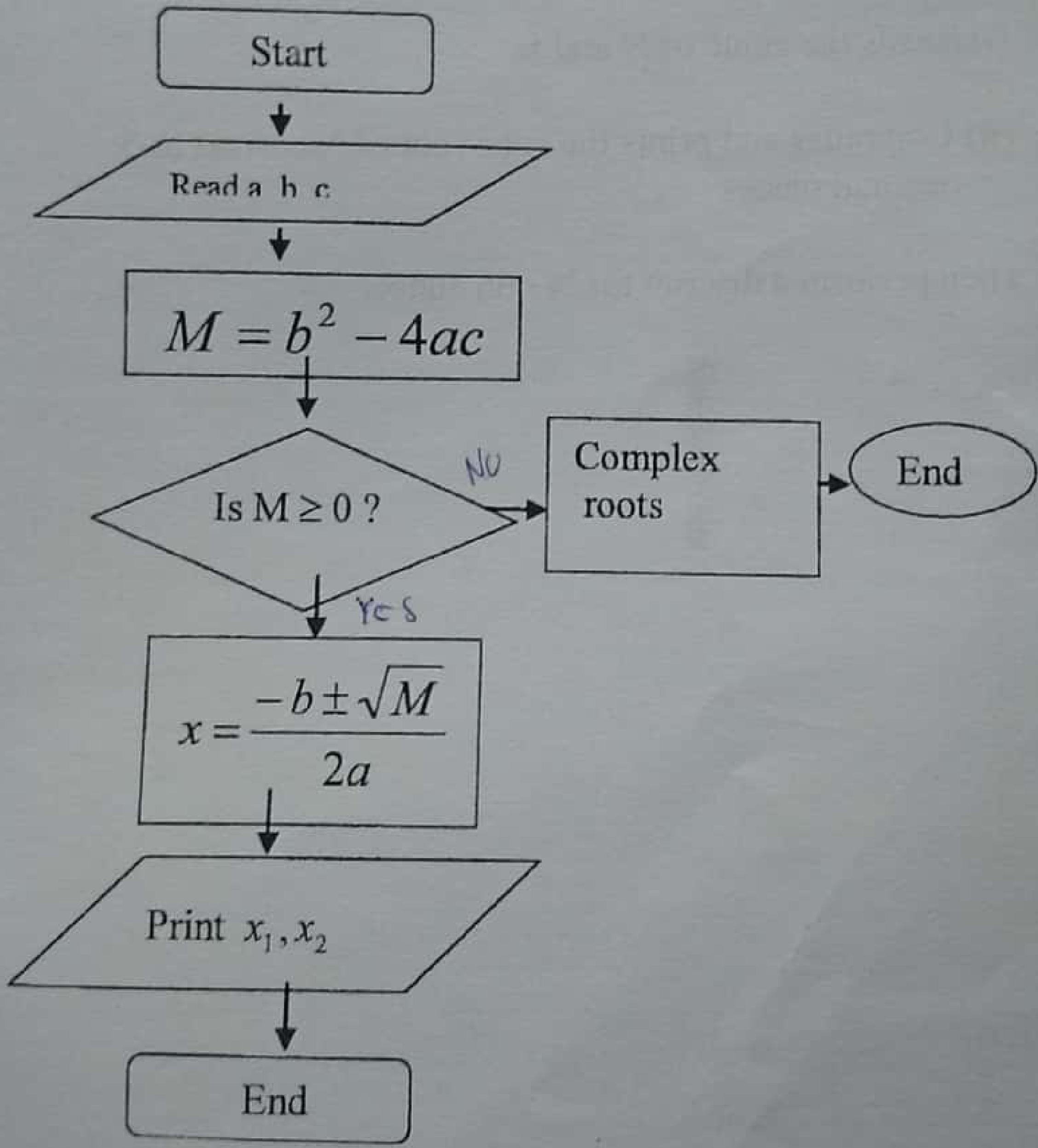
Example

Draw a flow chart that computes the real root of the equation ax^2+bx+c, where a $\neq 0$

Solution

$x=\dfrac{-b\pm\sqrt{b^2-4ac}}{2a}$ for real roots, $b^2-4ac\geq 0$

$M=b^2-4ac \quad \Rightarrow M\geq 0$



Example 9

An iterative method of finding the cube root of a number N , is given by

$$x_{n+1} = \frac{1}{3}(2x_n + \frac{N}{x_n^{\ 2}}); \ n = 0, 1, 2 \ldots$$

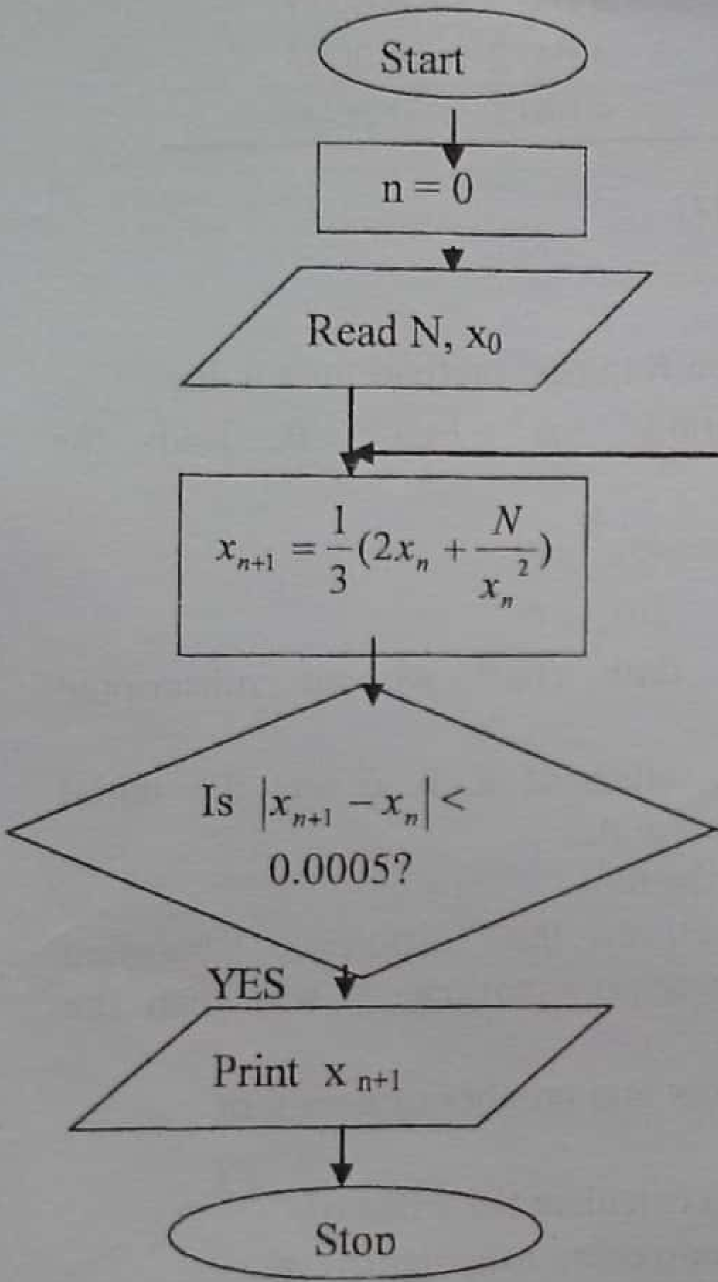
Draw a flow chart that

(i) Reads the value of N and x_0

(ii) Computes and prints the cube root of N correct to 3 decimal places

Then perform a dry run for N= 66 and $x_0 = 4$

Solution


Start
n = 0
Read N, x_0
$x_{n+1} = \frac{1}{3}(2x_n + \frac{N}{x_n^2})$
$x_n \leftarrow x_{n+1}$
n = n + 1
Is $|x_{n+1} - x_n| <$ 0.0005?
NO
YES
Print x_{n+1}
Stop

n	x_n	x_{n+1}	$\lvert x_{n+1} - x_n \rvert$
0	4	4.0417	0.0417
1	4.0417	4.0412	0.0005
2	4.0412	4.0412	0.0000

Hence the root = 4.041

Example

Show that the Newton-Raphson method for solving the quadratic equation $ax^2 + bx + c = 0$, leads the iterative formula

$$x_{n+1} = \frac{ax_n^2 - c}{2ax_n + b}$$

Hence construct a flow chart without subscripted variables to;

i. Read the values of a, b, c and the initial approximation A.

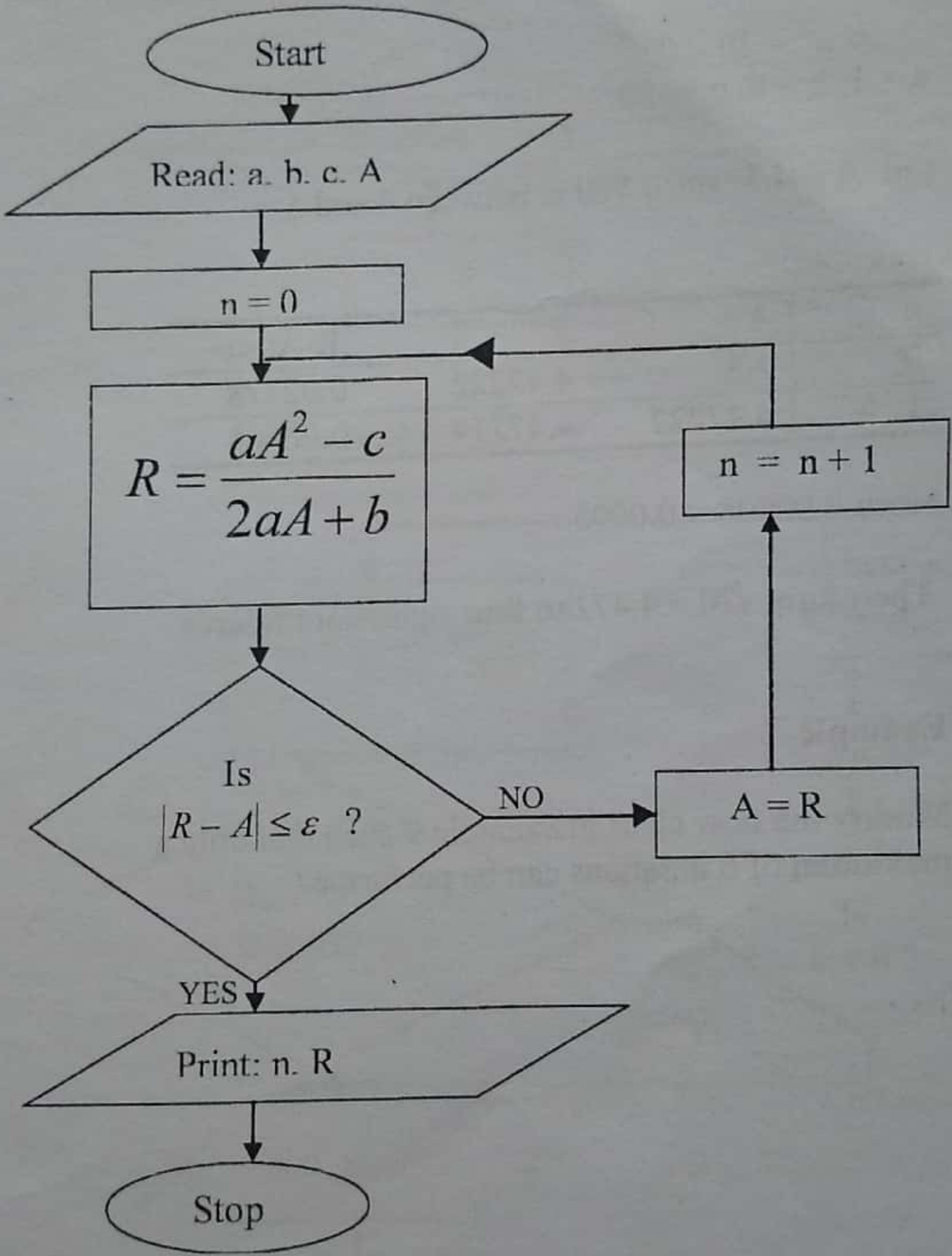
ii. Calculate the root

iii. Test whether the difference between successive approximations is less than the error limit ε ,

iv. Print the root and number of iterations

Use your flow chart to calculate the value of positive square of 20 correct to four significant figures.

Solution


Start
Read: a. b. c. A
n = 0
$R = \frac{aA^2 - c}{2aA + b}$
Is
$|R - A| \leq \varepsilon$?
NO
A = R
n = n + 1
YES
Print: n. R
Stop

Dry run

$x = \sqrt{20}$

$\Rightarrow x^2 - 20 = 0$

a = 1, b = 0, c = -20

and A = 4.5 since √20 is between 4 and 5

n	A	R	$\lvert R-A \rvert < \epsilon$
0	4.5	4.47222	0.02778
1	4.47222	4.47214	0.00008

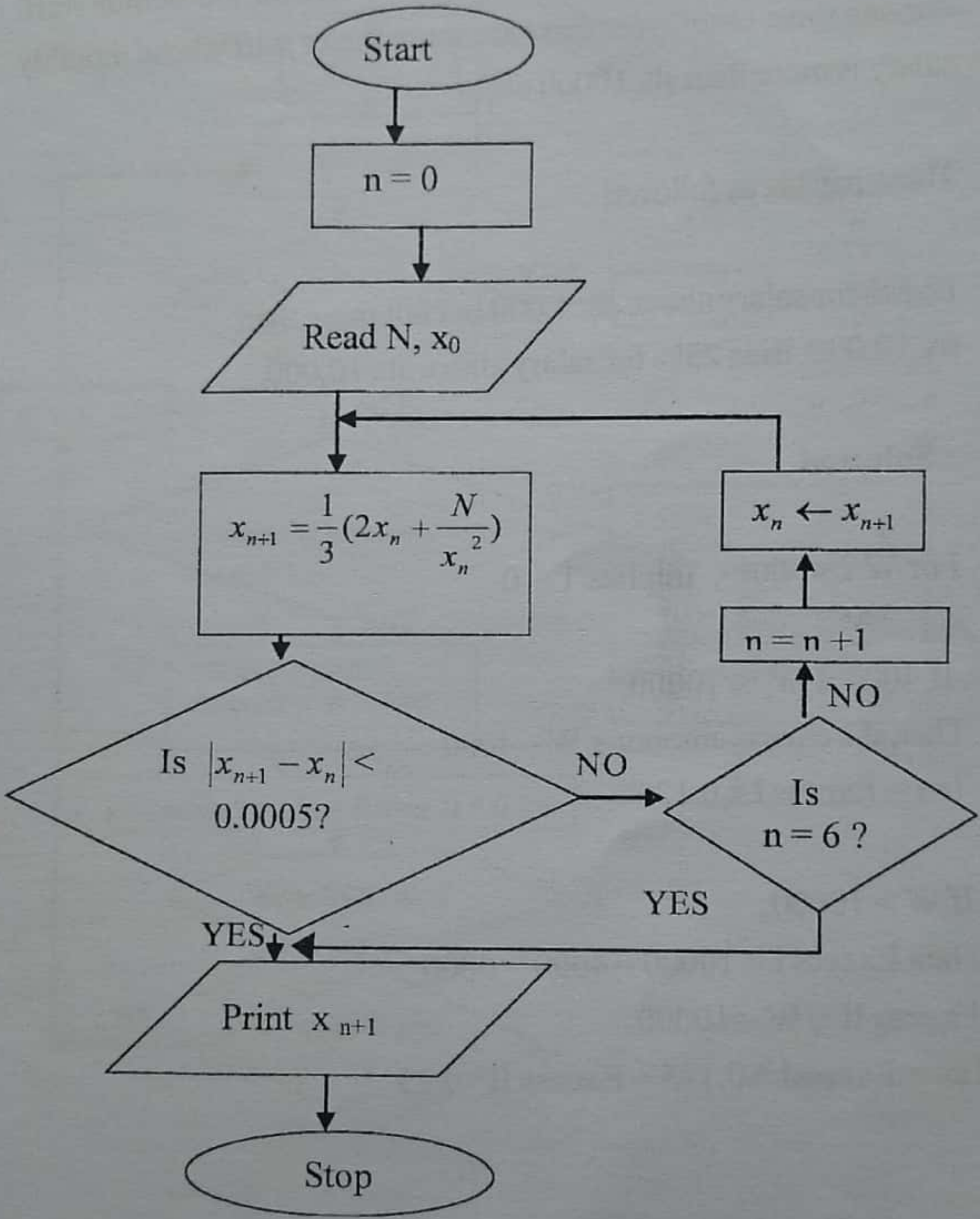
Since $0.00008 < 0.0005$

Therefore $\sqrt{20} = 4.472$ to four significant figures

Example

Modify the flow chart in example 9 such that only a maximum of 6 iterations can be performed

Modification



Example

Modify the flow chart in **example 4** to cater for the Senior staff. Among these employees there are some senior staff whose monthly salary is more than sh. 10000 each.

These pay tax as follows:

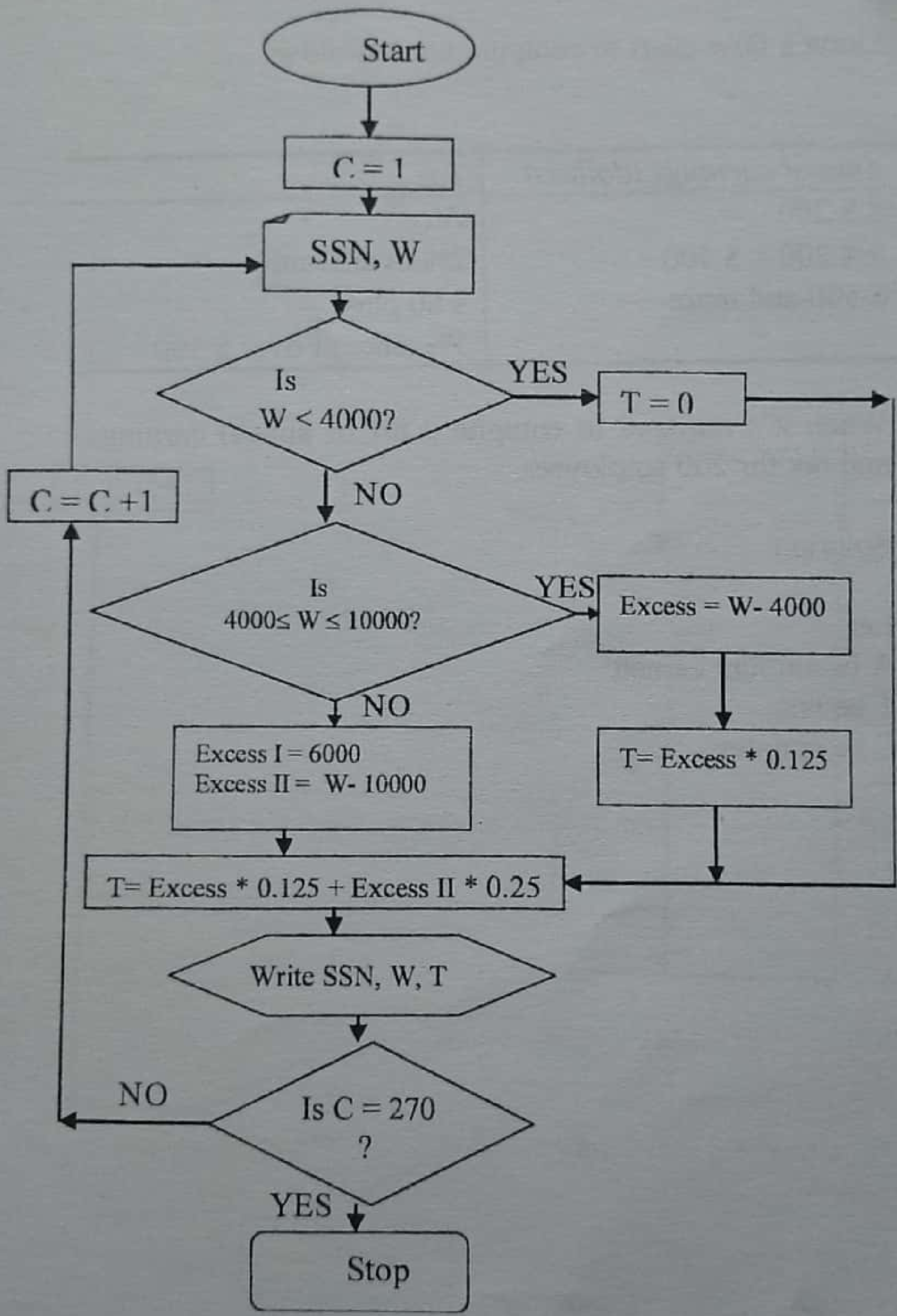
12.5% for salary above sh. 4,000 but not more than sh. 10,000; then 25% for salary above sh. 10,000.

Solution

For $W \leq 4000/=$, Implies $T = 0$

If $4000 \leq W \leq 10000$,
Then the excess amount $= W - 4000$
Tax $=$ Excess I x 0.125

If $W > 10000$,
Then Excess I $= 10000 - 4000 = 6000$
Excess II $= W - 10000$
Tax $=$ ExcessI * 0.125 + Excess II* 0.25


Start
C = 1
SSN, W
Is
W < 4000?
YES
T = 0
NO
C = C +1
Is
4000≤ W ≤ 10000?
YES
Excess = W- 4000
NO
Excess I = 6000
Excess II = W- 10000
T= Excess * 0.125
T= Excess * 0.125 + Excess II * 0.25
Write SSN, W, T
NO
Is C = 270
?
YES
Stop

Example

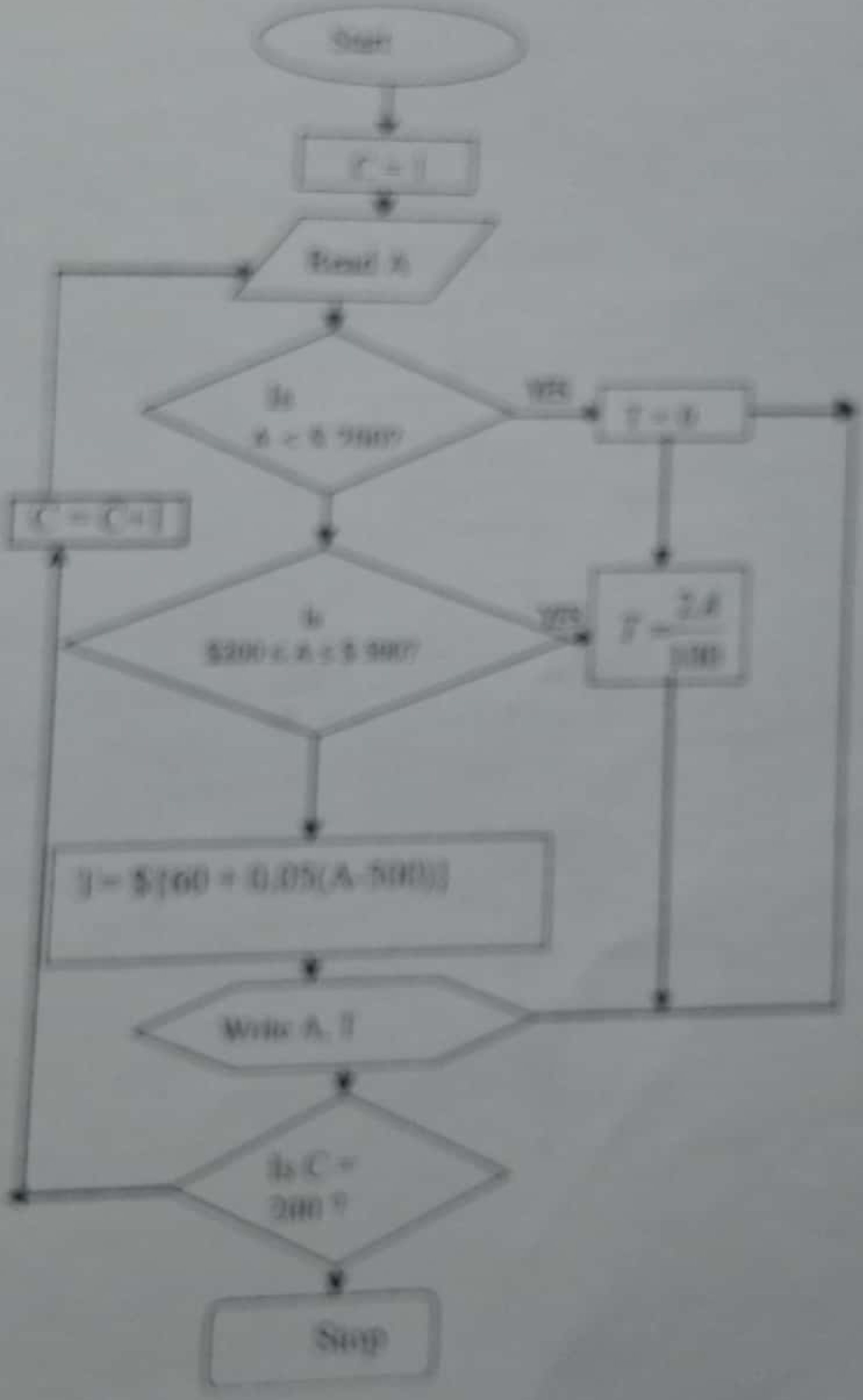
Draw a flow chart to compute tax as follows

Annual earnings (dollars)	*Tax*
< \$ 200	Zero
≥ \$ 200 < \$ 400	2% of amount
\$ 500 and more	\$ 60 plus 5% amount over \$ 500

When it's required to compile a list of annual earnings and tax for 200 employees.

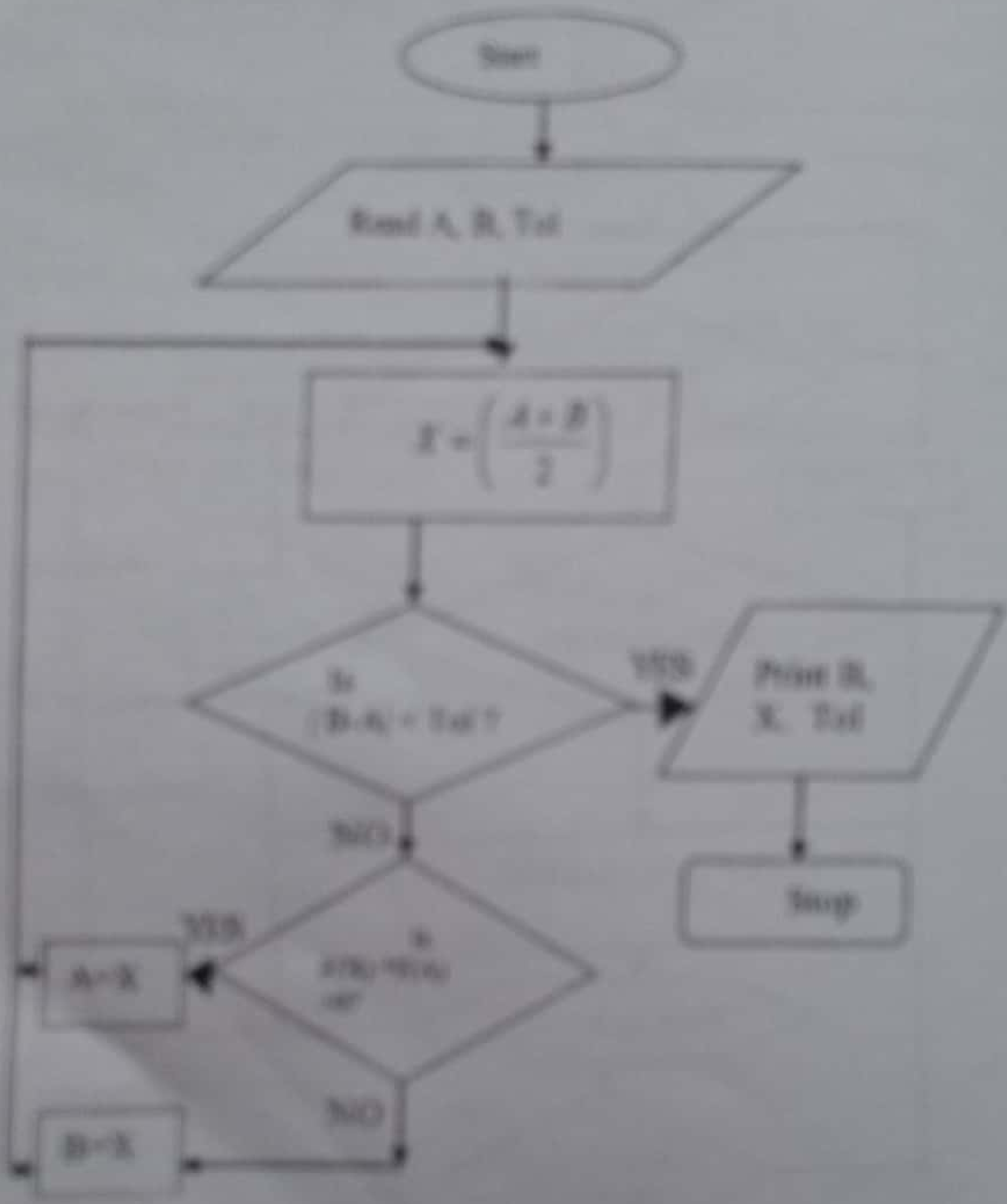
Solution

Let
A be amount earned
T be tax
C be counter



Start
Read A
YES
YES
Write A, T
Stop

An iteration method for approximating a root of the equation f(x)=0 is described in the follow chart below.



Given that A = 1.6875, B = 1.8750, TOL = 0.01 perform a dry run for the flow chart to determine $\sqrt[3]{5}$ tabulating the values of A,B and X at each stage.

Solution

$$f(x)=0 \Rightarrow \sqrt[3]{N}=x$$
$$f(x)=x^3-N=0$$
$$f(x)=x^3-N$$
$$When\ N=5$$
$$\Rightarrow f(1.6875)=1.6875^3-5=-0.1946$$
$$Then\ f(1.78125)=1.78125^3-5=0.6516$$

A	B	X	|B-A|	F(X)*F(A)
1.6875	1.8750	1.78125	0.1875	< 0
1.6875	1.78125	1.734375	0.099375	<0
1.6875	1.734375	1.7109375	0.046875	<0
1.6875	1.7109375	1.6992188	0.0224	<0
1.69922	1.7109375	1.7050788	0.0118	<0
1.705079	1.7109375	1.708007	0.0059	<0
1.708007	1.7109375	1.70947	0.002937	

$$\therefore \sqrt[3]{5}=1.70947=1.71$$

COMPOUND INTEREST

Sometimes people invest their money in the bank. It is important that the person get know how much he will earn after some time. If we let the amount be A and interest be 2% per year.

Then at end of the year he earns = $A+\frac{2A}{100}=1.02A$

The following year he earns =

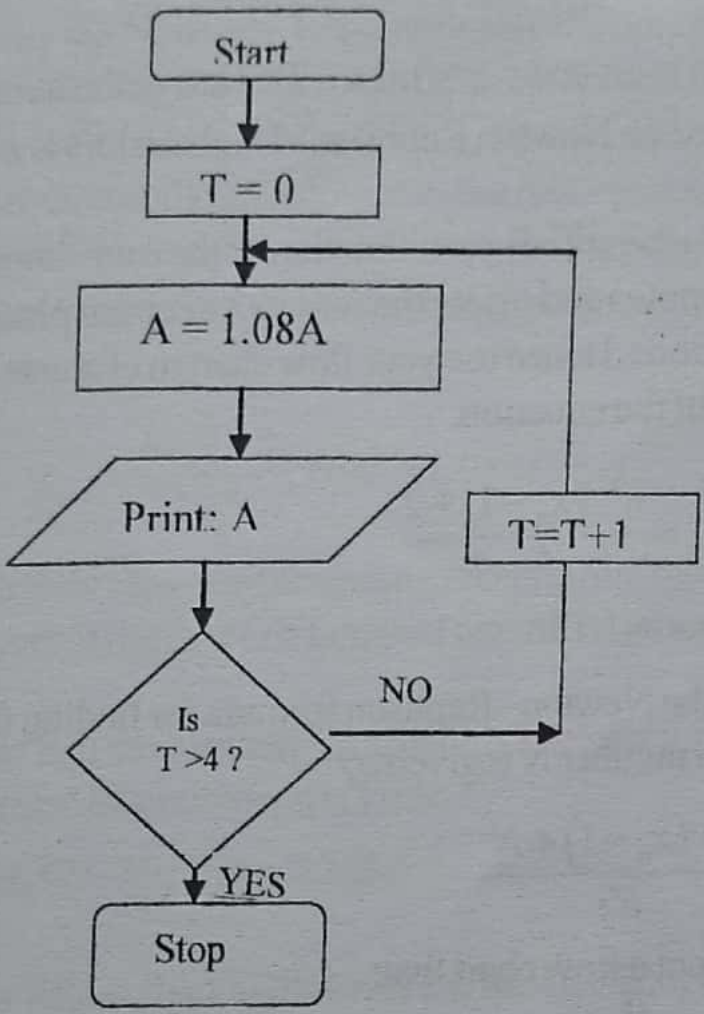
$$1.02A+\frac{2}{100}x1.02A=1.02^2A$$

So an expression can be obtained for any given number of years and flow chart can be constructed

$$P=1.02^nA$$

Example

Construct a flow chart for public servant who earns A millions and he/she invest in bank at rate of 8% compound interest for n years. Perform a dry run for
$n=4, A_0=2$ million



Dry run

T	$A(millions)$
0	2
1	2.16
2	2.3328
3	2.5194
4	2.939

Therefore A = 2.939 millions after four years

Exercise

1. An equation is given by $x = \ln(x+2)$. Derive the iterative formulae based on Newton Raphson's Method. Draw a flow chart that:

   i. Reads the initial approximation of the root
   ii. Computes and prints the root to 3 decimal places or after 4 iterations. Hence use your flow chart to evaluate the positive root of the equation.

Answer

$$x_{n+1} = \frac{e^{x_n}(x_n - 1) + 2}{e^{x_n} - 1}$$

$Root = 1.146$

2. Show that the Newton–Raphson formula for finding the natural logarithm of a number N is given by

$$x_{n+1} = \frac{e^{x_n}(x_n - 1) + N}{e^{x_n}}$$

Hence construct a flow chart that:

   i. Reads N and the initial approximation x_0
   ii. Computes the logarithm to 2 decimal places
   iii. Prints out the natural logarithm α

   Perform a dry run for N=1.7, $x_0 = 0.5$

Answer

n	x_n	x_{n+1}	$\lvert x_{n+1} - x_n \rvert$
0	0.5	0.5675	0.0675
1	0.5675	0.5653	0.0022

3.Derive by the Newton – Raphson's formula, the iterative that can be used for finding the fifth root of a number N. Write a simple algorithm that can be used for finding the root. Construct a simple flow chart that reads and computes the root. Taking N= 37 and initial approximation $x_0 = 2$, perform a dry for your flow chart to read the root correct to three decimal places.

Answer $x_{n+1} = \frac{4x_n^5 + 5}{5x_n^4}$, $root = 2.059(3d.p)$

4. Draw a flow chart for computing and printing the mean of the square roots of the first one hundred natural numbers

5. (i)Show that the Newton – Raphson's formula for approximating the reciprocal of a number n is given by

$x_{n+1} = x_n(2 - Nx_n);\ n = 0,1,2...$

(ii) Draw a flow chart to illustrate the use of the algorithm for computing andprinting of a close approximation to α. Take the initial approximation to be a and ensure that the computation ceases as soon as the absolute value of the difference between consecutive iterates is less than 0.005 or twenty iterations have been performed.

(iii) Perform a dry run for your flow chart for

N = 55.7; a = 0.02

Answer 0.02

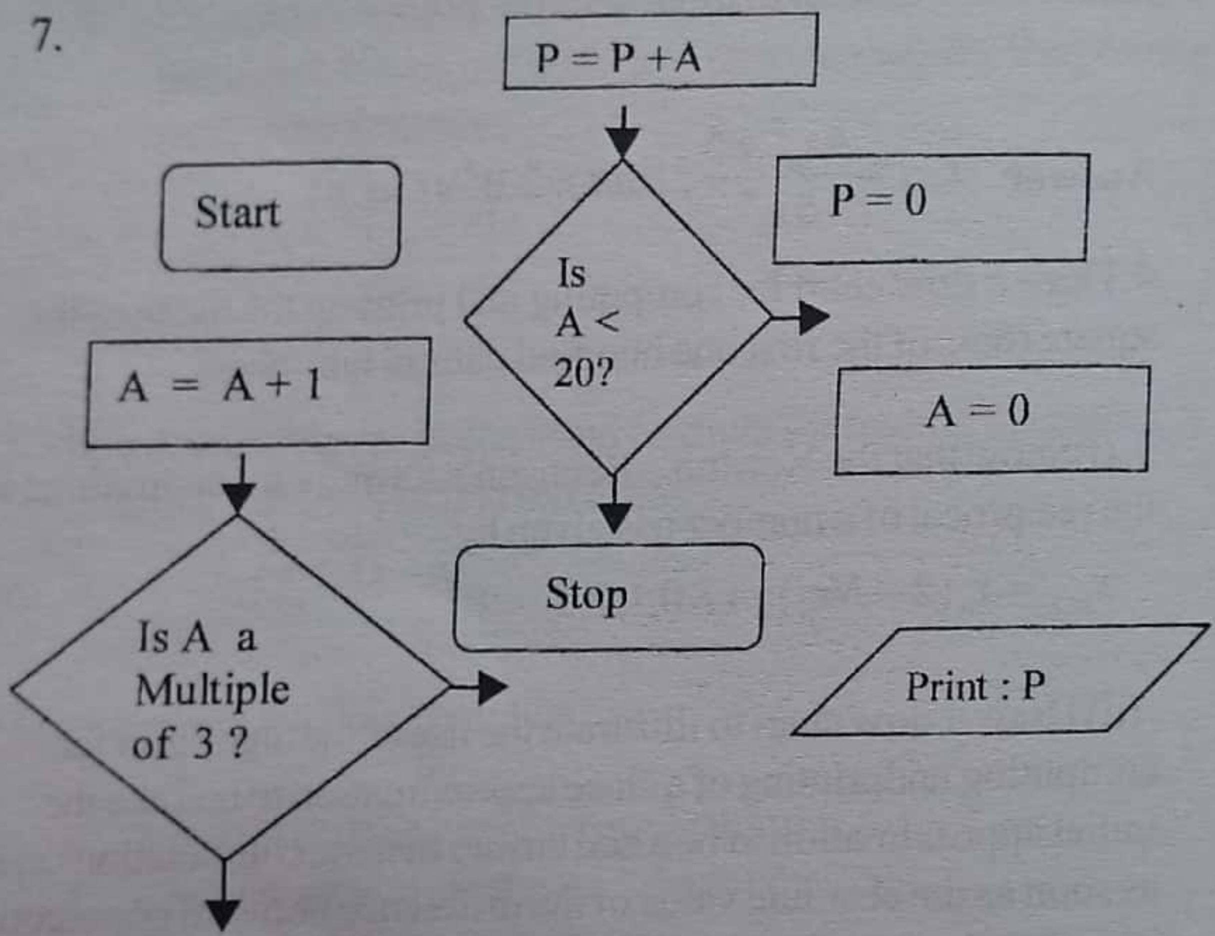
6. Show that the Newton-Rahson formula for finding the root of the equation $2x^3 + 5x - 8 = 0$ is

$$\frac{4x_n^3 + 8}{6x_n^2 + 5}$$

Taking the first approximation to the root of the above equation as 1.2. Draw a flow chart that reads and prints the number of iterations and root. Carry out a dry run of the flow chart and obtain the root with an error of less than 0.001

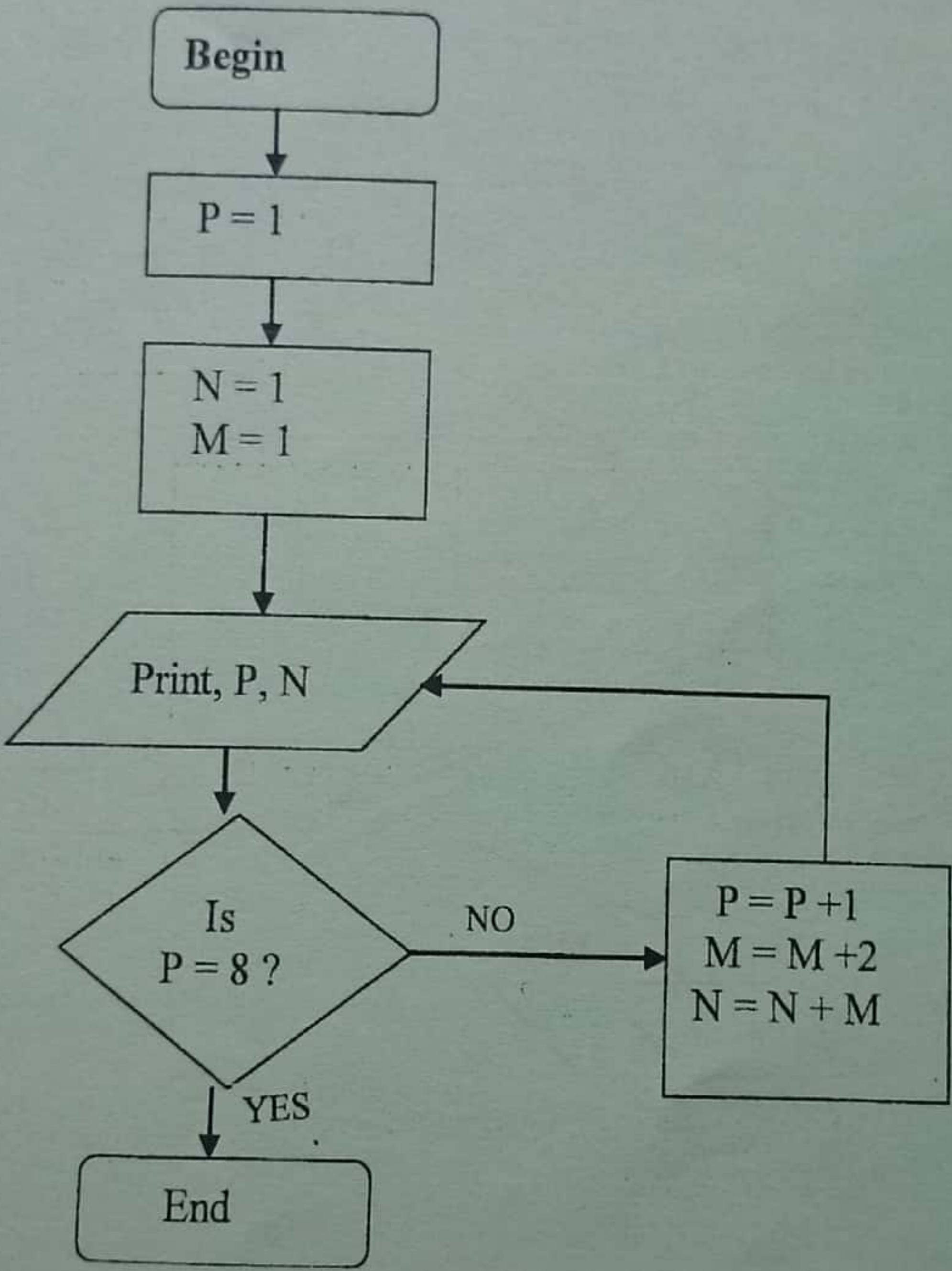
Answer 1.087

7.



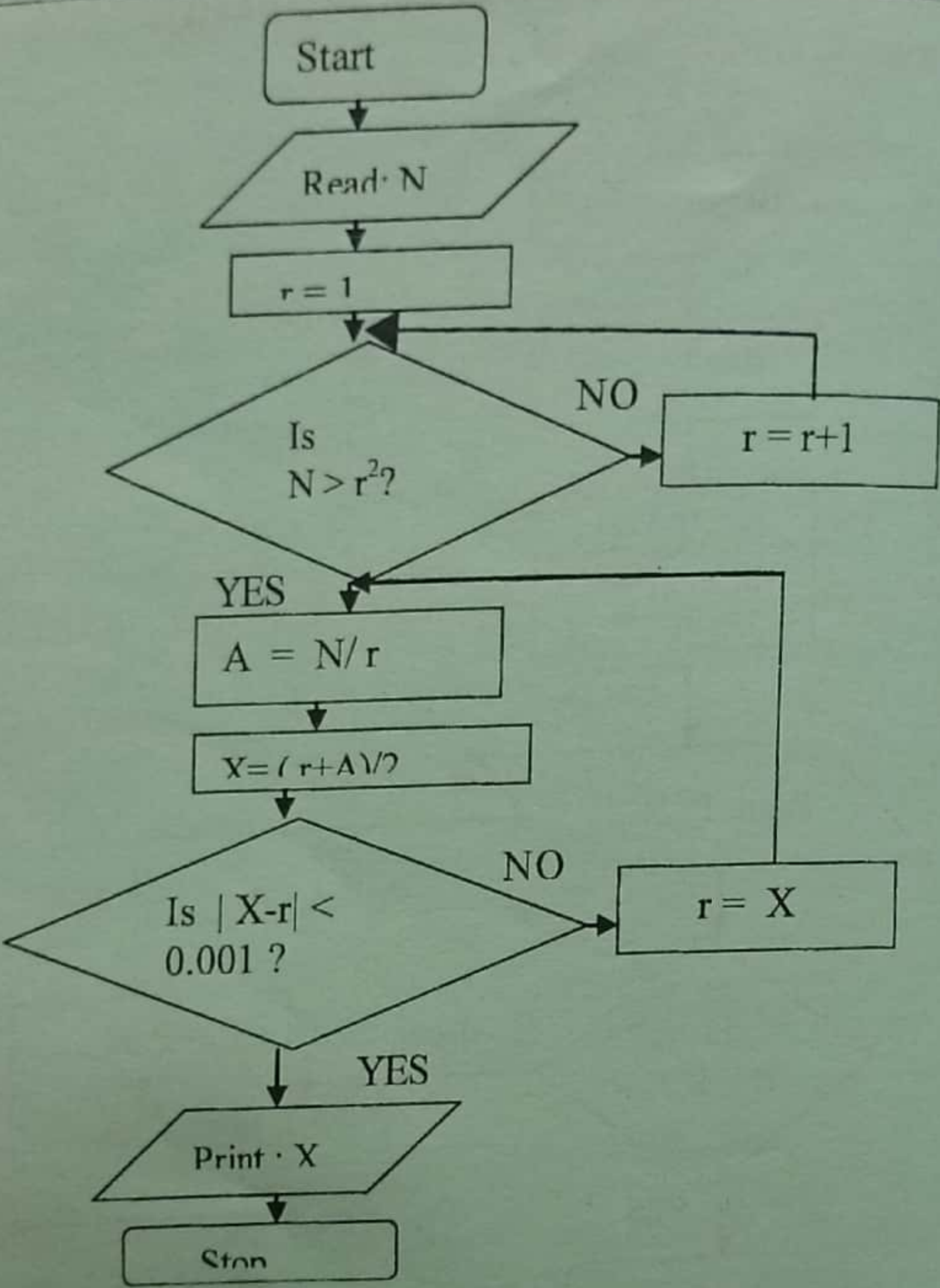
Rearrange them and draw a complete logical flow chart Perform a dry run for the flow chart

8. Study the flow chart below:-



(i.) Perform a dry run of the flow chart

(ii) State the relationship between P and N



i. Perform a dry run for N=20, showing clearly the contents of each store and values of X printed out.

ii. What process does the flow chart represent?

FORMULAE

DESCRIPTIVE STATISTICS

Mean $\bar{x} = \frac{\sum x}{n}$ or $\bar{x} = \frac{\sum fx}{\sum f}$

Or $\bar{x} = A + \left[\frac{\sum ft}{\sum f}\right]c$ Let $t = \frac{x - A}{c}$

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

$\left[\frac{\sum fd^2}{\sum f} - \left(\frac{\sum fd}{\sum f}\right)^2\right]$ Note: $d = x - A$

Note A is working mean

Variance $(\sigma^2) = c^2\left[\frac{\sum ft^2}{\sum f} - \left(\frac{\sum ft}{\sum f}\right)^2\right]$

Standard Deviation $(s) = \left[\frac{\sum fx^2}{\sum f} - \left(\frac{\sum fx}{\sum f}\right)^2\right]^{0.5}$

$$\text{Median} = L_1 + \left(\frac{N/_2 - fb}{fm}\right)c$$

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

PROBABILITY THEORY

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

$P(A \cap B) = P(A) + P(B)$ (mutually exclusive)

$P(A \cap B) = P(A)\, P(B \mid A)$ (multiplication rule)

$P(A \cap B) = P(A)\ P(B)$ (for independent events)

Conditional probability

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

Bayes' theorem $P(B_1|A) = \dfrac{P(B_1 \cap B)}{P(A)}$

$\Rightarrow P(B_1|A) =$

$$\frac{P(B_1)P(A \mid B_1)}{P(B_1)P(A \mid B_2) + P(B_2)P(A \mid B_2) + \ldots + P(B_k)P(A \mid B_k)}$$

DISCRETE PROBABILITY DISTRIBUTION

$\sum(X = x) = 1$

Mean $E(x) = \sum xP(X = x)$

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

$= \sum x^2 \quad P(X = x) - [\sum xP(X = x)]^2$

Note for series

$1 + 2 + 3 + 4 + 5 \ldots\ldots + n = ½\ n(n + 1)$

$1^2 + 2^2 \quad + 3^2 + \ldots + n^2 = 1/6\ n(n + 1)(2n + 1)$

If $Sn = a + ar + ar^2 + \ldots ar^{n-1}$

$$S_n = \frac{a(r^2 - 1)}{r - 1} = \frac{a(1 - r^n}{1 - r}$$

If $-1 < r > 1$ then $S_n \quad = \frac{a}{1 - r}$

BINOMIAL DISTRIBUTION n cr or $\binom{n}{n-r} = \binom{n}{r}$

n	0	1	2	3	4	5	6	7	8	9	10	11
1	1	1										
2	1	2	1									
3	1	3	3	1								
4	1	4	6	4	1							
5	1	5	10	10	5	1						
6	1	6	15	20	15	6	1					
7	1	7	21	35	35	21	7	1				
8	1	8	28	56	70	56	28	8	1			
9	1	9	36	84	126	126	84	36	9	1		
10	1	10	45	120	210	252	210	120	45	10	1	
11	1	11	55	165	330	462	462	330	165	55	11	1
12	1	12	66	220	495	792	924	792	495	220	66	12
13	1	13	78	286	715	1287	1716	1716	1287	715	286	78
14	1	14	91	364	1001	2002	3003	3432	3003	2002	1001	364
15	1	15	105	455	1365	3003	5005	6435	6435	5005	3003	1365
16	1	16	120	560	1820	4368	8008	11440	12870	11440	8008	4368
17	1	17	136	680	2380	6188	12376	19448	24310	24310	19448	123%
18	1	18	153	816	3060	8568	18564	31824	43758	48620	43758	31824
19	1	19	171	969	3876	11628	27132	50388	75582	92378	92378	75582
20	1	20	190	1140	4845	15504	38760	77520	125970	167960	184756	167960
21	1	210	1330	5985	20349	54264	116280	203490	293930	352716	352716	352716
22	1	22	231	1540	7315	26334	74613	170544	319770	497420	646646	705432
23	1	23	253	1771	8855	33649	100947	245157	490314	817190	1144066	1352078

BINOMIAL DISTRIBUTION B (n,p) INDIVIDUAL TERMS

n	r	0.01	0.050	0.10	0.15	0.20	0.25	0.30	6.35	040	045	050
2	0	0.9801	9025	8100	7225	6400	5625	4900	4225	3600	3025	2500
	1	0.0198	0950	1800	2550	3200	3750	4200	4550	4800	4950	5000
	2	0.0001	0025	0100	0225	0400	0625	0900	1225	1600	2025	2500
3	0	0.9703	8574	7290	6141	5120	4219	3430	2746	2160	1664	1250
	1	0.0294	1354	2430	3251	3840	4219	4410	4436	4320	4084	3750
	2	0.0003	0071	0270	0574	0960	1406	1890	2389	2880	3341	3750
	3		0001	0010	0034	0080	0156	0270	0429	0640	0911	1250
4	0	0.9606	8145	6561	5220	4096	3164	2401	1785	1296	0915	0625
	1	0.0388	1715	2916	3685	4096	4219	4116	3845	3456	2995	2500
	2	0.0006	0135	0486	0975	1536	2109	2646	3105	3456	3675	3750
	3		0005	0036	0115	0256	0469	0756	1115	1536	2005	2500
	4			0001	0005	0016	0039	0081	0150	0256	0410	0625
5	0	0.9510	7738	5905	4437	3277	2373	1681	1160	0778	0503	0312
	1	0.0480	2036	3280	3915	4096	3955	3602	3124	2592	2059	1562
	2	0.0010	0214	0729	1382	2048	2637	3087	3364	3456	3369	3125
	3		0011	0081	0244	0512	0879	1323	1811	2304	2757	3125
	4			0004	0022	0064	0146	0284	0488	0768	1128	1562
	5				0001	0003	0010	0024	0053	0102	0185	0312
6	0	0.9415	7351	5314	3771	2621	1780	1176	0754	0467	0277	0156
	1	0.0571	2321	3543	3993	3932	3560	3025	2437	1866	1359	0938
	2	0.0014	0305	0984	1762	2458	2966	3241	3280	3110	2780	2344
	3		0021	0146	0415	0819	1318	1852	2355	2765	3032	3125
	4		0001	0012	0055	0154	0330	0595	0953	1382	1861	2344
	5			0001	0004	0015	0044	0102	0205	0369	0609	0938
	6					0001	0002	0007	0018	0041	0083	0156
7	0	0.9321	6983	4783	3206	2097	1335	0824	0490	0280	0152	0078
	1	0.0659	2573	3720	3960	3670	3115	2471	1848	1306	0872	0547
	2	0.0020	0406	1240	2097	2753	3115	3177	2985	2613	2140	1641
	3		0036	0230	0617	1147	1730	2269	2679	2903	2918	2734
	4		0002	0026	0109	0287	0577	0972	1442	1935	2388	2734
	5			0002	0012	0043	0115	0250	0466	0774	1172	1641
	6				0001	0004	0013	0036	0084	0172	0320	0547
	7						0001	0002	0006	0016	0037	0078
8	0	0.9227	6634	4305	2725	1678	1001	0576	0319	0168	00&4	0039
	1	0.0746	2793	3826	3847	3355	2670	1977	1373	0896	0548	0312
	2	0.0026	0515	1488	2376	2936	3115	2965	2587	2090	1569	1094
	3.	0001	0054	0331	0839	1468	2076	2541	2786	2787	2568	2188
	4.		0004	0046	0185	0459	0865	1361	1875	2322	2627	2734
	5			0004	0026	0092	0231	0467	0808	1239	1719	2188
	6				0002	0011	0038	0109	0217	0413	0703	1094
	7					0001	0004	0012	0033	0079	0164	0312
	8.							0001	0002	0007	0017	0039

BINOMIALDI INDIVIDUAL TERMS

n	r	0.01	0.05	0.10	0.15	0.20	0.25	0.30	0.35	0.40	0.45	0.50
9	0	0.9135	6302	3874	2316	1342	0751	0404	0207	0101	0046	0020
	1	0.0830	2985	3874	3679	3020	2253	1556	1004	0605	0339	0176
	2	0.0034	0629	1722	2597	3020	3003	2668	2162	1612	1110	0703
	3	0.0001	0077	0446	1069	1762	2336	2668	2716	2508	2119	1641
	4		0006	0074	0283	0661	1168	1715	2194	2508	2600	2461
	5			0008	0050	0165	0389	0735	1181	1672	2128	2461
	6			0001	0006	0028	0087	0210	0424	0743	1160	1641
	7					0003	0012	0039	0098	0212	0407	0703
	8						0001	0004	0013	0035	0083	0176
	9								0001	0003	0008	0020
10	0	0.9044	5987	3487	1969	1074	0563	0282	0135	0060	0025	0010
	1	0.0914	3151	3874	3474	2684	1877	1211	0725	0403	0207	0098
	2	0.0042	0746	1937	2759	3020	2816	2335	1757	1209	0763	0439
	3	0.0001	0105	0574	1298	2013	2503	2668	1522	2150	1665	1172
	4		0010	0112	0401	0881	1460	2001	2377	2508	2384	2051
	5		0001	0015	0085	0264	0584	1029	1536	2007	2340	2461
	6			0001	0012	0055	0162	0368	0689	1115	1596	2051
	7				0001	0008	0031	0090	0212	0425	0746	1172
	8					0001	0004	0014	0043	0106	0229	0439
	9							0001	0005	0016	0042	0098
	10									0001	0003	0010
11	0	0.8953	5688	3138	1673	0859	0422	0198	0088	0036	0014	0005
	1.	0.0995	3293	3835	3248	2362	1549	0932	0518	0266	0125	0054
	2	0.0050	0867	2131	2866	2953	2581	1998	3395	0887	0513	0269
	3	0.0002	0137	0710	1517	2215	2581	2568	2254	1774	1259	0806
	4		0014	0158	0536	1107	1721	2201	2428	2365	2060	1611
	5		0001	0025	0132	0388	0803	1321	1830	2207	2360	2256
	6			0003	0023	0097	0268	0566	0985	1471	1931	2256
	7				0003	0017	0064	0173	0379	0701	1128	1611
	8					0002	0011	0037	0102	0234	0462	0806
	9						0001	0005	0018	0052	0126	0269
	10								0002	0007	0021	0054
	11										0002	0005
12	0	0.8864	5404	2824	1422	0687	0317	0138	0057	0022	0008	0002
	1	0.1074	3413	3766	3012	2062	1267	0712	0368	0174	0075	0029
	2	0.0060	0988	2301	2924	2835	2323	1678	1088	0639	0339	0161
	3	0.0002	0173	0852	1720	2362	2581	2397	1954	1419	0923	0537
	4		0021	0213	0683	1329	1936	2311	2367	2128	1700	1208
	5		0002	0038	0193	0532	1032	1585	2039	2270	2225	1934
	6.			0005	0040	0155	0401	0792	1281	1766	2124	2256
	7				0006	0033	0115	0291	0591	1009	1489	1934
	8				0001	0005	0024	0078	0199	0420	0762	1208

BINOMIAL DISTRIBUTION B(n, p) INDIVIDUAL TERMS

n	r	0.01	0.05	0.10	0.15	0.20	0.25	0.30	0.35	040	0.45	0.50
	9					0001	0004	0015	0048	0125	0277	0537
	10							0002	0008	0025	0068	0161
	11								0001	0003	0010	0029
	12										0001	0002
15	0	0.8601	4633	2059	0874	0352	0134	0047	0016	0005	0001	
	1	0.1301	3658	3432	2312	1319	0668	0305	0126	0047	0016	0005
	2	0.0092	1348	2669	2856	2309	1559	0916	0476	0219	0090	0032
	3	0.0004	0307	1285	2184	2501	2252	1700	1110	0634	0318	0139
	4		0049	0428	1156	1876	2252	2186	1792	1268	0780	0417
	5		0006	0105	0449	1032	1651	2061	2123	1859	1404	0916
	6.			0019	0132	0430	0197	1472	1906	2066	1914	1527
	7.			0003	0030	0138	0393	0811	1319	1771	2013	1964
	8.				0005	0035	0131	0348	0710	1181	1647	1964
	9.				0001	0007	0034	0116	0298	0612	1048	1527
	10.					0001	0007	0030	0096	0245	0515	0916
	11.						0001	0006	0024	0074	0191	0417
	12.							0001	0004	0016	0052	0139
	13.								0001	0003	0010	0032
	14.										0001	0005
20.	0	0.8179	3585	1216	0388	0115	0032	0008	0002			
	1	0.1652	3774	2702	1368	0576	0211	0068	0020	0005	0001	
	2.	0.0159	1887	2852	2293	1369	0669	0278	0100	0031	0008	0002
	3.	0.0010	0596	1901	2428	2054	1339	0716	0323	0123	0040	0011
	4.		0133	0898	1821	2182	1897	1304	0738	0350	0139	0046
	5.		0022	0319	1028	1746	2023	1789	1272	0746	0365	0148
	6.		0003	0089	0454	1091	1686	1916	1712	1244	0746	0370
	7.			0020	0160	0545	1124	1643	1844	1659	1221	0739
	8.			0004	0046	0222	0609	1144	1614	1797	1623	1201
	9.			0001	0011	0074	0271	0654	1158	1597	1771	1602
	10.				0002	0020	0099	0308	0686	1171	1593	1762
	11.					0005	0030	0120	0336	0710	1185	1605
	12.					0001	0008	0039	0136	0355	0727	1201
	13.						0002	0010	0045	0146	0366	0739
	14.							0002	0012	0049	0150	0370
	15.								0003	0013	0049	0148
	16.									0003	0013	0046
	17.										0002	0011
	18.											0002

CUMULATIVE BINOMIAL DISTRIBUTION $\sum b(x, n, p)$

n	r	0.01	0.05	0.10	0.15	0.20 P	0.25	0.30	0.35	040	0.45	0.50
2	1	0.0199	0975	1900	2775	3600	4375	5100	5775	6400	6975	7500
	2	0.0001	0025	0100	0225	0400	0625	0900	1225	1600	2025	2500
3	1	00297	1426	2710	3859	4880	5781	6570	7254	7840	8336	8750
	2	0.0803	0072	0280	0608	1040	1562	2160	2818	3520	4252	5000
	3		0001	0010	0034	0080	0156	0270	0429	0640	0911	1250
4	1	0.0394	1855	3439	4780	5904	6836	7599	8215	8704	9085	9375
	2	0.0006	0140	0523	1095	1808	2617	3483	4370	5248	6090	6875
	3		0005	0037	0120	0272	0508	0837	1265	1792	2415	3125
	4			0001	0005	0016	0039	0081	0150	0256	0410	0625
5	1	0.0490	2262	4095	5563	6723	7627	8319	8840	9222	9497	9688
	2	0.0010	0226	0815	1648	2627	3672	4718	5716	6630	7438	8125
	3		0012	0086	0266	0579	1035	1631	2352	3174	4069	5000
	4			0005	0022	0067	0156	0308	0540	0870	1312	1875
	5				0001	0003	0010	0024	0053	0102	0185	0312
6	1	0.0585	2649	4686	6229	7379	8230	8824	9246	9533	9723	9844
	2	0.0015	0328	1143	2235	3446	4661	5798	6809	7667	8364	8906
	3		0022	0158	0473	0989	1694	2557	3529	4557	5585	6562
	4		0001	0013	0059	0170	0376	0705	1174	1792	2553	3438
	5			0001	0004	0016	0046	0109	0223	0410	0692	1094
	6					0001	0002	0007	0018	0041	0083	0156
7	1	0.0679	3017	5217	6794	7903	805	9176	9510	9720	9848	9922
	2	0.0020	0444	1497	2834	4233	5551	6706	7662	8414	8976	9375
	3		0038	0257	0738	1480	2436	3529	4677	5801	6836	7734
	4		0002	0027	0121	0333	0706	1260	1998	2898	3917	5000
	5			0002	0012	0047	0129	0288	0556	0963	1529	2266
	6				0001	0004	0013	0038	0090	0188	1357	0625
	7						0001	0002	0006	0016	0037	0078
8	1	0.0773	3366	5695	7275	8322	8999	9424	9681	9832	9916	9961
	2	0.0027	0572	1869	3428	4967	6329	7447	8309	8936	9368	9648
	3	0.0001	0058	0381	1052	2031	3215	4482	5722	6846	7799	8555
	4		0004	0050	0214	0563	1138	1941	2936	4059	5230	6367
	5			0004	0029	0104	0273	0580	1061	1737	2604	3633
	6				0002	0012	0042	0113	0253	0498	0885	1445

CUMULATIVE BINOMIAL DISTRIBUTION $\sum b(x, n, p)$

n	r	0.01	0.05	0.10	0.15	0.20	0.25	0.30	0.35	040	0.45	0.50
	7					0001	0004	0013	0036	0085	0181	0352
	8							0001	0002	0007	0017	0039
9	1	0.0865	3698	6126	7684	8658	9249	9596	9793	9899	9954	9980
	2	0.0034	0712	2252	4005	5638	6997	8040	8789	9295	9615	9805
	3.	0.0001	0084	0530	1409	2618	3993	5372	6627	7682	8505	9102
	4		0006	0083	0339	0856	1657	2703	3911	5174	6386	7461
	5			0009	0056	0196	0489	0988	1717	2666	3786	5000
	6			0001	0006	0031	0100	0253	0536	0994	1658	2539
	7					0003	0013	0043	0112	0250	0498	0898
	8						0001	0004	0014	0038	0091	0195
	9								0001	0003	0008	0020
10	1	0.0956	4013	6513	8031	8926	9437	9718	9865	9940	9975	9990
	2	0.0043	0961	2639	4557	6242	7560	8507	9140	9536	9787	9893
	3	00001	0115	0702	1798	3222	4744	6172	7384	8327	9004	9453
	4		0010	0128	0500	1209	2241	3504	4862	6177	7340	8281
	5.		0001	0016	0099	0328	0781	1503	2485	3669	4956	6230
	6.			0001	0014	0064	0197	0473	0949	1662	2616	3770
	7.				0001	0009	0035	0106	0260	0548	1020	1719
	8					0001	0004	0016	0048	0123	0274	0547
	9.							0001	0005	0017	0045	0107
	10									0001	0003	0010
11	1	0.1047	4312	6862	8327	9141	9578	9802	9912	9964	9986	9995
	2	00052	1019	3006	5078	6779	8029	8870	9394	9698	9861	9941
	3	0.0002	0152	0896	2212	3826	5448	6873	7999	8811	9348	9673
	4		0016	0185	0694	1611	2867	4304	5744	7037	8089	8867
	5		0001	0028	0159	0504	1146	2103	3317	4672	6029	7256
	6			0003	0027	0117	0343	0782	1487	2465	3669	5000
	7				0003	0020	0076	0216	0501	0994	1738	2744
	8					0002	0012	0043	0122	0293	0610	1133
	9						0001	0006	0020	0059	0148	0327
	10								0002	0007	0022	0059
	11										0002	0005
12	1	0.1136	4596	7176	8578	9313	9683	9862	9943	9978	9992	9998
	2	0.0062	1184	3410	5565	7251	8416	9150	9576	9804	9917	9968
	3	0.0002	0196	1109	2642	4417	6093	7472	8487	9166	9579	9807
	4		0022	0256	0922	2054	3512	5075	6533	7747	8655	9270
	5		0002	0043	0239	0726	1576	2763	4167	5618	6956	8062

CUMULATIVE BINOMIAL DISTRIBUTION $\sum b(x, n, p)$

n	r	0.01	0.05	0.10	0.15	0.20	0.25	0.30	0.35	040	0.45	0.50
	6			0005	0046	0194	0544	1178	2127	3348	4731	6128
	7			0001	0007	0039	0143	0386	0846	1582	2607	3872
	8				0001	0006	0028	0095	0255	0573	1117	1938
	9					0001	0004	0017	0056	0153	0956	0730
	10							0002	0008	0028	0079	0193
	11								0001	0003	0011	0032
	12										0001	0002
15	1	0.1399	5367	7941	9126	9648	9866	9953	9984	9995	9999	1.0000
	2	0.0096	1710	4510	6814	8329	9198	9647	9858	9948	9983	9995
	3	0.0004	0362	1841	3958	6020	7639	8732	9383	9729	9893	9963
	4		0055	0556	1773	3518	5387	7031	8273	9095	9576	9824
	5		0006	0127	0617	1642	3135	4845	6481	7827	8796	9408
	6		0001	0022	0168	0611	1484	2784	4357	5968	7392	8491
	7			0003	0036	0181	0566	1311	2452	3902	5478	6964
	8				0006	0042	0173	0500	1132	2131	3465	5000
	9				0001	0008	0042	0152	0422	0950	1818	3036
	10					0001	0008	0037	0124	0338	0769	1509
	11						0001	0007	0028	0093	0255	0592
	12							0001	0005	0019	0063	0176
	13								0001	0003	0011	0037
	14										0001	0005
20	1	0.1821	6415	8784	9612	9885	9968	9992	9998	1.0000	1.0000	1.0000
	2	0.0169	2642	6083	8244	9308	9757	9934	9979	9995	9999	1.0000
	3	0.0010	0755	3231	5951	7939	9087	9645	9879	9964	9991	9998
	4		0159	1330	3523	5886	7748	8929	9556	9840	9951	9987
	5		0026	0432	1702	3704	5852	7625	8818	9490	9811	9941
	6		0003	0113	0673	1958	3828	5836	7546	8744	9447	9793
	7			0024	0219	0867	2142	3920	5834	7500	8701	9423
	8			0004	0059	0321	1018	2277	3990	5841	7480	8684
	9			0001	0013	0100	0409	1133	2376	4044	5857	7483
	10				0002	0026	0139	0480	1218	2447	4086	5881
	11					0006	0039	0171	0532	1275	2483	4119
	12					0001	0009	0051	0196	0565	1308	2517
	13						0002	0013	0060	0210	0580	1316
	14							0003	0015	0065	0214	0577
	15								0003	0016	0064	0207
	16									0003	0015	0069
	17										0003	0013
	18											0002

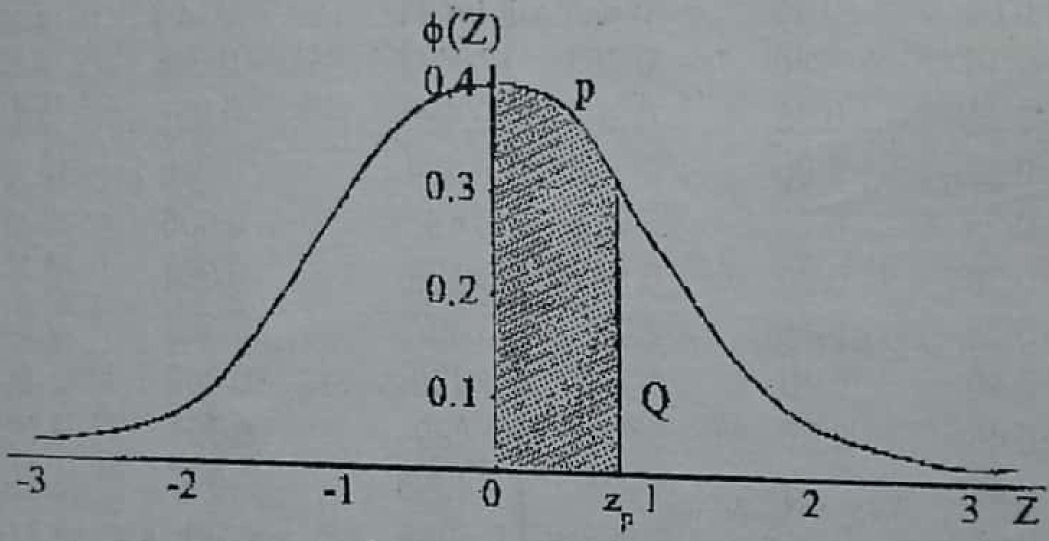
NOBMAL DISTRIBUTION NN(0,1) (p(Z)

Z	0	1	2	3	4	5	6	7	8	9	SUBTRACT 1	2	3	4	5	6	7	8	9
0.0	0.3989	3989	3989	3988	3986														
						3984	3982	3980	3977	3973									
0.1	0.3970	3965	3961	3956	3951						0	1	1	1	1	.2	2	2	3
											0	1	1	2	2	3	3	4	4
						3945	3939	3932	3925	3918	1	1	2	3	3	4	5	6	6
02	0.3910	3902	3894	3885	3876														
											1	2	3	4	4	5	6	7	5
						3867	3857	3847	3836	3825	1	2	3	4	5	7	8	9	10
0.3	0.3814	3802	3790	3778	3765														
											1	2	4	5	6	7	8	10	11
						3752	3739	3725	3712	3697									
0.4	0.3683	3668	3653	3637	3621	3605	3589	3572	3555	3538	1	3	4	6	7	8	10	11	13
0.5	0.3521	3503	3485	3467	3448	3429	3410	3391	3372	3352	2	3	5	6	8	10	11	13	14
0.6	0.3332	3312	3292	3271	3251	3230	3209	3187	3166	3144	2	4	6	8	10	11	13	15	17
0.7	0.3123	3101	3079	3056	3034	3011	2989	2966	2943	2920	2	4	6	8	10	13	15	17	19
0.8	0.2897	2874	2850	2827	2803	2780	2756	2732	2709	2685	2	5	7	9	11	14	16	18	21
0.9	0.2661	2637	2613	2589	2565	2541	2516	2492	2468	2444	2	5	7	10	12	14	17	19	22
1.0	0.2420	2396	2371	2347	2323	2299	2275	2251	2227	2203	2	5	7	10	12	14	17	19	22
1.1	0.2179	2155	2131	2107	2083	2059	2036	2012	1989	1965	2	5	7	10	12	14	17	19	22
1.2	0.1942	1919	1895	1872	1849	1826	1804	1781	1758	1736	2	5	7	10	12	14	17	19	22
1.3	0.1714	1691	1669	1647	1626	1604	1582	1561	1539	1518	2	5	7	9	12	14	16	18	21
1.4	0.1497	1476	1456	1435	1415	1394	1374	1354	1334	1315	2	4	7	9	11	13	15	18	30
1.5	0.1295	1276	1257	1238	1219	1200	1182	1163	1145	1127	2	4	6	8	10	12	14	16	18
1.6	0.1109	1092	1074	1057	1040	1023	1006	0989	0973	0957	2	4	6	8	9	11	13	15	17
1.7	0.0940	0925	0909	0893	0878	0863	0848	0833	0818	0804	2	3	5	7	8	10	12	14	15
1.8	0.0790	0775	0761	0748	0734						2	3	5	6	8	9	11	12	14
											I	3	4	6	7	8	10	11	11
						0721	0707	0694	0681	0669	1	3	4	5	7	8	9	10	12
1.9	0.0656	0644	0632	0620	0608	0596	0584	0.773	0561	0551	1	2	4	5	6	7	8	10	11
2.0	0.0540	0529	0519	0508	0498	0488	0478	0468	0459	0449	1	2	3	4	5	6	7	8	9
2.1	0.0440	0431	0422	0413	0404	0396	0389	0379	0371	0363	1	2	3	4	4	5	6	7	8
2.2	0.0355	0347	0339	0332	0325	0317	0310	0303	0297	0290	1	1	2	3	4	4	5	6	6
2.3	0.0283	0277	0270	0264	0258	0252	0246	0241	0235	0229	1	1	2	2	3	4	4	5	5
2.4	0.0244	0219	0213	0208	0203	0198	0194	0189	0184	0180	0	1	1	2	2	3	3	4	4
15	0.0175	0171	0167	0163	0158	0154	0151	0147	0143	0139	0	1	1	2	2	2	3	3	4
2.6	0.0136	0132	0129	0126	0122	0119	0116	0113	0110	0107	0	1	1	1	2	2	2	2	3
2.7	0.0104	0101	0099	0096	0093	0091	0088	0086	0084	0081	0	1	1	1	1	2	2	2	3
2.8	0.0079	0077	0075	0073	0071	0069	0067	0065	0063	0061									
2.9	0.0060	0058	0056	0055	0053	0051	0050	0048	0047	0046									
3.	0.0044	0033									1	2	3	4	5	6	7	8	9
			0024	0017							1	1	2	2	3	4	4	5	5
					0012	0009	0006	0004	0003	0002									

CUMULATIVE NORMAL DISTRIBUTION P(z)

Z	0	1	2	3	4	5	6	7	8	9	1	2	3	4	5	6	7	8	9
0.0	0.0000	0040	0080	0120	0160	0199	0239	0279	0319	0859	4	8	12	16	20	24	28	32	36
0.1	0.398	0438	0478	0517	0557	0596	0636	0675	0714	0753	4	8	12	15	19	22	27	31	35
0.2	0.0793	0832	0671	0910	0948	0987	1026	1064	1103	1141	4	8	12	15	19	22	27	31	35
0.3	0.1179	1217	1255	1293	1331	1366	1406	1443	1480	1517	4	8	11	15	19	22	26	30	34
0.4	0.1554	1591	1628	1664	1700	1736	1772	1808	1844	1879	4	7	11	14	18	22	25	29	32
0.5	0.1915	1950	1985	2019	2054	2088	2123	2157	2190	2224	3	7	10	14	17	21	24	27	31
0.6	0.2257	2291	2324	2357	2389	2422	2454	2486	2517	2549	3	6	10	13	16	19	23	36	29
0.7	0.2580	2611	2642	2673							3	6	9	12	15	19	22	25	28
					2704	2734	2764	2794	2823	2852	3	6	9	12	15	18	21	24	27
0.8	0.2881	2910	2939	2967	2995	3023					3	6	8	11	14	17	20	22	25
							3051	3078	3106	3133	3	5	8	11	13	16	19	22	24
0.9	0.3159	3186	3212	3238	3264	3289					3	5	8	10	13	16	18	21	23
							3315	3340	3365	3389	2	5	7	10	12	15	17	20	22
1.0	0.3413	3438	3461	3485	3508						2	5	7	10	12	14	17	19	22
						3531	3554	3577	3399	3621	2	4	7	9	11	13	15	18	20
1.1	0.366	3665	3686	3708							2	4	6	8	11	13	15	17	19
					3729	3749	3770	3790	3810	3830	2	4	6	8	10	12	14	16	18
1.2	0.3849	3869	3888	3907	3925						2	4	6	8	10	11	13	15	17
						3944	3962	3980	3997	4015	2	4	5	7	9	11	13	14	16
1.3	0.432	4049	4066	4082	4099	4115	4131	4147	4162	4177	2	3	5	6	8	10	11	13	14
1.4	0.4192	4207	4222	4236	4251	4265	4279	4292	4306	4319	1	3	4	6	7	8	10	11	13
1.5	0.4332	4345	4357	4370	4382	4394	4406	4418	4429	4441	1	2	4	5	6	7	8	10	11
1.6	0.4452	4463	4474	4484	4495	4905	4515	4535	4535	4545	1	2	3	4	5	6	7	8	9
1.7	0.4554	4564	4573	4382	4591	4599	4608	4616	4625	4633	1	2	3	3	4	5	6	7	8
1.8	0.4641	4619	4656	4661	4671	4678	4686	4693	4899	4706	1	1	2	3	4	4	5	6	6
1.9	0.4713	4719	4726	4732	4738	4744	4750	4756	4761	4767	1	1	2	2	3	4	4	5	5
2.0	0.4772	4778	4783	4788	4793	4798	4803	4808	4812	4817	0	1	1	2	2	3	3	4	4
2.1	0.4821	4836	4830	4834	4838	4842	4846	4850	4854	4857	0	1	1	2	2	2	3	3	4
2.2	0.4861	4864	4868	4871	4875	4878	4881	4884	4887	4890	0	1	1	1	2	2	2	3	3
2.3	0.4893	4896	4898	4901	4904	4906	4909	4911	4913	4916	0	0	1	1	1	2	2	2	2
2.4	0.4918	4920	4922	4925	4927	4929	4931	4932	4934	4936	0	0	1	1	1	1	1	2	2
2.5	0.4938	4940	4941	4943	4945	4946	4948	4949	4951	4952									
2.6	0.4953	4955	4956	4957	4959	4960	4961	4982	4963	4964									
2.7	0.4965	4966	4967	4968	4969	4970	4971	4972	4973	4974									
2.8	0.4974	4975	4976	4977	4977	4978	4979	4979	4980	4981									
2.9	0.4961	4982	4982	4983	4984	4984	4985	4985	4986	4986									
3.0	0.4987	4990	4993	4995	4997	4998	4998	4999	4999	5000									

Normal Curve



If the random variable Z is distributed as the standard normal distribution N(0, 1) then

1. $P(0< Z, <z_p) = P$ (shaded area)

2. $P(Z > z_P) = Q = 0.5 - P$

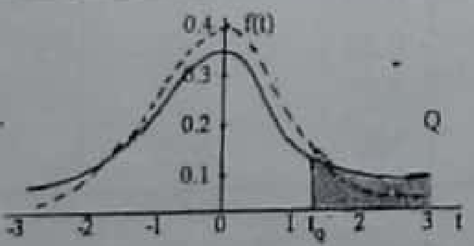
3. $P(Z > |Z_p|) = 1 - 2P = 2Q$

CRITICAL POINTS OF THE NORMAL DISTRIBUTION z_p

P	Q	z	P	Q	
0.00	0.50	0.000	0.490	0.010	2.326
0.05	0.45	0.126	0.491	0.009	2.366
0.10	0.40	0.253	0.492	0.008	2.409
0.15	0.35	0.385	0.493	0.007	2.457
0.20	0.30	0.524	0.494	0.006	2.512
0.25	0.25	0.674	0.495	0.005	2.575
0.30	0.20	0.842	0.496	0.004	2.652
0.35	0.15	1.036	0.497	0.003	2.748
0.40	0.10	1.282	0.498	0.002	2-878
0.45	0.05	1.645	0.499	0.001	3.090
0.450	0.050	1.645	0.4995	$0.0^{3}5$	3.291
0.452	0.048	1.665	0.4999	$0.0^{3}1$	3.719
0.454	0.046	1.685	0.49995	$0.0^{4}5$	3.891
0.456	0.044	1.706	0.49999	$0.0^{4}1$	4.265
0-458	0.042	1.728	0.499995	$0.0^{5}5$	4.417
0.460	0.040	1.751		$0.0^{5}1$	4.753
0.462	0.038	1.774		$0.0^{6}5$	4.892
0.464	0.036	1.799		$0.0^{6}1$	5.199
0.466	0.034	1.825		$0.0^{7}5$	5.327
0.468	0.032	1.852		$0.0^{7}1$	5.612
0.470	0.030	1.881		$0.0^{8}5$	5.731
0.472	0.028	1.911		$0.0^{8}1$	5.998
0,474	0.026	1.943		$0.0^{9}5$	6.109
0.476	0.024	1.977		25%	0.674
0.478	0.022	2.014		10%	1.282
0.480	0.020	2.054		5%	1.645
0.482	0.018	2.097		1%	2.326
0.484	0.016	2.144		0.5%	2.575
0.486	0.014	2.197		0.1%	3.090
0.488	0.012	2.257		0.05%	3.291

PERCENTAGE POINTS OF STUDENTS t- DISTRIBUTION t,
Probability*

v	0.25 0.50	0.10 0.20	0.05 0.10	0.025 0.050	0.01 0.02	0.005 0.010	0.0025 0.0050	0.001 0.002	0.0005 0.0010	Q 2Q
1	1.000	3.078	6.314	12.71	31.82	63.66	127.3	318.3	636.6	
2	0.816	1.886	2.920	4.303	6.965	9.925	14.09	22.33	31.60	
3	0.765	1.638	2.353	3.182	4.541	5.841	7,453	10.21	12.92	
4	0.741	1.533	2.132	2.776	3.747	4.604	5,598	7.173	8.610	
5	0.727	1.476	2.015	2.571	3.365	4.032	4.773	5.893	6.869	
6	0.718	1.440	1.943	2.447	3.143	3.707	4.317	5.208	5.959	
7	0.711	1.415	1.895	2.365	2.998	3.499	4.029	4.785	5.408	
8	0.706	1.397	1.860	2.306	2.896	3.355	3.833	4.501	5.041	
9	0.703	1.383	1.833	2.262	2.821	3.250	3,690	4297	4.781	
10	0.700	1.372	1.812	2.228	2.764	3.169	3.581	4.144	4.587	
11	0.697	1.363	1.796	2.201	2.718	3.106	3.497	4.025	4.437	
12	0.695	1.356	1.771	2.160	2.650	3.012	3.372	3.852	4.221	
14	0.692	1.345	1.761	2.145	2.624	2.977	3.326	3.787	4.140	
15	0.691	1.341	1.753	2.131	2.602	2.947	3.286	3.733	4.073	
16	0.690	1.337	1.746	2.120	2.583	2.921	3.252	3.686	4.015	
17	0.689	1.333	1.740	2.110	2.567	2.898	3.222	3.646	3.965	
18	0.688	1.330	1.734	2.101	2.552	2.878	3.197	3.610	3.922	
19	0.688	1.328	1.729	2.093	2.539	2.861	3.174	3.579	3.883	
20	0.687	1.325	1.725	2.086	2.528	2.845	3.153	3.552	3.850	
21	0.686	1.323	1.721	2.080	2.518	2.831	3.135	3.527	3.819	
22	0.686	1.321	1.717	2.074	2.508	2.819	3.119	3.505	3.792	
23	0.685	1.319	1.714	2.069	2.500	2.807	3.104	3.485	3.767	
24	0.685	1.318	1.711	2.064	2.492	2.797	3.091	3.467	3.745	
25	0.684	1.316	1.708	2.060	2.485	2.787	3.078	3.450	3.725	
26	0.684	1.315	1.706	2.056	2.479	2.779	3,067	3.435	3.707	
27	0.684	1.314	1.703	2.052	2.473	2.771	3,057	3.421	3,690	120
28	0.683	1.313	1.701	2.048	2.467	2.763	3.047	3.408	3.674	V
29	0.683	1.311	1.699	2.045	2.462	2.756	3.038	3.396	3.659	
30	0.683	1.310	1.697	2.042	2.457	2.750	3.030	3.385	3.646	4
40	0.681	1.303	1.684	2.021	2.423	2.704	2.971	3.307	3.551	3
60	0.679	1.296	1.671	2.000	2.390	2.660	2.915	3.232	3.460	2
120	0.677	1.289	1.658	1.980	2.358	2.617	2.860	3.160	3.373	1
α	0.674	1.282	1.645	1.960	2.326	2.576	2.807	3.090	3.291	0



SIGNIFICANCE LEVELS FOR CORRELATION COEFFICINTS

	Product-moment coefficient of correlation(r_{xy})		Spearman's rank correlation coefficient (ρ)		Kendall's rarik correlation coefficient (τ)	
No of Pairs	Signifant If $\lvert r_{xy} \rvert$ exceeds at 5%	at 1%	Significance if $\lvert \rho \rvert$ exceeds at 5%	at 1%	Significance if $\lvert \tau \rvert$ exceeds 5%	1%
3	1.00	1.00				
4	0.95	0.99				
5	0.88	0.96	1.00			
6	0.81	0.92	0.89	1.00	0.87	1.00
7	0.75	0.88	0.75	0.89	0.71	0.81
8	0.71	0.83	0.71	0.86	0.64	0.79
9	0.67	0.80	0.68	0.83	0.56	0.72
10	0.63	0.77	0.65	0.79	0.51	0.64
11	0.60	0.74	0.60	0.74	0.49	060
12	0.58	0.71	0.58	0.71	0.45	0.58
13	0.55	0.68	0.55	0.68		
14	0.53	0.66	0.53	0.66		
15	0.51	0.64	0.51	0.64		
16	0.50	0.62	0.50	0.62		
17	0.48	0.61	0.48	0.61		
18	0.47	0.59	047	0.59		
19	0.46	0.58	0.46	0.58		
20	0.44	0.56	0.44	0.56.	0.33	
30	0.35	0.45	0.35	0.45		
40	0.31	0.39	0.31	0.39		
50	0.27	0.35	0.27	0.35		
60	0.25	0.33	0.25	0.33		
70	0.23	0.31	0.23	0.31		
80	0.22	0.29	0.22	0.29		
90	0.21	0.27	0.21	0.27		
100	0.20	0.25	020	0.25		

This book cover Statistics, Probability and Numerical methods. It provides about Eighty five percent of Advanced Level Mathematics Paper II Course. The scope of Third Edition has been expanded to include Errors, Interpolation, Approximation methods and Flow charts.

The topics are presented in a logic sequence and are arranged in an orderly format to stimulate student's understanding. There are a variety of worked examples and numerous exercises, providing both routine practice and opportunities to develop the understanding.