APPLIED MATHEMATICS WITHOUT TEARS IN
Statistics, Probabilities AND Numerical Methods.

1^{st} **FIRST EDITION**

SSEGAWA YASINI
MSc. Mathematics (MUK)
BSC. EDUC (HONS) (KYU)
KAJJANSI PROGRESSIVE

CONTACTS
TEL. NO. : 0751081978
E. MAIL : ssegawaya@gmail.com

ACKNOWLEDGMENT
I extend my great appreciation to the Director KAJJANSI PROGRESSIVE Secondary School,

SSEGAWA YASINI

INTRODUCTION
I have personally used the notes of this book over many years and with students of a broad range of ability. I have always admired it for its common sense approach to these sections of the subject, for the large number and great variety of it's examples and for the scope and grading of its exercises.
The arrangement in this book is as follows:
Chapter $1-3$, Statistics,
Chapter $4-8$, Probabilities,
Chapter $9-13$ Numerical methods.
Therefore, it has been written in accordance to the demands of *UNEB*. The remaining section of mechanics is to be discussed in the next edition.
There are many questions to try in this book because mathematics is not a spectator sport. It is a game of participants., the only way to learn mathematics is by calculating many numbers.
As for the teachers, this book is designed to serve as a key teaching aid.

Contents

Chapter 1

STATISTICS

1.1 Introduction

This is a branch of mathematics that deals with the observation, collection,recording, analysis and interpretation of data gathered from the field of study.
Statistics can be divided into:

1. Descriptive statistics: This refers to statistics that deals with collection of data of things that have already happened such as collecting data of the population,harvests, age, rainfall, accidents, birth and others.

2. Inferential statistics: This refers to statistics that quantifies uncertainties such as predictions.

NB. **Data**: This refers to facts and figures collected together for a specific purpose.

Classification of Data:

Data can be classified/categorized into:

1. Quantitative Data **Vs** Qualitative Data.
   Qualitative data refers to the data that measure attributes that cannot be quantified such as sex(female or male), color(red,yellow or blue), responses(yes or no) and others, while quantitative data can be represented by numerical quantity like height, mass,time and others.

2. Discrete Data **Vs** Continuous Data: Some data collected can be discrete such as number of phones, number of people, colour of pens e.t.c., while other data can be continuous such as time, age, distance, height e.t.c.

3. Crude Data **Vs** Classified Data.

4. Population **Vs** sample:
   A population is the total set of items under consideration. It's characteristics such as mean, mode median are called parameters
   While a sample is a finite subset of the population. It's characteristics such as sample mean are estimates

Under this text, we are to concentrate on Descriptive statistics.

1.2 Descriptive Statistics.

.
Types of Descriptive Statistics. These include:

(i) Un-grouped Data, (ii) Grouped Data.

1.2.1 Un-grouped data:

This refers to the data where aech observation(data point) is considered independent of the other.E.g

1. The marks of 8 students: 3,4,7,4,8,6,3,5,
2. The height of 6 trees in the compound measured in meters: 1.3,2.4,4.2,3.3,4.4,6.0
3. Number of boys in a school,
4. Number of teachers in the school and others.

1.2.2 Terms used in ungrouped data

(i) Frequency: This refers to the number of times an observation has appeared/occurred.

(i) Frequency distribution table: This summarizes the number of times each observation has occurred. Therefore, it helps to remove repetitions in the data.

Mainly Data analysis is done using the following;

1.2.3 Measures of Central Tendency in Un-grouped Data

These are the parameters(X- tics of the population) that tend to locate the central values. These include mode, mean and median. All these are discussed as below:

- **Mode :** Is the observation (X) that has appeared most.I.e The value of data point (X) with the highest frequency.
- **Mean:** This is also known as the **Average** or **Expectation.**
  It is abbreviated as $\overline{X}$ or μ and it is defined basing on the following cases:
  Case 1: For small observations,

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

Where x = Observation and
n = Number of observations

Cases 2: For highly repeated number of observations:

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

where; Where

x = Observation

f = frequency, for $X = x_1, x_2, x_3, \cdots x_n$

Case3: For big observation:

$$\text{Mean, } \overline{X} = A + \frac{\sum fd}{\sum f}$$

where:

A = assumed / working mean and

$d = X - A$,i.e the deviation of X from working mean A

- **Median:** This is the value of X that lies in the middle of the data after it is arranged in either ascending or descending order.

  If the number data points/ frequency N is an odd number, then **Median** $= (\dfrac{N+1}{2})^{th}$ value

  While if the number of data points/ frequency N is an even number, then

  $$\textbf{Median} = \frac{\text{sum of 2 middle numbers}}{2}$$

  OR

  Median exists in data that has frequency, f by obtaining the cumulative frequency and reading that value of X where $(\frac{1}{2}\sum f)^{th}$ position is contained.This is identified basing on cumulative frequency.

1.2.4 Measures of Dispersion in Un-grouped Data

These are parameters(X-tics) of the population that are used to measure how the variables are spread away from the mean. These include:

(a) **Range:** This is the difference between the largest value and the smallest value of the given data. i.e

Range = Largest Value - smallest Value.

(b) **Deviation Mean.** This is abbreviated as d such that

$$d = \sum(\frac{X - \overline{X}}{n})$$

However; If frequency f is given, then

$$d = \frac{\sum f(X - \overline{X})}{\sum f}$$

Where:

$$X = \text{Data points/ values given.}$$
$$\overline{X} = \text{Calculated mean}$$
$$n = \text{Number of elements given}$$

(c) **Variance:** This is abbreviated as $Var(X)$ or σ^2 such that,
Case 1: For small observations,

$$Var(X) = \left(\frac{\sum x^2}{n} - \overline{x}^2\right) \iff Var(X) = \frac{(X-\overline{x})^2}{n}$$

Where n = Number observation
$\overline{x}$ = Calculated mean and
x = observation.

Cases 2: For highly repeated number of observations:

$$Var(X) = \frac{\sum fx^2}{\sum f} - (\frac{\sum fx}{\sum f})^2$$

Where f = frequency

Case 3: For big observations. If your are required to use the working mean A, then,

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

where $d = X - A$

NB:

(1) At senior six level, **the working mean** (A) can be given to you or not. If its not given please, I advice you to use any value of x with the highest frequency in the table as your A.

(2) For questions of **ESTIMATE in statistics**, your must use either A graph Or Working Mean.

(3) **Sample variance**(s^2) implies **Variance** while **population variance**($\hat{\sigma}^2$) is obtained from **Sample variance** by using the relationship $\hat{\sigma}^2 = \frac{n}{n-1} \times s^2$, which is known as the unbiased estimation for population variance. *This is to be fully covered later.*

1. (d)**The standard deviation:**, This is abbreviated as σ such that

$$\sigma = \sqrt{\text{Variance}}$$

$$\sigma = \sqrt{\frac{\sum fx^2}{\sum f} - (\frac{\sum fx}{\sum f})^2}$$

2. (e)**Quartiles:** This divides the data into four equal portions. They include:

(i) **Lower quartile**: This is abbreviated as q_1 . It covers $\frac{1}{4}$ of the data from the smallest value provided it's arranged in ascending order.

[(ii)] **Upper quartile**: This is abbreviated as q_3 . It covers $\frac{3}{4}$ of the data from the smallest value provided it's arranged in ascending order.

Procedure for obtaining q_1 and q_3

This is done as follows:
- Arrange the data in the ascending order,
- Obtain the median of the given data,
- Omit(drop) all the values of X that are used to obtain the median,
- Obtain the median of all values of X below the obtained median and you take it as your q_1,

However,
- Obtain the median of all values of X above the obtained median and you take it as your q_3.

(f) **Interquartile Range:** This is defined as Interquartile Range $= q_3 - q_1$

(g) **Semi - Interquartile Range:** This is defined as Semi-Interquartile Range $= \dfrac{q_3 - q_1}{2}$

Examples:

(1) The time in minutes taken by the different students on Monday to take break first, was recorded as 8, 9, 2, 1, 3, 2, 3, 4, 6, 2.

(a) calculate;

(i) mean,
(ii) variance,
(iii) mode,
(iv) median,
(v) Range,
(vi) standard deviation,
(vii) lower quartile,
(viii) upper quartile,
(ix) interquartile range,
(x) semi interquartile range,

(b) Estimate the variance.

Solution

(a)(i) mean, $\overline{X} = \frac{\sum X}{n} = \frac{(8+9+2+1+3+2+3+4+6+2)}{10} = 4.00$

$$(ii)\text{Variance}, \sigma^2 = \frac{\sum x^2}{n} - \left(\frac{\sum X}{n}\right)^2$$
$$= \frac{(8^2+9^2+2^2+1^2+3^2+2^2+3^2+4^2+6^2+2^2)}{10} - (4.00)^2$$
$$= 6.800$$

(iii) Mode = 2, B'se it has appeared most

(iv) Median(m): From $1, 2, 2, 2, 3, 3, 4, 6, 8, 9$

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

(v) Range = (largest - smallest) $= (9-1) = 8$

(Vi) Standard deviation$(\sigma) = \sqrt{Var(X)} = \sqrt{6.8} = 2.6077(4dps)$
(vii) and(viii) From $1, 2, 2, 2, 3, 3, 4, 6, 8, 9$

(vii) $q_1 = \dfrac{(2+2)}{2} = 2.0$

(viii) $q_3 = \dfrac{(6+8)}{2} = 7.0$

(ix) Interqurtile range $= (q_3 - q_1) = (7.0 - 2.0) = 5.0$

(x) Semi-Interquartile Range$= \dfrac{(q_3 - q_1)}{2} = \dfrac{(7.0 - 2.0)}{2} = 2.50$

(b) let $A = 3$ such that $d = X - A$
Using:

X	f	d	fd	fd^2
1	1	-2	-2	4
2	3	-1	-3	3
3	2	0	0	0
4	1	1	1	1
6	1	3	3	9
8	1	5	5	25
9	1	6	6	36
	$\sum f = 10$		$\sum fd = 10$	$\sum fd^2 = 78$

$$\begin{aligned}\therefore Var(X) &= \frac{\sum fd^2}{\sum f} - \left(\frac{\sum fd}{\sum f}\right)^2 \\ &= \frac{78}{10} - \left(\frac{10}{10}\right)^2 \\ &= 6.800.\end{aligned}$$

(2) The marks obtained by the students in a certain school for the paper marked out of 15 were as follows 10, 7, 12, 11, 13, 14, 10, 8, 10, 12
calculate;

(i) mean,
(ii) variance,
(iii) mode,
(iv) median,
(v) Range,
(vi) standard deviation,
(vii) lower quartile,
(viii) upper quartile,
(ix) interquartile range,
(x) semi interquartile range.

Solution:
(i) mean, $\overline{X} = \frac{\sum X}{n} = \frac{(10+7+12+11+13+14+10+8+10)}{9} = \dfrac{95}{9} = 10.556$

$$
\begin{aligned}
(ii) Var(X) &= \frac{\sum x^2}{n} - \overline{X}^2 \\
&= \frac{(10^2 + 7^2 + 12^2 + 11^2 + 13^2 + 14^2 + 10^2 + 8^2 + 10^2)}{9} - (\frac{95}{9})^2 \\
&= \frac{1043}{9} - (\frac{95}{9})^2 \\
&= 4.4691
\end{aligned}
$$

(iii) Mode = 10, B'se it has appeared most

(iv) Median(m): From $7, 8, 10, 10, 10, 11, 13, 14$

Median = 10

(v) Range = (largest - smallest) = $(14 - 7) = 7$

(Vi) Standard deviation(σ) $= \sqrt{Var(X)} = \sqrt{\frac{1043}{9} - (\frac{95}{9})^2} = 2.1140(4dps)$

(vii) and(viii) From $7, 8, 10, 10, 10, 11, 13, 14$

(vii) $q_1 = \frac{(8+10)}{2} = 9.0$

(viii) $q_3 = \frac{(12+13)}{2} = 12.5$

(ix) Interqurtile range = $(q_3 - q_1) = (12.5 - 9.0) = 3.5$

(x) Semi-Interquartile Range= $\frac{(q_3 - q_1)}{2} = \frac{(12.5 - 9.0)}{2} = 1.75$

(3) Given the data below showing the height of trees in centimeters: 400,350.460,500,690, 380,410,550 and $500cm$. Find the mean and variance.

solution:

Since we have very big values, the let us apply the formula for big values:

Take working mean $A = 500$

x	f	$d = x - A$	fd	fd^2
350	1	-150	-150	22500
380	1	-120	-120	14400
400	1	-100	-100	10000
410	1	-90	-90	8100
460	1	-40	-40	1600
500	2	0	0	0
550	1	50	50	2500
690	1	190	190	36100
	$\sum f = 9$		$\sum fd$= -260	$\sum fd^2$ =95200

$\Longrightarrow$ Mean $= A + \frac{\sum fd}{\sum f} = [500 + \frac{^-260}{9}] = 471.111$ and

Var(X)= $\frac{\sum fd^2}{\sum f} - \left(\frac{\sum fd}{\sum f}\right)^2 = \frac{95200}{9} - \left(\frac{^-260}{9}\right)^2 = 9743.210$

(4) The table below shows the marks of 40 $S.5$ students in a certain school.

Marks	10	11	12	15	17	18	25
Number of students	40	34	30	20	12	8	3

Find:

(i) Mean,

(ii) Var(X),

(iii) standard deviation,

(iii) mode,

(iv) median,

(v) 90^{th} percentile,

(vi) semi - interquartile range,

(vii) 10^{th} percentile,

(viii) 60^{th} percentile,

(ix) P_{40},

(x) P_{70}.

Solution:

X	f	cf	fX	fX^2
10	6	6	60	600
11	4	10	44	484
12	10	20	120	1440
15	8	28	120	1800
17	4	32	68	1156
18	5	37	90	1620
25	3	40	75	1875
	$\sum f = 40$		$\sum fX = 577$	$\sum fX^2 = 8975$

(i) Mean, $\overline{X} = \frac{\sum fX}{\sum f} = \frac{577}{40} = 14 \cdot 425$

(ii) Standard deviation, $\sigma = \sqrt{\text{variance}}$

But Variance, $\sigma^2 = \frac{\sum fX^2}{\sum f} - \overline{X}^2 = \frac{8975}{40} - 14 \cdot 425^2 = 16 \cdot 2944$

$\Longrightarrow$ Standard deviation, $\sigma = \sqrt{16 \cdot 2944} = 4 \cdot 0366$

(iii) Mode $= 12$

(iv) Median $= 12$

(v) $P_{90} = 18$

Position of $P_{90} = \frac{90}{100} \cdot 40 = 36^{th}$ position.

$\Longrightarrow P_{90} = 18$. I.e from the table.

(vi) Semi - interquartile range $= \frac{(\text{upper quartile}-\text{lower quartile})}{2}$

For upper quartile,

Position of $q_3 = \frac{3}{4} \cdot 40 = 30^{th}$ position.

$\Longrightarrow q_3 = 17$.

Also for lower quartile,

Position of $q_1 = \frac{1}{4} \cdot 40 = 10^{th}$ position.

$\Longrightarrow q_1 = 11$.

$\therefore$ Semi - interquartile range$= \frac{(17-11)}{2} = 3$

Calculate the remaining parts as you exercise.

4 For a particular set of observations $n = 15, \sum X^2 = 17301, \sum X = 400$. Find the values of the mean and the standard deviation.

Solution:

Mean$(\overline{X}) = \dfrac{\sum X}{n} = \dfrac{400}{15} = 26.6667$

Standard deviation$(\sigma) = \sqrt{Var(X)} = \sqrt{\left(\dfrac{17301}{15} - \left(\dfrac{400}{15}\right)^2\right)} = 21.0306(4dps)$

5 For a given frequency distribution $\sum(X - \overline{X})^2 = 182.3, \sum X^2 = 1000, n = 30$ Find the mean and standard deviation of the distribution.

Solution:

For Mean$(\overline{X})$, using $Var(X) = \frac{(X-\overline{X})^2}{n}$,Substituting,

$\dfrac{182.3}{30} = \dfrac{1000}{30} - (\overline{X})^2 \Longleftrightarrow \overline{X} = 5.2208.$

Standard deviation$(\sigma) = \sqrt{Var(X)} = \sqrt{182.3} = 13.5019(4dps)$

1.2.5 EXERCISE: 1.1

1. The marks obtained by the students in a certain school for the paper marked out of 12 were as follows 5, 7, 12, 9, 10, 7, 10, 8, 10, 6

   (a) calculate;

   (i) mean,
   (ii) variance,
   (iii) mode ,
   (iv) median,
   (v) Range,
   (vi) standard deviation,
   (vii) lower quartile,
   (viii) upper quartile,
   (ix) interquartile range,
   (x) semi interquartile range.

   Ans: (i) 8.40, (ii) 4.250,(iii) 10, (iv) 8.50 (v) 7, (vi) 2.0616, (vii) 6.50, (viii) 10, (ix) 4.50, (x) 2.250.

2. The marks obtained by the students in a certain school for the paper marked out of 30 were as follows 15, 27, 12, 29, 10, 17, 28, 10, 4

   (a) calculate;

   (i) mean ,
   (ii) variance ,
   (iii) mode ,
   (iv) median ,
   (v) Range,
   (vi) semi interquartile range.

   Ans: (i) 16.8889, (ii) 73.430, (iii) 10.0, (iv) 15.0, (v) 25, (vi) 8.750.

3. Given the following values:1,2, 3, 5, 4, 6, 7, 9, 8. calculate;

   (i) mean,
   (ii) Standard deviation,
   (iii) mode,
   (iv) median,
   (v) semi interquartile range.

Ans: (i) 5, (ii) 6.6667 (iii) Any, (iv) 5, (v) 2.50.

4. For a particular set of observations $n = 19, \sum X^2 = 17000, \sum x = 563$. Find the values of the mean and the standard deviation.
Ans: $\overline{X}$ = 29.630 , σ = 4.0873

5. For a given frequency distribution $\sum(X - \overline{X})^2 = 121.16, \sum X^2 = 1200, n = 30$ Find the mean and standard deviation of the distribution.
Ans: $\overline{X}$ = 5.9968 , σ = 2.0096

6. The numbers $a, b, 8, 5, 7$ have mean 6 and variance 2. Find the values of a and b , if $a > b$. Hence find the semi- interquartile range.
Ans: a=6,b=4 , 1.5

7. Given the table below, showing the marks of 100 applicants for a certain job.

marks	1	2	3	4	5	6	7
Number of applicants	10	25	50	75	80	93	100

(i) mean,
(ii) Standard deviation,
(iii) mode,
(iv) median,
(v) semi interquartile range.

Ans. (i) 3.607, (ii) 1.6845, (iii) 3 or 4, (iv) 3.5, (v) 1.00

8. The information below is showing marks obtained in test by applicants seeking to become bank managers.

Mark	10	12	13	16	20	26	28
Number of applicants	10	15	25	20	5	17	3

Find

(i) Mean,
(ii) semi - interquartile range,
(iii) variance,
(iv) median.

Ans: (i) (ii)

9. The table below shows the number of children for 100 families.

Children	1	2	3	4	5	6	7	8
Families	8	9	x	25	y	12	6	4

Given that the average number of children per family is 4.20, find the;

(i) values of x and y
(ii) median of the number of children,
(iii) Variance.

Ans. (i) x **=16,** y **= 20 (ii) 4**

10 The numbers $3, 6, 4, 10, 14, 12, a, b$ have mode 6 and mean 7. Calculate;

(i) The value of a and b,

(ii) The variance of these numbers,

(iii) median.

Ans. (i) a **= 6,** b **= 1 (ii) 18.25**

1.3 Grouped data.

Here data is arranged into different classes / groups/ intervals. The data is taken to be continuous .
The classes can be already formulated, partially formulated in form of a table **or** you can be instructed to group them following given intervals.

1.3.1 Special Terms Used in Grouped Data:

(i) Class boundaries: These are the true classes of the given interval obtained by adjusting the class limits using $\pm 0.5 \times 10^{-n}$ where $n =$ number of decimal places.
It's done by subtracting it from the lower class limit and adding it on the upper class limit.
The examples are to be discussed in the section of table interpretation.

(ii) Class width or Class interval: This is obtained by subtracting the class boundaries of a given class. It's the size of the class.

(iii) Frequency: This refers to the number of times the data value/point occurs.The frequency of data value is often represented by f.

(iv) Frequency density($f.d$): This refers to the ratio of frequency of the given class to the interval of that same class. i.e $\boxed{f.d = \frac{Frquency(f)}{Interval(i)} = \frac{f}{i}}$

(v) Cumulative frequency. This is denoted as F. This involves continuously adding the frequency from either the top of the table or bottom of the table. F is either continuously **increasing** or **decreasing**.E.g

Frequency:
4
6
5
10
2

$\Longrightarrow$

F
4
10
15
25
27

OR

F
27
23
17
12
2

1.3.2 Measures of Central Tendency in Grouped Data

.These include;

1. Mean $(\overline{X}) = \frac{\sum fx}{\sum f}$ Where: f = frequency

OR

Mean$\overline{X} = A + \frac{\sum fd}{\sum f}$ Where $d = X - A$ such that $A =$ assumed / working mean

2. Median: This is defined as

$$\text{Median} = L_o + \left(\frac{\frac{1}{2}\sum f - cf_b}{f_m}\right)i,$$

where:
i = Interval of the median class
L_0 = lower class boundary of the median class
f_m = Frequency of the median class
cf_b= cumulative frequency before the median class.
NB $\left(\frac{1}{2}\sum f\right)^{th}$ defines the position of the median class. It's identified from the column of the cumulative frequency.

3. Mode:
CASE 1: **Constant interval**(i).
Here mode is defined by

$$\text{mode} = L_o + \left(\frac{\Delta_1}{\Delta_1 + \Delta_2}\right)i,$$

where:
i = Interval of the modal class
L_0 = lower class boundary of the modal class
$\Delta 1$ = Difference between the modal frequency and frequency before the modal class
Δ_2= Difference between the modal frequency and frequency after the modal class.
NB The modal class in this case is the one with the **highest frequency** in the given table.
CASE 2: **Un constant interval**(i) **or Unequal class intervals**.
Here mode is defined by

$$\text{mode} = L_o + \left(\frac{|\Delta_1|}{|\Delta_1| + |\Delta_2|}\right)i,$$

where:
i = Interval of the modal class
L_0 = lower class boundary of the modal class
$|\Delta 1|$ = Difference between the modal frequency density and frequency density before the modal class
$|\Delta_2|$= Difference between the modal frequency density and frequency density after the modal class.
NB The modal class in this case is the one with the **highest frequency density** in the given table.

1.3.3 Measures of Dispersion in Grouped Data

These include;

(i) **Variance**: This is defined by; Variance, $(\sigma^2) = \frac{\sum fx^2}{\sum f} - (\frac{\sum fx}{\sum f})^2$
However; if your required to use **working mean**, then we use
Variance $= \frac{\sum fd^2}{\sum f} - \left(\frac{\sum fd}{\sum f}\right)^2$: where:

$d = X - A$
$A =$ Assumed mean /Working mean.

(ii) Standard deviation$(\sigma) = \sqrt{\text{Variance}} = \sqrt{\frac{\sum fx^2}{\sum f} - (\frac{\sum fx}{\sum f})^2}$

(iii) **Quartile**: This divides the data into four equal parts: i.e

(a) **Lower Quartile:** This is abbreviated as q_1 such that

$$\text{Lower Quartile}(q_1) = L_o + \left(\frac{\frac{1}{4}\sum f - cf_b}{f_q}\right)i,$$

where:
$i =$ Interval of the lower quartile class
$L_o =$ lower class boundary of the lower quartile class
$f_q =$ Frequency of the lower quartile class
cf_b= cumulative frequency before the lower quartile class.
NB $\left(\frac{1}{4}\sum f\right)^{th}$ defines the position of the lower quartile class. It's identified from the column of the cumulative frequency.

(b) **Upper Quartile:** This is abbreviated as q_3 such that

$$\text{Upper Quartile}(q_3) = L_o + \left(\frac{\frac{3}{4}\sum f - cf_b}{f_3}\right)i,$$

where:
$i =$ Interval of the lower quartile class
$L_o =$ lower class boundary of the lower quartile class
$f_3 =$ Frequency of the lower quartile class
cf_b= cumulative frequency before the lower quartile class.
NB $\left(\frac{3}{4}\sum f\right)^{th}$ defines the position of the upper quartile class. It's identified from the column of the cumulative frequency.

(c) **Interquartile Range:** This is defined as Interquartile Range $= q_3 - q_1$

(d) **Semi - Interquartile Range:**
This is defined as Semi-Interquartile Range $= \dfrac{q_3 - q_1}{2}$

(iii) Percentile: This divides the data into 100 equal parts. It's abbriviated as P_n where n is the n^{th} percentile such that

$$\text{Percentile } (p_n) = L_o + \left(\frac{\frac{n}{100}\sum f - cf_b}{f_n}\right)i,$$

where:
$i =$ Interval of the n^{th} percentile class

L_o = lower class boundary of the n^{th} percentile class
f_3 = Frequency of the n^{th} percentile class
cf_b= cumulative frequency before the n^{th} percentile class.
NB $\left(\frac{n}{100}\sum f\right)^{th}$ defines the position of the n^{th} percentile class. It's identified from the column of the cumulative frequency.

(iv) Decile: This divides the data into 10 equal parts. It's abbriviated as D_n where n is the n^{th} decile such that

$$\text{Decile } (D_n) = L_o + \left(\frac{\frac{n}{10}\sum f - cf_b}{f_n}\right)i,$$

where:
i = Interval of the decile class
L_o = lower class boundary of the decile class
f_n = Frequency of the decile class
cf_b= cumulative frequency before the decile class.
NB $\left(\frac{n}{10}\sum f\right)^{th}$ defines the position of the decile class. It's identified from the column of the cumulative frequency.

(v) **middle n% percentiles:** This gazettes n% percentile in between 0 and 100. This indicates that it is a range of percentile for which the answer must be n% percentile. E.g For middle 60% percentile, we have

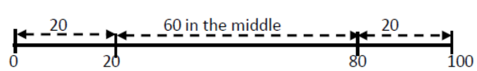


Figure 1.1

$\Longrightarrow$ middle 60% percentile= $p_{80} - p_{20}$
This is similar to middle n^{th} decile.

1.3.4 Table Interpretations in Grouped Data

A learner must be careful as you are handling the information in the table. This is because the information in the table can be in form of:
- Marks Vs Frequency
- Marks Vs cumulative frequecy
- Marks Vs frequency density
- Class boundary Vs Frequency
- Class boundaries Vs cumulative frequecy
- Class boundary Vs frequency density
All these can be interpreted as below;

1.

Marks	Number of students
-10	0
-20	4
-25	10
-38	20
-42	30
-50	36
-56	40

$\Longrightarrow$

Class boundaries	F	f	i
10 - 20	4	4	10
20 - 25	10	6	5
25 - 38	20	10	8
38 - 42	30	10	4
42 - 50	36	6	8
50 - 56	40	4	6

2,

Marks	Number of students
<10	0
<20	4
<25	10
<38	20
<42	30
<50	36
<56	40

$\Longrightarrow$

Class boundaries	F	f	i
10 - 20	4	4	10
20 - 25	10	6	5
25 - 38	20	10	8
38 - 42	30	10	4
42 - 50	36	6	8
50 - 56	40	4	6

3

Marks	Number of students
≤ 10	3
≤ 20	4
≤ 25	6
≤ 38	10
≤ 42	10
≤ 50	6
≤ 55	4

$\Longrightarrow$

Class boundaries	f	i
0 - 10	3	5
10 - 20	4	10
20 - 25	6	5
25 - 38	10	8
38 - 42	10	4
42 - 50	6	8
50 - 55	4	5

I.e **For less, you take** 0, **because it is the smallest relevant value in applied mathematics.**

3

Marks(X)	Number of student.
20 - 29	6
30 - 38	8
39	7
40 - 44	5
45 - 60	9
61 - 80	4

$\Longrightarrow$

Class boundaries	f	i
19.5 - 29.5	3	10
29.5 - 38.5	4	9
38.5 - 39.5	7	1
39.5 - 44.5	5	5
44.5 - 60.5	9	16
60.5 - 80.5	4	20

4

Mass(X)	No. of trees.
24 -	12
36 -	8
56 -	14
60 -	10
70 -	6
84 -	16
95 -	0

=

Mass(X)	No. of trees.
$24 \geq$	12
$36 \geq$	8
$56 \geq$	14
$60 \geq$	10
$70 \geq$	6
$84 \geq$	16
$95 \geq$	0

=

Mass(X)	No. of trees.
$24 >$	12
$36 >$	8
$56 >$	14
$60 >$	10
$70 >$	6
$84 >$	16
$95 >$	0

$\Longrightarrow$

classboundaries	f.	i
24 -36	12	12
36 -56	8	20
56 -60	14	4
60 -70	10	10
70 -84	6	14
84 -95	16	11

I.e you drop the class with frequency zero

5

Marks	Cumulative frequency
Above 5	35
Above 10	32
Above 20	25
Above 28	20
Above 33	10
Above 50	2
Above 78	0

$\Longrightarrow$

classboundaries	F	f	i	
5 - 10	35	3		
10 - 20	32	7		
20 - 28	25	5		
28- 33	20	10		
33 - 50	10	8		
50 - 78	2	2		

6

Marks	Number of students
$5 \leq X \leq 10$	3
$10 \leq X \leq 20$	7
$20 \leq X \leq 30$	5
$30 \leq X \leq 45$	10
$45 \leq X \leq 50$	8
$50 \leq X \leq 70$	2

$\Longrightarrow$

classboundaries	f	i
$5 - 10$	3	
$10 - 20$	7	
$20 - 30$	5	
$30 - 45$	10	
$45 - 50$	8	
$50 - 70$	2	

7

Marks	frequency
5 and below 10	3
below 20	7
below 30	5
below 45	10
below 50	8
below 70	2

=

Marks	frequency
5 up to 10	3
up to 20	7
up to 30	5
up to 45	10
up to 50	8
up to 70	2

$\Longrightarrow$

Classboundary	f	i
5 - 10	3	
10- 20	7	
20 - 30	5	
30 - 45	10	
45 - 50	8	
50 - 70	2	

Marks	Frequency Density
>5	1
> 10	1.5
> 20	0.2
> 30	1
> 45	1
> 50	2
> 60	0

$\Longrightarrow$

Classboundary	f.d	i	f
5 - 10	1	5	5
10- 20	1.5	10	15
20 - 30	0.2	10	2
30 - 45	1	5	5
45 - 50	1	5	5
50 - 60	2	10	20

9

Marks	Cumulative frequency
Above 5	35
Above 10	32
Above 15	25
Above 30	20
Above 35	10
Above 50	6
Above 70	2

$\Longrightarrow$

class boundaries	F	f	i	
5 - 10	35	3	5	
10 - 15	32	7	5	
15 - 30	25	5	15	
30 - 35	20	10	5	
35 - 50	10	4	15	
50 - 70	6	4	20	
70 - 75	2	2	5	

I.e you take the most common interval to obtain the upper class boundary of the last class.

10

Marks(X)	Number of student.
0 - 9	6
10 - 28	8
29 - 39	7
40 - 44	5
45 - 60	9
61 - 80	4

$\Longrightarrow$

Class boundaries	f	i
-0.5 - 9.5	3	10
9.5 - 28.5	4	9
28.5 - 39.5	7	1
39.5 - 44.5	5	5
44.5 - 60.5	9	16
60.5 - 80.5	4	20

I.e You subtract from zero(o) to go to the negative side.

Examples

(1) The table below shows the ages of people on a certain village of census that was carried out in Masaka district in the early 90's,

Age(years)	Number of people
0 – 9	4
10 – 19	10
20 – 29	5
30 – 39	15
40 – 49	8
50 – 59	2
60 – 69	6

Calculate;

(i) Mean

(ii) Mode

(iii) Variace

(iv) Median

(v) lower quartile

(vi) Upper quartile

(vii) semi interqurtile Range

(viii) 3^{th} decile and 9^{th} decile

(ix) Middle 70% percentile range

(x) Number of people whose age is above the modal age

(xi) Number of people with age between 15.6783 years and 50.4217 years.

(xii) Middle 45% percentile range

(xiii) Middle 65% percentile range

(xiv) Standard deviation

Solution

Class Boundaries	f	i	X	fX	fX^2	cf
-0.5 – 9.5	4	10	4.5	18	81	4
9.5 – 19.5	10	10	14.5	145	2102.5	14
19.5 – 29.5	5	10	24.5	122.5	3001.25	19
29.5 – 39.5	15	10	34.5	517.5	17853.75	34
39.5 – 49.5	8	10	44.5	356	15842	42
49.5 – 59.5	2	10	54.5	109	5940.5	44
59.5 – 69.5	6	10	64.5	387	24961.5	50
	$\sum f = 50$			$\sum fX = 1655$	$\sum fX^2 = 69782.5$	

(i) Mean, $(\overline{X}) = \frac{\sum fX}{\sum f} \;=\; \frac{1655}{50} \;=\; 33.100$

(ii) Mode $= L_0 + \left(\frac{\Delta_1}{\Delta_1+\Delta_2}\right)i \;=\; 29.5 + \left(\frac{10}{10+7}\right)10 \;=\; 35.3824$. **I.e By case** 1

(iii) Variance, $(\sigma^2) = \frac{\sum fX^2}{\sum f} - \left(\frac{fx}{\sum f}\right)^2 \;=\; \frac{69782.5}{50} - \left(\frac{1655}{50}\right)^2 \;=\; 300.040$

(iv) Median $= L_0 + \left(\frac{\frac{1}{2}\sum f - cf_b}{f_m}\right)i$

Position of median $= \frac{1}{2}\sum f = \frac{1}{2} \times (50) = 25^{th}$ position

$\Longrightarrow$ Median $= 29.5 + \left(\frac{25-19}{15}\right)10 \;=\; 33.50$

(v) Lower quartile $(Q_1) = L_0 + \left(\frac{\frac{1}{4}\sum f - cf_b}{f_m}\right)i$

Position of Lower quartile $= \frac{1}{4}\sum f = \frac{1}{4} \times (50) = 12.5^{th}$ position

$\Longrightarrow$ Lower quartile $(Q_1) = 9.5 + \left(\frac{12.5-4}{10}\right)10 \;=\; 18$

(vi) Upper quartile, $(Q_3) = L_0 + \left(\frac{\frac{3}{4}\sum f - cf_b}{f_q}\right)i$

Position of Lower quartile $= \frac{3}{4}\sum f = \frac{3}{4} \times (50) = 37.5^{th}$ position

$\Longrightarrow$Upper quartile $(Q_3) = 39.5 + \left(\frac{37.5-34}{8}\right)10 \;=\; 43.875$

(vii) Semi interquartile range $= \dfrac{Q_3 - Q_1}{2} = \dfrac{43.875 - 18}{2} = 12.9375$

(viii) 3^{th} decile $d_3 \;=\; L_0 + \left(\frac{\frac{3}{10}\sum f - cf_b}{f_d}\right)i$

Position of $d_3 = \frac{3}{10}\sum f = \frac{3}{10} \times (50) = 15^{th}$ position

$\Longrightarrow d_3 = L_0 + \left(\frac{\frac{3}{10}\sum f - cf_b}{f_d}\right)i \ = \ 19.5 + \left(\frac{15-14}{5}\right)10 \ = \ 21.5$

Also 9^{th} decile $d_9 \ = \ L_0 + \left(\frac{\frac{9}{10}\sum f - cf_b}{f_d}\right)i$

Position of $d_9 = \frac{9}{10}\sum f = \frac{9}{10} \times (50) = 35^{th}$ position

$\Longrightarrow d_9 = 59.5 + \left(\frac{35-44}{6}\right)10 \ = \ 61.1667$

(ix) Middle 70% percentile = $\mathrm{P}_{85} - \mathrm{P}_{15}$

$\mathrm{P}_{85} = L_0 + \left(\frac{\frac{85}{100}\sum f - cf_b}{f_p}\right)i \ = \ 49.5 + \left(\frac{\frac{85}{100}(50)-42}{2}\right)10 \ = \ 52$

$\mathrm{P}_{15} = L_0 + \left(\frac{\frac{15}{100}\sum f - cf_b}{f_p}\right)i \ = \ 9.5 + \left(\frac{\frac{85}{100}(50)-4}{10}\right)10 \ = \ 13$

$\Longrightarrow$Middle 70% percentile = $52 - 13 \ = \ 39$ years or 13 to 52

(x) Number of people with age above 35.3824 years.

Note: The answer must be a value from the frequency
Method: 1 (Additions)

$\mathrm{N} = 6 + 2 + 8 + \left(\frac{39.5-35.3824}{10}\right)15 \ = \ 22.1764$ people

Method: 2 (Using linear interpolation)

29.5	35.3824	39.5
19	K	34

Where K = Number below modal value

$$\frac{34-19}{39.5-35.3824} \ = \ \frac{K-19}{35.3824-29.5}$$

$$K = 27.8236$$

$\Longrightarrow N = 50 - 27.8236 \ = \ 22.1764$

NB: Much about Linear interpolation is to be covered in NUMERICAL METHODS in this same book.

(xi) Number of people with age between 15.6783 years and 50.4217 years.

Method: 1 (Additions)

Number above 15.6783 years $= 6 + 2 + 8 + 15 + 5 + \left(\frac{19.5-15.6783}{10}\right) \times 10 = 39.8217$

Number above 50.4217 years $= 6 + \left(\frac{59.5-50.4217}{10}\right) \times 2 = 7.81566$

$\Longrightarrow$Those in the range $= 39.8217 - 7.81566 = 32$ people

Alternatively ; Please apply Method 2
♢ **Other parts are left as Exercise.**

(2) The table below shows the height in centimeters 0f 25 students in a certain school.

Height (cm)	< 10	< 20	< 25	< 30	< 50	< 55	< 65
Number of students	0	3	7	15	17	23	25

Calculate:

(i) Mode,
(ii) Mean height,
(iii) Variance,
(iv) Median height,
(v) lower quartile,
(vi) Upper quartile,
(vii) Interqurtile Range,
(viii) 4^{th} decile and 7^{th} decile height,
(ix) Middle 65% percentile range,
(x) Number of students whose height is above the 28 cm,
(xi) Number of student whose height is between 23 cm and 52 cm,
(xii) Middle 45% percentile range,
(xiii) Middle 65% percentile range,
(xiv) Standard deviation.

Solution:

Class Boundaries	cf	f	i	$f.d{=}\frac{f}{i}$	X	fX	fX^2
$10 - 20$	3	3	10	0.3	15	45	625
$20 - 25$	7	4	5	0.8	22.5	90	2025
$25 - 30$	15	8	5	1.6	27.5	220	6050
$30 - 50$	17	2	20	0.1	40	80	3200
$50 - 55$	23	6	5	1.2	52.5	315	16537.5
$55 - 65$	25	2	10	0.2	60	120	7200
		$\sum f = 25$				$\sum fX = 870$	$\sum fX^2 = 35687.5$

(i) Mode $= L_b + \left(\frac{\Delta_1}{\Delta_1+\Delta_2}\right)i \ = \ 25 + \left(\frac{0.8}{0.8+1.5}\right)5 \ = \ 26.7391$ **I.e By case 2**

(ii) Mean, $\overline{X} = \frac{\sum fX}{\sum f} \ = \ \frac{870}{25} \ = \ 34.800$

(iii) Variance, $\sigma^2 = \frac{\sum fX^2}{\sum f} - \overline{X}^2 \ = \ \frac{35687.5}{25} - (34.800)^2 \ = \ 216.46$

(iv) Median $= L_b + \left(\frac{\frac{1}{2}\sum f - cf_b}{f_m}\right)i \ = \ 25 + \left(\frac{\frac{1}{2}(25)-7}{8}\right)5 \ = \ 28.4375$

♢ **Other parts are left as Exercise.**

(3) Given the data below showing the marks of students:

34	40	50	56	71	60
40	88	71	36	73	80
66	44	66	55	81	87
30	89	40	31	42	69
52	50	63	73	85	39
57	46	52	80	53	59

Taking the class interval of 10, construct the frequency distribution table, and hence calculate:

(i) Mode,

(ii) Mean ,

(iii) Variance,

(iv) Median mark,

Solution:

Class limits	f	X	fX	fX^2	Class Boundaries	cf
$30-39$	5	34.5	172.5	5951.250	29.5 - 39.5	5
$40-49$	6	44.5	267.0	11881.50	39.5 - 49.5	11
$50-59$	9	54.5	490.5	26732.250	49.5 - 59.5	20
$60-69$	5	64.5	322.5	20801.250	59.5 - 69.5	25
$70-79$	4	74.5	298.0	22201.0	69.5 - 79.5	29
$80-89$	7	84.5	591.5	49981.750	79.5 - 89.5	36
	$\sum f = 36$		$\sum fX = 2142$	$\sum fX^2 = 137549.0$		

(i) Mode $= 49.5 + \left(\frac{3}{3+4}\right) \times 10 = 53.786.$

(ii) Mean $= \frac{2142}{36} = 59.500$

(iii) Var$(X) = \frac{137549.0}{36} - \left(\frac{2142}{36}\right)^2 = 280.5556.$

(iv) Median $=?$

Position of median $= \frac{1}{2} \times 36 = 18^{th} position.$

$\Longrightarrow$ Median $= 49.5 + \frac{(18-11)}{9} \times 10 = 57.2778$

1.3.5 Graphs in grouped data

In this text,we are to concentrate on only the histogram and Orgive.
These two graphs must be plotted on the graph paper only.

Histogram:

The histogram is plotted using the following two cases:
NB: In each case, make sure that at the origin, you have two starting values one is for the vertical axis while the other is for the horizontal axis.

Case 1: **For constant class intervals:**
Here, a histogram is plotted as frequency against class boundaries. I.e

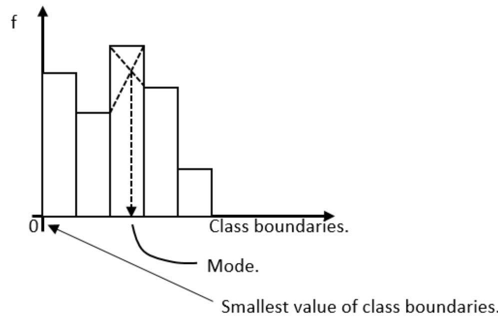


Figure 1.2: All bars are of the same size but of different heights.

Case 2: **For un constant class intervals:**
Here, a histogram is plotted as frequency density against class boundaries. I.e

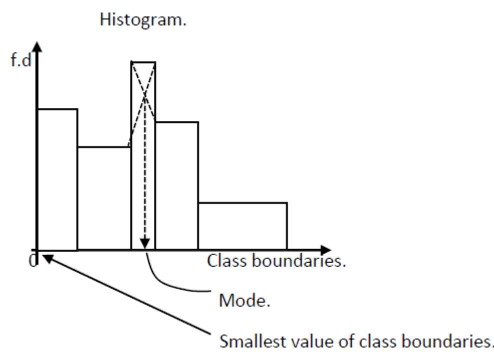


Figure 1.3: All bars are of different sizes except those of the same frequency density and of different heights.

From the histogram, we can estimate mode as shown above.
However, there are special cases of obtaining the mode from the graphs as follows:
(i)

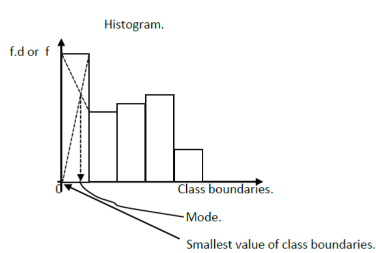


Figure 1.4: Here you use the frequency of the class before the first class as zero.

(ii)

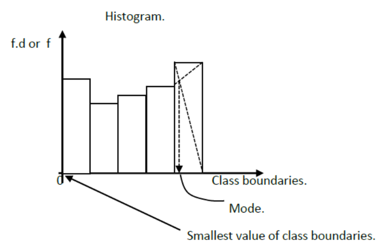


Figure 1.5: Here you use the frequency of the class after the last class as zero.

Orgive:

This is also known as the **cumulative frequency curve**. It's plotted as cumulative frequency against class boundaries.I.e

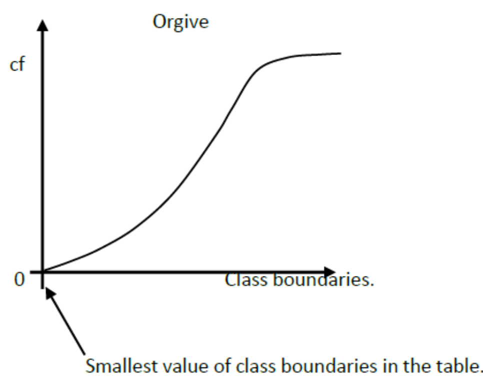


Figure 1.6: Make sure that at the origin, you have two starting values one is for the vertical axis while the other is for the horizontal axis.

From Orgive, we can estimate the following;

- Median,
- Number of items above or below a certain value,
- Quartiles,
- Deciles,
- Percentiles,
- Middle / Central portions.

All these can be obtained as illustrated below:
Example:

1. Given the information below showing marks students.

Height (cm)	> 20	> 25	> 32	> 40	> 45	> 55	> 75	> 80
Number of students	1	3	2	3	12	14	5	0

(a) Draw a histogram and use it to obtain the mode.
(b) Draw an Ogive and hence use it to find

(i) median

(ii) semi - interquartile range

(iii) 80^{th} percentile

(iv) Percentage of students who got below 60 marks

Class Boundaries	f	i	f.d	cf
20 – 25	1	5	0.2	1
25 – 32	3	7	0.43	4
32 – 40	2	8	0.25	6
40 – 45	3	5	0.6	9
45 – 55	12	10	1.2	21
55 – 75	14	20	0.7	35
75 – 80	5	5	1	40
	$\sum f = 40$			

(a)

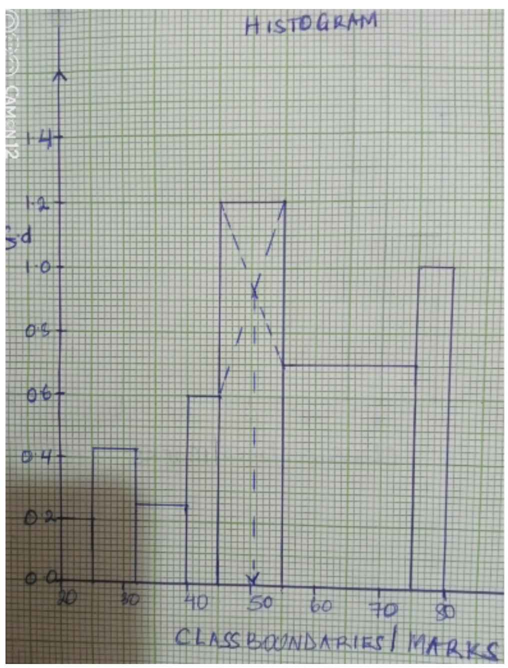


Figure 1.7

GRAPH

(i) Median = 54.5

(ii) Semi - interquartile = $\frac{Q_3 - Q_1}{2} = \frac{63-47}{2} = 8$

(iii) $P_{80} = 71$

(iv) $\frac{24}{40} \times 100 = 60\%$

More examples

1.3.6 Exercise 1.2

(1) The masses measured to the nearest kilogram of 50 boys are noted as below

Marks	Number of students
< 39.5	0
<54.5	4
<69.5	16
< 76.5	40
< 79.5	68
< 84.5	88
< 90.5	100

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

(b) Represent the above information on an orgive and use it to estimate

(i) Median,

(ii) Interquartile range,

(iii) Middle 60% percentile,

(c) Calculate :

(i) Mean,

(ii) Median,

(iii) Mode,

(iv) Variance,

(v) Interquartile range,

(vi) 4^{th} decile and 7^{th} decile,

(vii) 60% percentile ,

(viii) Middle 70% percentile,

(ix) Middle 20% percentile,

(x) 10 to 90$-$ percentile range.

Ans: C(i) 75.580, (ii) 77.560,(iii) 78.076, (iv) 83.336, (v) 9.125, (vi) 76.5 and 80.00, (viii) 15.500

(2) Given the marks scored by some senior five students in $B.O.T$

Marks	Number of students
$10-$	2
$15-$	8
$20-$	17
$30-$	26
$35-$	24
$40-$	16
$50-$	6
$60- < 65$	1

(a)Calculate:

(i) Median,

(ii) Mean,

(iii) mode,

(iv) Probability that a student has a mark between 24% and 38%.

(b) Represent the above information on cumulative frequency curve and use it to estimate the semi intequrtile range.
(c) Represent the above information a histogram and use it to estimate the mode.
Ans. (i) 34.423, (ii) 34.475 (iii) 34.487 (iii) 0.506

(3) Given the information in the table below showing the distribution of the interest paid to 100 shareholders in a certain Association by the end of 2019.

Interest in 1000$ushs$.	25−	30−	40−	60−	80−	110−	120−
Number of share holders	10	25	50	75	80	93	100

(i) Represent the above information on a histogram hence estimate the mode.

(ii) Find the average interest each share holder receives.

(iii) Calculate the standard deviation of the distribution above.

(iv) Using a cumulative frequency curve, find the range of the interest obtained by the middle 80% of the share holders.

Ans. (ii) 54.550 (iii) 95.634

(4) Given the information below showing the marks of test of S.6 mathematics students from the $COVID\_19$ holiday.

Marks (%)	Number of students.
25 − 38	12
39	10
40 − 46	17
47 − 49	13
50 − 59	18
60 − 64	14
65 − 75	8

(a) Represent the information above on a histogram and use it to estimate the modal mark,
(b)Find the;

(i) Mean mark,

(ii) Standard deviation.

(c) Represent the information above on an Orgive and use it to estimate the number of students who obtained above 62%.
Ans. (b)(i) 49.261 (ii) 11.293

(5) Given the information below showing the number of people in millions in a certain country.

Marks	0 – 10	10 – 22	22 – 30	30 – 45	45 – 50	50 – 60	60–
Frequency density	0.2	0.75	1.25	1.5	2	0.5	0.1

(a) Draw a histogram for the above data and use it to estimate the modal age.
(b) Calculate;

(i) Mean age
(ii) Median age
(iii) Standard deviation.
(iv) state the modal age interval

(c) Represent the information above on the orgive, hence use it to estimate the number of people between 20 and 55 year of age.
Ans. (b)(i) (ii) (iii)

(6) The table below shows the masses of 55 girls in a certain school.

Mass (kg)	–40	–49	–59	–65	–74	–78	–80
Number of girls	0	12	20	33	42	50	55

(a) Draw a cumulative frequency curve and estimate:
(i)the range of mass of the central 80% of the girls,
(ii)the number of girls whose mass exceeded the mean mass.
(b) Calculate the;

(i) Mean mass,
(ii) Variance,
(iii) Median,
(iii) Mode.

(c) Draw a histogram for the above data and use it to estimate the mode.
(d) Estimate the standard deviation.
Ans. (b)(i) 53.36 (ii) 55

(8) Given the table below showing the mass in Kg of S.5 science girls.

Mass (kg)	5 – 9	10 – 14	15 – 19	20 – 24	25 – 29	30 – 39
Number of girls	5	14	30	17	11	8

(a) Draw a histogram and from it estimate the mode.
(b) Calculate the;

(i) median,

(ii) Mean,

(iii) Mode,

(iv) Standard deviation.

Ans. (b)(i) (ii) (iii)

(9) The times to the nearest minute taken by 100 students to solve a mathematics problem are shown below.

Time (min)	Number of students
30−	10
50−	30
65−	25
70−	20
75−	15
87−	0

(a)Find the;

(i) Mean

(ii) variance

(ii) median

(b) Draw a histogram and use it to find the modal time of the distribution.
Ans. (i) 62.25 mins (ii) 68.5 mins

(10) 100 players from big football clubs were interviewed about their weekly earnings and the results tabulated as below.

Earnings in 100's of €	Frequency
45−	1
55−	1
65−	2
75−	6
85−	21
95−	29
105−	24
115−	12
125−	4

(a) Calculate the mean earning of the players and standard deviation.

(b) Draw an Ogive and use it to estimate;

(i) interqurtile range,

(ii) Number of players whose earnings are within the middle 65% percentile.

(iii) median

(c)Represent the above information on a histogram and use it to estimate the modal mark. **Ans.(a) 101.2 , 14.5107 (b)(i) 19.5 (ii) 63 players**

Ans. (a)(i) 3.63 (ii) 1

(11) Measurements of the time intervals between successive arrivals of telephone calls at an office exchange were taken. The first 50 time intervals were recorded and the following grouped frequency distribution was obtained.

Time interval (X minutes)	Number of calls.
$0 \leq X \leq 0.5$	50
$0.5 \leq X \leq 1.0$	35
$1.0 \leq X \leq 2.0$	28
$2.0 \leq X \leq 3.0$	14
$3.0 \leq X \leq 6.0$	6

(a) Draw a histogram and use it estimate the mode.

(b) Calculate the mean.

Ans. (a) 0.2 mins (b) 1.11 mins

(12) Given the information below showing the heights in *cm* of the students.

Heights (cm)	Number of students
$50- < 55$	4
$55- < 75$	20
$75- < 85$	15
$85- < 90$	12
$90- < 100$	7
$100- < 120$	2

(a) State the modal height interval.

(b) Construct an Ogive and use it to estimate;

(i) Median height.

(ii) The number of students with height between 60 and 80 cm

(c) Calculate the mean and variance. **Ans.(a) 85 - 90 (b)(i) 79 cm (ii) 22 students**

(13) Given the table below showing the time taken by the students to answer a certain question.

Time(minutes)	Number of students.
≤ 1	0
≤ 4	8
≤ 8	10
≤ 10	6
≤ 13	4
≤ 15	5
≤ 20	2

(a)Calculate the;

(i) Mean ,

(ii) Median,

(iii) Mode,

(iv) variance,

(v) p_{45},

(vi) $D_{6.4}$,

(vii) Interqurtile range,

(viii) middle 27% percentile of the data

(b) Construct an Orgive and use it to estimate;

(i) Median.

(ii) The number of students who passed if the minimum pass mark time is 11 minutes.

Ans. (a)(i) 5.1351 (ii) 3.3293 (iii) 1.0942 (b) mode

(14) The table shows the age distribution of a population of a certain country in a certain year.

Age	Population (Millions)
Under 20	18
-26	16
-30	13
-44	10
-54	14
-64	12
-74	10
-80	6
80 and above	0

(a) Estimate the mean and standard deviation age.

(b) Draw an cumulative frequency curve and use it to determine:

(i) the median,

(ii) if the retirement age of this population is 50 years, how many people are in the age of retirement.

(c) Represent the information above on a histogram and use it to estimate the mode.

Ans. (a)(i) 35.5556 years, 22.4776 years (b)(i) 34 years (ii) 60.2%

(15) The table below shows the marks of mathematics for a group of students.

Marks	Number of students
$- < 15$	5
$- < 25$	10
$- < 33$	14
$- < 45$	18
$- < 69$	7
$- < 75$	2
$- < 85$	4
$- < 95$	1

(a) Calculate the mean and modal mark.

(b) Draw a cumulative frequency curve, hence estimate:

(i) the Median,

(ii) the 20^{th} to 80^{th} percentile range,

(iii) if the pass mark is 40%, how many people passed.

(iv) if only 8 students passed, what was the pass mark.

(c) Represent the above information on a histogram and use it to estimate the modal mark.

Ans.(a) 38.7692, 36.4286 (b)(i) 36.5 (ii) 23

(16) The table below shows masses of objects in kg obtained.

Mass (kg)	$20 - 24$	$25 - 29$	$30 - 34$	$35 - 39$	$40 - 49$	$50 - 54$	$55 - 64$
Frequency density	0.4	1.2	1.4	2.2	1.8	1.6	1.2

(a) Calculate the;

(i) Modal mass,

(ii) Mean mass,

(iii) Standard deviation,

(iv) Median.

(b) Represent the above information on a histogram and use it to estimate the modal mark.

(c) Represent the above information on an orgive and use it to estimate the minimum mass within which the top 10% objects lie.

Ans. (i) 37.833 (ii) 43.25

(17) The histogram below represents the masses of monkeys (kg) and their respective frequency densities.

GRAPH

(a) Construct the frequency distribution table.

(b) Calculate the;

(i) median mass

(ii) upper quartile

(iii) mean mass

(iv) percentage of the monkeys within 8 kg of the median mas

Ans. (b)(i) 28.5 kg (ii) 36.5 kg (iii) 31.419 kg (iv) 68.55%

(17) The table below shows the number of people in millions in different age groups in a certain country.

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

Calculate

(i) the mean age

(ii) standard deviation

(b) Draw a histogram to represent the above data and use it to estimate the mode.

Ans. (a)(i) 34 years (ii) 15.556 years (b) 35 years

(18) The table below shows the blood pressure of s.6 students at SCOBA.

Blood pressure	95−	105−	110−	115−	120−	125−	130−	140−
Frequency	2	5	6	9	14	3	6	5

(a) Calculate

(i) Mean blood pressure.

(ii) variance,

(ii) mode,

(ii) median,

(b) Draw an Ogive and estimate the lower and upper quartile blood pressure. (c) Draw a histogram and use it to estimate the mode.

Ans. (a)(i) 122.05 (ii) 28 students (b) 115, 128 Given the data below showing the marks of students:

44	46	57	80	71	60
40	68	71	49	64	71
47	64	66	75	80	67
65	89	60	61	42	59
73	40	63	73	85	59
51	56	52	79	53	69

Taking the class interval of 8, construct the frequency distribution table, and hence (a) calculate:

(i) Mode,

(ii) Mean ,

(iii) Standard deviation,

(b) Represent the above data on an cumulative frequency curve and use it to estimate the median.

NOTE: More Exercise on this Chapter refer to ***Wonderful touch.***

Chapter 2

INDEX NUMBERS /THE PRICE INDICES

2.1 INDEX NUMBERS

2.1.1 Introduction

In economics and finance, an index is a statistical measure of change in a representative group of individual data points. This data may be derived from a number of sources, including company performance, prices, productivity, employment. and others.
Index numbers are statistical measures designed to show changes in a variable or a group of related variables with respect to time, geographical location or other characteristics such as income, profession and others.
They provide a measures of the relative change in some variable or group of variables at a specified date when compared to with some fixed period in the past.
Index numbers may be classified in terms of the variables that they are intended to measure. In business, different groups of variables in the measurement of which index number techniques are commonly used are: price,quantity, standards of living (values) and others.
Base Year: In Economics, this is the year or period when the prices of the item(s) is/are said to be relatively stable. while;
In mathematics, this is the year or period against which all other years or periods are compared.
The base year is normally given a standard statistical value or measure of 100
NB: The base year price is also known as the 100% price.
Current Year: This is the year for which the index is to be computed. It's also known as the **given Year**.

2.1.2 Terms Used in index numbers

(i) **Price relative:** This is also known as Simple index number or price index or simple price index. It's abbreviated as P_r such that;

$$\text{Price relative } (P_r) = \frac{P_1}{P_0} \times 100 \text{ or } \frac{P_1}{P_0}$$

where:
P_1 = Current Year price of the item.
P_0 = Base Year price of the item.
NB: The percentage sign (%) can be dropped or maintained.
Procedure for Drawing Conclusion;
- If the price relative is greater than 100, then it implies that the price of an item (commodity) in the question increased over a given period of time. This increase can be obtained by using price relative value that is multiplied with 100 minus 100
- If the price relative is less than 100, then it implies that the price of an item (commodities) in the question decreased over a given period of time. This decrease can be obtained by using 100 - price relative value that is multiplied with 100

(ii) Simple Aggregate Index: This is mainly divided into three categories. I.e
- Simple Aggregate price index = $\frac{\sum p_1}{\sum p_0} \times 100$. This is used when the **price** is given.

- Simple Aggregate quantum index = $\frac{\sum Q_1}{\sum Q_0} \times 100$. This is used when the **quantity** is given.
- Simple Aggregate wage index = $\frac{\sum W_1}{\sum W_0} \times 100$. This is used when the **Wage** is given.

(iii) Simple average price index/ Simple index number/ Cost of living index: This is applicable when weights are **not** given. It's given as;

$$\text{Simple average price index} = \frac{\sum \frac{p_1}{p_0} \times 100}{n}$$

where n is number of items.

(iv) Weighted Aggregate price index/composite index

$$\text{composite index} = \frac{\sum WP_1}{\sum WP_0} \times 100$$

(v) Weighted Average price index/ cost of living index/ weighted index.

$$\text{Weighted Average price index} = \frac{\sum \frac{p_1}{p_0} \times w}{\sum w} \times 100$$

where w is weight of items
NB: Weight of items: This refers to the relative importance of the goods and services (item) as measured by their shares in the total consumption of households.
NB: Cost of living index means **Weighted Average price index** when the weights are **given** while it means **Simple average price index** when weights are **not given**.
NB: If the term **VALUE** or **QUANTITY** is used in addition of the above terms, then you must include quantity whenever you use price P like

$$\text{Value index/ composite index} = \frac{\sum p_1 q_1}{\sum p_0 q_1} \times 100 \text{ OR } \frac{\sum p_1 q_0}{\sum p_0 q_0} \times 100$$

NB: If two different weights w_o and w_1 are given and you are required to find the Average weighted index, then you choose one of the weights and move with it in the calculation. I.e

$$\text{Average weighted index} = \frac{\sum p_1 w_1}{\sum p_0 w_1} \times 100 \text{ OR } \frac{\sum p_1 w_0}{\sum p_0 w_0} \times 100$$

In this case, the two answers will be different but both answers are correct.

2.1.3 Examples

(1) In April 2011, the price of a kilogram of sugar was $shs.4310$. In April 2013, the price was $shs.5000$. Taking 2011 as the base year, find the price relative and comment on your answer.

Solution:

Price relative $= \frac{p_{2013}}{p_{2011}} \times 100 = \frac{5000}{4310} \times 100 = 139.211$

Comment/Conclusion: The price of sugar increased by 39.211% between 2011 to 2013.

(2) In April 2011, the price of a kilogram of salt was $shs.3750$. In April 2013, the price was $shs.3000$. Taking 2011 as the base year, find the price relative and comment on your answer.

Solution:

Price relative $= \frac{p_{2013}}{p_{2011}} \times 100 = \frac{3000}{3750} \times 100 = 80.0$

Comment/Conclusion: The price of salt decreased by 20.0% between 2011 to 2013.

(3) The price of suck of grasshoppers in 2016 was $sh.80,000$, and in 2018 it was $sh.95,000$. Calculate the simple price index using 2016 as a base year. Comment on your results.

Solution:

Simple price index $= \frac{p_1}{p_0} \times 100 = \frac{95{,}000}{80{,}000} \times 100 = 118.750$

Comment: The price of grasshoppers increased by 18.750% over the time.

(4) The price of a sanitizer in 2019 was $sh.4000$, and in 2020 it was $sh.10,000$. Calculate the price index taking 2019 as a 100%. Comment on your results.

Solution:

Simple price index $= \frac{p_1}{p_0} \times 100 = \frac{10{,}000}{4{,}000} \times 100 = 250.00$

Comment: The price of a sanitizer increased by 150.00% over the time.

(5) The wages of a certain group of workers in 2014 was $shs.200,000$ and in 2018 it was $240,000$. Using 2014 as the 100%, calculate the workers' wage index in 2018

Solution:

Simple wage index $= \frac{W_1}{W_0} \times 100 = \frac{200{,}000}{240{,}000} \times 100 = 120.00$

Comment: The wages of a certain group of workers increased by 20.00% between 2014 and 2018.

(6) In 2000, the price of an item using 1999 as the base year was 166. In 2006, the index using 2000 as the base year was 123. what is the index in2006 using 1999 as the base year?

Solution:

$\frac{p_{2000}}{p_{1999}} \times 100 \;=\; \frac{166}{100} \cdots\cdots\cdots\cdots\cdots\cdots\cdots\cdots\cdots\cdots\cdots\cdots\cdots (i)$

$\frac{p_{2006}}{p_{2000}} \times 100 \;=\; \frac{123}{100} \cdots\cdots\cdots\cdots\cdots\cdots\cdots\cdots\cdots\cdots\cdots\cdots\cdots\cdots (ii)$

Now,from (i) and (ii), we get;

$\Longrightarrow \frac{p_{2006}}{p_{2001}} \times 100 \;=\; \left(\frac{p_{2006}}{p_{2000}} \times \frac{p_{2000}}{p_{2001}}\right) = \left(\frac{123}{100} \times \frac{100}{166}\right) = 204.180$

Comment: The price of the item increased by 104.180 between 1999 and 2006

(7) The table below shows the price in shilling of items A, B and C and their weights in 2010 and 2014.

Items	price 2010	price 2014	weight
A	$shs.$1500	$shs.$1800	4
B	$shs.$2500	$shs.$2800	6
C	$shs.$900	$shs.$800	5

Taking 2010 as the base year, calculate;

(a) simple aggregate price index for 2009.

(b) Price relative for each item for 2009

(c) Find weighted price index for 2009.

(d) Find weighted aggregate price index for 2009.

Solution:

(a) Simple aggregate price index $=\frac{\sum p_1}{\sum p_0} \times 100 \;=\; \frac{1800+2800+800}{1500+2500+900} \times \;=\; \frac{5400}{4900} \times 100 \;=\; 110.2$

(b) For A, $\text{P}_r = \frac{p_1}{p_0} \times 100 \;=\; \frac{1800}{1500} \times 100 \;=\; 120.00\%$

For B, $\text{P}_r = \frac{p_1}{p_0} \times 100 \;=\; \frac{2800}{2500} \times 100 \;=\; 112.00\%$

For$C$, $\text{P}_r = \frac{p_1}{p_0} \times 100 \;=\; \frac{800}{900} \times 100 \;=\; 88.890\%$

(c) Weighted index $= \frac{\sum \frac{p_1}{p_0} \times 100 \times w}{\sum w} \;=\; \frac{(120\times4)+(112\times6)+(88.89\times5)}{4+6+5} \;=\; 106.43$

(d) Weighted aggregate price index $= \frac{\sum WP_1}{\sum WP_0} \times 100$

$= \dfrac{[(1800 \times 4) + (2800 \times 6) + (800 \times 5)]}{[(1500 \times 4) + (2500 \times 6) + (900 \times 5)]} \times 100 = 109.804\%$

(8) The table below shows the cost of components A, B, C and D used in making a complete refrigerator in 2014 and 2017 and quantities required.

Item	COSTS (SHS)		QUANTITY	
	2014	2017	2014	2017
Bees Honey(tin)	7000	8500	7	9
Soap (Bar)	4500	4000	5	4
Maize flour(kg)	2800	3500	6	8
Milk flour(0.25kg)	5600	6000	4	6

$$\text{“2014”} = 100\%$$

(a)Calculate

(i) simple aggregate quantity index.

(ii) Wieghted aggregate price index.

(iii) Wieghted average price index , comment on your results

(b) If the cost of complete refrigerator was 650,000/= in 2017, using the Index (a)(ii), find the cost in 2014.

Solution

(a)(i) Simple aggregate quantity index $=\frac{\sum q_1}{\sum q_0}\times 100 \ = \ \frac{(9+4+8+6)}{(7+5+6+4)}\times \ = \ \frac{27}{22}\times 100 \ = \ 122.727$
The quantities increased by 22.727% over the time.

(ii)
$$\begin{aligned}\text{Wieghted aggregate price index} &= \frac{\sum p_1q_1}{\sum p_0q_0}\times 100\\ &= \frac{(8500\times 9)+(4000\times 4)+(3500\times 8)+(6000\times 6)}{(7000\times 7)+(4500\times 5)+(2800\times 6)+(5600\times 4)}\times 100\\ &= 141.373\%\end{aligned}$$

The composite index increased by 41.373 over the period of time.

$$\begin{aligned}(iii)\ \text{Wieghted average price index} &= \frac{\sum p_1q_1}{\sum p_0q_1}\times 100\\ &= \frac{(8500\times 9)+(4000\times 4)+(3500\times 8)+(6000\times 6)}{(7000\times 9)+(4500\times 4)+(2800\times 8)+(5600\times 6)}\times 100\\ &= \frac{156500}{137000}\times 100\\ &= 114.234.\end{aligned}$$

The standards of living increased by by 14.234% in 2017 Or

(iii) Wieghted average price index $= \dfrac{\sum p_1 q_0}{\sum p_0 q_0} \times 100$

$$= \frac{(8500 \times 7) + (4000 \times 5) + (3500 \times 6) + (6000 \times 4)}{(7000 \times 7) + (4500 \times 5) + (2800 \times 6) + (5600 \times 4)} \times 100$$
$$= \frac{124500}{110700} \times 100$$
$$= 112.466.$$

The standards of living increased by by 12.466% in 2017

(b) Index $= \frac{c_1}{c_0} \times 100 = 114.23$

$\frac{650000}{c_0} \times 100 = 114.23$

$\Longrightarrow c_0 = 569027.4/=$

(9) The price index of a tray of eggs in 2002 based on 2000 was 125, the price index for tray of eggs in 2006 based on 2002 was 90. calculate the;

(a) Price index of the tray of eggs in 2006 based on 2000.

(b) the price of the tray of eggs in 2006 if the price of a tray of eggs in 2000 was 11000.

Solution

Let $P_{2000} = A, \qquad P_{2002} = B, \qquad P_{2006} = C$

$\frac{B}{A} \times 100 = 125$

$B = 1.25A \cdots\cdots\cdots\cdots\cdots\cdots\cdots\cdots\cdots\cdots\cdots\cdots (\text{i})$

$\frac{C}{B} \times 100 = 90$

$B = \frac{C}{0.9} \cdots\cdots\cdots\cdots\cdots\cdots\cdots\cdots\cdots\cdots\cdots\cdots\cdots (\text{ii})$

(i) = (ii)

$\frac{C}{0.9} = 1.25A$

$\frac{C}{A} \times 100 = 112.5$

$\frac{C}{11000} \times 100 = 112.5$

$C = 12375/=$

2.1.4 Exercise 2

(1) The table shows price of elements and their quantities in 2003 and 2005.

Item	Price per unit(U)		Quantities	
	2003	2005	2003	2005
A	1000	1200	36	42
B	1100	1000	69	88
C	500	650	10	12
D	800	850	11	10

Using“2003” = 100

(a)Calculate

(i) The price relatives for 2005 for each item.

(ii) The wieghted aggregate price index.

(iii) value index .

(b) If complete assembled radio costs 70,000shs in 2005, using index in (a)(ii) above, estimate the price of a radio in 2003.

Ans. (a)(i) 120 , 90.91, 130, 106.25 (ii) 101.24 (iii) 123.07 (b) 69142.63/=

(2) In 2002, the price of an item using 1999 as the base year was 96. In 2007, the index using 2002 as the base year was 120. what is the index in2007 using 1999 as the base year?

(3) The following information relates to three products sold by a company in the year 2000 and 2004.

Product	2000		2004	
	Quantity(Kg)	Selling price per unit in $	Quantity (Kg)	Selling price per unit in $
P	76	0.60	72	0.18
Q	52	0.75	60	1.00
R	28	1.10	40	1.32

Taking 2000 as the base year,
(a)Calculate the:

(i) Percentage increase in sales over the period (Price relative of each item),

(ii) Simple aggregate price index.

(b) If the total expenditure in 2004 was 200$ and the Simple aggregate price index is 88.588%, what would be the total expenditure in 2000? **Ans. (a) 8.98%**

(4) The table below shows the expenditure of a certain family for the months April and November 2019.

Item	Expenditure (shs)		weight
	April	November	
Food	200,000	225,0000	2
Bills	160,000	260,0000	7
Others	500,000	625,000	1

(a) Calculate the price index of each item taking April as the base year.
(b) Calculate the cost of living index for the month of November, comment on your results.

Ans. 148.75, The cost of living of the family increased by 48.8% in November

(5) The table below shows Marion's expenditure on break time during School time in term one and term two 2015.

Item	Term one	Term two
sodas	800	1200
chapatti	1000	1500
samosa	900	1200
sweets	300	400

(a) Calculate the simple price index of each item in term two,
(a) Calculate Marion's cost of living index in term two,
(b) Calculate Marion's simple aggregate index for term two taking term one as the 100% year.
Ans. (a) 141.67 (b) 143.33

(6) The cost of living index is calculated using weights as shown in the table below with base year 2010.

Item	Weight
Food	4
Housing	3
Clothing and fuel	2
Others	1

It is known that the value of index for 2012 is 110 and for 2015 is 121. In 2012 the cost of food had risen by 9% and the cost of housing had risen by 8% above their 2010 values. In 2015, the cost of food and housing had risen by 16% and 20% respectively above their 2010 values. Clothing and fuel increased by α% and other items by β% respectively above their previous year's values. Find the possible values of α and β.

(7) The table below shows the expenditure of restaurants for the years 2014 and 2016.

Items	Price(Shs)		Weight
	2014	2016	
Milk (per litre)	1000	1300	0.5
Eggs (per tray)	6500	8300	1
Sugar (per kg)	3000	3800	2
Blue band	7000	9000	1

Taking 2014 as the base year, calculate for 2016 the;

(a) Price relative for each item.

(b) Simple aggregate Price index.

(c) Weighted aggregate Price index and comment on your results.

(d) In 2016, the restaurant spent $Shs.45,000$ on buying these items. Using the index obtained in (c), find how much money the restaurant could have spent in 2014.

Ans. (a) 130, 127.69 , 126.67 , 128.57 (b) 128 (c) 127.75, the expenditure increased by 27.8% (d) 35225.05/=

(8) The table below shows the cost of ingredients used for making chapatti for two different birthday parties for 2015 and 2017.

Ingredients	Cost 2015	Cost 2017
Salt	200	350
Baking flour	3800	4600
Cooking oil	1500	1800

(i) Calculate the price relative for each ingredient and hence obtain the average index number.

Ans. (i) 175 , 121.05 , 120, average index = 138.68

(9) The table below shows the average monthly wage in thousands of shillings of workers in category in 2012 and 2014 a certain soft drinks factory.

Category	Monthly wage 2012	Monthly wage 2014	Index number	Number of workers
1	120	192	160	180
2	150	285	Z	165
3	X	330	200	100
4	170	Y	250	55

Taking 2012 = 100%, find the;

(i) values of X, Y and Z.

(ii) weighted index number for the monthly wage of the whole factory in 2014.

Ans. (i) 165, 425 , 190 (ii) 187

(10) A manufacturer makes a cereal which is a mixture of three different elements. The table below shows the cost per kg of these elements for the year 2012 and 2014.

Element	Cost per KG 2012	Cost per KG 2014	Weight
A	5,000	6,250	1
B	2,000	2,250	2
C	1,600	2,400	7

Taking 2012 as base year, calculate the;

(i) simple aggregate index number.

(ii) simple price index for each element, hence determine the composite index number for the cost of the cereal in 2014.

Ans. (i) 126.7442 (ii) 125, 112.5, 150, composite index = 140

(11) The table shows the allowances of workers in thousands of shillings in centenary bank in 2014 and 2018.

Department	Monthly allowance 2014	2018	Number of workers
Credit section	1200	1920	180
Loans section	1500	2850	165
Supervisors	1650	3300	100
Manager	1700	4250	55

(i) Calculate the weighted aggregate index number for the monthly allowance in 2018. Comment on your results.

Ans. (i) 191.1, allowances increased by 91.1%

(12) The table below shows price relatives (in percentage) foe three types of detegents sold in packets of small, medium and large size by two super markets, Shoprite and Tuskeys.

Detergent	small	medium	large
Shoprite	90	110	98
Tuskeys	85	115	82

A trader decides to buy small, medium and large from both super markets in the ratio 1 : 2 : 2 respectively.
If she spent $shs.95000/=$ in Tuskeys, how much did the trader spend in Shoprite?

Ans. 96446.7

(13) The table below shows the average retail prices in shillings of a kilogram of sugar during the year 1983 – 1988

Year	1983	1984	1985	1986	1987	1988
Retail price	110	120	130	150	165	185

(a) Using 1983 as the base year, find the price index corresponding to the years 1986 and 1988. By how much would a family have reduced their consumption of sugar in 1988 if they had to spend the same amount as they did in 1983?

(b) Using 1983 – 1985 as a base year, find the retail price in 1989 if the price index was 160.

Ans. (a)136.364, 168.182, by 68.182 (b) 208

Chapter 3

Correlation:

This deals with the relationship between two variables where one variable is said to be independent while the other is dependent. This form of relationship is called **Bi-variates**.

3.1 RANK CORRELATION

This is an approach of determining the degree of the relationship between two variables by ranking them.
It involves showing the relation between two variables by using a formula.I.e by calculations.

Introduction

A correlation is a relation between a pair of observations.
To rank means to give positions to data points either in descending or ascending order. There are two common ways of ranking data i.e Spearman's and Kendall's rank correlation.

3.1.1 Spearman's Rank correlation

This is used the find the rank correlation coefficients and it's a abbreviated as ρ.
The spearman's rank correlation co-efficient (ρ), ranks data by giving positions to data in either descending or ascending order. In the cases where we have repeating values of data, we give such values the same positions by getting the average of the positions that they would have assumed and consider that.
The spearman's rank correlation co-efficient $(\rho) = 1 - \dfrac{6\sum d^2}{n(n^2-1)}$
where
$n =$ number of pair of observations or pairs, and
$d =$ the difference between rankings.
Drawing Conclusion: This is also referred to as to comment on the answer.
This involves **interpreting of the correlation** or relation between the variables using the correlation coefficients obtained from the calculation. We have majorly two ways of concluding and these include:

(a) In case your told "to comment on your answer", we use the tables below.
For positive answer,

Spearman's value range	comment
$0.0 -^{+} 0.19$	Very low positive correlation
$0.2 - 0.39$	Low(weak) positive correlation
$0.4 - 0.59$	Moderate positive correlation
$0.6 - 0.79$	High(substantial) positive correlation
$0.8 - 0.99$	Very high positive correlation
$^{+}1$	perfect positive correlation

For negative answer,

Spearman's value range	comment
$0.0 -^{-} 0.19$	Very low negative correlation
$^{-}0.2 -^{-} 0.39$	Low(weak) negative correlation
$^{-}0.4 -^{-} 0.59$	Moderate negative correlation
$^{-}0.6 -^{-} 0.79$	High(substantial) negative correlation
$^{-}0.8 -^{-} 0.99$	Very high negative correlation
$^{-}1$	perfect negative correlation

NB: We use the idea of Size, Sign correlation. While the work *coefficient* should not appear in your conclusion.

NB: If your to comment using the above tables, put your answer to **two** decimal places.

(b) In case your are at a given level of significant.
Here you will be required to comment at either 1% or 5% level of significance. Here we use the significance levels for correlation coefficient. This is a the back of this same book.
I.e *If magnitude of the value calculated is more than/ exceeds that in the mathematical tables at that vales of n at either 1% or 5%, then we say that its significant at that level otherwise its not significant.*

Examples:

(1) The table below shows marks obtained in Mathematics and Physics tests by 10 students.

Maths	50	60	55	88	72	68	40	86	65	80
Physics	62	80	35	70	48	85	53	68	75	90

(a) Calculate the Spearman's rank correlation,
(b) Comment on your results,
(c) Comment on your results on your result at 5% level of significant.

Solution:

NB: In this text, we are going to give *rank 1* to the largest value unless otherwise.

Maths	Physics	R_m	R_p	$d = (R_m - R_p)$	d^2
50	62	9	7	2	4
60	80	7	3	4	16
55	35	8	10	-2	4
88	70	1	5	-4	16
72	48	4	9	-5	25
68	85	5	2	3	9
40	53	10	8	2	4
86	68	2	6	-4	16
65	75	6	4	2	4
80	90	3	1	2	4
					$\sum d^2 = 102$

Using $\rho = 1 - \dfrac{6\sum d^2}{n(n^2-1)} \; = \; 1 - \dfrac{6(102)}{10(10^2-1)} \; = \; 0.3818.$

(b)Comment: There is low positive correlation.

(c) Since $|0.3818| < 0.65$, then the correlation is not significant at 5%

(2) The table below shows marks obtained in History and Geography tests by 10 students.

History	80	80	70	60	65	80	68	90	95	50
Geography	50	45	70	80	70	90	70	80	70	95

(a) Calculate the Spearman's rank correlation,

(b) Comment on your results,

(c) Comment on your results on your result at 1% level of significant.

Solution:

Remember, for similar marks, we take the average of positions.

History	Geography	R_H	R_G	$d = (R_H - R_G)$	d^2
70	50	5.5	9	-3.5	12.25
80	45	3.5	10	-6.5	42.250
70	70	5.5	6.5	-1	1
60	80	9	3.5	5.5	30.250
65	70	8	6.5	1.5	2.25
80	90	3.5	2	1.5	2.25
68	70	7	6.5	0.5	0.25
90	80	2	3.5	-1.5	2.25
95	70	1	6.5	-5.5	30.25
50	95	10	1	9	81
					$\sum d^2 = 204$

Using $\rho = \; 1 - \dfrac{6(204)}{10(100-1)} \; = \; -0.2364.$

(b)Comment: There is low negative correlation.

(c) Since $|-0.2364| < 0.65$, then the correlation is not significant at 5%.

(3) The table shows grades obtained in general paper and Economics by 8 students.

G.P	P7	C3	D1	C4	C3	D2	F9	C6
Econ	F	C	A	D	B	A	E	O

Calculate the rank correlation. Comment on your results. **Solution:**
Here we base on the known a warding of $UNEB$ to rank them.

History	Geography	R_m	R_p	d^2
P7	F	7	8	1
C3	C	3.5	4	0.25
D1	A	1	1.5	0.25
C4	D	5	5	0
C3	B	3.5	3	0.25
D2	A	2	1.5	0.25
F9	E	8	6	4
C6	0	6	7	1
				$\sum d^2 = 9$

Using $\rho = \; 1 - \dfrac{6(7)}{8(64-1)} \; = \; 0.80.$

Comment: There is very high positive correlation.

3.1.2 Kendall's Rank Correlation

For Kendall's rank correlation coefficient (τ), this is such that

$$\tau = \frac{\text{agreements} - \text{disagreements}}{\text{total number of pairs}} = \frac{S}{\frac{1}{2}n(n-1)}$$

$$\implies \tau = \frac{2S}{n(n-1)}$$

Where S is the total sum of all the scores and $\frac{1}{2}n(n-1)$ is the total number of pairs for a set of n objects.

Procedure for Obtaining Score:

1. Arrange the data in two rows with the first row first row in the descending order.
2. Arrange the data in accordance with that of the first row.
3. Obtain the 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 given to the pair of object in the right order and $^{-}1$ for a pair not in the right order.

5. Add all the scores which is taken as S.

Examples.
Two examiners Y and Z each marked the scripts of ten students who sat a test. The table below shows the examiner's ranking of the students.

Examiner	A	B	C	D	E	F	G	H	I	J
Y	5	3	6	1	4	7	2	10	8	9
X	6	3	7	2	5	4	1	10	9	8

Calculate the Kendell's correlation co-efficient for the two examiners.

Solution:

Examiner	D	G	B	E	A	C	F	I	J	H
R_Y	1	2	3	4	5	6	7	8	9	10
R_X	2	1	3	5	6	7	4	9	8	10

Then

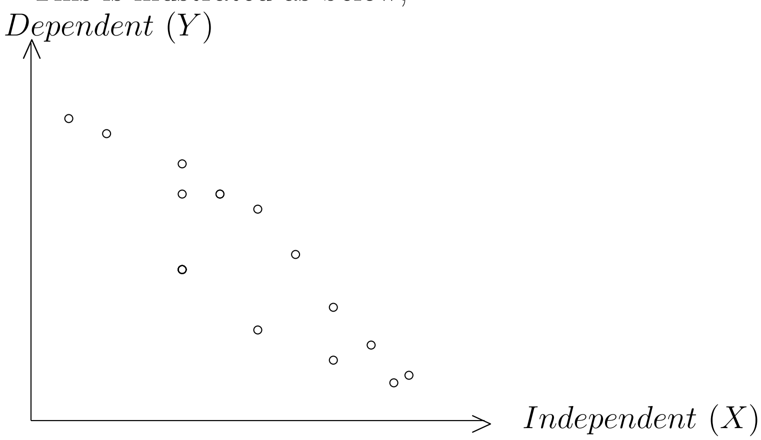
DG $^{-}1$	DB 1	DE 1	DA 1	DC 1	DF 1	DI 1	DJ 1	DH 1	Score 7
	GB 1	GE 1	GA 1	GC 1	GF 1	GI 1	GJ 1	GH 1	8
		BE 1	BA 1	BC 1	BF 1	BI 1	BJ 1	BH 1	7
			EA 1	EC 1	EF $^{-}1$	EI 1	EJ 1	EH 1	4
				AC 1	AF $^{-}1$	AI 1	AJ 1	AH 1	3
					CF $^{-}1$	CI 1	CJ 1	CH 1	2
						FI 1	FJ 1	FH 1	3
							IJ $^{-}1$	IH 1	0
								JH 1	1
Total Score									35

$\therefore \tau = \dfrac{2S}{n(n-1)} = \dfrac{2 \times 35}{10(10-1)} = \dfrac{35}{45} = 0.78$ (to $2dps$)

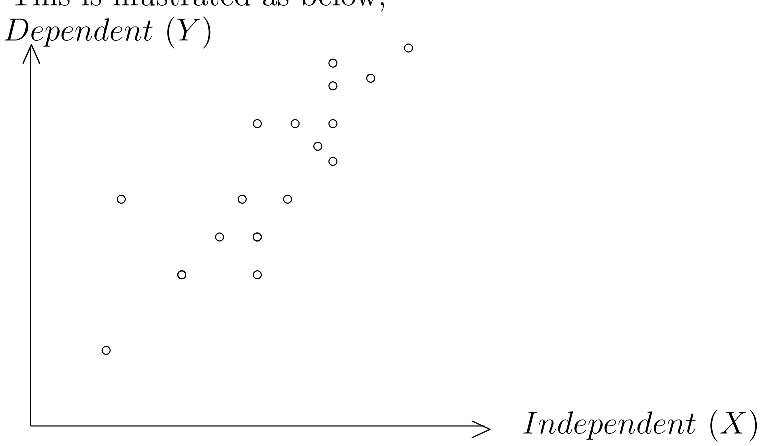
3.2 Scatter Diagram/Scatter Graph

This refers to the graphic representation of relationship between two variables. The title, labelled axes and accurate scale on the drawn graphs are very important.
The scater plots helps us to visualise the appparent relationship between the two variables that are ploted in pairs.
The independent variable X is ploted on the horizantal axis while the dependent variable Y is ploted on the vertical axis.
Types of Correlation; These helps to explain the relationship between the two ploted

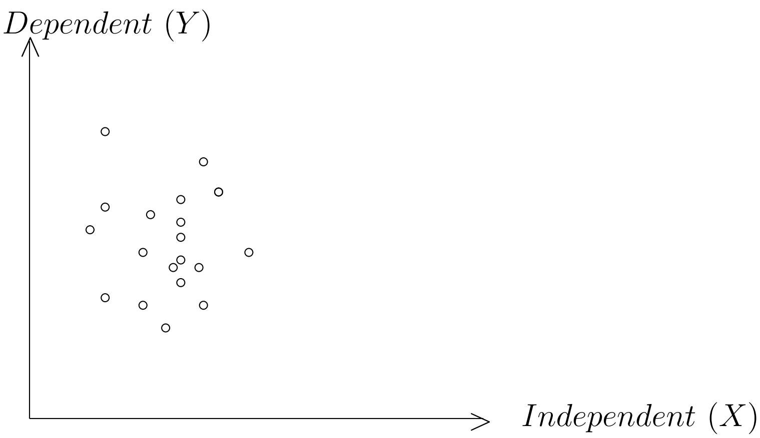
variables. Say X and Y.
The types include:
(a) **Negative correlation;** Here the increase in the Independent variable (X) results into a decrease in the Dependent variable (Y) and
This is illustrated as below;



(b) **Positive correlation;** Here the increase in the Independent variable (X) results into a increase in the Dependent variable (Y) and
This is illustrated as below;



(c) **Zero/No/Neutral correlation;** Here the increase/decrease in the Independent variable (X) results into no/very small effect in the Dependent variable (Y) and vice versa and This is illustrated as below;



NB: To comment on the graph , you do this by observing the trend of the scattered points and give the comment as either a positive correlation, a negative correlation or no correlation. i.e consider the above illustrations to make conlusions.

Drawing a line of best fit.

This is done by ensuring that the mean point for the given data $M_0(\overline{x}, \overline{y})$ is obtained and making sure that the line passes through this point.

The line therefore must pass through the mean point of the given data $M_0(\overline{x}, \overline{y})$ such that $\overline{x} = \dfrac{\sum x}{n}$ and $\overline{y} = \dfrac{\sum y}{n}$ where n = Number of observation. The line of best fit can be drawn generally in two formats; I.e

The line of best fit of x on y by eye: this can be done in the following ways;

- Plot the mean$M_0(\overline{x}, \overline{y})$ on the scatter,
- Draw a line parallel to $x-$ axis, (Apply the imaginary eye idea)
- Find the M_a of the points above,
- Find the M_b of the points below,
- Draw a solid line passing through M_0, M_a and M_b which becomes the line of best fit.

NB; This line is used when y values are more accurate.

While,

The line of best fit of y on x by eye: this can be done in the following ways;

- Plot the mean$M_0(\overline{x}, \overline{y})$ on the scatter,
- Draw a line parallel to $y-$ axis, (Apply the imaginary eye idea)
- Find the M_l of the points on left,
- Find the M_r of the points on right,

- Draw a solid line passing through M_0, M_a and M_b which becomes the line of best fit.

NB; This line is used when x values are more accurate.
To improve on the accuracy of you answer, Still ensure that your line of best fit pass through the mean of all points

Equation of the Line of Best Fit.

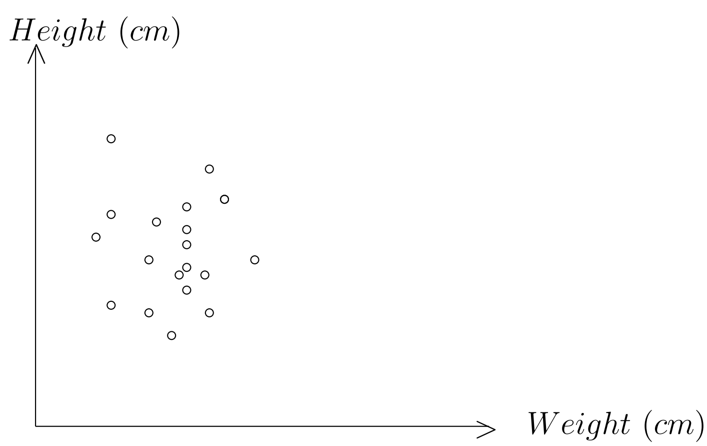
This is obtained by reading any two points from the drawn line of best fit, there after you use them to find the equation of the line in form of $y = mx + c$ using the analytical approach.I.e the pure mathematics approach

Example:

(1) The table below shows the height and weight of 8 students in certain school.

Weight (kg)	60	70	40	35	52	60	55	44
Height (cm)	150	120	95	105	170	128	160	100

(a) Draw a scatter diagram and comment on your graph.
(b) Draw a line of best fit and hence obtain the height of the 9$^{\text{th}}$ student whose height was 140cm.
(c) Find the equation of the line of best fit in form of $y = mx + c$
(d) Calculate the rank correlation coefficient and comment on your answer. **Solution:**



GRAPH

(a) There is Positive correlation

(b) 59

(c)

Weight	Height	R_w	R_h	d^2
60	150	2.5	3	0.25
70	120	1	5	16
40	95	7	8	1
35	105	8	6	4
52	170	5	1	16
60	128	4	1.5	6.25
55	160	4	2	4
44	100	6	7	1
				$\sum d^2 = 48.50$

Using $\rho = \ 1 - \dfrac{6(48.50)}{8(64-1)} \ = \ 0.4226.$
Comment: There is moderate positive correlation.

Some one more example on negative correlation

3.2.1 Exercise 3

(1) The weighing scales from three different stalls X, Y and Z in Kajjansi market were used to weigh 11 bags of beans $A, B, C.........K$ and the results (in Kgs) were as given in the table below.

	A	B	C	D	E	F	G	H	I	J	K
X	56	70	56	69	70	80	66	49	74	40	63
Y	55	68	63	64	74	75	72	66	60	68	60
Z	63	74	78	73	64	73	79	67	67	79	63

(a) Determine the rank correlation, for the performance of scale:

(i) Z and X,

(ii) X and Y.

(b) Which of the three scales: Y, X and Z were in good working conditions.

(2) The table below shows the percentage of sand Y, in the soil at different depths, Xin cm.

Sand depth,X (cm)	35	65	55	25	45	75	20	90	51	60
% of sand,Y	86	70	84	92	79	68	96	58	86	77

(a) (i) Plot a scatter diagram of the data. Comment on the relationship between the depths of the soil and the percentage of sand in the soil.($UNEB$)

(ii) Draw a line of best fit through the points of the scatter diagram. Use your graph to estimate;
- the percentage of the sand in the soil at a depth of 31cm
- depth of the soil with 54% sand.

(b) Calculate the rank correlation coefficient between the percentage of sand in the soil and the depth of the soil.

Ans. (b)-0.9485, very high negative correlation

(3) The following were scores by 8 schools in 100m and 200m race.

School	A	B	C	D	E	F	G	H
100m race	48	45	35	35	42	15	45	33
200m race	68	54	48	58	75	40	60	57

(a) Calculate the rank correlation coefficient of the school's performance. Give a comment to your result.
(b) Plot a scatter diagram and use it to obtain the equation of the line of best fit.
(c) Use your equation of the line of best fit to find 100m when 200m is 20
Ans. ρ = 0.6310, high positive correlation

(4) The marks of 10 candidates in English and Mathematics are;

Candidates	1	2	3	4	5	6	7	8	9	10
English	50	58	35	86	76	43	40	60	76	86
Mathematics	65	72	54	82	32	74	40	53	35	65

(a) Plot a scatter diagram for the data and comment on your graph.

(b) Draw a line of best fit and use it to estimate a score of the 11$^{\text{th}}$ candidate in mathematics who scored 70 in English.

Calculate a rank correlation coefficient between English and Mathematics and comment about the performance in the two subjects at 5% level of significance.

Ans. (c) ρ = 0.0667, not significant at 5%

(5) Eight students took an examination in Physics and Mathematics and their grades were as shown in the table below. Calculate the rank correlation coefficient for the student. Comment on the results.

Maths (M)	A	B	A	D	O	D	B	F
Physic (P)	C	B	C	E	E	C	A	F

Ans. (b) ρ = 0.7083, high positive correlation or not significant at 5% and 1%

(6) In certain A - level school, the following grades were attained in National Examinations Mathematics and Physics.

Mathematics	A	O	B	F	E	C	D	B
physics	C3	D2	D1	P8	P8	D2	C3	D2

Calculate a rank correlation coefficient for the grades and comment on your results.

Ans. ρ = 0.4940, moderate positive correlation

(7) In a certain commercial institution, a speed and error typing examination was administered to 12 randomly selected candidates$A, B, C.\ldots\ldots L$ of the institution. The table below shows their speeds(y) in seconds and number of errors in the typed script (x)

	A	B	C	D	E	F	G	H	I	J	K	L
X	11	24	20	9	32	30	28	15	18	40	27	40
Y	140	137	124	150	153	160	139	142	145	172	140	157

(a) (i) Plot the data on a scatter diagram.

(ii)Draw the line best fit on your diagram and comment on the likely association between speed and error made.

(iii)Determine the equation of your line in the form $y = xk + b$ where k and b are constants.

(b) By giving $Rank1$ to the fastest student and the student with the fewest errors, rank the above data and use it to calculate the correlation coefficient. Comment on your results.

(8) At ABC high school, there was inter houses music competitions and there were two judges. Judge A ranked the houses as follows starting from best; Charles, Balikudembe, Kizito and John. According to judge B, there was interchange of positions as follows; the 1$^{\text{st}}$ house with the 4$^{\text{th}}$ house and also the 2$^{\text{nd}}$ with 3$^{\text{rd}}$ house. Using the data above, calculate the rank correlation coefficient. And comment on your results.

Ans. ρ = -1.0, very high negative correlation, significant at 5% and 1%

(9) The table below shows the marks obtained by 10 pupils in two Mathematics tests.

Pupils	A	B	C	D	E	F	G	H	I	J
Test 1	50	43	56	52	68	66	74	70	80	76
Test 2	65	55	56	53	52	49	48	45	42	40

(a) Illustrate the two sets of marks by means of a scatter diagram plotting the Test 1 marks on the x- axis. Calculate the mean marks for each test $(\overline{X}, \overline{Y})$ and plot the result on the scatter diagram.

(b) Draw the line of best fit and find its equation in the form $y = a + bx$, where y represent Test 2 marks.

(c) Determine the mark of Test 1 given that the pupil scored 60 in Test 2.

(d) Find the Rank correlation coefficiet between the performance in Test 1 and Test 2 and comment on the results.

Ans. (d) $\rho =$ -0.9030, very high negative correlation or significant at 5% and 1%

(10) The table below shows the marks scored by ten students in Mathematics (X) and Economics (Y) tests.

	A	B	C	D	E	F	G	H	I	J
X	45	65	75	37	35	50	31	79	58	45
Y	61	62	70	49	46	41	65	31	55	69

Calculate the rank correlation coefficient for the student's performance in the two subjects. Comment on your answer.

Ans. (b) $\rho =$ -0.0697, very low negative correlation or not significant at 5% and 1%

(11) Given the information below showing the preferences of senior six boys and girls with respect to the meals they want to take during lunch

Girls	B	C	A	F	E	C	G
Boys	F	A	B	E	C	G	D

Calculate the coefficient of rank correlation and comment on your results.

Ans.

(12) The table below shows the lost and gained scores in the two games A and B which were played on different days by Shakira during the competitions at Lugogo stadium.

DAYS	Mon	Tue	Wed	Thurs	Frid	Sat	Sun
A	-4	5	-10	6	-3	-4	8
B	9	-5	11	-15	4	-5	-6

Calculate the rank correlation coefficient and comment on your results.

Ans. $\rho =$ **-0.83036, very high negative correlation or significant at 5% or not significant at 1%**

Chapter 4

PROBABILITY THEORY

4.1 PROBABILITY THEORY

4.1.1 Introduction

Probability is a chance that an event will occur/ happen. The term probability arose from the games of chance. For example tossing a coin, rolling a die, producing a baby boy,and others.

Terms Used in Probability Theory.

1. **Sample space(S):** This refers to the set of all possible outcomes of an experiment. Each possible outcome is called an **event**.
   E.g,
   (1) Rolling a die, Sample space$(S) = \{1, 2, 3, 4, 5, 6\}$ while sample points are $1, 2, 3, 4, 5, 6$
   (2) Rolling a tetrahedron , Sample space$(S) = \{1, 2, 3, 4\}$ while sample points are $1, 2, 3, 4$
   The sample space can be generated majorly in three ways and these include:

   - Table of outcomes.
   - Tree diagram.
   - Permutation and combination.

   All these are to be covered later!

2. **An event:** This refers to any possible outcome.
   E.g
   (a) Tossing a die, Sample space$(S) = \{1, 2, 3, 4, 5, 6\}$. If the desired event is that an 'odd' number show up, then Odd is the event E such that $E = \{1, 3, 5\}$
   (b) When a fair coin is tossed three times, Sample space $(S) = \{HHH, HHT, HTH, HTT, THH, TH$
   If your interested in picking at least two heads out of the three tosses, then this becomes the event E such that $E = \{HHH, HHT, HTH, THH\}$

3. **Probability function:** The Probability function of an event A denoted by $P(A)$ is the sum of probabilities of the sample points in A.
   It's given by

$$\text{Probability of event } A = \frac{n(A)}{n(S)}$$

   where:
   $n(A)$ = Number of elements in event$A$,
   $n(S)$ = Number of elements in sample space$S$,

4. **Properties of probabilities:**
   Let the Sample space(S) be having the event(A) then:

   - $P(A) \geq 0$. I.e Probabilities can not be negative.
   - $P(A) \leq 1$. I.e Probabilities don't go above 1.
   - $P(S) = 1$. I.e The sum of all probabilities add up to 1.

   NB: If $P(A)$ is the probability that an event, then $0 \leq P(A) \leq 1$

Types of Events and their Conditions:

1. **Compound events:** These are events which are formed after joining two or more events. They are formed by using the words like **and, either, neither, or,** etc.

2. **Intersection of events:** If A and B are any two events, we define a new event called the intersection of A and B and consists of outcomes that are in both A and B or $(A \cap B)$.
   $\Longrightarrow$ The probability that A and B occur is denoted by $P(A \cap B)$.
   This is also known as the and situation I.e $P(A \text{ and } B) = P(A \cap B)$.
   NB: $P(A \cap B) = P(B \cap A)$

3. **Union of events:** If A and B are any two events, we define a new event called the union of A and B i.e $(A \cup B)$ and consists of all sample points in either A or B or both .
   $\Longrightarrow$ The probability that A or B occur is denoted by $P(A \cup B)$.
   For undefined events, we define it as

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

   This is also known as the Or situation I.e $P(A \text{ or } B) = P(A \cup B)$.
   NB:

The probability that A or B but not both occur is denoted by $P(A \cup B) - P(A \cap B)$.
I.e

$$\begin{aligned} P(A \text{ or } B \text{ but not both}) &= P(A \cup B) - P(A \cap B) \\ &= P(A) + P(B) - P(A \cap B) - P(A \cap B) \\ &= P(A) + P(B) - 2P(A \cap B). \end{aligned}$$

NB: $P(A \cup B) = P(B \cup A)$

4. **Complement events:** These are events where if one does not occur, the other has to occur. e.g day and night, boy and girl, passing and failing, etc.
I.e
If A is an event of a sample spaces$S$, the complement of A is the set containing all the sample points in S that are not in$A$. It's denoted by$\overline{A}$ or A'.
Given,

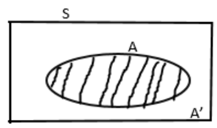


Figure 4.1

From the diagram, $\Longrightarrow P(S) = P(A) + P(A')$ but $P(S) = 1$
$\therefore 1 = P(A) + P(A') \Longleftrightarrow P(A) + P(A') = 1$ and $P(A \cap B) = 0$
Also we can conclude that for two events A and B,
$P(A \cup B) + P(A' \cap B') = 1$ and $P(A \cap B) + P(A' \cup B') = 1$
NB: $P(A' \cap B')$ can be written as $P(\overline{A \cup B})$ or $P(A \cup B)'$ similarly $P(A' \cup B')$ can be written as $P(\overline{A \cap B})$ or $P(A \cap B).'$

5. **Exclusive(or Mutually Exclusive) events:** These are events that cannot together.I.e If the two events A and B have no sample points in common,i.e $(A \cap B) = \{\}$, then we say that A and B are mutually exclusive events. Such events are said to be disjoint.Given

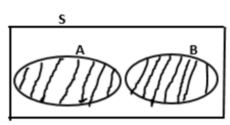


Figure 4.2

From the diagram, $(A \cap B) = \{\} \Longrightarrow n(A \cap B) = 0, \ \therefore P(A \cap B) = 0.$
For mutually exclusive events, $P(A \cup B)$ becomes $P(A \cup B) = P(A) + P(B)$.
Generally, Extending this to n events:
If $A_1, A_2, A_3, \cdots A_n$, are n mutually events, then
$P(A_1 \cup A_2 \cup A_3 \cup \cdots \cup A_n) = P(A_1) + P(A_2) + P(A_3) + P(A_4) + \cdots + P(A_n)$.
NB: Complete events are mutually exclusive while the reverse is not true.

6. **Independent events:**
Two events A and B are said to be independent events if either of the events A and B can occur without being affected by the other.
Therefore, for independent events, $P(A \cap B) = P(A) \times P(B)$.
For independent exclusive events, $P(A \cup B)$ becomes:
$P(A \cup B) = P(A) + P(B) - P(A) \times P(B)$.
Generally, Extending this to n events:
If $A_1, A_2, A_3, \cdots A_n$, are n independent events, then
$P(A_1 \cap A_2 \cap A_3 \cap \cdots \cap A_n) = P(A_1) \times P(A_2) \times P(A_3) \times P(A_4) \times \cdots \times P(A_n)$.

7. **Exhaustive events:**
Two events A and B are said to be exhaustive events if $P(A \cup B) = 1$. I.e If two events A and B are such that between them they make up the whole of the probability space, then they exhaustive events.
For example, If
$S =$ (the integers from 1 to 10 inclusive),
$A =$ (the integers below 7) $= \{1, 2, 3, 4, 5, 6\}$,
$B =$ (the integers above 5),$= \{6, 7, 8, 9, 10\}$
then $(A \cup B) = \{1, 2, 3, 4, 5, 6, 7, 8, 9, 10\} = S \iff P(A \cup B) = P(S) = 1$.
However for a special case:
Consider an event A and it's complement event A',

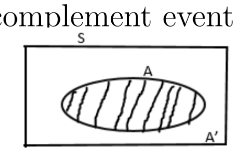


Figure 4.3

From the diagram, $P(A \cap A') = 0$, and

$$\begin{aligned} P(A \cup A') &= P(A) + P(A') \\ 1 &= P(A) + P(A') \\ \implies P(A) + P(A') = 1, \end{aligned}$$

In this case, the event A and it's complement A' are both mutually exclusive and exhaustive.
Genenally, Extending this to n events:
If $A_1, A_2, A_3, \cdots A_n$, are n events between them make up the whole probability space without overlapping, then $P(A_1) + P(A_2) + P(A_3) + P(A_4) + \cdots + P(A_n) = 1$ and the n events are both mutually exclusive and exhaustive.

Examples.

(1) Events A and B are such that $P(A) = \frac{19}{30}$, $P(B) = \frac{2}{5}$ and $P(A \cup B) = \frac{4}{5}$. Find $P(A \cap B)$.
Solution.

$P(A)=\frac{19}{30}, P(B)=\frac{2}{5}, P(A\cup B)=\frac{4}{5}, P(A\cap B)=?$
Using

$$\begin{aligned} P(A\cup B) &= P(A)+P(B)-P(A\cap B)\\ \frac{4}{5} &= \frac{19}{30}+\frac{2}{5}-P(A\cap B)\\ \implies P(A\cap B) &= \frac{19}{30}+\frac{2}{5}-\frac{4}{5}\\ &= \frac{7}{30}. \end{aligned}$$

(2) Events A and B are such that $P(A)=0.4, P(B)=x, P(A\cup B)=0.58.$ and $P(A\cap B)=0.12$.
(a) Find the value of x.
(b) Show that A and B and independent events.Hence find $P(A'\cup B')$.
Solution.
(a) $P(A)=0.4, P(B)=x, P(A\cup B)=0.58, P(A\cap B)=0.12$
Using

$$\begin{aligned} P(A\cup B) &= P(A)+P(B)-P(A\cap B)\\ 0.58 &= 0.4+x-0.12\\ \implies x &= 0.58-0.4+0.12\\ &= 0.3. \end{aligned}$$

(b) For independent events,

$$\begin{aligned} P(A\cap B) &= P(A)\times P(B)\\ 0.12 &= 0.4\times 0.3\\ 0.12 &= 0.12 \end{aligned}$$

Since $P(A\cap B)=P(A)\times P(B)=0.120$, then the two events are independent.
Hence

$$P(A'\cup B') = P(A')+P(B')-P(A'\cap B')$$
$$\text{Since } A \text{ nd } B \text{ are independent, then so are } A' \text{ and } B'$$
$$\text{Similary,} P(A')=1-P(A)=1-0.4=0.6 \text{ and } P(B')=1-P(B)=1-0.3=0.7$$
$$\begin{aligned} \implies P(A'\cup B') &= 0.6+0.7-0.6.7\\ &= 1.30-0.42\\ &= 0.88. \end{aligned}$$

(3) In a survey, 15% of the participants said that they had never bought lottery tickets or premium bonds,73% had bought lottery tickets and 49% had bought premium bonds.Find the probability that a person chosen at random from those taking part in the survey:

(a) had bought lottery tickets or premium,
(b) had bought lottery ticket and premium,
(c) had bought lottery ticket only,
(d) had bought lottery tickets or premium but not both.

Solution.

Let L be a person who has bought lottery tickets
Let B be a person who has bought premium
$\therefore P(L) = 0.73,\ P(B) = 0.49,$ and $P(\text{neither } L \text{ or } B) = P(L' \cap B') = 0.15$

(a)
$$\begin{aligned} P(L \text{ or } B) = P(L \cup B) &= 1 - P(L' \cap B') \\ &= 1 - 0.15 \\ \Longrightarrow P(L \text{ or } B) &= 0.85 \end{aligned}$$

(b)$P(L \text{ and } B) = P(L \cap B)$
Thus using
$$\begin{aligned} P(L \cup B) &= P(L) + P(B) - P(L \cap B) \\ 0.85 &= 0.73 + 0.49 - P(L \cap B) \\ \Longrightarrow P(L \cap B) &= 0.73 + 0.49 - 0.85 \\ &= 0.37. \end{aligned}$$

(c)
$$\begin{aligned} P(L \text{ only}) &= P(L) - P(L \text{ and } B) \text{ I.e from the venn diagrm.} \\ \Longrightarrow P(L\text{only}) &= 0.73 - 0.37 \\ &= 0.36. \end{aligned}$$

(d)
$$\begin{aligned} P(L \text{ or } B \text{ but not both}) &= P(L \cup B) - P(L \cap B) \\ &= 0.85 - 0.37 \\ \Longrightarrow P(L \text{ or } B \text{ but not both}) &= 0.48 \end{aligned}$$

(4) Events A and B are independent such that $P(A \cup B) = 0.8,\ P(A) = 0.5$. Find $P(B)$

Solution.

From:
$$\begin{aligned} P(A \cup B) &= P(A) + P(B) - P(A \cap B) \\ 0.8 &= 0.5 + P(B) - P(B) \times 0.5 \quad \text{I.e because they are independent evets} \\ \Longrightarrow 0.8 &= 0.5 + P(B) - 0.5P(B) \\ \therefore P(B) &= \frac{(0.8 - 0.5)}{0.5} = 0.6. \end{aligned}$$

(5) Events A and B are independent such that $P(A \cup B) = 0.8,\ P(A \cap B) = 0.1$. Find the possible values of $P(B)$ and $P(A)$.

Solution.

Since they indepenent events, then

$$
\begin{aligned}
P(A\cap B) &= P(B)\times P(A)\\
0.1 &= P(B)\times P(A)\\
\implies P(A) &= \frac{0.1}{P(B)}\cdots\cdots\cdots\cdots\cdots\cdots\cdots\cdots\cdot (i)\\
P(A\cup B) &= P(A)+P(B)-P(A\cap B)\\
0.8 &= P(A)+P(B)-0.1
\end{aligned}
$$

Substituting for $P(A)$ in(i)

$$
\begin{aligned}
0.8+0.1 &= \frac{0.1}{P(B)}+P(B)\text{ Let }P(B)\text{be }x,\\
0.9 &= \frac{0.1}{x}+x \iff 0.9x=0.1+x^2 \Leftrightarrow x^2-0.9x+0.1=0\\
x &= \frac{0.9\pm\sqrt{((0.9^2)-4\times 0.1)}}{2}\\
\text{Either } x &= \frac{0.9+0.64}{2}=0.77 \implies P(B)=0.77.\\
\text{Or } x = \frac{0.9-0.64}{2} &= 0.13 \implies P(B)=0.13.\\
\text{When } P(B)=0.77, P(A) &= \frac{0.1}{P(B)}=\frac{0.1}{0.77}=0.13\\
\text{When } P(B)=0.13, P(A) &= \frac{0.1}{P(B)}=\frac{0.1}{0.13}=0.77.
\end{aligned}
$$

(6) Events A and B are mutually exclusive such that $P(A\cup B)=0.88,\ P(A)=0.5$. Find $P(B)$
Solution.
From:

$$
\begin{aligned}
P(A\cup B) &= P(A)+P(B)-P(A\cap B)\\
0.88 &= 0.5+P(B)-0\ \text{ I.e because they are mutually exclusive evets}\\
\implies 0.88 &= 0.5+P(B)\\
\therefore P(B) &= (0.88-0.5)=0.38.
\end{aligned}
$$

(7) In a race in which there are no dead heats, the probability that John wins is 0.3, the probability that Paul wins is 0.2 and that of Mark wins is 0.4. Find the probability that:
(a) John or Mark wins,
(b) John or Paul or Mark wins,
(c) Someone else wins,
(d) Neither John nor Mark wins.
Solution.
Since in this game, there is only one person winning,then the events are mutually exclusive.
(a) $P(\text{John or Mark wins}) = P(\text{John wins}) + P(\text{Mark wins}) = 0.3+0.4 = 0.70.$

(b)P(John or Paul or Mark wins) $= P$(John wins) $+ P$(Paul wins) $+ P$(Mark wins) $= 0.3 + 0.2 + 0.4 = 0.90$,
(c) P(someone else wins) $= 1 - P$(John or Paul or Mark wins) $= 1 - 0.9 = 0.1$.
(d) P(Neither John nor Mark wins) $= P(\overline{J \cup M}) = 1 - P$(John or Mark wins) $= 1 - 0.7 = 0.30$.

4.1.2 The Contingency Table:

This is the that is used to all the different probability equations derived from the intersection with the set theory. It's is illustrated as below:

	A	A'	
B	$P(A \cap B)$	$P(A' \cap B)$	$P(B)$
B'	$P(A \cap B')$	$P(A' \cap \mathrm{B}')$	$P(B')$
	$P(A)$	$P(A')$	1

From the above contingency table, we have the following:

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

Also from, $P(A \cup B) = P(A) + P(B) - P(A \cap B)$ we can generalize the following without proofs.

- $P(A \cup B') = P(A) + P(B') - P(A \cap B')$
- $P(A' \cup B') = P(A') + P(B') - P(A' \cap B')$
- $P(A \cup B) + P(A \cup B)' = 1$
- $P(A \cap B) + P(A \cap B)' = 1$

All the above equations can clearly be obtained using the set theory idea as follows.
Consider the Venn diagram below.

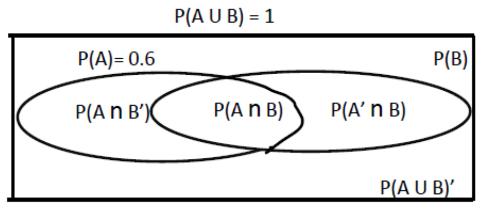


Figure 4.4

Then from the diagram above, we have the following important expressions;

- $P(A) = P(A \cap B) + P(A \cap B')$
- $P(B) = P(A \cap B) + P(A' \cap B)$
- $P(A \cup B) = P(A \cap B) + P(A' \cap B) + P(A'B')$.
  $\Longrightarrow P(AuB) = P(A) + P(B) - P(A \cap B)$

Demorgan's laws:

For any two events A and B;
(i) $P(A \cup B)' = P(A' \cap B')$
(ii) $P(A \cap B)' = P(A' \cup B')$

Distributive laws:

For any 3 events A, B and C;
(i) $P\big[A \cap (B \cup C)\big] = P\big[(A \cap B) \cup (A \cap C)\big]$
(ii) $P\big[A \cup (B \cap C)\big] = P\big[(A \cup B) \cap (A \cup C)\big]$

4.1.3 Conditional Probability:

For any two events A and B, then the conditional probability that A occurs, given that B has already occurred is written as $P(A, \text{given } B)$ or $P(A/B)$.
It's defined as $P(A/B) = \frac{P(A \cap B)}{P(B)}$ for $P(B) \neq 0$ I.e $P(A, \text{given } B) = \frac{P(A \text{ and } B)}{P(B)}$
It refers to the probability of A occurring given that B has already occurred.
Examples:

(1) Events A and B are such that $P(A) = 0.6, P(B) = 0.7$ and $P(A \cap B) = 0.4$, Find

(i) $P(A \cup B)$ (iii) $P(A' \cup B)$ (v) $P(A/B')$ (vii) $P(A \cap B/B)$.

(ii) $P(A \cap B')$ (iv) $P(A/B)$ (vi) $P(B'/A')$

Solution

(i) $P(A \cup B) = P(A) + P(B) - P(A \cap B) = 0.6 + 0.7 - 0.4 = 0.9$

(ii) $P(A \cap B')$,
From $P(A) = P(A \cap B) + P(A \cap B'), \iff P(A \cap B') = P(A) - P(A \cap B) = 0.6 - 0.4 = 0.2.$

(iii) $P(A' \cup B) = P(A') + P(B) - P(A' \cap B) = (0.4 + 0.7 - 0.4) = 0.7$ (iv) $P(A/B) = \dfrac{P(A \cap B)}{P(B)} = \dfrac{0.4}{0.7} = \dfrac{4}{7}$

(iv) $P(A/B) = \dfrac{P(A \cap B)}{P(B)} = \dfrac{0.4}{0.7} = \dfrac{4}{7}$

(v) $P(A/B') = \dfrac{P(A \cap B')}{P(B)} = \dfrac{0.2}{0.3} = \dfrac{2}{3}$

(vi) $P(B'/A') = \dfrac{P(A' \cap B')}{P(B')}$ But $P(A' \cap B') = P(B') - P(A \cap B') = 0.3 - 0.2 = 0.1$
$\implies P(B'/A') = \dfrac{0.1}{0.3} = \dfrac{1}{3}.$

(vii) $P(A \cap B/B) = \dfrac{P(A \cap B \cap B)}{P(B)} = \dfrac{P(A \cap B)}{P(B)} = \dfrac{0.4}{0.7} = \dfrac{4}{7}.$

(2) Events R and S are such that $P(R') = \frac{2}{5}$, $P(S/R) = \frac{1}{3}$ and $P(R \cup S) = \frac{4}{5}$. Find

(i) $P(S)$, (iii) $P(R' \cap S')$, (v) $P(R'/S')$.

(ii) $P(RnS')$, (iv) $P(R' \cup S')$. (vi) $P(R' \cup S'/S)$.

Solution

(i) $P(S)$
From $P(S/R) = \dfrac{P(R \cap S)}{P(R)} \iff \frac{1}{3} = \dfrac{P(R \cap S)}{P(R)}$, But $P(R) = \dfrac{3}{5}$
$\therefore P(R \cap S) = P(R) \times \dfrac{1}{3} = \dfrac{3}{5} \times \dfrac{1}{3} = \dfrac{1}{5}.$
For $P(R \cup S) = P(R) + P(S) - P(R \cap S)$
$\implies \frac{4}{5} = \frac{3}{5} + P(S) - \frac{1}{5}$
$\therefore P(S) = \frac{2}{5}$

(ii) $P(R \cap S') = P(R) - P(R \cap S) = \dfrac{3}{5} - \dfrac{1}{5} = \dfrac{2}{5}$

(iii) $P(R' \cap S') = P(R \cup S)' = 1 - P(R \cup S) = 1 - \dfrac{4}{5} = \dfrac{1}{5}$
Or from $P(S') = P(R' \cap S') + P(R \cap S') \iff P(R' \cap S') = P(S') - P(R \cap S')$
$\therefore P(R' \cap S') = (\dfrac{3}{5} - \dfrac{2}{5}) = \dfrac{1}{5}.$

(iv) $P(R' \cup S') = \frac{2}{5} + \frac{3}{5} - \frac{1}{5} = \frac{4}{5}$

(v) $P(R'/S') = \frac{P(R' \cap S')}{P(S')} = \frac{1/5}{3/5} = \frac{1}{3}.$

(vi) $P(R \cup S/S) = \frac{P((R \cup S) \cap S)}{P(S)} = \frac{P(S)}{P(S)} = 1$

(3) Given that $P(B/A) = 0.6, p(B/A') = 0.5$ and $P(B) = 0.52$. Find $P(A)$ and $P(AUB)$.
Solution:
From $P(B/A) = 0.6 \iff P(B \cap A) = 0.6P(A) \cdots\cdots\cdots\cdots\cdots (i)$
From $P(B/A') = 0.5 \iff P(B \cap A') = 0.5P(A') \cdots\cdots\cdots\cdots\cdots (ii)$
Taking $(i) + (ii)$

$$P(B \cap A) + P(B \cap A') = 0.6P(A) + 0.5P(A')$$
$$P(B) = 0.6P(A) + 0.5[1 - P(A)]$$
$$0.52 - 0.5 = 0.1P(A) \iff P(A) = \frac{0.02}{0.1} = 0.20$$

(4) Events A, B and C are such that $P(A) = x, P(B) = y$ and $P(C) = x+y$. If $P(A \cup B) = 0.7$ and $P(A/B) = 0.4$.
(a) Show that $10x - 3y = 7$.
(b) Given B and C are mutually exclusive and that $P(B \cap C) = 0.9$, determine the values of x and y.
Solution:
(a) From $P(A/B) = 0.4 \iff P(B \cap A) = 0.4P(B) \cdots\cdots\cdots\cdots\cdots\cdots\cdots\cdots (i)$

$$P(A \cup B) = P(A) + P(B) - P(A \cap B)$$
$$0.7 = x + y - 0.4P(B) \text{I.e from equation}(i)$$
$$\frac{7}{10} = x + y - 0.4y$$
$$\frac{7}{10} = x - \frac{3}{10}y \text{ On simplification,}$$
$$7 = 10x - 3y \iff 10x - 3y = 7$$

(b) **Please complete the question.**

(5) Show that for three events A, B and C,
(i) $P(A/B) = P(A)$, if A and B are independent.
(ii) $P(A/B) = 0$, if A and B are mutually exclusive.
(iii) $P(A/B) + P(A'/B) = 1$.
Solution:

(i) $P(A/B) =?$,
For A and B to be independent,$P(A \cap B) = P(A).P(B)$.

$$\Longrightarrow P(A/B) = \frac{P(A\cap B)}{P(B)} = \frac{P(A).P(B)}{P(B)} = P(A).\text{ Hence shown.}$$

(ii) $P(A/B) = ?$,
For A and B to be mutually exclusive, $P(A\cap B) = 0$.
$$\Longrightarrow P(A/B) = \frac{P(A\cap B)}{P(B)} = \frac{0}{P(B)} = 0.\text{ Hence shown.}$$

(iii) $P(A/B) + P(A'/B) = ?$

$$\begin{aligned}\Longrightarrow P(A/B) + P(A'/B) &= \frac{P(A\cap B)}{P(B)} + \frac{P(A'\cap B)}{P(B)}\\ &= \frac{P(A\cap B) + P(A'\cap B)}{P(B)}\\ &= \frac{P(B)}{P(B)}\\ \therefore P(A/B) + P(A'/B) &= 1.\text{ Hence shown.}\end{aligned}$$

(6) Given that events A and B are independent, show that also the following are independent:

(i) $P(A\cap B')$, (ii) $P(A'\cap B)$, (iii) $P(A'\cap B')$,

Solution:
(i) Required $P(A\cap B') = P(A)\times P(B')$

$$\begin{aligned}\text{From } P(A) &= P(A\cap B) + P(A\cap B')\\ \Longrightarrow P(A\cap B') &= P(A) - P(A\cap B)\\ &= P(A) - P(A)\times P(B)\\ &= P(A)[1 - P(B)]\\ \therefore P(A\cap B') &= P(A)\times P(B'). \qquad \text{Hence shown.}\end{aligned}$$

(ii) Required $P(A'\cap B) = P(A')\times P(B)$

$$\begin{aligned}\text{From } P(B) &= P(A\cap B) + P(A'\cap B)\\ \Longrightarrow P(A'\cap B) &= P(B) - P(A\cap B)\\ &= P(B) - P(A)\times P(B)\\ &= P(B)[1 - P(A)]\\ \therefore P(A'\cap B) &= P(A)\times P(B'). \qquad \text{Hence shown.}\end{aligned}$$

(iii) Required P(A′nB′) = P(A′) × P(B′)

$$
\begin{aligned}
&\text{From } P(B') = P(A\cap B') + P(A'\cap B') \\
\Longrightarrow P(A'\cap B') &= P(B') - P(A\cap B') \\
&= P(B') - [P(A) - P(A\cap B)] \\
&= P(B') - [P(A) - P(A)\times P(B)] \\
&= P(B') - P(A)[1 - P(B)] \\
&= P(B') - P(A)\times P(B')) \\
&= P(B')[1 - P(A)] \\
\therefore P(A'\cap B') &= P(B')\times P(A'). \qquad \text{Hence shown.}
\end{aligned}
$$

(7) Show that for three events A, B and C,
(i) $P(A\cup B) = P(A) + P(B) - P(A\cap B)$,
(ii) $P(A\cup B\cup C) = P(A)+P(B)+P(C)-P(A\cap B)-P(A\cap C)-P(C\cap B)+P(A\cap B\cap C)$.
Solution.
(a)Given

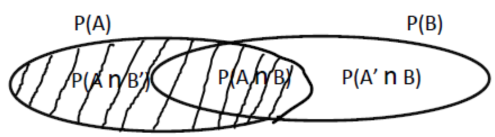


Figure 4.5

$$
\begin{aligned}
\text{From the diagram, } P(A\cup B) &= P(A) + P(A'\cap B) \\
\Longrightarrow P(A\cup B) &= P(A) + P(A'\cap B) \\
&= P(A) + [P(B) - P(A\cap B)] \\
&= P(A) + P(B) - P(A\cap B) \\
\therefore P(A\cup B) &= P(A) + P(B) - P(A\cap B). \qquad \text{Hence shown.}
\end{aligned}
$$

(b) $P(A\cup B\cup C) = ?$

Let $B\cup C = X$

$$
\begin{aligned}
\Longrightarrow P(A\cup B\cup C) &= P(A\cup X) \\
&= P(A) + P(X) - P(A\cap X) \text{I.e by exapanding.}
\end{aligned}
$$

Substituing for X

$$
\begin{aligned}
P(A\cup B\cup C) &= P(A) + P(B\cup C) - P(A\cap(B\cup C)) \\
&= P(A) + P(B) + P(C) - P(B\cap C) - P(A\cap B\cup A\cup C) \\
&= P(A) + P(B) + P(C) - P(B\cap C) - [P(A\cap B) + P(A\cap C) - P(A\cup B\cup C)] \\
&= P(A) + P(B) + P(C) - P(B\cap C) - P(A\cap B) - P(A\cap C) + P(A\cup B\cup C) \\
\therefore P(A\cup B\cup C) &= P(A) + P(B) + P(C) - P(B\cap C) - P(A\cap B) - P(A\cap C) + P(A\cup B\cup C).
\end{aligned}
$$

Hence shown.

4.1.4 The " One " and "Only One " situation .

This is the case concerning competing of participants all together at once. It's a situation in which we are interested in only a certain number of the series of events to happen. E.g only one winner, only two, only th three etc.
I.e here you can hit the target as a group or otherwise.
NB: All the events are taken to be independent.
This can be illustrated as below: Given events A, B and C, with their respective probabilities of hitting the target(winning) as $P(A), P(B)$ and $P(C)$, while those of respective failure as $P(A'), P(B')$ and $P(C')$.
Then:

(i) $P($ Non of them wins$) = P(A' \cap B' \cap C')$.

(ii) $P($Exactly one of them wins$) = P(A \cap B' \cap C') + P(A' \cap B \cap C') + P(A' \cap B' \cap C)$.

(iii) $P($Only two of them win $) = P(A \cap B \cap C') + P(A \cap B \cap C) + P(A \cap B' \cap C)$

(iv) $P($All of them win $) = P(A \cap B \cap C)$

(v) $P($At least two $) = P($Only two of them win $) + P($All of them win $)$.

(vi) $P($At most one win $) = P($Exactly one of them wins$) + P($ None of them wins$)$

(vii) $P($one of them takes wins $) = P(A$ or B or $C) = P(A \cup B \cup C)$
Examples:

(1) Two boys A and B are competing in a game with their respective chances of winning as 0.4 and 0.3. If the boys are competing at once, Find the probability that:

(a) all of them win,
(b) Only one wins,
(c) none of them wins.
(d) at least one wins,
(e) at most one wins.

Solution:
Let $P(A) = 0.4 \Longrightarrow P(A') = 0.6$, and $P(B) = 0.3 \Longrightarrow P(B') = 0.7$.
(a) $P($All of them win$) = P(A \cap B) = P(A) \times P(B) = 0.4 \times 0.3 = 0.120$

$$\begin{aligned}
\text{(b) } P(\text{Only one wins}) &= P(\text{Only A wins}) + P(\text{Only B wins}) \\
&= P(A \cap B') + P(A' \cap B) \\
&= (0.4 \times 0.7) + (0.6 \times 0.3) \\
&= 0.280 + 0.180 = 0.460.
\end{aligned}$$

(c) $P($None of them win$) = P(A' \cap B') = P(A') \times P(B') = 0.6 \times 0.7 = 0.420$
(d)

$$\begin{aligned}
P(\text{at least one wins}) &= P(\text{Only one wins}) + P(\text{Only two wins}) \\
&= 0.460 + 0.120 \\
&= 0.540.
\end{aligned}$$

(e)

$$P(\text{at most one wins}) = P(\text{None of them win}) + P(\text{Only one wins}) \\ = 0.420 + 0.460 \\ = 0.880.$$

(2) A, B and C are three events completing all together with their respective probabilities of winning as $0.4, 0.7$ and 0.5 . Find the probability that;

(a) all of them win,
(b) only one wins,
(c) only two win,
(d) none of them wins,
(e) at most two them win,
(f) at least one of them wins.

Solution:
Let $P(A) = 0.4 \Longrightarrow P(A') = 0.6$, $P(B) = 0.7 \Longrightarrow P(B') = 0.3$ and $P(C) = 0.5 \Longrightarrow P(C') = 0.5$.
(a) $P(\text{All of them win}) = P(A \cap B \cap C) = P(A) \times P(B) \times C = 0.4 \times 0.7 \times 0.5 = 0.140$

$$\text{(b) } P(\text{Only one wins}) = P(\text{Only A wins}) + P(\text{Only B wins}) + P(\text{Only C wins}) \\ = P(A \cap B' \cap C') + P(A' \cap B \cap C') + P(A' \cap B' \cap C) \\ = (0.4 \times 0.3 \times 0.5) + (0.6 \times 0.7 \times 0.5) + (0.6 \times 0.3 \times 0.5) \\ = 0.06 + 0.21 + 0.09 = 0.360.$$

$$\text{(c) } P(\text{Only two win}) = P(A \cap B \cap C') + P(A \cap B' \cap C) + P(A' \cap B \cap C) \\ = (0.4 \times 0.7 \times 0.5) + (0.4 \times 0.3 \times 0.5) + (0.6 \times 0.7 \times 0.5) \\ = 0.140 + 0.06 + 0.21 = 0.410.$$

(d) $P(\text{none of them win}) = P(A' \cap B' \cap C') = P(A') \times P(B') \times P(C') = 0.6 \times 0.3 \times 0.5 = 0.090$

$$\text{(e) } P(\text{at most two of them win}) = P(\text{None of them win}) + P(\text{Only one wins}) + P(\text{Only two wins}) \\ = 0.090 + 0.360 + 0.410 \\ = 0.860.$$

Or $P(\text{at most two of them win}) = 1 - P(\text{three of them win}) = (1 - 0.140) = 0.860$

$$\text{(f) } P(\text{at least one of them wins}) = P(\text{Only one wins}) + P(\text{Only two wins}) + P(\text{three of them win}) \\ = 0.360 + 0.410 + 0.140 \\ = 0.910.$$

Or $P(\text{at least one of them wins}) = 1 - P(\text{none of them wins}) = (1 - 0.090) = 0.910.$

(3) Three football stars, Ronaldo, Messi and Salah take part in final FIFA awards of 2017 with chances of winning it as $\frac{2}{5}, \frac{3}{8}$ and $\frac{2}{3}$. If they are playing a knock out, find the probability that;

(i) one and only one of them takes it. (ii) The award is taken

Solution:
$P(R) = \frac{2}{5}$, $P(R') = \frac{3}{5}$
$P(M)= \frac{3}{8}$, $P(M') = \frac{5}{8}$
$P(S)= \frac{2}{3}$, $P(S') = \frac{1}{3}$

$$\begin{aligned}\text{(i)}P(\text{one and only one of them takes}) &= P(R\cap M'\cap S') + P(R'\cap M\cap S') + P(R'\cap M'\cap S)\\ &= \left(\frac{2}{5}\times\frac{5}{8}\times\frac{1}{3}\right) + \left(\frac{3}{5}\times\frac{3}{8}\times\frac{1}{3}\right) + \left(\frac{3}{5}\times\frac{5}{8}\times\frac{2}{3}\right)\\ &= \frac{49}{120}.\end{aligned}$$

$$\begin{aligned}\text{(i)}P(\text{The award is taken}) &= P(R \text{ or } M \text{ or } S) = P(R\cup M\cup S)\\ &= \left(\frac{2}{5}\right) + \left(\frac{3}{8}\right) + \left(\frac{2}{3}\right) - \left(\frac{2}{5}\times\frac{3}{8}\right) - \left(\frac{2}{5}\times\frac{2}{3}\right) - \left(\frac{3}{8}\times\frac{2}{3}\right) + \left(\frac{2}{5}\times\frac{3}{8}\times\frac{2}{3}\right)\\ &= \frac{7}{8}.\end{aligned}$$

(4) There are 3 custom's road blocks from Malaba boarder to Kampala. Shakirah transports her container from Malaba to Kampala and the chances of her container being checked at any of the 3 road blocks are 0.4, 0.8, 0.7 respectively. Find the probability that her container will be checked at;

(i) one and only one road block (iii) all the 3 road blocks

(ii) atmost two road block

Solution

$P(A) = 0.4, P(B) = 0.8$, and $P(C) = 0.7$

$P(A') = 0.6, P(B') = 0.2$, and $P(C') = 0.3$

(i)$P(\text{one and only one road block}) = P(A\cap B'\cap C') + P(A'\cap B\cap C') + P(A'\cap B'\cap C)$
$= (0.4\times 0.2\times 0.3) + (0.6\times 0.8\times 0.3) + (0.6\times 0.2\times 0.7) = 0.252$
(ii)

$$\begin{aligned}P(\text{at most two road block}) &= P(A\cap B\cap C') + P(A'\cap B\cap C) + P(A\cap B'\cap C) + P(A\cap B'\cap C')\\ &+ P(A'\cap B\cap C') + P(A'\cap B'\cap C) + P(A'\cap B'\cap C')\\ &= (0.4\times 0.8\times 0.3) + (0.6\times 0.8\times 0.7) + (0.4\times 0.2\times 0.7)\\ &+ (0.4\times 0.2\times 0.3) + (0.6\times 0.8\times 0.3) + (0.6\times 0.2\times 0.7)\\ &+ (0.6\times 0.2\times 0.3)\\ &= 0.776\end{aligned}$$

(iii)$P(\text{all the 3 road blocks}) = P(A\cap B\cap C) = (0.4\times 0.8\times 0.7) = 0.224$

4.1.5 Cases Concerning Competing in Turns.

Under this case,participants take turns to hit at the target and whoever hits the target first becomes the winner and the competition ends. Here the participants are competing for the same item but doing it in turns and once the target is hit, the game comes to end.
In this case we must end up with a single winner.
Here you must be careful to be able to identify the participants' order of participating.I.e who should start,and who fallows.
Examples:

(1) A and B are two events completing in turns with their respective probabilities of winning as 0.4 and 0.7. Given that A starts first followed by B, find the probability that;

(a) B wins on the first trial,
(b) A wins on the second trial,
(c) A wins on the third trial,
(d) A wins,
(e) B wins.

Solution:
Let $P(A) = 0.4 \iff P(A') = 0.6$ and $P(B) = 0.7 \iff P(B') = 0.3$

$$\begin{aligned}(a)P(B \text{ wins on the first trial}) &= P(A \text{ fails on the first trial}) \times P(B \text{ wins on the first trial})\\ &= P(A' \cap B)\\ &= 0.6 \times 0.7 = 0.42.\end{aligned}$$

$$\begin{aligned}(b)P(A \text{ wins on the second trial}) &= P(A \text{ fails on the first trial}) \times P(B \text{ fails on the first trial}) \times\\ &\quad P(A \text{ wins on the first trial})\\ &= P(A' \cap B' \cap A)\\ &= 0.6 \times 0.3 \times 0.4 = 0.072.\end{aligned}$$

(c)$P(A \text{ wins on } 3^{rd} \text{ trial}) = P(A' \cap B' \cap A' \cap B' \cap A) = (0.6 \times 0.3 \times 0.6 \times 0.3 \times 0.4) =$ 0.01296.

$$\begin{aligned}(d)P(A \text{ wins}) &= P(A \text{ wins on the first trial}) + P(A \text{ wins on } 2^{rd} \text{ trial}) + P(A \text{ wins on } 3^{rd} \text{ trial})\\ &\quad + \cdots + P(A \text{ wins on infinity trial})\\ &= 0.4 + 0.072 + 0.01296 + \cdots\\ &= \frac{a}{1-r}. \text{ I.e Apply sum to infinity of an GP}\end{aligned}$$

Where $a = 0.4$, and $r = \dfrac{0.072}{0.4} = 0.180$,

$$= \frac{20}{41} = 0.4878.$$

(e) $P(B \text{ wins}) = P(B \text{ wins on the first trial}) + P(B \text{ wins on } 2^{rd} \text{ trial}) + P(B \text{ wins on } 3^{rd} \text{ trial})$

$$\begin{aligned}
&+ \cdots + P(B \text{ wins on infinity trial}) \\
&= 0.42 + (0.6 \times 0.3 \times 0.6 \times 0.7) + \cdots \\
&= 0.42 + 0.0756 + \cdots \\
&= \frac{a}{1-r}. \text{ I.e Apply sum to infinity of an GP}
\end{aligned}$$

Where $a = 0.42$, and $r = \dfrac{0.0756}{0.42} = 0.180$,

$$= \frac{21}{41} = 0.5122.$$

(2) A, B and C are three events completing in turns with their respective probabilities of winning as $0.8, 0.6$ and 0.7. Given that A starts first followed by C and finally B, find the probability that;

(a) B wins in the first trial,

(b) A wins on the second trial,

(c) A wins on the third trial,

(d) A wins,

(e) C wins.

Solution:
Let $P(A) = 0.8 \Longleftrightarrow P(A') = 0.2$ $P(B) = 0.6 \Longleftrightarrow P(B') = 0.4$ and
$P(C) = 0.7 \Longleftrightarrow P(C') = 0.3$
The order of competition: $A \mapsto C \mapsto B$.

(a)
$$\begin{aligned}
P(B \text{ wins on the first trial}) &= P(A' \cap C' \cap B) \\
&= 0.2 \times 0.3 \times 0.6 \\
&= 0.036.
\end{aligned}$$

(b)
$$\begin{aligned}
P(A \text{ wins on the second trial}) &= P(A' \cap C' \cap B' \cap A) \\
&= 0.2 \times 0.3 \times 0.4 \times 0.8 \\
&= 0.0192.
\end{aligned}$$

(c)
$$\begin{aligned}
P(A \text{ wins on the third trial}) &= P(A' \cap C' \cap B' \cap A' \cap C' \cap B' \cap A) \\
&= 0.2 \times 0.3 \times 0.4 \times 0.2 \times 0.3 \times 0.4 \times 0.8 \\
&= 0.0004608.
\end{aligned}$$

(d) $P(A \text{ wins}) = P(A \text{ wins on the first trial}) + P(A \text{ wins on } 2^{rd} \text{ trial}) + P(A \text{ wins on } 3^{rd} \text{ trial})$

$$\begin{aligned}
&+ \cdots + P(A \text{ wins on infinity trial}) \\
&= 0.8 + 0.0192 + 0.0004608 + \cdots
\end{aligned}$$

Where $a = 0.8$, and $r = \dfrac{0.0192}{0.80} = 0.024$,

$$= \frac{50}{61} = 0.8197.$$

(e) **Please do it.**

(3) Jane, Joan and James are playing a game in turns with Joan first followed by Jane and James respectively. The probabilities of Jane Joan and James winning are $0.25, 0.3$ and 0.15 respectively. Find the probability of that James wins.

Solution:

Let $P(Jane) = 0.25 \iff P((Jane)') = 0.75$ $P(Joan) = 0.3 \iff P((Joan)') = 0.7$ and

$P(James) = 0.15 \iff P((James)') = 0.85$

The order of competition: Joan $\mapsto$ Jane $\mapsto$ James.

$$\begin{aligned}
P(\text{ James wins}) &= P(\text{James wins on the first trial}) + P(\text{James wins on } 2^{rd} \text{ trial}) + P(\text{James wi}\\
&+ \cdots + P(\text{James wins on infinity trial})\\
&= (0.7 \times 0.75 \times 0.15) + (0.7 \times 0.75 \times 0.85 \times 0.7 \times 0.75 \times 0.15) + \cdots\\
&= 0.07875 + 0.035142187 + \cdots\\
&= \frac{a}{1-r}.
\end{aligned}$$

Where $a = 0.07875$, and $r = \dfrac{0.035142187}{0.07875} = 0.44625$,

$$= \frac{63}{443} = 0.1422.$$

4.1.6 Probability Tree Diagram:

Tree diagrams is used to combine more than one situation(conditions) in the probability. We obtain probabilities by multiplying them along the same branch but as we shift from from one branch to another, we add the probabilities and therefore selections are independent. The selections can be done with or without replacement.

NB:

(1)The total probability for any one set of branches = 1,

(2) The sum of the final probabilities is 1.

(3) The tree diagram must be drawn based on the given conditions in the question.

NB: Before drawing the tree diagram, you must first check if it involves Picking <u>with</u> or <u>with out</u> Replacement.

Probability Tree Diagram Concerning Picking with Replacement:

Under this case, the number of elements(objects) in the given box do not change and therefore the probability of picking an element remains <u>constant</u>.

Examples:

(1) A box contains 4 green and 5 red sweets, two sweets are picked at random one after the other <u>with replacement.</u> Find the probability that;

(i) all sweets are red,

(ii) all of them are of different colours,

(iii) all of them are of the same colour,

(iv) at least a green sweet is picked.

Solution:

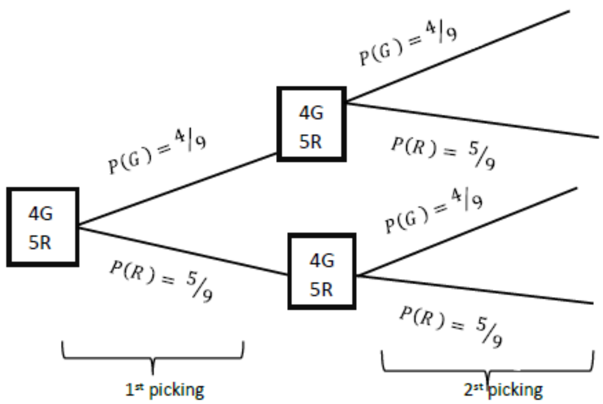


Figure 4.6: Case concerning picking with replacement.

(i) $P(\text{all sweets are red}) = P(R \cap R) = (\frac{5}{9} \times \frac{5}{9}) = \frac{25}{81}$

(iii)

$$P(\text{all are of different colours}) = (\frac{4}{9} \times \frac{5}{9}) + (\frac{5}{9} \times \frac{4}{9}) \\ = \frac{40}{81}$$

(iv)

$$P(\text{all are of same colour}) = (\frac{4}{9} \times \frac{4}{9}) + (\frac{5}{9} \times \frac{5}{9}) \\ = \frac{41}{81}$$

(iv)

$$P(\text{at least a green sweet}) = P(G \cap G) + P(R \cap G) + P(G \cap R) \\ = (\frac{4}{9} \times \frac{4}{9}) + (\frac{5}{9} \times \frac{4}{9}) + (\frac{4}{9} \times \frac{5}{9}) \\ = \frac{56}{81}$$

(2) Bag A contains 2 green and 3 red marbles, while Bag B contains 4 green and 1 red marbles. A box was picked at random and from a marble was picked, it's colour noted and then returned into the same bag. This was repeated once more.Find the probability that;

(i) all marbles are green,

(ii) all marbles are of different colours,

(iii) all marbles are of the same colour,

(iv) the second marble is red,

(v) the first marble is green given that second marble is red.

Solution:

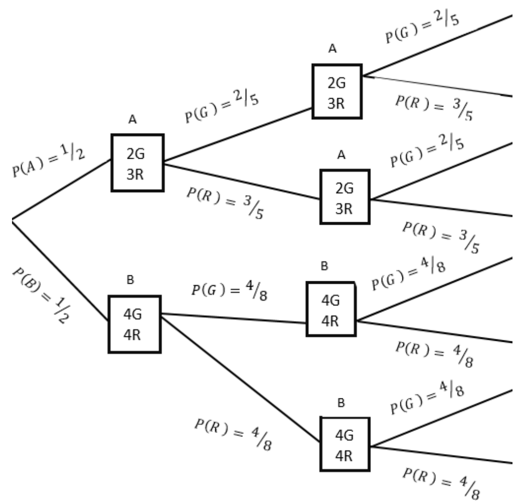


Figure 4.7: Case concerning fair picking of a box and there after picking with replacement.

(i) $P(\text{all marbles are green}) = (\frac{1}{2} \times \frac{2}{5} \times \frac{2}{5}) + (\frac{1}{2} \times \frac{4}{8} \times \frac{4}{8}) = \frac{13}{50} = 0.205$

$$(ii) P(\text{all marbles are of different colours}) = (\frac{1}{2} \times \frac{2}{5} \times \frac{3}{5}) + (\frac{1}{2} \times \frac{3}{5} \times \frac{2}{5}) + (\frac{1}{2} \times \frac{4}{8} \times \frac{4}{8}) + (\frac{1}{2} \times \frac{4}{8} \times \frac{4}{8}) = 0.490$$

(iii)

$$\begin{aligned} P(\text{all are of same colour}) &= (\frac{1}{2} \times \frac{2}{5} \times \frac{2}{5}) + (\frac{1}{2} \times \frac{4}{8} \times \frac{4}{8}) + (\frac{1}{2} \times \frac{3}{5} \times \frac{3}{5}) \\ &+ (\frac{1}{2} \times \frac{4}{8} \times \frac{4}{8}) \\ &= \frac{51}{100} = 0.510 \\ &Or \\ P(\text{all are of same colour}) &= 1 - P(\text{all marbles are of different colours}) \\ &= 1 - 0.490 = 0.510 \end{aligned}$$

(iv)

$$\begin{aligned} P(\text{second marble is red}) &= (\frac{1}{2} \times \frac{2}{5} \times \frac{3}{5}) + (\frac{1}{2} \times \frac{3}{5} \times \frac{3}{5}) \\ &+ (\frac{1}{2} \times \frac{4}{8} \times \frac{4}{8}) + (\frac{1}{2} \times \frac{4}{8} \times \frac{4}{8}) \\ &= \frac{11}{20} = 0.550 \end{aligned}$$

(iv)

$$P(G_1/R_2) = \frac{P(G_1 \cap R_2)}{P(R_2)} = \frac{((\frac{1}{2} \times \frac{2}{5} \times \frac{3}{5}) + (\frac{1}{2} \times \frac{4}{8} \times \frac{4}{8}))}{P(\text{second marble is red})} = \frac{0.245}{0.550} = 0.4455.$$

Probability Tree Diagram Concerning Picking with out Replacement:

Under this case, the number of elements(objects) in the given box keeps on reducing by 1 every time of a random sample is taken and therefore the probability of picking an element keeps on changing.

Examples:

(1) A box contains 4 green and 5 red sweets, two sweets are picked at random one after the other without replacement. Find the probability that;

(i) all sweets are red,

(ii) all of them are of different colours,

(iii) all of them are of the same colour,

(iv) at least a green sweet is picked.

Solution

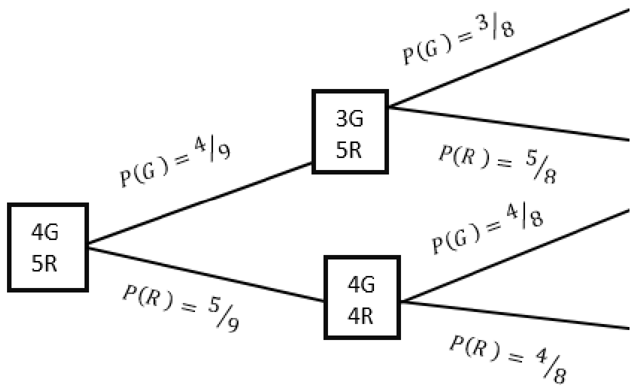


Figure 4.8: Case concerning picking without replacement.

(i) $P(\text{all sweets are red}) = P(R \cap R) = (\frac{5}{9} \times \frac{4}{8}) = \frac{5}{18}$

(ii)
$$P(\text{all are of different colours}) = (\frac{4}{9} \times \frac{5}{8}) + (\frac{5}{9} \times \frac{4}{8}) = \frac{5}{9}$$

(iii)
$$\begin{aligned} P(\text{all are of same colour}) &= (\frac{4}{9} \times \frac{3}{8}) + (\frac{5}{9} \times \frac{4}{8}) \\ &= \frac{4}{9} \end{aligned}$$

Or

$$P(\text{all are of same colour}) = 1 - P(\text{all are of different colours}) = 1 - \frac{5}{9} = \frac{4}{9}.$$

(iv)
$$\begin{aligned} P(\text{at least a green sweet}) &= P(G \cap G) + P(R \cap G) + P(G \cap R) \\ &= (\frac{4}{9} \times \frac{3}{8}) + (\frac{4}{9} \times \frac{5}{8}) + (\frac{5}{9} \times \frac{4}{8}) \\ &= \frac{13}{18} \end{aligned}$$

(3) Box A contains 3 red and 5 black balls, while Box B contains 4 red and 6 black balls. A box is chosen at random and two balls are picked at random one after the other from it without replacement. Find the probability that;

(i) balls are of the same colour,
(ii) the second black ball is the first black ball picked,
(iii)the probability density function for Red(x) balls picked.
Solution:

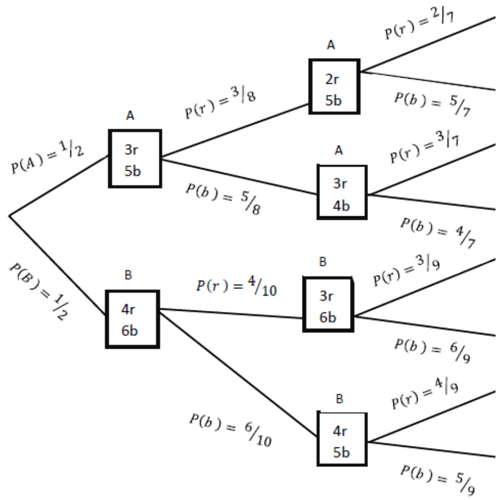


Figure 4.9: Case concerning the selection of the box first there after selection without replacement.

$$
\begin{aligned}
(i)P(\text{Balls are of the same colour}) &= \left(\frac{1}{2}\times\frac{3}{8}\times\frac{2}{7}\right)+\left(\frac{1}{2}\times\frac{5}{8}\times\frac{4}{7}\right)+\left(\frac{1}{2}\times\frac{4}{10}\times\frac{3}{9}\right) \\
&+ (\frac{1}{2}\times\frac{6}{10}\times\frac{5}{9}\Big) \\
&= \frac{391}{840}
\end{aligned}
$$

$$
\begin{aligned}
(ii)P(2^{rd} black is the 1^{st} black picked) &= \left(\frac{1}{2}\times\frac{3}{8}\times\frac{5}{7}\right)+\left(\frac{1}{2}\times\frac{4}{10}\times\frac{6}{9}\right) \\
&= \frac{449}{1680}.
\end{aligned}
$$

(iii)

x	$P(X = x)$
0	$\frac{29}{84}$
1	$\frac{449}{840}$
2	$\frac{101}{840}$

(4) Bag A contains 4 green and 4 red balls, Bag B contains 5 green and 5 red balls. A bag is chosen and two balls are withdrawn from it without replacement, if B is thrice as likely to be chosen as bag A. Find the probability that;
(i) balls are of different colours.
(ii) the two red balls picked came from bag A.
Solution:

Let $P(B) = 3P(A)$
From $P(A) + P(B) = 1 \Longleftrightarrow P(A) + 3P(A) = 1 \Longrightarrow P(A) = \frac{1}{4}$
$\therefore P(A) = \frac{1}{4}, P(B) = \frac{3}{4}$

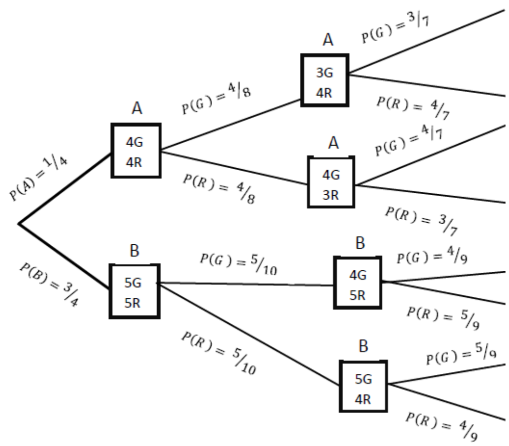


Figure 4.10: Case concerning the biased/unfair selection of boxes.

$$
\begin{aligned}
(i)P(\text{ different colours}) &= 1 - P(\text{balls are of same colour})\\
&= 1 - \Big((\frac{1}{4} \times \frac{4}{8} \times \frac{3}{7}) + (\frac{1}{4} \times \frac{4}{8} \times \frac{3}{7}) + (\frac{3}{4} \times \frac{5}{10} \times \frac{4}{9}) + (\frac{3}{4} \times \frac{5}{10} \times \frac{4}{9}) \Big)\\
&= 1 - \frac{37}{84}\\
&= \frac{47}{84}.
\end{aligned}
$$

$$
\begin{aligned}
(ii)P(A/2R) = \frac{P(A \cap 2R)}{P(2R)} &= \frac{P(A \cap R \cap R)}{P(A \cap R \cap R) + P(B \cap R \cap R)}\\
&= \frac{(\frac{1}{4} \times \frac{4}{8} \times \frac{3}{7})}{(\frac{1}{4} \times \frac{4}{8} \times \frac{3}{7}) + (\frac{3}{4} \times \frac{5}{10} \times \frac{4}{9})}\\
&= \frac{9}{37}
\end{aligned}
$$

(6) Kirabo is out of school and her probability of coming back is 0.35. The probability that is back and his team wins the game is 0.88 otherwise its 0.46. Find the probability that she is a round , given that their loses.
Solution:

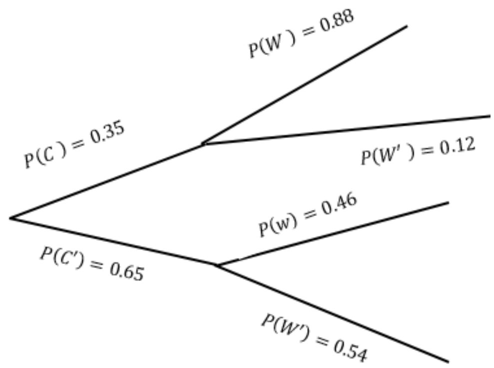


Figure 4.11: Case concerning the application of tree diagrams.

$$P(C/W') = \frac{P(C \cap W')}{P(W')} = \frac{(0.35 \times 0.12)}{(0.35 \times 0.12) + (0.65 \times 0.54)} = \frac{14}{131} = 0.1069.$$

(7) Bag A contains 4 black marbles and 2white marbles. The second bag Bcontains 3 black marbles and 5 white marbles. A marble was drawn at random from bag A and placed in B. A marble was then drawn at random from B into A.

(i) Find the probability that A contains exactly the same number of marbles as it had initially,

(ii) Find the probability of picking a black marble now from A,

(iii) Find the probability of picking a white marble now from A.

Solution:

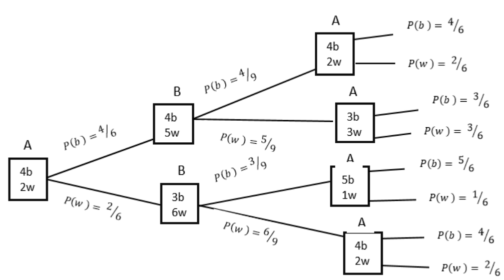


Figure 4.12: Case of transferring from one bag to another.

(i) $P(\text{same number as before}) = \left(\frac{4}{6} \times \frac{4}{9}\right) + \left(\frac{2}{6} \times \frac{6}{9}\right) = \frac{14}{27}.$

$$(ii) P(\text{black marble now from } A) = \left(\frac{4}{6} \times \frac{4}{9} \times \frac{4}{6}\right) + \left(\frac{4}{6} \times \frac{5}{9} \times \frac{3}{6}\right) + \left(\frac{2}{6} \times \frac{3}{9} \times \frac{5}{6}\right) + \left(\frac{2}{6} \times \frac{6}{9} \times \frac{4}{6}\right)$$
$$= \frac{101}{162} = 0.6235.$$

$(iii) P(\text{black marblenow from } A) = 1 - P(\text{black marble now from } A)$

$$= 1 - \left[\left(\frac{4}{6} \times \frac{4}{9} \times \frac{4}{6}\right) + \left(\frac{4}{6} \times \frac{5}{9} \times \frac{3}{6}\right) + \left(\frac{2}{6} \times \frac{3}{9} \times \frac{5}{6}\right) + \left(\frac{2}{6} \times \frac{6}{9} \times \frac{4}{6}\right)\right]$$

$$= 1 - \frac{101}{162}$$

$$= \frac{61}{162}$$

$$= 0.3765.$$

(8) Bag X contains 5 blue and 7 green balls, Bag Y contains 7 blue and 5 green balls. A ball is picked from bag X and placed in bag Y, a ball is then picked from bag Y and placed into bag X.
(a) Find the probability that;
(i) Each bag contains the same number of each colour as it was at the start.
(ii) Each bag now contains 6 blue and 6 green balls.
(b) if now a ball is picked again from bag X, what is the probability that a green ball is picked.

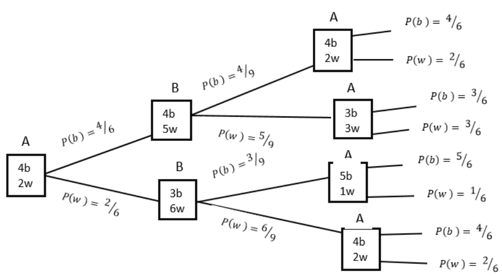


Figure 4.13

(i) $P(\text{same number as before}) = \left((\frac{5}{12} \times \frac{8}{13}\right) + \left(\frac{7}{12} \times \frac{6}{13}\right) = \frac{82}{156}$

(ii) $P(6 \text{ Blue and } 6 \text{ green}) = \left(\frac{7}{12} \times \frac{7}{13}\right) = \frac{49}{156}$

$(iii) P(\text{green ball is picked}) = \left(\frac{5}{12} \times \frac{8}{13} \times \frac{7}{12}\right) + \left(\frac{5}{12} \times \frac{5}{13} \times \frac{8}{12}\right) + \left(\frac{7}{12} \times \frac{7}{13} \times \frac{6}{12}\right) + \left(\frac{7}{12} \times \frac{6}{13} \times\right.$

$$= \frac{89}{156} = 0.5705.$$

(9) A bag contains 7 green, 4 blue and 6 red marbles. Three marbles were picked at random without replacement. Find the probability that;

(i) The first marble is blue and the third marble is green.

(ii) The first marble is green and the third marble is green.

(iii) The first marble is blue and the third marble is red.

(iv) The first marble is red and the third marble is green.

Solution:

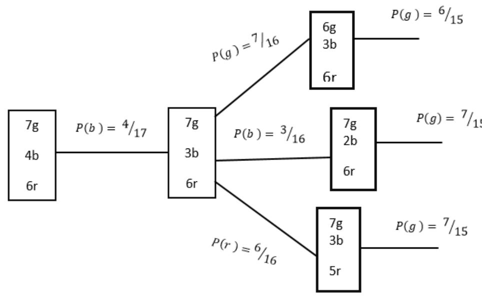


Figure 4.14: For simplicity,here you only consider the branches of the tree diagram you need to obtain the answer.

(i) $P(b_1 \text{ and } g_3) = \left(\dfrac{4}{17}\times\dfrac{7}{16}\times\dfrac{6}{15}\right)+\left(\dfrac{4}{17}\times\dfrac{3}{16}\times\dfrac{7}{15}\right)+\left(\dfrac{4}{17}\times\dfrac{6}{16}\times\dfrac{7}{15}\right) = \dfrac{7}{68} = 0.1029$

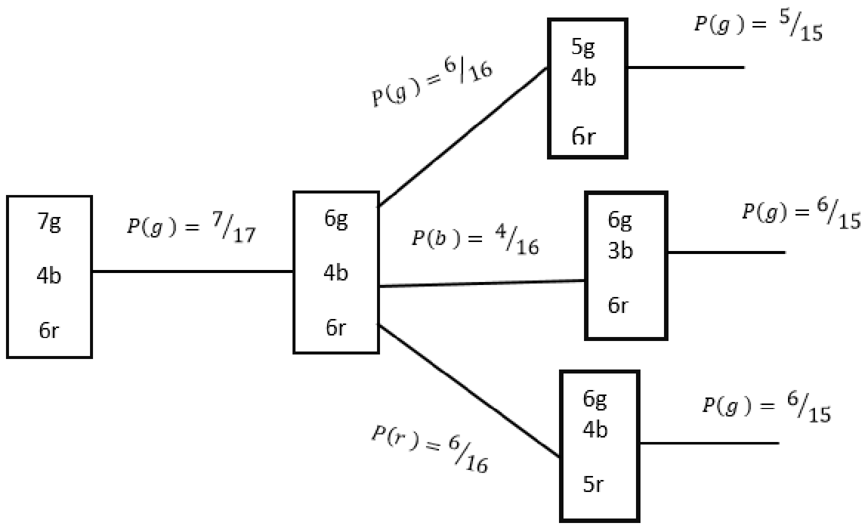


Figure 4.15: Here you need the branch og green only.

(ii) $P(g_1 \text{ and } g_3) = \left(\dfrac{7}{17}\times\dfrac{6}{16}\times\dfrac{5}{15}\right)+\left(\dfrac{7}{17}\times\dfrac{4}{16}\times\dfrac{6}{15}\right)+\left(\dfrac{7}{17}\times\dfrac{6}{16}\times\dfrac{6}{15}\right) = \dfrac{21}{136} = 0.1544$

Please complete the question.

(10) The table below shows the number of apples in boxes A, B and C.

Apples	A	B	C
Red	5	5	8
Green	4	5	7

A box is randomly selected and two apples are randomly selected from it without replacement. Box A is thrice as likely to be picked as B while A and C have the same chance of being selected. Determine the probability that;

(i) All apples are of different colors,

(ii) From box A given that they are of the same color.

Solution:

Let $P(A) = 3P(B)$ while $P(A) = P(C)$

But $P(A) + P(B) + P(C) = 1 \Longleftrightarrow 3P(B) + P(B) + 3P(B) = 1 \Longleftrightarrow P(B) = \frac{1}{7}$

$\therefore P(B) = \frac{1}{7}, P(A) = \frac{3}{7}$, and $P(C) = \frac{3}{7}$.

Please complete the question.

(11) At KAPROSS, 40% of the students offering mathematics in S.5 Biological are Girls. On a certain weekend 78% of the boys and 85% of the girls attended Mr. Peter's lesson. Of those who attended 80% 0f boys and 75% of girls understood the concepts taught about the topic, while among those who never attended, 38% of the boys and 30% of girls also understood the concepts. Find the probability that a student chosen at random from this class;

(i) attended the lesson

(ii) a girl who attended the lesson and understood concepts taught?

(iii) understood the concept?

(iv) is a boy who attended the lesson given that he did not understand the concept?

(v) did not attend the lesson given that understood the concept?

Solution

Draw the tree diagram for this question for easy understanding.

(i) $P(A) = P(B \cap A) + P(G \cap A) = (0.6 \times 0.78) + (0.4 \times 0.85) = 0.808$

(ii) $P(G \cap A \cap U) = P(0.4 \times 0.85 \times 0.75) = 0.255$

$$\begin{aligned}(iii)P(\text{understood}) &= P(B \cap A \cap U) + P(B \cap A' \cap U) + P(GnAnU) + P(G \cap A' \cap U)\\ &= (0.6 \times 0.78 \times 0.8) + (0.6 \times 0.22 \times 0.38)\\ &+ (0.4 \times 0.85 \times 0.75) + (0.4 \times 0.15 \times 0.3)\\ &= 0.69756\end{aligned}$$

(iv) $\mathrm{P(BnA/U')} = \frac{\mathrm{P(BnAnU')}}{\mathrm{P(U')}} = \frac{(0.6\times0.78\times0.2)}{1-\mathrm{P(U)}} = \frac{0.0936}{0.30244} = \frac{2340}{7561}$

(v) $\mathrm{P(A'/U)} = \frac{\mathrm{P(A'nU)}}{\mathrm{P(U)}} = \frac{\mathrm{P(BnA'nU)+P(GnA'nU)}}{\mathrm{P(U)}} = \frac{(0.6\times0.22\times0.38)+(0.4\times0.15\times0.3)}{0.69756} = \frac{568}{5813}$

4.1.7 Probabilities Involving Combinations:

This is the alternative way of finding probabilities other than use of tree diagram. A combination refers to the number of ways of selecting r items(things) from the n unlike items. This is denoted by $\begin{pmatrix} n \\ r \end{pmatrix}$ or ${}^{n}\mathrm{C}_{r}$ such that

$\begin{pmatrix} n \\ r \end{pmatrix} = \frac{n!}{(n-r)!r!}$

Here we obtain the probabilities by using the idea of

$$\text{probabilities} = \frac{n(events)}{n(sample\ size.)}$$

NB: Under combination, the order of arrangement of items is not important

NB: Finding probabilities by combinations is appreciate when we are having many selections

and the picking must be without replacement.
Examples

(1) A basket contains 6 green, and 7 red mangoes. Four Mangoes are picked at random one after the other without replacement. Find the probability that;

(i) All are red,

(ii) All are green,

(iii) all are of the same colour,

(iv) there is only one green mango,

(v) At most one mango picked is green,

(vi) At least three mangoes picked are red.

Solution:

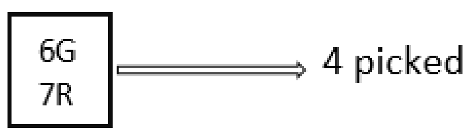


Figure 4.16

$$\text{(i) } P(\text{all are red}) = \frac{\binom{7}{4}\binom{6}{0}}{\binom{11}{4}} = \frac{7}{66} = 0.1061.$$

$$\text{(ii)} P(\text{all are green}) = \frac{\binom{7}{0}\binom{6}{4}}{\binom{11}{4}} = \frac{1}{22} = 0.0455.$$

$$\begin{aligned}(iii)P(\text{ all are of the same colour}) &= P(\text{all are red}) + P(\text{all are green})\\ &= \frac{\binom{7}{4}\binom{6}{0}}{\binom{11}{4}} + \frac{\binom{7}{0}\binom{6}{4}}{\binom{11}{4}}\\ &= \frac{7}{66} + \frac{1}{22}\\ &= \frac{5}{33}.\end{aligned}$$

$$\text{(iv)}P(\text{Only one green}) = \frac{\binom{6}{1}\binom{7}{(4-1)}}{\binom{11}{4}} = \frac{\binom{6}{1}\binom{7}{3}}{\binom{11}{4}} = \frac{7}{11} = 0.6364.$$

$$\begin{aligned}(v)P(\text{ At most one mango picked is green}) &= P(\text{zero green}) + P(\text{one green})\\ &= \frac{\binom{6}{0}\binom{7}{4}}{\binom{11}{4}} + \frac{\binom{6}{1}\binom{7}{(4-1)}}{\binom{11}{4}}\\ &= \frac{7}{66} + \frac{7}{11}\\ &= \frac{49}{66}.\end{aligned}$$

$$\begin{aligned}(vi)P(\text{ At least three mangoes picked are red}) &= P(\text{three red}) + P(\text{four red})\\ &= \frac{\binom{7}{3}\binom{6}{1}}{\binom{11}{4}} + \frac{\binom{7}{4}\binom{6}{0}}{\binom{11}{4}}\\ &= \frac{7}{11} + \frac{7}{66}\\ &= \frac{49}{66}.\end{aligned}$$

(2) A bag contains 6 green, 5 black and 6 red balls. Five balls were withdrawn at random from it without replacement, find the probability that;

(i) all are black,

(ii) all are green,

(iii) all are of the same colour,

(iv) there are only two red balls,

(v) more than half the picked ball are green,

(vi) At least three balls picked are red.

Solution:
Hint: For simplicity, whenever they give you more than two items or colours, all the items(colours) they are not interested in just combine them. I.e Add them together.
Total $= 6g + 5b + 6r = 17$items and Number selected $= 5$

$$\text{(i) } P(\text{all are black}) = \frac{\binom{5}{5}\binom{12}{0}}{\binom{17}{5}} = \frac{1}{6,188}.$$

(ii) $P(\text{all are green}) = \dfrac{\binom{6}{5}\binom{11}{0}}{\binom{17}{5}} = \dfrac{3}{3094}.$

$$
\begin{aligned}
(iii) P(\text{ all are of the same colour}) &= P(\text{all are black}) + P(\text{all are green}) + P(\text{all are red}) \\
&= \frac{\binom{5}{5}\binom{12}{0}}{\binom{17}{5}} + \frac{\binom{6}{5}\binom{11}{0}}{\binom{17}{5}} + \frac{\binom{6}{5}\binom{11}{0}}{\binom{17}{5}} \\
&= \frac{1}{6,188} + \frac{3}{3094} + \frac{3}{3094} \\
&= \frac{1}{476}.
\end{aligned}
$$

(iv) $P(\text{there are only two red balls}) = \dfrac{\binom{6}{2}\binom{11}{3}}{\binom{17}{5}} = \dfrac{2475}{6188} = 0.400.$

$$
\begin{aligned}
(v) P(3,4,5 \text{ green}) &= P(\text{ three green}) + P(\text{ only four green}) + P(\text{ only five green}) \\
&= \frac{\binom{6}{3}\binom{11}{2}}{\binom{17}{5}} + \frac{\binom{6}{4}\binom{11}{1}}{\binom{17}{5}} + \frac{\binom{6}{5}\binom{11}{0}}{\binom{17}{5}} \\
&= \frac{275}{1547} + \frac{165}{6188} + \frac{3}{3094} \\
&= \frac{1271}{6188} = 0.2054.
\end{aligned}
$$

(vi) **Please do the last part!**

4.1.8 Total Theorem and Bayes' theorem.

This theorem is an extension of conditional probability.It's defined as:

$$P(B_i/A) = \frac{P(B_i)P(A/B_i)}{\sum_{k=1}^{n}\Big(P(B_k)P(A/B_k)\Big)},$$

where $i = 1, 2, 3, \cdots, n$

Proof:
By conditional probability, $P(A/B) = \frac{P(A\cap B)}{P(B)}$ for $P(B) \neq 0$, then
$P(A/B) = \frac{P(A\cap B)}{P(B)} \iff P(A \cap B) = P(B) \times P(A/B) \cdots\cdots\cdots\cdots (i)$
Also;
From $P(B/A) = \frac{P(A\cap B)}{P(A)} \iff P(A \cap B) = P(A) \times P(B/A) \cdots\cdots\cdots\cdots (ii)$
From (i) and (ii), we can have that
$P(B_i \cap A) = P(A) \times P(B_i/A) \cdots\cdots\cdots\cdots (iii)$

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

From (iii)

$$P(B_i/A) = \frac{P(A).P(B_i/A)}{P(A)} \cdots\cdots\cdots\cdots (iv)$$

For $P(A)$, using,

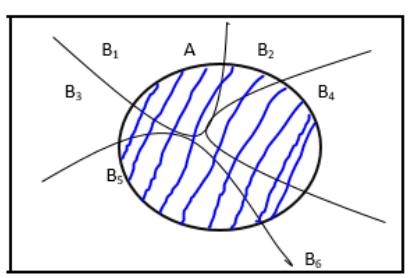


Figure 4.17

From the diagram,
$\implies P(A) = P(A \cap B_1) + P(A \cap B_2) + P(A \cap B_3) + \cdots + P(A \cap B_n)$
Substituting for $P(A)$ in equation (iv)

$$
\begin{aligned}
P(B_i/A) &= \frac{P(A).P(B_i/A)}{P(A)} \\
&= \frac{P(A).P(B_i/A)}{P(A\cap B_1)+P(A\cap B_2)+P(A\cap B_3)+\cdots+P(A\cap B_n)} \\
\implies P(B_i/A) &= \frac{P(B_i)P(A/B_i)}{\sum_{k=1}^{n}\Big(P(B_k)P(A/B_k)\Big)}.\text{Hence proved}
\end{aligned}
$$

. **Examples**

(1) The probability of a person catching a certain disease when exposed to it is 0.2 If a person has been immunized against the disease. The probability of catching the disease is 0.9 if the person is not immunized. If 70% of the population has been immunized against the disease, what is the probability that a person who caught the disease when exposed to it had been immunized.

Solution:

Let D denote catching a disease, and C denote being moving a round the city.
$P(D/I) = 0.2$, $P(D/C') = 0.95$
$P(I) = 0.7, P(I') = 0.3$

$$
\begin{aligned}
\text{Required } P(I/D) &= \frac{P(D\cap I)}{P(D)} \\
&= \frac{P(D\cap I)}{P(D\cap I)+P(D\cap I')} \\
&= \frac{(0.7\times 0.2)}{(0.7\times 0.2)+(0.3\times 0.9)} \\
&= \frac{0.14}{0.41} = \frac{14}{41} \\
&= 0.3415.
\end{aligned}
$$

Or
Apply the tree diagram

(2) During going for a tour, Katumba may go by road, air or water. The probability of using road, air and water are $0.3, 0.8$ and 0.6 respectively. The corresponding probabilities of arriving on time are $0.2, 0.8$ and 0.1 respectively. Find the probability of;
(i) Arriving on time,
(ii) having used the road given that she arrived on time,
(iii) using air if she did not arrive on time.

Solution:

(i) Let T denote arriving on time
$P(R) = 0.3, P(A) = 0.8, P(W) = 0.6$
$P(T/R) = 0.2, P(T/A) = 0.8, P(T/W) = 0.1$

$$(i)P(\text{Arriving on time }) = P(R \cap T) + P(A \cap T) + P(W \cap T)$$
$$= (0.3 \times 0.2) + (0.8 \times 0.8) + (0.6 \times 0.1)$$
$$= 0.760.$$

$$(i)P(\text{having used the road given that she arrived on time }) = \frac{P(R \cap T)}{P(T)}$$
$$= \frac{(0.3 \times 0.2)}{0.76}$$
$$= \frac{3}{38} = 0.0789$$

(iii)P(using air if she did not arrive on time)
$P(A/T') = \dfrac{P(A \cap T')}{\mathrm{P(T')}}$
But $P(A \cap T') = P(A) \times P(T'/A) = 0.8 \times 0.2 = 0.16$
$\Longrightarrow P(A/T') = \dfrac{0.16}{0.24} = \dfrac{2}{3}$

Repeat this qn with the use of a tree diagram.

(3) In a certain district, 30% of the people are conservative, 50% are Liberal and 20% are independents. Records indicate that in an election, 65% of the conservatives voted, 85% of the liberals voted and 50% of the independents voted. A person in the district is selected at random.

(i) Determine the probability that he voted,

(ii) Given that he didn't vote, determine the probability that he is a conservative.

Solution:
Let $C =$ A person is conservative
$L =$A person is Liberal
$I =$A person is Independent
$V =$A person is voted
$\Longrightarrow P(L) = 0.5, P(C) = 0.3, P(I) = 0.2, P(V/L) = 0.85, P(V/C) = 0.65$ and $P(V/I) = 0.5$,

(i) $$P(\text{A person voted}) = P(V) = P(C \cap V) + P(L \cap V) + P(I \cap V)$$
$$= P(C).P(V/C) + P(L).P(V/L) + P(I).P(V/I)$$
$$= (0.3 \times 0.65) + (0.5 \times 0.85) + (0.25 \times 0.5)$$
$$= 0.720.$$

(ii) $$P(\text{conservative given that didn't vote}) = P(C/V') = \frac{P(C \cap V')}{P(V')}$$
$$= \frac{P(C).P(V'/C)}{1 - P(V)}$$

$$\text{But}P(V'/C) + P(V/C) = 1 \Longleftrightarrow P(V'/C) = 1 - P(V/C) = 1 - 0.65 = 0.35$$

$$\therefore P(\text{conservative given that didn't vote,}) = \frac{(0.3 \times 0.35)}{1 - 0.72}$$
$$= 0.3750.$$

4.1.9 Exercise 4

(1) A bag contains 6 black marbles and 8 white marbles. A second bag contains 5 black marbles and 5 white marbles. A marble is drawn at random from the first bag and placed in the second bag. Given that, the probability of choosing the first bag is thrice that of the second bag. Find the probability that the first bag contains the same number of marbles as it had initially.
Find the probability of picking a white marble now.
Ans:

(2) A box contains two types of balls, red and black. When a ball is picked from the box, the probability that it is red is $\frac{7}{12}$. Two balls are selected at random from the box without replacement. Find the probability that;
(i) The second ball is black.
(ii) The first ball is red given that the second one is black.
Ans: (i) $\frac{5}{12}$ **(ii)** $\frac{7}{11}$

(3) An interview involves written, oral and practical test. The probability that an interviewee passes the written test is 0.8, the oral test is 0.6 and the practical test is 0.7 . What is the probability that the interviewee will pass;

(i) the entire interview?
(ii) Exactly two of the interview tests?
(iii) at most one of the tests?
(iv) at least one of the tests?

Ans: (i) 0.336 (ii) 0.452 (iii) 0.212

(4) Events A and B are independent such that $P(A) = y, 5P(B) = 5y+1$ and $20P(A\cap B) = 3$.
(i) Find the value of y.
(ii) Determine $P(A \cup B)$ and$P(A'/B')$.
Ans: (i) 0.3 (ii) 0.65, 0.7

(5) Events $C$and D are exhaustive events, such that $P(C/D) = 0.25$ and $P(D) = 0.67$. Find $P(C)$. and P(D/C)
Ans: 0.4975

(6) Events X and Y are such that $P(X) = 4/7, P(X \cap Y') = 1/3$ and $P(X/Y) = 5/14$. Find

(i) $P(X \cap Y)$,
(ii) $P(Y)$,
(iiI) $P(Y/X)$.

Ans: (i)5/21, (ii) 2/3, (iii) 5/12

(7) A fair coin is tossed four times. Find the probability of obtaining at least two heads.
Ans: 11/16

(8) Two ordinary dice are thrown, find the probability that;

(i) at least one six appears

(ii) at least one three appears

(iii) at least one six or one three appears

Ans: (i) 11/36 (ii) 11/36 (iii) 5/9

(9) Two independent events A and B are such that there is probability $\frac{1}{6}$ that they will both occur and a probability $\frac{1}{3}$ that neither will occur. Find $P(A)$ and $P(B)$.
Ans; $P(A) = 1/2$ **or** $1/3$ **and** $P(B) = 1/3$ **0r** $1/2$

(10) Miriam's probabilities of passing physics, economics and mathematics are $0.6, 0.75$ and 0.80 respectively.
(i) Find the probability that she passes at least two subjects.
(ii) If it is known that she passed at least two subjects, what is the probability that she failed Economics?
Ans; (i) 0.81 (ii) 4/27

(11) At a certain fuel station 30% of the customer buy Supper (S), 60% buy Regular (R) and the remainder Diesel (D). Of those who buy Supper, 25% fill the tank, 20% fill the tank with Diesel, 30% do not fill their tank with Regular.
(i) Find the probability that when a vehicle leaves the station, it has a full tank.
(ii) Given that the vehicle has a full tank, what is the probability that the tank contains Diesel/
Ans; (i) 0.515 (ii) 4/103

(12) 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 $P(C \cap B) = 0.1$, find;

(i) $P(A \cap B')$ (ii) $P(A/B \cap C)$ (iii) $P(A \cup B \cup C)$ (iv) $P(A/B' \cup C)$

Ans; (i) 0.88 (ii) 0.2

(13) A shop stocks tinned food of three moves. A, B and C of two sizes, large and small. Of the stock 60% is of brand A, 30% of brand B. Of the tins of brand A 30% are small, while for brand B and C, 40% and 70% are small respectively. Find the probability that;
(i) a tin chosen at random from the stock will be of small size.
(ii) a small tin chosen at random from the stock will be of brand A.
Ans; (i) 0.37 (ii) 18/37

(14) The events C and D are such that $P(C) = 1/3, P(C \text{ or } D \text{ but not both}) = 5/12, P(D) = 1/4$. Determine

(i) $P(C \cap D)$ (ii) $P(C \cap D')$ (iii) $P(C/D)$

Ans; (i) 1/12 (ii) 1/4 (iii) 1/3

(15) The table shows the likelihood of where teacher Isoba and teacher Tom spend their weekends.

Teacher	Isoba	Tom
Teach extra lessons	$\frac{1}{2}$	$\frac{1}{3}$
Attend a party	$\frac{1}{3}$	$\frac{5}{6}$
Stay at home	$\frac{1}{6}$	$\frac{2}{3}$

(i) Find the probability that both go out of home
(ii) If we know that they both go out, what is the probability that they both went to attend a party.
Ans; (i) 25/36 (ii) 2/25

(16) A box of 5 red balls and 6 blue balls. Three balls are selected at random without replacement. Find the probability that;

(i) they are of the same colour.

(ii) atmost two are of the same colour.

Ans; (i) 2/11 (ii) 9/11

(17) The probability that it will rain on any day next week is a half of the probability that it will rain on the same day this week; otherwise the probability it will rain on that day next week remain the same. The probability that it will rain on Friday this week is 1/4. What is the probability that
(i) it will not rain next friday.
(ii) it will rain on only two Fridays out o three consective Fridays starting this week.
Ans; (i) 7/8 (ii) 25/512

(18) Given that A and B are two independent events such that $P(A) = 0.2, P(B) = 0.15$, evaluate,

(i) $P(A/B)$

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

(iii) $P(A \cup B)$

Ans; (i) 0.2 (ii) 0.03 (iii) 0.32

(19) Two events A and B are such that $P(A) = 8/15, P(B) = 1/3, P(A/B) = 1/5$. Calculate the probabilities that;

(i) both events occur.

(ii) only one of the events occur.

(iii) neither events occur.

Ans; (i) 1/15 (ii) 11/15 (iii) 1/5

(20) Given that the events A and B are such that $P(A) = 1/3, P(B) = 2/5$ and $P(B/A') = 11/20$, find

(a) $P(A \cap B)$ (b) $P(B \cup A)$ (c) $P(A'/B)$ (d) $P(A/B)$

(e) State whether A and B are

(i) independent (ii) mutually exclusive

Ans; (a) 11/30 (b) 7/10 (c) 11/12 (d) 1/12

(21) Events A and B are such that $P(B) = 7/20, P(A/B) = 3/7$ and $2P(A) = 3P(A \cap B')$. Find

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

(ii) $P(A)$

(iii) $P(A/B')$

(iv) $P(A$ or B but not both$)$

Ans; (i) 3/20 (ii) 9/20

(22) Events A and B are such that $P(A) = 4/7, P(A \cap B') = 1/3$ and $P(A/B) = 5/14$, find

(i) $P(B)$

(ii) $P(B/A')$

(iii) $P(B \cap A/A')$

(iv) $P(B' \cup A')$

Ans; (i) 2/3 (ii) 1

(23) The probability that three players A, B and C score in a netball game are $1/3, 2/5$, and $3/4$ respectively. If they play together in a game, what is the probability that,

(i) only C scores,

(ii) at least one player scores,

(iii) at most two player scores,

(iv) two and only two players score.

Ans; (i) 3/10 (ii) 9/10 (iii) 23/60

(24) Two events A and B are independent. Given that $P(A \cap B') = 1/4$ and $P(A'/B) = 1/6$
.
(a) Show that also A' and B' are independent,
(b) Use a Venn-diagram to find the probabilities;

(i) $P(A)$ (ii) $P(B)$ (iii) $P(A \cap B)$ (iv) $P(A \cup B)$

Ans; (i) (ii) (iii) (iv)

(25) P, Q and C are three events completing all together with their respective probabilities of winning as $0.4, 0.38$ and 0.61. Find the probability that;

(i) All of them win, (ii) Only one wins, (iii) Only two win, (iv) None of them wins.

Ans; (i) (ii) (iii) (iv)

(26) A, B, C and D are three events completing in turns with their respective probabilities of winning as $0.32, 0.460.7$ and 0.54. Given that D starts first followed by A , then C and finally B, find the probability that;

(i) B wins, (ii) A wins, (iii) C wins on the second trial.

(27) Two biased tetrahedrons are such that an even number is twice as likely to occur as an odd one. If they are tossed once, find the probability that they show the same number.
Ans; 5/18

(28) In a particular government hospital 80% of the medical personnel are nurses and 20% are doctors. Of the nurses 40% are female, whilst of the doctors 10% are female. $33\frac{1}{2}\%$ of the men and 50% of the women are resident in government houses; residence in a government house is independent of whether the medical personnel is a doctor or a nurse. Find;
(i) the probability that a medical personnel chosen at random is a male doctor.
(ii) the probability that a medical personnel chosen at random is a male.
(iii) the probability that male medical personnel chosen at random is a doctor.
(iv) the probability that a doctor chosen at random is a female and a resident in a government house
Ans; (i) 0.18 (ii) 0.66 (iii) 3/11 (iv) 0.01

(29) The probability that it will be rainy on a November morning is 1/3 . The probability that Mr.Khalid will be late for work when it is foggy is 1/2 . The probability that he will be late if it is not rainy is 1/8 . If on a particular November morning Mr.Khalid is late, find the probability that it is rainy.
Ans; 2/3

(30) A bag contains 6 blue and 4 red counters, 3 counters are drawn at random and not replaced.
(i) Find the probability distribution for the number of red counters drawn.
(ii) Find the probability that there are at most 2 red counter books picked.
Ans; (i) (ii)

(31) A box contains b black and r red pens. One pen is drawn at random, but when it is put back in the box c additional pens of the same colour are put with it. Now suppose that we draw another pen. Show that the probability that the first pen drawn was black given that the second pen drawn was red is $\frac{b}{b+r+c}$

(32) A, B and C are three events completing in turns with their respective probabilities of winning as $0.5, 0.4$ and 0.47. Given that A starts first followed by B and finally C, find the probability that;
(i) B wins in the first trial,

(ii) C wins on the second trial,
(iii) A wins on the third trial,
(iv) A wins,
(v) B wins.

(33) A restaurant has three times as many male customers as females; 40% of men and 70% of women take the set lunch, the remainder choosing from among the optional items of the menu. Of men choosing the set lunch , 10% drink wine, 50% beer and the remainder a soft drink, whilst of those choosing the optional items the corresponding proportions are 60% and 30%. Among the women customers the corresponding proportions are 30% and 10% of those taking the set lunch and 40% and 20% of those choosing optional items. What proportion of;

(i) customers take the set lunch?
(ii) these drink wine?
(iii) women customers who drink beer?
(iv) wine-drinking customers who are men?

Ans; (i) 47.5% (ii) 38.25% (iii) 3.25% (iv) 78.43%

(34) The population of three towns B, C and D are in the ratio $5 : 7 : 9$, and the proportions of those who support Liverpool FC in these towns are 65%, 10% and 40% respectively. If a person at random from those living in the three towns and is found to support Liverpool, what is the probability that he;

(i) lives in town D,
(ii) is not a resident of C?

Ans; (i) 12/23 (ii) 62/69

(35) Mr. Tom is the senior inspector in an electronics-manufacturing firm. One of his roles is to inspect in coming lots of memory sticks produced by two terminals daily and determine their effectiveness. In a tray containing 8 sticks, two were found to be defective. He randomly selects three sticks for inspection without replacement. Form a probability distribution showing the number of defective sticks. Hence, find;
i. Probability that at most two of the three are defective
ii. The expected number of defective sticks

(36) Last year the employees of a firm either recieved no pay rise, a small pay rise or a large pay rise. The following table shows the number of each category, classified by whether they were weekly paid or monthly paid.

	No pay rise	Small pay rise	Large pay rise
Weekly paid	25	85	5
Monthly	4	8	23

(37) Joan has injury problems. When he is playing the probability that Barcelona will win is 3/4 but otherwise it is only 1/2. The probability that Joan will be fit this weekend is 1/3 . Determine the probability that his team will win the match.
Ans; 7/12

(38) Kato wishes to send a message to Mary. The probabilities that he uses e-mail, letter or personnel contact are 0.4, 0.1 and 0.5 respectively. He uses only one method; the probabilities of Mary receiving the message if John uses e-mail, letter or personnel contact are $0.6, 0.8$ and 0.1 respectively.
(i) Find the probability that Mary receives the message.
(ii) Given that Mary receives the message, find the probability that she receives it via e-mail.
Ans; (i)0.37 (ii) 24/37

(39) For two events A and B, such that $P(A/B) = 5/11, P(A \cup B) = 9/10, P(B) = X$. If $P(A \cap B) = 2P(A \cap B')$
(a) Show that $P(A) = \frac{9}{10} - \frac{6X}{11}$
(a) find the equation of X and deduce that $X = \frac{11}{15}$

(40) A bag contains 8 green, 12 blue and 5 red marbles. Three marbles were picked at random without replacement. Find the probability that;

(i) The first marble is blue and the third marble is blue,

(ii)The first marble is green and the third marble is green,

(iii)The first marble is blue and the third marble is red,

(iv)The first marble is red and the third marble is green.

(41) Events A and B are such that $P(A/B) = \dfrac{1}{3}, P(B/A') = \dfrac{5}{8}, P(B) = 0.08$ and $P(A' \cap B') = 3/20$. Find ;

(i) $P(A \cap B')$,

(ii) $P(A'/B')$,

(iii) $P(A \cup B')$.

(42) A box contains 18 green, 28 red marbles. 15 marbles were selected at random without replacement . find the probability that;

(i) All of them are green,

(ii) All of them are red,

(iii) None of them is red,

(iv) At most three are green,

(v) All of them are of the same colour,

(vi) Between 4 to 5 inclusive are red,

(vii) Exactly 10 are green.

Chapter 5

PROBABILITY DENSITY FUNCTIONS (pdf)

A function is said to be a pdf if it's random variable, say X that takes on either specific(whole) values or continuous (values within a given range)
The **pdf** is also known as a **probability distribution**
A probability distribution is a statistical function that describes all the possible values and likelihoods that a random variable can take within a given range. This range will be bounded between the minimum and maximum possible values, but precisely where the possible value is likely to be plotted on the probability distribution depends on a number of factors. These factors include the distribution mean (average),variance, standard deviation, skewness, and kurtosis.
Types of pdf:
There are two types of pdf and these include;
− Discrete pdf.
− Continuous pdf. (This is to be covered later)

5.1 DISCRETE PROBABILITY DISTRIBUTION

Introduction

A discrete random variable is probability distribution whose probability density function takes on whole numbers or countable domains.
It is also known as the probability density function (pdf).
It's abbreviated as $P(X = x)$ or $f(x)$ where X is a discrete random variable.
NB: Whole number or countable domain implies the integers only i.e $x = \cdots, -2, -1, 0, 1, 2, 3, \cdots$

5.1.1 Properties of a Discrete Random Variables.

If X is a discrete random variable, then

(i) $\sum_{\text{All } x} P(X = x) = 1.$, i.e Total probability for all values of x is equal to **one**,
NB: $\sum_{\textbf{All } x} P(X = x) = \sum_{\textbf{All x}} \mathbf{f(x)} = 1$

(ii) $\sum_{\text{All } x} \text{P(X=}x) \geq 0$. i.e probabilities can not be negative.

NB. We apply the first property to find unknowns.

5.1.2 Finding Probabilities in discrete pdf.

This done by substituting the random variable x directly into the function P(X=x) = f(x)
Here we must interpret the inequality signs i.e $>, \geq, <, \leq, or =$ in order to obtain the required probabilities such that:

(i) $P(X > a) = P(X = \underline{a+1}) + P(X = a+2) + \cdots$

(ii) $P(X \geq a) = P(X = a) + P(X = a+1) + \cdots$

(iii) $P(X < a) = P(X = \underline{a-1}) + P(X = a-2) + \cdots$

(iv) $P(X \leq a) = P(X = a) + P(X = a-1) + \cdots$

(v) $P(a < x < b) = P(X = \underline{a+1}) + P(X = a+2) + \cdots + P(X = b-2) + P(X = \underline{b-1})$.

(vi) $P(a \leq x \leq b) = P(X = a) + P(X = a+1) + \cdots + P(X = b-1) + P(X = b)$

Where a and b are constants.

5.1.3 Mean and Variance

Mean of X. This is defined as Mean $= \sum_{\text{All } x} xP(X = x)$
Mean is also known as Expectation of X denoted by $E(x)$ such that
$E(x) = \sum_{\text{All } x} xP(X = x)$
Properties of mean. These include:
Let a and b be constants, then

(i) $E(a) = a$

(ii) $E(aX) = aE(X)$

(iii) $E(aX + b) = aE(X) + b$

(iv) $E(X^{\text{n}}) = \sum_{\text{All } x} x^{\text{n}}\text{P}(\text{X} = x)$ such that $E(X^2) = \sum_{\text{All } x} x^2\text{P}(\text{X} = x)$

Variance of X. This is abbreviated as $Var(X)$ such that $Var(X) = E(X^2) - [E(X)]^2$
where $E(X^2) = \sum_{\text{All } x} x^2\text{P}(\text{X} = x)$ and
$E(x) = \sum_{\text{All } x} xP(X = x)$.

Properties of $Var(X)$ Let a and b be constants, then

(i) $Var(a) = 0$

(ii) $Var(aX) = a^2Var(X)$

(iii) $Var(aX + b) = a^2Var(X)$

NB. Standard deviation, $\sigma = \sqrt{\text{variance}}$
Examples:

(1) A discrete probability function is given as P(X = 1) = 0.3, P(X = 2) = 0.1, P(X = 3) = 0.4, P(X = 4) = 0.2 . Find

(i) Show that the distribution above is a pdf.

(ii) mean (iv) Var(X) (vi) P(X > 2) (viii) P(X = 2/ X ≤ 3)

(iii) E(X + 2)^2 (v) Var(3X - 4) (vii) P(1 < X ≤ 3) (ix) P(X ≤ 3/ X ≥ 1)

Solution

x	$P(X = x)$	$xP(X = x)$	$x^2P(X = x)$
1	0.3	0.3	0.30
2	0.1	0.2	0.40
3	0.4	1.20	3.60
4	0.2	0.80	3.20
	$\sum P(X = x) = 1$	$\sum xP(X = x) = 2.50$	$\sum x^2P(X = x) = 7.50$

(i) Since $\sum P(X = x) = 1$ and all $P(X = x) \geq 0$, then it's a pdf.

(ii) $E(X) = \sum_{\text{All } x} x$ P(X = x) = 2.50

(iii) E(X + 2)^2 = E(X^2 + 4X + 4) = E(X^2) + 4E(X) + 4
where E(X^2) = $\sum_{\text{All } x} x^2$P(X = x) = 7.50
$\Longrightarrow E(X + 2)^2 = 7.5 + 4(2.5) + 4 = 21.5$

(iii) $Var(X) = E(X^2) - [E(X)]^2 = 7.5 - (2.5)^2 = 1.25$

(iv) $Var(3X - 4) = 3^2Var(X) = 9 \times 1.25 = 11.25$

(v) $P(X > 2) = P(X = 3) + P(X = 4) = 0.4 + 0.2 = 0.6$

(vi) $P(1 < X \leq 3) = P(X = 2) + P(X = 3) = 0.1 + 0.4 = 0.5$

(vii) $P(X = 2/X \leq 3) = \frac{\text{P(X = 2)}}{\text{P(X = 3) + P(X = 2) + P(X = 1)}} = \frac{0.1}{0.4+0.1+0.3} = \frac{1}{8}$

(ix) $P(X \leq 3/X \geq 1) = \dfrac{P(1 \leq X \leq 3)}{P(X \geq 1)} = \dfrac{0.3 + 0.1 + 0.4}{(0.3 + 0.1 + 0.4 + 0.2)} = 0.8$

(2) A random variable X of a discrete pdf is given as

$$f(x) = \begin{cases} k(x+2) & ; \quad x = -1, 0, 1, 2, 3, 4 \\ 0 & ; \quad \text{Otherwise} \end{cases} \quad \text{where k is a constant}$$

Find:

(a) The value of K

(b) $P(X=3)$

(c) E(X)

(d) Standard deviation

(e) $P(X>1/X\leq 3)$

(f) $P(X\leq 3/X\geq 2)$

Solution

Using

$$f(x)=\begin{cases} k(x+2) & ; \quad x=-1,0,1,2,3,4 \\ 0 & ; \quad \text{Otherwise} \end{cases}$$

x	$P(X=x)=f(x)$
-1	K
0	2K
1	3K
2	4K
3	5K
4	6K
	$\sum P(X=x)=1$

(a). From $\sum P(X=x)=1$,

$\Longrightarrow K+2K+3K+4K+5K+6k=1 \Longleftrightarrow K=\dfrac{1}{21}$

x	$P(X=x)=f(x)$	$xP(X=x)$	$x^2P(X=x)$
-1	$\frac{1}{21}$	$\frac{-1}{21}$	$\frac{1}{21}$
0	$\frac{2}{21}$	0.0	0.0
1	$\frac{3}{21}$	$\frac{3}{21}$	$\frac{3}{21}$
2	$\frac{4}{21}$	$\frac{8}{21}$	$\frac{16}{21}$
3	$\frac{5}{21}$	$\frac{15}{21}$	$\frac{45}{21}$
4	$\frac{6}{21}$	$\frac{24}{21}$	$\frac{96}{21}$
	1	$\sum xP(X=x)=\frac{49}{21}$	$\sum x^2P(X=x)=\frac{161}{21}$

(ii) $P(X=3)=\dfrac{5}{21}$

(c) $E(X)=\sum xP(X=x)=\dfrac{49}{21}=2.3333$

(d) Standard deviation $=\sqrt{Var(X)}=\sqrt{\left(\dfrac{161}{21}-\dfrac{49}{21}\right)}=1.4907$

(e) $P(X>1/X\leq 3)=\dfrac{\left(\frac{4}{21}+\frac{5}{21}\right)}{\left(\frac{1}{21}+\frac{2}{21}+\frac{3}{21}+\frac{4}{21}+\frac{5}{21}\right)}=0.60$

(f) $P(X\leq 3/X\geq 2)=\dfrac{\left(\frac{4}{21}+\frac{5}{21}\right)}{\left(\frac{4}{21}+\frac{5}{21}+\frac{6}{21}\right)}=0.60$

(3) A random variable X has got the following probability distribution

$$f(x) = \begin{cases} kx^2 & ; \quad x = 1, 2 \\ k(8-x) & ; \quad x = 3, 4, 5 \\ 2k & ; \quad x = 6, 7 \\ 0 & ; \quad \text{elsewhere} \end{cases}$$

(i) Determine the value of the constant k
(ii) Var(X)
(iii) Find $P(X < 5)$ and $P(2 \leq x < 6)$
Solution:
Using

$$f(x) = \begin{cases} kx^2 & ; \quad x = 1, 2 \\ k(8-x) & ; \quad x = 3, 4, 5 \\ 2k & ; \quad x = 6, 7 \\ 0 & ; \quad \text{elsewhere} \end{cases}$$

x	$P(X = x)$
1	K
2	4K
2	5K
4	4K
5	3K
6	2K
7	2K
	$\sum P(X = x) = 1$

(a). From $\sum P(X = x) = 1$,
$\Longrightarrow K + 4K + 5K + 4K + 3K + 2k + 2K = 1 \Longleftrightarrow K = \dfrac{1}{21}$

x	$P(X = x) = f(x)$	$xP(X = x)$	$x^2P(X = x)$
1	$\frac{1}{21}$	$\frac{1}{21}$	$\frac{1}{21}$
2	$\frac{4}{21}$	$\frac{8}{21}$	$\frac{16}{21}$
3	$\frac{5}{21}$	$\frac{15}{21}$	$\frac{45}{21}$
4	$\frac{4}{21}$	$\frac{16}{21}$	$\frac{64}{21}$
5	$\frac{3}{21}$	$\frac{15}{21}$	$\frac{75}{21}$
6	$\frac{2}{21}$	$\frac{12}{21}$	$\frac{72}{21}$
7	$\frac{2}{21}$	$\frac{14}{21}$	$\frac{98}{21}$
	1	$\sum xP(X = x) = \dfrac{81}{21}$	$\sum x^2P(X = x) = \dfrac{371}{21}$

(ii) $Var(X) = E(X^2) - E(X) = \dfrac{371}{21} - \left(\dfrac{81}{21}\right)^2 = 2.7891.$

(iii) $P(X < 5) = \left(\dfrac{1}{21} + \dfrac{4}{21} + \dfrac{5}{21} + \dfrac{4}{21}\right) = \dfrac{2}{3}.$ and

$P(2 \leq x < 6) = \left(\dfrac{4}{21} + \dfrac{5}{21} + \dfrac{4}{21} + \dfrac{3}{21}\right) = \dfrac{16}{21}.$

(4) A family is planning to have 3 children. If X represents the number of girls in that family. Find

(a) probability distribution of X

(b) E(X)

(c) standard deviation

Solution

(a) P(B) = $\frac{1}{2}$, P(G) = $\frac{1}{2}$

For x = 0, P(BBB) = $\left(\frac{1}{2} \times \frac{1}{2} \times \frac{1}{2}\right) = \frac{1}{8}$

For x = 1, P(BBG) + P(BGB) + P(GBB)

$$= \left(\frac{1}{2} \times \frac{1}{2} \times \frac{1}{2}\right) + \left(\frac{1}{2} \times \frac{1}{2} \times \frac{1}{2}\right) + \left(\frac{1}{2} \times \frac{1}{2} \times \frac{1}{2}\right) = \frac{3}{8}$$

For x = 2, P(BGG) + P(GGB) + P(GBG)

$$= \left(\frac{1}{2} \times \frac{1}{2} \times \frac{1}{2}\right) + \left(\frac{1}{2} \times \frac{1}{2} \times \frac{1}{2}\right) + \left(\frac{1}{2} \times \frac{1}{2} \times \frac{1}{2}\right) = \frac{3}{8}$$

For x = 3, P(GGG) = $\left(\frac{1}{2} \times \frac{1}{2} \times \frac{1}{2}\right) = \frac{1}{8}$

x	0	1	2	3	
P(X = x)	$\frac{1}{8}$	$\frac{3}{8}$	$\frac{3}{8}$	$\frac{1}{8}$	$\sum P(X = x) = 1$

(b) (i) $E(X) = \sum_{\text{All x}} \text{x P(X = x)}$

$$= \left(0 \times \frac{1}{8}\right) + \left(1 \times \frac{3}{8}\right) + \left(2 \times \frac{3}{8}\right) + \left(3 \times \frac{1}{8}\right) = \frac{12}{8} = \frac{3}{2}$$

(c) Standard deviation, $\delta = \sqrt{\text{variance}}$
But $\text{E}(\text{X}^2) = \sum_{\text{All x}} \text{x}^2\text{P(X = x)}$

$$= \left(0^2 \times \frac{1}{8}\right) + \left(1^2 \times \frac{3}{8}\right) + \left(2^2 \times \frac{3}{8}\right) + \left(3^2 \times \frac{1}{8}\right) = \frac{24}{8} = 3$$

Var(X) = 3 - 1.5^2 = 0.75
$\delta = \sqrt{0.75} = 0.8660$

(5) A biased coin is such that a head is twice as likely to show a tail, is tossed four times and Y represents the number of heads that show. (a) Write down the ;

(i) sample space outcome

(ii) probability distribution of Y

(b) Find the expectation of Y + 1

Solution

P(H) = 2P(T)
P(H) + P(T) = 1
2P(T) + P(T) = 1
P(T) = $\frac{1}{3}$
P(H) = $\frac{2}{3}$

(i) Sample space(S) = x^n where x is the number of sides and n is the number of times an oblect is thrown/ tossed.

S =
$$\left\{\begin{array}{r} (HHHH), (HHHT), (HHTH), (HTHH), (THHH) \\ (HHTT), (HTTH), (TTHH), (THTH), (HTHT), (THHT) \\ (TTTH), (TTHT), (THTT), (HTTT), (TTTT) \end{array}\right\}$$

$n(S) = 2^4 = 16$

(ii)
For no head (y = 0) = $\left(\frac{1}{3} \times \frac{1}{3} \times \frac{1}{3} \times \frac{1}{3}\right) = \frac{1}{81}$
For one head (y = 1) = $4 \times \left(\frac{2}{3} \times \frac{1}{3} \times \frac{1}{3} \times \frac{1}{3}\right) = \frac{8}{81}$
For two head (y = 2) = $6 \times \left(\frac{2}{3} \times \frac{2}{3} \times \frac{1}{3} \times \frac{1}{3}\right) = \frac{24}{81}$
For three head (y = 3) = $4 \times \left(\frac{2}{3} \times \frac{2}{3} \times \frac{2}{3} \times \frac{1}{3}\right) = \frac{32}{81}$
For four head (y = 4) = $\left(\frac{2}{3} \times \frac{2}{3} \times \frac{2}{3} \times \frac{2}{3}\right) = \frac{14}{81}$

y	0	1	2	3	4	
P(Y = y)	$\frac{1}{81}$	$\frac{8}{81}$	$\frac{24}{81}$	$\frac{32}{81}$	$\frac{16}{81}$	$\sum$ P(Y = y) = 1

(a) $E(Y+1) = E(Y) + 1 = \sum \mathrm{P(Y = y)} + 1$
$\mathrm{E(Y + 1)} = \left(0 \times \frac{1}{81}\right) + \left(1 \times \frac{8}{81}\right) + \left(2 \times \frac{24}{81}\right) + \left(3 \times \frac{32}{81}\right) + \left(4 \times \frac{16}{81}\right) + 1 = \frac{11}{3}$

(5) Four boxes contain green and yellow mangoes as follows.

Box	A	B	C	D
Green	3	2	6	4
Yellow	1	3	3	5

A box is selected and a mango is picked from it, if the probability of choosing box C is thrice any other;
(a) Find the probability distribution for picking a green mango.
(b) Obtain the mean of this distribution.

Solution

(a) $P(A) = P(B) = P(D) = k, P(C) = 3k$

$P(A) + P(B) + P(C) + P(D) = 1$
$k + k + 3k + k = 1$
$k = \frac{1}{6}$

For no green (X = 0) = P(AnY) + P(BnY) + P(CnY) + P(DnY)

$$= \left(\frac{1}{6} \times \frac{1}{4}\right) + \left(\frac{1}{6} \times \frac{3}{5}\right) + \left(\frac{3}{6} \times \frac{3}{9}\right) + \left(\frac{1}{6} \times \frac{5}{9}\right) = \frac{433}{1080}$$

For one green (X = 1) = P(AnG) + P(BnG) + P(CnG) + P(DnG)

$$= \left(\frac{1}{6} \times \frac{3}{4}\right) + \left(\frac{1}{6} \times \frac{2}{5}\right) + \left(\frac{3}{6} \times \frac{6}{9}\right) + \left(\frac{1}{6} \times \frac{4}{9}\right) = \frac{647}{1080}$$

x	0	1	
$P(X = x)$	$\frac{433}{1080}$	$\frac{647}{1080}$	$\sum$P(X = x) = 1

(b) $E(X) = \sum_{\text{All x}} \text{x P(X = x)} = \left(0 \times \frac{433}{1080}\right) + \left(1 \times \frac{647}{1080}\right) = \frac{647}{1080}$

(6) The table below shows the number of apples in boxes A, B and C.

Apples	A	B	C
Red	6	5	8
Green	6	6	7

A box is randomly selected and two apples are randomly selected from it without replacement. Box A is twice as likely to be picked as B while B and C have the same chance of being selected. If X is the number of green apples taken, draw a probability distribution table for X. hence calculate the mean and standard deviation of X.
Solution:
Let $P(A) = 2P(B)$ while $P(B) = P(C)$
But $P(A) + P(B) + P(C) = 1 \iff 2P(B) + P(B) + P(B) = 1 \iff P(B) = \dfrac{1}{4}$
$\therefore P(B) = \dfrac{1}{4}, P(A) = \dfrac{2}{4} = \dfrac{1}{2}$, and $P(C) = \dfrac{1}{4}$.
Please complete the question.

(7) A biased tetrahedron is such that the chance of any face showing uppermost is inversely proportional to the cube of the number x on it.
(a) Find the probability distribution of X.
(b) obtain

(i) $E(2X - 1)$ (ii) $P(X \geq 3)$

Solution:

$P(X=x) \propto \frac{1}{x^3} \iff P(X=x) = \frac{K}{x^3}, x = 1, 2, 3, 4$ Where K is a constant.

Then using

x	P(X= x)
1	$\frac{K}{1}$
2	$\frac{K}{8}$
3	$\frac{K}{27}$
4	$\frac{K}{64}$
	$\sum P(X=x) = 1$

$$\Longrightarrow \sum P(X=x) = \frac{K}{1} + \frac{K}{8} + \frac{K}{27} + \frac{K}{64} = 1$$

$$\Longrightarrow \text{k} = \frac{1728}{2053}$$

$$\therefore P(X=x) = \frac{1728}{2053x^3}, x = 1, 2, 3, 4$$

x	1	2	3	4	
$P(X=x)$	$\frac{1728}{2035}$	$\frac{216}{2035}$	$\frac{64}{2035}$	$\frac{27}{2035}$	

(b)$E(2X-1) = 2E(X) - 1$

$= \left[\left(1 \times \frac{1728}{2035}\right) + \left(2 \times \frac{216}{2035}\right) + \left(3 \times \frac{64}{2035}\right) + \left(4 \times \frac{27}{2035}\right)\right] - 1$

$= \frac{85}{407} = 0.2088$

(ii) $P(X \geq 3) = P(X=3) + P(X=4) = \frac{64}{2035} + \frac{27}{2035} = \frac{91}{2035}$

(6) A radioactive materials emit radiations every week and this follows a probability distribution given by;

$$f(x) = \left\{ \begin{array}{rl} k\left(\frac{2}{3}\right)^x & ; \quad x = 0, 1, 2, 3, \cdots \\ 0 & ; \quad \text{otherwise} \end{array} \right\}.$$

(i) Find the value of constant k

(ii) Probability that at most two radiations will be emitted a week.

Solution

(i) $\sum_{\text{All x}} \text{P(X = x)} = \text{k}\left(\frac{2}{3}\right)^0 + \text{k}\left(\frac{2}{3}\right)^1 + \text{k}\left(\frac{2}{3}\right)^2 + \text{k}\left(\frac{2}{3}\right)^3 \cdots = 1$

$\text{k}\left[1 + \frac{2}{3} + \frac{4}{9} + \frac{8}{27} + \cdots\right] = 1$

We obseve a series going to infinity; $\text{S}_\infty = \frac{a}{1-r} = \frac{1}{1-\frac{2}{3}} = 3$

$\text{k}\left(\text{S}_\infty\right) = 1$

$3\text{k} = 1$, $\text{k} = \frac{1}{3}$

(ii) $\text{P}(\text{X} \leq 2) = \text{P}(\text{X} = 2) + \text{P}(\text{X} = 1) + \text{P}(\text{X} = 0) = \text{k}\left(1 + \frac{2}{3} + \frac{4}{9}\right) = \frac{1}{3}\left(\frac{19}{9}\right) = \frac{19}{27}$

5.1.4 Cumulative Probability Density Function (c.d.f)

This is done by continuously adding the probabilities up to one. It's abbreviated as $F(X)$ such that $F(+\infty) = 1$ where the $+\infty$ is the upper limit of the distribution.
It's is defined as $F(X) = P(X \leq x)$.

5.1.5 Mode and Median

Mode: This is the value of x corresponding to the highest probability in the given pdf. It's also known as the most likely value of x.

Median: This is the value of x corresponding to the first cumulative probability value that is at least 0.5, I.e $F(Median) \geq 0.5 \iff F(m) \geq 0.5$
For quartiles, use $F(q_1) \geq 0.25$ and $F(q_3) \geq 0.75$
NB: m, q_1 and q_3 are all obtained from the x values.

5.1.6 Graph of f(x):

This is sketched by drawing vertical lines parallel to the y-axis drawn at every specified value of x. This is done by plotting $f(x)$ against x.

5.1.7 Graph of F(x):

This is sketched by drawing horizontal straight lines parallel to the x - axis. This is done by plotting $F(x)$ against x.
The lines are drawn from the a given x value to the right up to below the next values.

Examples.

(1) A discrete random variable X has its distribution as given in the table below.

x	1	2	3	4
P(X = x)	0.3	$2a$	0.1	a

(a) Find

(i) value of a

(ii) mode

(iii) median

(iv) 8^{th} deciles

(v) lower quartile

(vi) middle 80%

(vii) Mean

(viii) Standard deviation

(ix) The probability of obtaining at most $x = 3$

(b) If $Y = 2X + 1$, find the probability distribution of Y and hence find $E(Y)$

(c) sketch $f(x)$ and $F(X)$

Solution

(a)(i) $\sum_{\text{All x}} P(X = x) = 1 \iff 0.3 + 2a + 0.1 + a = 1 \iff 3a = 0.6, \implies a = 0.2$

x	$P(X = x)$	$F(x) = P(X \leq \text{x})$
1	0.3	0.3
2	0.4	0.7
3	0.1	0.8
4	0.2	1
	$\sum P(X = x) = 1$	

(ii) Mode = 2

(iii) Median = 2

(iv) 8^{th} deciles = ?

Position of 8^{th} deciles $= \dfrac{8}{10} = 0.8^{th}$ position, $\implies 8^{th}$ deciles = 3

(v) lower quartile = ? Position of lower quartile $= \dfrac{25}{100} = 0.25^{th}$ position, $\implies$ lower quartile = 1

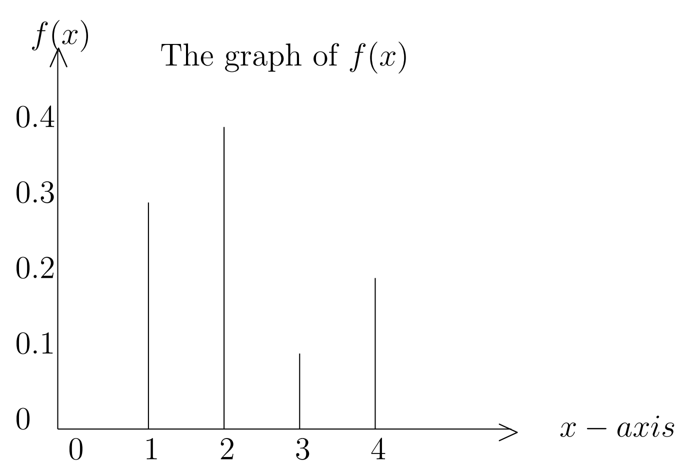
(vi) middle 80% = P_{90} - P_{10} = 4 - 1 = 3

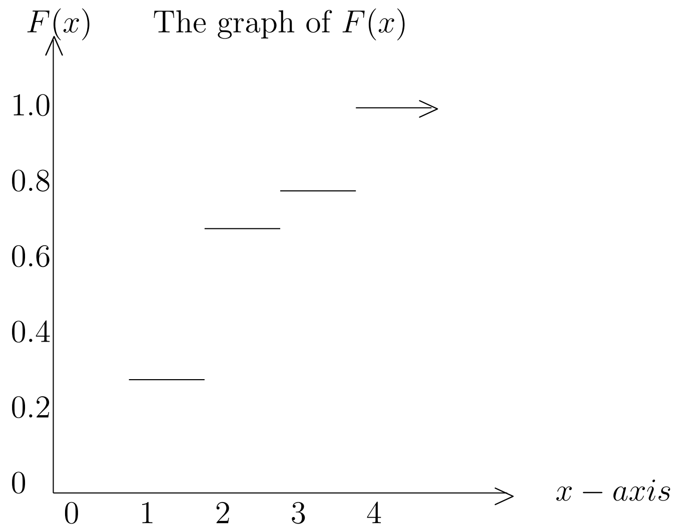
(b) $Y = 2X + 1$

Y	3	5	7	9
$P(Y)$	0.3	0.4	0.1	0.2

$E(Y)$ =(3 × 0.3) + (5 × 0.4) + (7 × 0.1) + (9 × 0.2) = 5.4

(c)





The other parts are left as exercise for you.

(2) A random variable X has its c.d.f given as in the table below;

x	1	3	4	5	8
$P(X \leq \text{x})$	k	4k	5k	7k	10k

(a) Find the value of constant k
(b) Sketch the graphs of;

(i) $f(x)$ (ii) $F(x)$

Solution

(a) Using $F(+\infty) = 1$
$\Longrightarrow 10k = 1, \Longleftrightarrow k = 0.1$

x	1	3	4	5	8
$f(x) = P(X = x)$	0.1	0.3	0.1	0.2	0.3
$F(x) = P(\text{X} \leq \text{x})$	0.1	0.4	0.5	0.7	1

Part (b) above is left as your exercise.

(3) Given that:

$$P(X \leq x) = \begin{cases} kx & ; \quad x = 1, 2, 3, , 4, 5 \\ 1 & ; \quad \text{Otherwise} \end{cases} \quad \text{where } k \text{ is a constant}$$

(a) Find

(i) value of k

(ii) pdf

(iii) median

(iv) Sketch of $f(x)$ and $F(x)$

(v) Semi interqurtile range

(vi) Mode

(vii) Mean

(viii) Standard deviation

(ix) The probability of obtaining at least $x = 2$

Solution

Solution
(i) Using

$$f(x) = \begin{cases} k(x) & ; \quad x = 1, 2, 3, 4, 5 \\ 0 & ; \quad \text{Otherwise} \end{cases}$$

x	$P(X \leq x) = F(x)$
1	K
2	2K
3	3K
4	4K
5	5K

From $F(+\infty) = 1$

$\Longrightarrow 5K = 1 \Longleftrightarrow k = \dfrac{1}{5}$ (ii)

x	$P(X \leq x) = F(x)$	$p(X = x)$
1	0.2	0.2
2	0.4	0.2
3	0.6	0.2
4	0.8	0.2
5	1	0.2
		1

The remaining parts are left as exercise.

(4) The discrete random variable X has probability function

$$P(X = x) = \begin{cases} \dfrac{kx}{(x^2+1)} & : \quad x = 2, 3 \\ \dfrac{2kx}{(x^2-1)} & : \quad x = 4, 5 \\ 0 & : \quad \text{Otherwise} \end{cases} \quad \text{where k is a constant}$$

(a) Show that the value of k is $\dfrac{20}{33}$
(b) Find the probability that X is less than 2 or greater than 4
(c) Find $F(4)$
(d) Find $i)E(X),\ ii)Var(X)$

Solution

x	$P(X=x)=f(x)$
2	$\frac{2k}{5}$
3	$\frac{3k}{10}$
4	$\frac{8k}{15}$
5	$\frac{10k}{24}$
	$\sum P(X=x) = \frac{33k}{20}$

Using $\sum P(X=x) = 1 \Longrightarrow \dfrac{33k}{20} = 1 \Longleftrightarrow k = \dfrac{20}{33}$

x	$P(X=x)=f(x)$	$F(X)$	
2	$\frac{8}{33}$	$\frac{8}{33}$	
3	$\frac{2}{11}$	$\frac{14}{33}$	
4	$\frac{32}{99}$	$\frac{74}{99}$	
5	$\frac{25}{99}$	1	
	$\sum P(X=x) = 1$		

(b) $P(X = 2 \text{ } or \text{ } X = 4) = \dfrac{8}{33} + \dfrac{32}{99} = \dfrac{56}{99} = 0.5657$

(c) $F(X) = \frac{74}{99} = 0.7475$

(d) **Left as exercise for you.**

(5) A news agent stocks 15 copies of a magazine each week. He has regular order for 11, and the additional copies sold varies from week to week. The news agent uses previous sales data to estimate the probability for each possible total number of copies sold as follows:

Number of copies	11	12	13	14	15
probability	0.4	0.2	0.2	0.15	0.05

(a) Calculate the expected number of copies that he sells in a week.

(b) The news a gent buys the magazines at $shs.$1500 each and sells them at $shs.$2000 each. Any copies left unsold are destroyed.

(i) Find the profit on these magazines in a week when he sells 13 copies.

(ii) Construct a probability distribution table for the news agent's weekly profit from the sale of these magazines. Hence or otherwise calculate the expected profits.

Solution

Let X = Number of copies sold per week

X	$P(X=x)$	$xP(X=x)$
11	0.4	4.40
12	0.2	2.40
13	0.2	2.60
14	0.15	2.10
15	0.05	0.75
		12.250

(a) $E(X) = \sum xP(X=x) = 12.250$

(b) (i) profit = selling price - Buying price = $(13 \times 2000) - (15 \times 1500) = shs : 3500$

(ii) Let the profits be Y

X	$P(X=x)$	Y	$P(X=y)$	$yP(X=y)$
11	0.4	-500	0.4	-200
12	0.2	1500	0.2	300
13	0.2	3500	0.2	700
14	0.15	5500	0.15	825
15	0.05	7500	0.05	375
			$\sum P(X=y) = 1$	$\sum yP(X=y) = 2000$

$\Longrightarrow$ The expected weekly profits $= shs : 2000$

(6) A news agent stocks 13 copies of a magazine each week. He has regular order for 9, and the additional copies sold varies from week to week. The news agent uses previous sales data to estimate the probability for each possible total number of copies sold as follows:

Number of copies	9	10	11	12	13
probability	0.3	0.35	0.2	0.1	0.05

(a) Calculate the expected number of copies that he sells in a week.
(b) The news a gent buys the magazines at $shs.1000$ each and sells them at $shs.1200$ each. Any copies left unsold are returned to the company and therefore he does not pay for them.
(i) Find the profit on these magazines in a week when he sells 9 copies.
(ii) Construct a probability distribution table for the news agent's weekly profit from the sale of these magazines. Hence or otherwise calculate the expected profits.

Solution

Let X = Number of copies sold per week

X	$P(X=x)$	$xP(X=x)$
9	0.3	2.70
10	0.35	3.50
11	0.2	2.20
12	0.15	1.80
13	0.05	0.65
		10.850

(a) $E(X) = \sum xP(X=x) = 10.850$

(b) (i) profit

$$= \text{selling price - Buying price} = (13 \times 1200) - (13 \times 1000) = shs : 2600$$

(ii) Let the profits be Y

X	$P(X=x)$	Y	$P(X=y)$	$yP(X=y)$
9	0.3	1800	0.3	540
10	0.35	2000	0.35	700
11	0.2	2200	0.2	440
12	0.15	2400	0.15	360
13	0.05	2600	0.05	130
			$\sum P(X=y)=1$	$\sum yP(X=y)=2170$

$\Longrightarrow$ The expected weekly profits = $shs : 2170$

5.1.8 Exercise 5

(1) The number of times a machine breaks down every month is a discrete random variable X with the probability distribution.

$$P(X=x) = \begin{cases} k0.25^x & ; \quad x = 0,1,2,3,\cdots \\ 0 & ; \quad \text{Otherwise} \end{cases} \quad \text{where } k \text{ is a constant}$$

Determine the probability that the machine will break down not more than two times a month.
Ans; k = 3/4, 63/64

(2) A random variable has P.D.F given by;

$$P(X=x) = k|x-1|, \ x = -2, -1$$

$$P(X=X) = k\left(\frac{1}{4}\right)^x, \ x = 0,1,2,\cdots$$

Find

(i) the value of k

(ii) $P(-4 < 2X < 4/X > -1)$

(iii) Sketch $f(x)$ and $F(X)$.

Ans; (i) 3/19 (ii) 15/16

(3) Two discs are drawn at random without replacement from a bag containing 3 blue and 4 yellow discs. If 'X' is the random variable the number of blue discs drawn;

(i) Construct a probability distribution for X

(ii) Determine $E(x)$

Ans (a)

x	0	1	2
$P(X = x)$	$\frac{2}{7}$	$\frac{4}{7}$	$\frac{1}{7}$

Ans (b) $\frac{6}{7}$

(4) A random variable x has its distribution function given by

x	1	2	5	7	10
$F(x)$	$2c + k$	$2(c + k)$	$5c + k$	$k + 5c + 0.1$	$6c + 0.1$

(i) Given that $P(X = 1) = P(X = 5)$, find the values of c and k.
(ii) Derive the probability mass function of x and find
(iii) Mean and mode of x
(iv) The standard deviation of x
Ans; (i)

(5) The probability of a discrete random variable X is given by;
$P(X = r) = kr; r = 1, 2, 3, \cdots, n$; where k is a constant
(a) Show that $k = \frac{2}{n(n+1)}$
(b) Find in terms of n, the expected value of x.
(c) Show that $Var(X) = \frac{(n-1)(n+2)}{18}$
Ans; (b) $E(X) = \frac{2n+1}{3}$

(6) A supermarket sells three different types of plates, X, Y and Z. On some day, the stock in such that the ratio of $X : Y : Z$ is equal to $6 : 5 : 1$. The costs of the three types are in the ratio $X : Y : Z = 2000 : 4000 : 3500$ respectively. 25 plates were sold randomly on this day and the total of the sales was S shillings. Find

(i) the expected value of S.

(ii) the standard deviation of S.

Ans; (i) 73958.35 (ii) 24183.69695

(7) A random variable X has got the following probability distribution.

$$f(x) = \begin{cases} kx & ; \quad x = 1,2,3 \\ k(8-x) & ; \quad x = 4,5,6,7 \\ 0 & ; \quad \text{elsewhere} \end{cases}$$

(i) Determine the value of the constant k
(ii) $E(X-1)^2$
(iii) Find $P(X < 5)$ and $P(2 \leq x < 6)$
(iv) Obtain the cumulative distribution function of x and hence find the median and upper quartile of x.
Ans; (i) 1/16 (ii) 11.5 (iii) 5/8, 3/4

(8) A box contains 4 pink counters, 3 green counters and 3 yellow counters. Three counters are drawn at random one after the other without replacement.
(a) Find the probability that the third counter drawn is green and the first two are of the same colour.
(b) Find the expected number of pink counters drawn.
Ans; (a) 1/12 (b) 1.2

(9) A random variable$X$ can assume values $0, 1, 2$ and 3 only. Given that $P(X \leq 2) = 0.9, P(X \leq 1) = 0.5$ and $E(X) = 1.4$. Find

(i) $P(X = 1)$ (ii) $P(X = 0)$ (iii) the mode

Ans; (i) 0.3 **(ii)** 0.2 **(iii) 2**

(10) A discrete random variable R has a p.d.f given by;

$$f(r) = \begin{cases} \frac{a^2 r}{4} & ; \quad r = 1,2,3,\cdots,n \\ 0 & ; \quad \text{elsewhere} \end{cases}$$

If $a > 0$ and $E(R) = 17/3$, find the value of a.
Ans; a = $\pm\frac{1}{3}$

(11) A discrete random variable R has a cumulative mass function given as;

$$F(r) = \begin{cases} kr(r+1) & ; \quad r = 1,2,3,4 \\ 0 & ; \quad \text{elsewhere} \end{cases}$$

Determine the

(i) value of k (ii) $Var(2R+3)$

Ans ; (i) k = 1/20 (ii) 4

(12) A random variable X assumes values -3, 0, 2 and 3 only as shown in the table.

x	-3	0	2	3
$P(X=x)$	$\frac{1}{4}$	$\frac{1}{4}$	$\frac{1}{4}$	$\frac{1}{4}$

Given that $Y = X^2$ and $Z = X(X-1)$.
(i) Draw a table to show the probability distribution of Y and Z.

(ii) Determine $E(Y)$ and $E(Z)$ (iii) Show that $E(Z) = E(X^2)$ - $E(X)$

Ans; (i)

y	9	0	4	9
z	12	0	2	6
$P(Y)$	$\frac{1}{4}$	$\frac{1}{4}$	$\frac{1}{4}$	$\frac{1}{4}$
$P(Z)$	$\frac{1}{4}$	$\frac{1}{4}$	$\frac{1}{4}$	$\frac{1}{4}$

(ii) $E(Y) = 5.5,\ E(Z) = 5$

(13) The probability mass function of a variable x is given by

$$P(X=x) = \begin{cases} cx + a & ; \quad x = -2, -1, 0, 1, 2 \\ 0 & ; \quad \text{elsewhere} \end{cases}$$ Where c and a are constants.

(i) If $P(X=2) = 2P(X=-2)$ determine a and c
(ii) Compute the mean and variance of x. (iii) What is $P(X \neq 1)$
(iv) Obtain the cumulative distribution function and sketch it
Ans; (i) a = 1/5 , c = 1/30 (ii) E(X) = 1/3, Var(X) = 61/45 (iii) 23/30

(14) The discrete random variable X has probability function

$$P(X=x) = \begin{cases} \dfrac{kx}{(x+1)} & : \quad x = 2, 3, 4 \\ \dfrac{2kx}{(x^2-1)} & : \quad x = 5, 6 \\ 0 & : \quad \text{Otherwise} \end{cases}$$ where k is a constant

(a) Find the pdf $f(x) = P(X=x)$
(b) Sketch $f(x)$ and $F(X)$.
(C) Find the standard deviation.

(15) The discrete random variable X has probability function

$$P(X=x) = \begin{cases} \dfrac{cx}{(x+1)} & : \quad x = 2, 3, 4 \\ c & : \quad x = 5 \\ 2c & : \quad x = 6 \\ 0 & : \quad \text{Otherwise} \end{cases}$$ where c is a constant

(a) Find the pdf of $f(x)$
(b) Sketch $f(x)$ and $F(X)$.
(C) Find the standard deviation.

Chapter 6

BINOMIAL PROBABILITY DISTRIBUTION

6.1 BINOMIAL PROBABILITY DISTRIBUTION

6.1.1 Introduction

A binomial distribution is a type of discrete probability distribution a rising from the binomial experiment.
A binomial experiment is characterized with repeated number trials ($n \leq 20$), called Bernoulli trials with each number of trial two possible outcomes called failure and success.
The number of successes (r)where $r = 0, 1, 2, 3, \cdots, n$ is our interest. While
The probability of success in any trial as p and that of failure as q such that $p + q = 1$.

6.1.2 Properties of Binomial Distribution

- It involves repeated number of trials($n \leq 20$).
- Each trial results into two possible outcomes i.e success (p) and failure (q).
- The trials are independent to each other.
- Probability of success (p) is a constant.

A random variable X is said to follow a binomial distribution abbreviated as $X \sim B(n, p)$ or $X \sim Bin(n, p)$, with the corresponding probability distribution as $P(X = x)$ where

- $p =$ The probability of success.
- $n =$ Number of repeated trials.

If $X \sim B(n, p)$, then
$\mathrm{P(X = r)} = \left(\begin{array}{c} n \\ r \end{array} \right) p^r q^{n-r}$. where $q = 1 - p$ and $\left(\begin{array}{c} n \\ r \end{array} \right) =^n \mathrm{C}_r = \dfrac{n!}{(n-r)!r!}$ with

- $p =$ The probability of success.

- $n =$ Number of repeated trials.
- $r =$ Number of successes out of n trials
- $q =$ The probability of failure such that $p + q = 1$

Examples

(1) In a basket containing 8 tomatoes, the probability of getting a ripe tomato is 0.2 . Find the probanility that this sample will contain;

(i) exactly four ripe tomatoes.

(ii) no ripe tomato

(iii) 2 or 3 ripe tomatoes

(iv) at least 7 ripe tomatoes

(v) at most 2 ripe tomatoes,

(vi) between 3 and 7 ripe tomatoes,

(vii) between 2 and 5 ripe tomatoes inclusive,

Solution

$p = 0.2, q = 0.8, n = 8$ and take $P(X = r) = \begin{pmatrix} n \\ r \end{pmatrix} p^r q^{n-r}$

(i) P(X = 4), r = 4

Using $P(X = r) = \begin{pmatrix} n \\ r \end{pmatrix} p^r q^{n-r} \Longleftrightarrow P(X = 4) = \begin{pmatrix} 8 \\ 4 \end{pmatrix} \times 0.2^4 \times 0.8^4 = 0.0459$

(ii)P(X = 0), r = 0

Using $P(X = r) = \begin{pmatrix} n \\ r \end{pmatrix} p^r q^{n-r} \Longleftrightarrow P(X = 0) = \begin{pmatrix} 8 \\ 0 \end{pmatrix} \times 0.2^0 \times 0.8^8 = 0.1678$

(iii)
$$\begin{aligned} P(X = \ 2 \ or \ 3) &= P(X = 2) + P(X = 3) \\ &= \begin{pmatrix} 8 \\ 2 \end{pmatrix} \times 0.2^2 \times 0.8^6 + \begin{pmatrix} 8 \\ 3 \end{pmatrix} \times 0.2^3 \times 0.8^5 \\ &= 0.2936 + 0.1468 = 0.4404 \end{aligned}$$

(iv)
$$\begin{aligned} P(X \geq 7) &= P(X = 7) + P(X = 8) \\ &= \begin{pmatrix} 8 \\ 7 \end{pmatrix} \times 0.2^7 \times 0.8^1 + \begin{pmatrix} 8 \\ 8 \end{pmatrix} \times 0.2^8 \times 0.8^0 \\ &= 0.000082 + 0.000003 = 0.000085 \end{aligned}$$

(v)
$$\begin{aligned} P(X \leq 2) &= P(X = 0) + P(X = 1) + P(X = 2) \\ &= \begin{pmatrix} 8 \\ 0 \end{pmatrix} \times 0.2^0 \times 0.8^8 + \begin{pmatrix} 8 \\ 1 \end{pmatrix} \times 0.2^1 \times 0.8^7 + \begin{pmatrix} 8 \\ 2 \end{pmatrix} \times 0.2^2 \times 0.8^6 \\ &= 0.000082 + 0.000003 + 0.2936 = 0.2937. \end{aligned}$$

(vi)

$$
\begin{aligned}
P(3 \leq x \leq 7) &= P(X=4)+P(X=5)+P(X=6) \\
&= \begin{pmatrix} 8 \\ 4 \end{pmatrix} \times 0.2^4 \times 0.8^4 + \begin{pmatrix} 8 \\ 5 \end{pmatrix} \times 0.2^5 \times 0.8^3 + \begin{pmatrix} 8 \\ 6 \end{pmatrix} \times 0.2^6 \times 0.8^2 \\
&= 0.0459 + 0.0092 + 0.0011 = 0.0562.
\end{aligned}
$$

(vii)

$$
\begin{aligned}
P(2 \leq x \leq 5) &= P(X=2)+P(X=3)+P(X=4)+P(X=5) \\
&= \begin{pmatrix} 8 \\ 4 \end{pmatrix} \times 0.2^4 \times 0.8^4 + \begin{pmatrix} 8 \\ 5 \end{pmatrix} \times 0.2^5 \times 0.8^3 + \begin{pmatrix} 8 \\ 6 \end{pmatrix} 0.2^6 \times 0.8^2 \\
&+ \begin{pmatrix} 8 \\ 6 \end{pmatrix} \times 0.2^6 \times 0.8^2 \\
&= 0.2936 + 0.1468 + 0.0459 + 0.0092 = 0.4955.
\end{aligned}
$$

(2) Given that $X \sim B(10, 0.64)$. Find:

(i) $P(X < 3)$

(ii) $P(X \geq 7)$

Solution:
From $X \sim B(10, 0.64), \Longleftrightarrow n = 10, p = 0.64$ and $q = 0.36$
(i)

$$
\begin{aligned}
P(X<3) &= P(X=0)+P(X=1)+P(X=2) \\
&= \begin{pmatrix} 10 \\ 0 \end{pmatrix} \times 0.64^0 \times 0.36^{10} + \begin{pmatrix} 10 \\ 1 \end{pmatrix} \times 0.64^1 \times 0.36^9 + \begin{pmatrix} 10 \\ 2 \end{pmatrix} \times 0.64^2 \times 0.36^8 \\
&= 0.0000 + 0.0006 + 0.0052 = 0.0058.
\end{aligned}
$$

(ii)

$$
\begin{aligned}
P(X \geq 7) &= P(X=7)+P(X=8)+P(X=9)+P(X=10) \\
&= \begin{pmatrix} 10 \\ 7 \end{pmatrix} \times 0.64^7 \times 0.36^3 + \begin{pmatrix} 10 \\ 8 \end{pmatrix} \times 0.64^8 \times 0.36^2 + \begin{pmatrix} 10 \\ 9 \end{pmatrix} \times 0.64^9 \times 0.36^1 \\
&+ \begin{pmatrix} 10 \\ 10 \end{pmatrix} \times 0.64^{10} \times 0.36^0 \\
&= 0.2462 + 0.1642 + 0.0649 + 0.0115 = 0.4868.
\end{aligned}
$$

6.1.3 Expectation and Variance of Binomial Distribution.

The expectation, $E(X)$ is the same as mean of a binomial distribution such that , $E(X) = np$ and
The variance of a binomial distribution, $Var(X) = npq$,
i.e
If $X \sim B(n, P)$, then $E(X) = np$ and $\sigma = np(1-p) = npq$.

NB: Standard deviation, $\sigma = \sqrt{\mathrm{npq}}$

6.1.4 The Most Likely Value of X to Occur.

This is also known as the most probable value of X to occur or mode, It refers to the value of X with the highest probability. However, this value is so close to the mean of X, and therefore, we take the value of X that is close to the mean but with the highest probability.

Examples:

(1) In a certain village, $\frac{1}{5}$ of people are tall. If a sample of 16 people is taken,
(a) what is expectation and variance of tall people in this village?
(b) Find the most likely number of people.
Solution

$n = 16, p= \frac{1}{5}, q = \frac{4}{5}$

(a) $\Longrightarrow$Expectation $E(X) = np = \left(\frac{1}{5} \times 16\right) = 3.20$

Similarly, Variance $= npq = \left(\frac{1}{5} \times 16 \times \frac{4}{5}\right) = 2.560$

(b) Since $E(X) = 3.20$, Test for $P(X = 3)$ and $P(X = 4)$

When $X = 3, P(X = 3) = \begin{pmatrix} 16 \\ 3 \end{pmatrix} \times (\frac{1}{5})^3 \times (\frac{4}{5})^{13} = 0.2463$. similarly,

When $X = 4, P(X = 4) = \begin{pmatrix} 16 \\ 4 \end{pmatrix} \times (\frac{1}{5})^4 \times (\frac{4}{5})^{12} = 0.2001$.

Since $P(X = 3) > P(X = 4)$, then mode $= 3$.

(2) The secretary of an NGO works 5 days a week from Monday to Friday and her probability of being late on any given day is 0.2.

(a) Calculate the mean and variance of number of days she will be late in a month consisting of 4 weeks.

(b) What is the probability of her being punctual for a whole week.

(c) Calculate expected number of completely punctual weeks in year.

Solution:

(a) $n = 20, p = 0.2, q = 0.8$

$\Longrightarrow E(X) = np = 20 \times 0.2 = 4$ days

Similarly $Var(X) = npq = 20 \times 0.2 \times 0.8 = 3.2$ (days)^2

(b) $n = 5, p = 0.2, q = 0.8$

P(X = 0), r = 0

P(no day she is late) $= P(X = 0) = \begin{pmatrix} 5 \\ 0 \end{pmatrix} \times 0.2^0 \times 0.8^5 = 0.3277$

(a) $n = 48$ weeks, $p = 0.3277, q = 0.6723$

$\therefore E(X) = np = 48 \times 0.3277 = 15.7286\text{weeks} = 78.6430$ days.

(3) A random variable X is such that $X \sim B(n, 0.55)$. Given that the mean of this distribution is 7.150, find;

(i) value of n,

(ii) standard deviation,

(iii) Mode.

Solution.
(i) $E(X) = 7.150, p = 0.55, q = 0.45$

Using $E(X) = np \iff 7.150 = n \times 0.550 \iff n = \dfrac{7.150}{0.550} \implies n = 13.$

(ii) Standard deviation $= \sqrt{Var(X)} = \sqrt{npq} = \sqrt{(16 \times 0.2 \times 0.8)} = 1.60$
(iii) Mode = ?
Since $E(X) = 7.150$, Test for $P(X = 7)$ and $P(X = 8)$
When $X = 7, P(X = 7) = \begin{pmatrix} 13 \\ 7 \end{pmatrix} \times (0.55)^7 \times (0.45)^6 = 0.2169$. similarly,
When $X = 8, P(X = 8) = \begin{pmatrix} 13 \\ 8 \end{pmatrix} \times (0.55)^8 \times (0.45)^5 = 0.1989.$
Since $P(X = 7) > P(X = 8), \implies$ mode $= 7$.

(4) Of the students of S.5 Sciences, 70% are known to have got distinctions in mathematics in their UCE. If a sample of 14 students is taken from this class,

(a) Find the most likely number of them to have got distinctions in their UCE,
(b) Var(X),
(c) How many students must be picked so that the probability that there is at least one student who got a distinction is above 0.95?

Solution
(a) $n = 14, p = 0.7, q = 0.3$

$E(X) = np = 14 \times 0.7 = 9.8$

Test for$x = 9$ and $x = 10$

$P(X = 9) = \begin{pmatrix} 14 \\ 9 \end{pmatrix} 0.7^9 0.3^5 = 0.1963$

$P(X = 10) = \begin{pmatrix} 14 \\ 10 \end{pmatrix} 0.7^{10} 0.3^4 = 0.2290$

$\implies$ the most likely is 10 students
(b) $Var(X) = npq = (14 \times 0.7 \times 0.3) = 2.940$
(c) $X \sim B(n, p)$ and $P(X \geq 1) > 0.95$
$\implies 1 - P(X = 0) > 0.95 \iff 1 - \begin{pmatrix} n \\ 0 \end{pmatrix} (0.7)^0 (0.3)^{(n-0)} > 0.95$
$1 - 0.3^n > 0.95$
$0.05 > 0.3^n \iff log(0.05) = nlog(0.3)$
$n \approx 2$

(5) In a certain African village, 80% of the villagers are known to have a particular eye disorder. Twelve people are waiting to see the nurse.

(i) What is the most likely number to the disorder?

(ii) Find the probability that fewer than half have the eye disorder,

(iii) Standard deviation.

(i) Most likely value =?
Let $p = 0.8, q = 0.2$ and $n = 12$. Let $E(X) = 0.8 \times 12 = 9.6$ so test for $P(X = 9)$ and $P(X = 10)$
When $X = 9, P(X = 9) = \begin{pmatrix} 12 \\ 9 \end{pmatrix} \times (0.8)^9 \times (0.2)^3 = 0.2362$. similarly,
When $X = 10, P(X = 10) = \begin{pmatrix} 12 \\ 10 \end{pmatrix} \times (0.8)^{10} \times (0.2)^2 = 0.2835$.
Since $P(X = 10) > P(X = 9)$, $\Longrightarrow$ Most likely value $= 10$.
(ii) and (iii) are left as exercise.

(7) The random variable X follows a binomial distribution with mean 2 and variance 1.6. find

(i) The probability that X is less than 6,

(ii) The most likely value of X. **Solution:**
Let $E(X) = np \Longleftrightarrow 2 = np \cdots\cdots\cdots\cdots\cdots (i)$
Similarly, $Var(X) = npq \Longleftrightarrow 1.6 = npq \cdots\cdots\cdots\cdots\cdots (ii)$
From (i) and (ii), then $1.6 = (np)q \Longleftrightarrow 1.6 = 2q \Longleftrightarrow q = \dfrac{1.6}{2} \Longrightarrow q = 0.8$
From $p + q = 1, \Longleftrightarrow p = (1 - 0.8) = 0.2$
From $(i), E(X) = np \Longleftrightarrow 2 = n(0.2) \Longleftrightarrow n = 10$
(i) $n = 10, p = 0.2$ and $q = 0.8$
$\Longrightarrow P(X < 4) =?$
Please complete the qn.

(6) **Task.**
In a bag, there are six red counters, eight yellow counters and six green counters. An experiment consists of taking a counter at random from the bag, noting its color and then replacing it in the bag. This procedure is carried out ten times in all. Find

(a) The expected number of red counters drawn,

(b) The most likely number of green counters drawn,

(c) The probability that no more than four yellow counters are drawn

.

6.1.5 Working with tables

To work with tables of binomial, the probability of success (p) and number of repeated trials (n) **must** be in the tables, otherwise we must stick to the use of binomial formula discussed earlier.
There are two types of binomial tables that we use i.e individual terms and cumulative binomial tables.

Individual terms tables

These tables are used when getting probabilities at exact i.e P(X = r) or for some few values of r from at least and at most cases.
Examples

(1) A box containing 15 green and red apples. Given that the chance of picking a green apples is 0.3 . Find the probability that;

(i) exactly 5 green apples will be picked.

(ii) exactly 9 red apples will be picked.

(iii) less than 3 green apples

Solution
$n = 15, p = 0.3, q = 0.7$

(i) $P(X = 5), r = 5$ $P(X = 5) = 0.2061$

(ii) $P(8red) = P(7green), r = 7$ $P(X = 7) = 0.0811$

(iii) $P(X < 3) = P(X = 2) + P(X = 1) + P(X = 0)$

$$0.0916 + 0.0305 + 0.0047 = 0.1268$$

6.1.5.1 Cumulative / sums tables

These tables are used basically when we need to find probabilities for several values of r and these probabilities need to be summed up. This is common with at-least (x $\geq$ r) and at most (x $\leq$ r) cases. Though the cumulative tables use only at least scenario (x $\geq$ r), one can express any other scenario in to at least form equivalence. i.e if

- $P(\text{At least r}) = P(X \geq r)$
- $P(\text{More than r}) = P(X > r) = P(X \geq r + 1)$
- $P(\text{At most r}) = P(X \leq r) = 1 - P(X \geq r + 1)$
- $P(\text{Less than r}) = P(X < r) = 1 - P(X \geq r + 1)$
- $P(Morethanr_1andlessthanr_2) = P(betweenr_1andr2)$
  $P(r_1 < x < r_2) = P(r_1 + 1 \leq x \leq r_2 - 1)$
  But for $P(r_1 \leq x \leq r_2) = P(X \geq r_1) - P(X \geq r_1 + 1)$

Example:

(1) In a certain town 20% of the people are known to have hypetitis B virus, a sample of 12 people is taken from this town, find the probability that;

(i) at least 7 will have the virus.

(ii) less than 4 will have the virus.

(iii) more than 8 will have the virus.

(iv) more than 3 and at most 8 will have the virus.

Solution

$p = 0.2, q = 0.8, n = 12$
(i) $P(X \geq 7) = 0.0039$

(ii) $P(X < 4) = 1 - P(X \geq 5) = 1 - 0.0726 = 0.9274$

(iii) $P(X > 8) = P(X \geq 9) = 0.0001$

(iv) $P(3 < x \leq 8) = P(4 \leq x \leq 8) = P(X \geq 4) - P(X \geq 9) = 0.2054 - 0.0001 = 0.2053$

6.1.6 Exercise 6

(1) Given that Y ∼ B(n, 0.3), find the least value of n such that P(Y ≥ 1) = 0.8
Ans; n = 5

(2) The random variable X ∼ B(10, p). Given that p is less than a half and the variance of X is 1.875 . Find p, E(X) and the mode.
Ans; p = 0.25, E(X) = 2.5, mode = 2

(3) The probability that it will be a sunny day is 0.4, out of 15 days. What is the probability

(i) Exactly seven will be sunny days,

(ii) exactly eight are not sunny days.

(iii) Atleast ten are sunny days

(iv) from three to eight are sunny days

Ans; (i) 0.1771 (ii) 0.1771 (iii) 0.0338 (iv) 0.8779

(4) A bag contains 6 blue and 4 red balls. Three balls are drawn at random without replacement. Find the probability distribution and the mean for the number of red balls drawn.

Ans;

x	0	1	2	3
P(X = x)	$\frac{1}{6}$	$\frac{1}{2}$	$\frac{3}{10}$	$\frac{1}{30}$

E(x) = 1.2

(5) An experiment consists of tossing two tetrahedral dice at once. If the experiment is repeated 12 times and X is the event that "the sum on both faces of the dice is less than 5."
(a) Find the probability that X occurs;

(i) exactly 7 times. (ii) between 5 and 8 times

Ans; p = 6/16 (i) 0.0788 (ii) 0.2320

(6) In a certain clan, the probability of getting a baby girl is 3 in 5. If a random sample of 10 children is taken, find
(i) expected number and variance of girls in the family
(ii) probability of getting atleast two girls
Ans; (i) E(X) = 6, 2.4 (ii) 0.9983

(7) The basket ball team of Lubiri S.S. is known to have a probability of 4/5 in wining a match. If 15 matches are played, What is the

(i) mean number of success (ii) the variance

(iii) the probability of atleast 10 successes in the 15 games
Ans; (i) 12 (ii) 2.4 (iii) 0.9389

(8) Of the students in a certain school in Wakiso, 30% are day-scholars. On a given day the headmaster picked 10 students at random;
(i) Find the probability that out of these atmost 8 were day-scholars.
(ii) Calculate the most probable number of boarders from the sample
Ans; (i) 0.9999 (ii) 7

(9) A coin is biased such that the probability that a head occurs is three times that of a tail occuring. What is the least number of times that the coin should be tossed so that the probability a tail occurs atleast once exceeds 0.995
Ans; 18

(10) The probability that a target is hit is 0.3 . Find the least number of shots which should be fired if the probability that the target is hit atleast once is atleast 0.95
Ans; 8

(11) On a certain farm, 20% of the cows are infected by a tick disease. If a random sample of 18 cows is selected from the farm, find;
(i) the probability that not more than 10% of the cows are infected.
(ii) the modal number of cows infected
Ans; (i) 0.2714 (ii) 3

(12) The random variable R has mean 7.2 and standard deviation 1.2 . Given that R has a Binomial distribution. Find P(R = 6)
Ans; 0.1716

(13) Two boxes A and B contain marbles. Box A contains 2 red and 3 green marbles while Box B contains 3 red and 1 green marble. Two marbles are picked at random such that if the first marble comes from A, the second should come from B and vice versa. Given that the pickings of these marbles are with replacement and repeated 5 times,

find the probability that the number of times that the two marbles were not red exceeded 2.
Ans; 0.0266

(14) Two boys play a game in which each throws a balanced tetrahedral dice. The game is a success if both girls get the same value. Find the probability that;
(i) The successes will occur in the first trial.
(ii) In 10 trials, the success will occur atmost 2 times.
Ans; (i) 0.25 (ii) 0.52559

(15) Two fair die are tossed 8 times, find the probability that a sum of five is obtained atleast 3 times.
Ans; 0.0501

Chapter 7

THE CONTINUOUS PROBABILITY FUNCTION

7.1 THE CONTINUOUS PROBABILITY FUNCTION (C. R. V)

7.1.1 Introduction

A function $f(x)$ is said to follow a continuous probability function if it takes on a continuous domain. These include; age, time, height, and others.
Its function is f(x), also known as its probability density function (p.d.f)

7.1.2 Properties of a Continuous Random Variables

- It's domain x extends from $-\infty \leq x \leq +\infty$
- $f(x) \geq 0$. I.e all probabilities are in the range of zero to one.
- The total probability of the curve $f(x)$ is one (1), i.e $\int_{-\infty}^{+\infty} f(x)dx = 1$

NB: $+\infty$ we refer to the upper limit(b) while $-\infty$ is the lower limit(a)
NB: We mainly apply the **third** property to find the unknown.

7.1.3 How to obtain Probabilities for Continuous Random Variables

This is done by finding the area under the curve $f(x)$. I.e By integrating.
Given

$$y = \begin{cases} f(x) & ; \quad a \leq x \leq b \\ 0 & ; \quad \text{elsewhere} \end{cases}$$

Then:

(i) $P(X > a_1) = P(X \geq a_1) = \int_{a_1}^{+\infty \ or \ b} f(x)dx$

(ii) $P(X < a_1) = P(X \leq a_1) = \int_{-\infty \ or \ a}^{a_1} f(x)dx$

(iii) $P(a_1 < X < b_1) = P(a_1 \leq X \leq b_1) = P(a_1 < X \leq b_1) = P(a_1 \leq X < b_1) = \int_{a_1}^{b_1} f(x)dx$

(iv) $P(|x - a_1| < b_1) = P(a_1 - b_1 < X < a_1 + b_1) = \int_{a_1-b_1}^{a_1+b_1} f(x)dx$

(v) $P(|x - a_1| > b_1) = 1 - P(|x - a_1| < b_1) = 1 - \int_{a_1-b_1}^{a_1+b_1} f(x)dx$
Where a and b are constants with $a_1, b_1 \in [a, b]$. I.e $a \leq a_1 < b_1 \leq b$
NB:

(1) It is sometimes possible to find an area by geometry, for example by using formulae for area of a triangle or trapezium.

(2) The only difference between a continuous and a discrete random variable is only seen in the domain i.e

consider

(i)

$$f(x) = \begin{cases} kx & ; \quad 0 \leq x \leq 3 \\ 0 & ; \quad \text{elsewhere} \end{cases} \quad \textbf{is a continuous}$$

(ii)

$$f(x) = \begin{cases} kx & ; \quad x = 0, 1, 2, 3. \\ 0 & ; \quad \text{elsewhere} \end{cases} \quad \textbf{is a discrete}$$

7.1.4 Mean and Variance of a Continuous Random Variables

7.1.4.1 Expectation, $E(x)$:

This is also known as Mean such that $E(x) = \int_{-\infty}^{+\infty} xf(x)dx$
Properties of Mean For constants a and b, -$E(a) = a$
$- E(aX) = aE(X)$
$- E(aX + b) = aE(X) + b$
$- E(X^{\text{n}}) = \int_{-\infty}^{+\infty} x^n f(x)dx$

7.1.4.2 Variance, σ^2:

This is abbreviated as $var(x)$ such that $Var(X) = E(X^2) - [E(X)]^2$
where $\text{E}(\text{X}^2) = \int_{-\infty}^{+\infty} x^2 f(x)dx$ and $E(x) = \int_{-\infty}^{+\infty} xf(x)dx$

Properties of variance.
For constants a and b,

- $Var(a) = 0$
- $Var(aX) = a^2 Var(X)$

- $Var(aX + b) = a^2 Var(X)$

NB: Standard deviation, $\sigma = \sqrt{\text{variance}}$
Examples:

(1) A continuos random variable X has its probability density function given by;

$$f(x) = \begin{cases} kx & ; \quad 0 \leq x \leq 3 \\ 0 & ; \quad \text{otherwise} \end{cases}$$

Find

(i) value of the constant k

(ii) E(X)

(iii) variance

(iv) P(X $\geq$ 2)

(v) P(X $\leq$ 2.5)

(vi) P(1 $\leq$ x $\leq$ 2)

(vii) P(1 $<$ x $\leq$ 1.8)

(viii) $E(3X - 1)$

(ix) $Var(2X + 10)$

Solution

(i) $\int_{-\infty}^{+\infty} f(x)dx = 1 \iff \int_0^3 kxdx = 1 \iff \left[\frac{k}{2}x^2\right]_0^3 = 1 \iff \frac{k}{2}(3)^2 - \frac{k}{2}(0)^2 = 1$
$\implies k = \frac{2}{9}$

$$\therefore f(x) = \begin{cases} \frac{2}{9}x & ; \quad 0 \leq x \leq 3 \\ 0 & ; \quad \text{otherwise} \end{cases}$$

(ii) $E(X) = \int_{-\infty}^{+\infty} xf(x)dx = \int_0^3 \frac{2}{9}x^2 dx = \left[\frac{2}{27}x^3\right]_0^3 = \frac{2}{27}(3)^3 - \frac{2}{27}(0)^3 = 2.00$

(iii) $Var(X) = E(X^2) - [E(X)]^2$
But $E(X^2) = \int_{-\infty}^{+\infty} x^2 f(x)dx = \int_0^3 \frac{2}{9}x^3 dx = \left[\frac{2}{36}x^4\right]_0^3 = \frac{2}{36}(3)^4 - \frac{2}{36}(0)^4 = 4.5$
$\implies Var(x) = 4.5 - 2^2 = 0.50$

(iv) $P(X > 2) = \int_2^3 \frac{2}{9}xdx = \left[\frac{1}{9}x^2\right]_2^3 = \frac{1}{9}(3)^2 - \frac{1}{9}(2)^2 = \frac{5}{9}$

(v) $P(X < 2.5) = \int_0^{2.5} \frac{2}{9}xdx = \left[\frac{1}{9}x^2\right]_0^{2.5} = \frac{1}{9}(2.5)^2 - \frac{1}{9}(0)^2 = 0.6944$

(vi) $P(1 \leq x \leq 2) = \int_1^2 \frac{2}{9}xdx = \left[\frac{1}{9}x^2\right]_1^2 = \frac{1}{9}(2)^2 - \frac{1}{9}(1)^2 = 0.3333$

(vii) $P(1 \leq x \leq 1.8) = \int_1^{1.8} \frac{2}{9}xdx = \left[\frac{1}{9}x^2\right]_1^{1.8} = \frac{1}{9}(1.8)^2 - \frac{1}{9}(1)^2 = 0.2489$

(viii) $E(3X-1) = 3E(X) - 1 = (3 \times 2) - 1 = 5.0$

(ix) $Var(2X+10) = 2^2 Var(X) = (4 \times 0.5) = 2.0$

(2) A continuous random variable X has its probability density function given by;

$$f(x) = \begin{cases} \frac{1}{18}x & ; \quad 0 \leq x \leq 6 \\ 0 & ; \quad \text{otherwise} \end{cases}$$

(a) Show that $f(x)$ is a p.d.f
(b) Find:

(i) P(X < 1) (iv) $P(X \geq 2/X \leq 5)$ (vi) $E(X+2)$

(ii) $P(1 < x \leq 5)$ (iv) $P(|x-3| < 2)$

(iii) $P(3 < x < 4)$ (v) $P(|x-2| > 1.5)$ (vii) $Var(3X-20)$

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

So, $\int_0^6 \frac{1}{18}xdx = \left[\frac{1}{18} \times \frac{x^2}{2}\right]_0^6 = \frac{1}{26}.6^2 - 0 = 1$, Since $\int_0^6 \frac{1}{18}xdx = 1$, then it's a pdf.

b(i) $P(X < 1) = \int_0^1 f(x)dx = \int_0^1 \frac{x}{18}dx = \left[\frac{x^2}{36}\right]_0^1 = \left(\frac{1}{36} - 0\right) = \frac{1}{36}$.

(ii) $P(1 < x \leq 5) = \int_1^5 f(x)dx = \int_1^5 \frac{x}{18}dx = \left[\frac{x^2}{36}\right]_1^5 = \left(\frac{5^2}{36}\right) - \left(\frac{1^2}{36}\right) = \frac{2}{3}$.

(iii) $P(3 < x < 4) = \int_3^4 f(x)dx = \int_3^4 \frac{x}{18}dx = \left[\frac{x^2}{36}\right]_3^4 = \left(\frac{4^2}{36}\right) - \left(\frac{3^2}{36}\right) = \frac{7}{36}$.

(iv) $P(X \geq 1/X \leq 5) = \dfrac{P(X \leq 1 \cap X \leq 5)}{P(X \leq 5)} = \dfrac{P(1 \geq x \leq 5)}{P(X \leq 5)}$

But $P(1 \geq x \leq 5) = \frac{2}{3}$ and $P(X < 5) = \int_0^5 \frac{x}{18}dx = \left[\frac{x^2}{36}\right]_0^5 = \left(\frac{25}{36} - 0\right) = \frac{25}{36}$.

$\Longrightarrow P(X \geq 1/X \leq 5) = \dfrac{\frac{2}{3}}{\frac{25}{36}} = \dfrac{24}{25} = 0.960$.

(iv) $P(|x-3| < 2) \Longleftrightarrow P(-2+3 < x < 2+3) \Longleftrightarrow P(1 < x < 5)$
$\therefore P(|x-3| < 2) = P(1 < x < 5) = \frac{2}{3}$ I.e as in part (ii) above.

(v) $P(|x-2| > 1.5) = 1 - P(|x-2| \leq 1.5) \Longleftrightarrow 1 - P(-1.5+2 < x < 1.5+2) \Longleftrightarrow$
$1 - P(0.5 < x < 3.5)$
$\therefore P(|x-2| > 1.5) = 1 - P(0.5 < x < 3.5)$
But $P(0.5 < x < 3.5) = \int_{0.5}^{3.5} \frac{x}{18}dx = \left[\frac{x^2}{36}\right]_{0.5}^{3.5} = \left(\frac{(3.5)^2}{36}\right) - \left(\frac{(0.5)^2}{36}\right) = \frac{1}{3}$
$\Longrightarrow P(|x-2| > 1.5) = 1 - P(0.5 < x < 3.5) = 1 - \frac{1}{3} = \frac{2}{3}$.

(vi) $E(X+2) = E(X) + 2$
But $E(X) = \int_a^b xf(x)dx = \int_0^6 \frac{1}{18}x^2dx = \left[\frac{1}{54}x^3\right]_0^6 = \frac{1}{54}(6)^3 - \frac{2}{54}(0)^3 = 4.00$
$\Longrightarrow E(X+2) = 4.00 + 2 = 6.00$

(vii) $Var(3X-20) = 3^2Var(X) = 9Var(X)$
But $Var(X) = E(X^2) - [E(X)]^2$
Where $E(X^2) = \int_a^b x^2f(x)dx = \int_0^6 \frac{1}{18}x^3dx = \left[\frac{1}{72}x^4\right]_0^6 = \frac{1}{72}(6)^4 - \frac{1}{72}(0)^4 = 18.0$
$\therefore Var(x) = 18.0 - 4.00^2 = 2.00$
$\Longrightarrow Var(3X-20) = 9 \times 2.00 = 18.00$

(3) A continuous random variable X has its probability density function given by;

$$f(x) = \begin{cases} kx & ; \quad 0 \le x \le 2 \\ k(4-x) & ; \quad 2 \le x \le 4 \\ 0 & ; \quad \text{otherwise} \end{cases}$$

Find

(i) value of constant k
(ii) $E(X+2)$
(iii) Standard deviation
(iv) $P(X < 1)$
(v) $P(X < 3)$
(vi) $P(X \ge 1)$
(vii) $P(|x-2| < 1.5)$
(viii) $P(|x-2| > 1.5)$

Solution
(i) $\int_{-\infty}^{+\infty} f(x)dx = 1 \Longleftrightarrow \int_0^2 kxdx + \int_2^4 k(4-x)dx = 1$

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

$\left[\frac{k}{2}(2)^2 - \frac{k}{2}(0)^2\right] + \left[k\left(4(4) - \frac{1}{2}(4)^2\right) - k\left(4(2) - \frac{1}{2}(2)^2\right)\right] = 1 \Longleftrightarrow 4k = 1$

$\Longrightarrow k = \frac{1}{4}$

(ii) $E(X+3) = E(X) + 3$

$E(X) = \int_{-\infty}^{+\infty} xf(x)dx = \int_0^2 kx^2dx + \int_2^4 k(4x - x^2)dx = \frac{k}{3}\left[x^3\right]_0^2 + k\left[2x^2 - \frac{1}{3}x^3\right]_2^4$

$$= \frac{1}{12}\left[(2)^3 - (0)^3\right] + \frac{1}{4}\left[\left(2(4)^2 - \frac{1}{3}(4)^3\right) - \left(2(2)^2 - \frac{1}{3}(2)^3\right)\right] = 2$$

$E(X+3) = 2 + 3 = 5$

(iii) standard deviation, $\sigma = \sqrt{\text{variance}}$

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

where $\mathrm{E(X^2)} = \int_{-\infty}^{+\infty} x^2 f(x)dx = \int_0^2 \frac{1}{4}x^3 dx + \int_2^4 \frac{1}{4}(4x^2 - x^3)dx = \frac{1}{16}\Big[x^4\Big]_0^2 + \frac{1}{4}\Big[\frac{4}{3}x^3 - \frac{1}{4}x^4\Big]_2^4$

$$= \frac{1}{16}\Big[(2)^4 - (0)^4\Big] + \frac{1}{4}\Big[\Big(\frac{4}{3}(4)^3 - \frac{1}{4}(4)^4\Big) - \Big(\frac{4}{3}(2)^3 - \frac{1}{4}(2)^4\Big)\Big] = \frac{17}{3}$$

$Var(X) = \frac{17}{3} - 2^2 = \frac{8}{3}$

$\delta = \sqrt{\text{variance}} = \sqrt{\frac{8}{3}} = 1.633$

(iv) $P(x < 1) = \int_0^1 \frac{1}{4}xdx = \frac{1}{8}\Big[x^2\Big]_0^1 = \frac{1}{8}\Big[(1)^2 - (0)^2\Big] = \frac{1}{8}$

(v)] $P(X < 3) = ?$

$P(X < 3) = \displaystyle\int_0^3 f(x)dx = \int_0^2 f(x)dx + \int_2^3 f(x)dx$. I.e because of more than one interval.

$$\begin{aligned}
&= \int_0^2 \frac{1}{4}xdx + \int_2^3 \frac{1}{4}(4 - x)dx \\
&= \frac{1}{8}\Big[x^2\Big]_0^2 + \Big[\frac{1}{4}\Big(4x - \frac{1}{2}x^2\Big)\Big]_2^3 \\
&= \frac{1}{8}\Big[(2)^2 - (0)^2\Big] + \frac{1}{4}\Big[\Big(4(3) - \frac{1}{2}(3)^2\Big) - \Big(4(2) - \frac{1}{2}(2)^2\Big)\Big] \\
&= \frac{7}{8} = 0.875.
\end{aligned}$$

$(vi)P(|x - 2| < 1.5) = P(0.5 < x < 3.5) = \displaystyle\int_{0.5}^{3.5} f(x)dx$ I.e because covers two interval.

$$\begin{aligned}
&= \int_{0.5}^2 \frac{1}{4}xdx + \int_2^{3.5} \frac{1}{4}(4 - x)dx \\
&= \frac{1}{8}\Big[x^2\Big]_{0.5}^2 + \Big[\frac{1}{4}\Big(4x - \frac{1}{2}x^2\Big)\Big]_2^{3.5} \\
&= \frac{1}{8}\Big[(2)^2 - (0.5)^2\Big] + \frac{1}{4}\Big[\Big(4(3.5) - \frac{1}{2}(3.5)^2\Big) - \Big(4(2) - \frac{1}{2}(2)^2\Big)\Big] \\
&= \frac{15}{16} = 0.9375.
\end{aligned}$$

(vi) $P(|x - 2| > 1.5) = 1 - P(|x - 2| < 1.5) = 1 - 0.9375 = 0.0625$

7.1.5 Mode of Continuous Random Variables

The mode of is the value of x corresponding to the maximum value of the probability function (p.d.f).I.e The mode is the value of x for which $f(x)$ is the greatest in the given range of x.
It is obtained basing on the following **cases** of the function $f(x)$ given:

1. Case 1: When $f(x)$ is a line : Under this case, you must sketch the function $f(x)$ over the given interval. There after, read the value of x corresponding to highest point(peak) of the graph and take it as the mode.
   For example:
   Given $f(x)$ such that
   (1)
   $$f(x) = \begin{cases} \dfrac{1}{18}x & ; \quad 0 \leq x \leq 6 \\ 0 & ; \quad \text{otherwise} \end{cases}$$
   Find the mode of $f(x)$.
   Solution.
   Since $f(x)$ is a line, then we sketch it.
   Let $y = \frac{1}{18}x$,

x	0	6
y	0	$\frac{1}{3}$

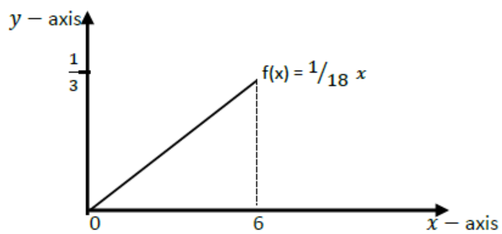


Figure 7.1: We sketch $f(x)$ because we can not obtain the turning point(s) by a analytical approach.

Since the peak of $f(x)$ is at $x = 6, \Longrightarrow$ mode $= 6$.
(2)
$$f(x) = \begin{cases} \dfrac{1}{8}(4-x) & ; \quad 0 \leq x \leq 4 \\ 0 & ; \quad \text{otherwise} \end{cases}$$
Find the mode of $f(x)$.
Solution.
Since $f(x)$ is a line, then we sketch it.
Let $y = \frac{1}{8}(4-x)$,

x	0	4
y	0.5	0

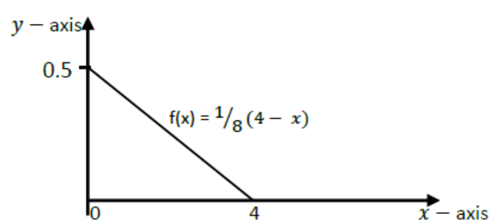


Figure 7.2: We sketch $f(x)$ because we can not obtain the turning point(s) by a analytical approach.

Since the peak of $f(x)$ is at $x = 0, \Longrightarrow$ mode $= 0$.

2. Case 2: When $f(x)$ is a function of degree two or more. Under this case, you differentiate $f(x)$ and you equate it to zero in obtain the x value which becomes our mode.I.e mode is the value of x for which $f'(x) = 0$
For example:
(1) Given $f(x)$ such that

$$f(x) = \begin{cases} \frac{3}{80}(x+2)(4-x) & ; \quad 0 \leq x \leq 4 \\ 0 & ; \quad \text{otherwise} \end{cases}$$

Find the mode of $f(x)$.
Solution
$f(x) = \frac{3}{80}(x+2)(4-x) = \frac{3}{80}(8+2x-x^2)$
mode is the value of x for which $f'(x) = 0 \Longrightarrow f'(x) = \frac{3}{80}(2-2x) \Longleftrightarrow \frac{3}{80}(2-2x) = 0$
$\Longrightarrow x = 1$
Testing if it's indeed maximimum
$\Longrightarrow f''(x) = \frac{-6}{80}$. Since $f''(x) \leq 0$, then $\Longrightarrow mode = 1$

(2) Given $f(x)$ such that

$$f(x) = \begin{cases} \frac{1}{108}x(6-x)^2 & ; \quad 0 \leq x \leq 4 \\ 0 & ; \quad \text{otherwise} \end{cases}$$

Find the mode of $f(x)$.
Solution
$f(x) = \frac{1}{108}x(6-x)^2 = \frac{1}{108}(36x - 24x^2 + x^3)$
mode is the value of x for which $f'(x) = 0 \Longrightarrow f'(x) = \frac{1}{108}(36 - 48x + 3x^2) \Longleftrightarrow$
$\frac{1}{108}(36 - 48x + 3x^2) = 0$
$\Longrightarrow$ Either, $x = 6$ or $x = 2$
Testing for the maximum value of x . Here you can apply the sign change approach or the second derivative approach. By the later,
$\Longrightarrow f''(x) = \frac{3}{108}(2x-8)$.
When $x = 6, f''(x) = \frac{3}{108}((2 \times 6) - 8) = \frac{12}{108} > 0$, then 6 is not the mode
When $x = 2, f''(x) = \frac{3}{108}((2 \times 2) - 8) = \frac{-12}{108} < 0$, then 2 is the mode
$\Longrightarrow mode = 2$

3. Case 3: When $f(x)$ is having more than one interval.
Here irrespective of the nature of the functions given, you should sketch the function $f(x)$ over the given as it is soon to be discussed.

Then read the value of x corresponding to highest peak of the graph and take it as the mode.
For example:
Given $f(x)$ such that

For example:
Given $f(x)$ such that

$$f(x) = \begin{cases} \frac{1}{4}x & ; \quad 0 \leq x \leq 2 \\ \frac{1}{4}(4-x) & ; \quad 2 \leq x \leq 4 \\ 0 & ; \quad \text{Elsewhere} \end{cases}$$

Find the mode of $f(x)$.
Solution.
Since $f(x)$ is a composition of lines, then we sketch it.

For $y = \frac{1}{4}x, 0 \leq x \leq 2$:

x	0	2
y	0	0.5

For $y = \frac{1}{4}(4-x), 2 \leq x \leq 4$:

x	2	4
y	0.5	0

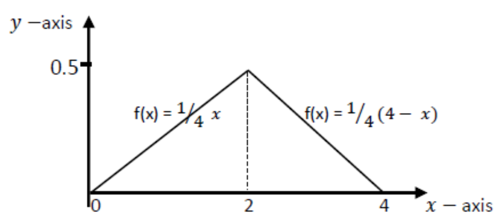


Figure 7.3

Since the peak of $f(x)$ is at $x = 2, \Longrightarrow$ mode $= 2$. Given $f(x)$ such that

$$f(x) = \begin{cases} \frac{1}{4}x & ; \quad 0 \leq x \leq 2 \\ \frac{1}{4}(4-x) & ; \quad 2 \leq x \leq 4 \\ 0 & ; \quad \text{Elsewhere} \end{cases}$$

Find the mode of $f(x)$.
Solution.
Since $f(x)$ is a composition of lines, then we sketch it.

For $y = \frac{1}{4}x, -1 \leq x \leq 0$:

x	-1	0
y	0	$\frac{2}{3}$

For $y = \frac{1}{4}(4-x), 0 \leq x \leq 2$:

x	0	2
y	$\frac{2}{3}$	0

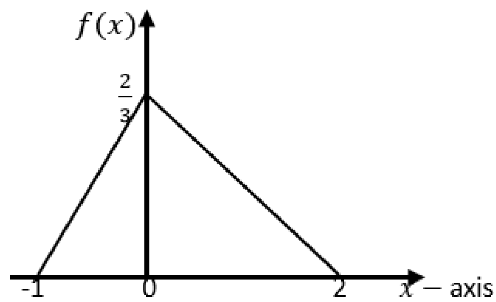


Figure 7.4

Since the peak of $f(x)$ is at $x = 0, \Longrightarrow$ mode $= 0$.

More Examples.

(1) A Continuous random variable, X has its p.d.f $f(x)$ given by;

$$f(x) = \begin{cases} \frac{2}{27}(5x - x^2) & ; \quad 0 \leq x \leq 3 \\ 0 & ; \quad \text{elsewhere} \end{cases}$$

(a) Find the mode.

(b) Calculate the mean and standard deviation

Solution

(a) For mode; $\frac{d}{dx}\big(f(x)\big) = 0, \Longrightarrow \frac{d}{dx}\Big(\frac{2}{27}(5x - x^2)\Big) = \frac{2}{27}(5 - 2x)$

$\frac{2}{27}(5 - 2x) = 0 \Longleftrightarrow x = 2.5$

So $\frac{d^2}{dx^2}\Big(\frac{2}{27}(5x - x^2)\Big) = \frac{2}{27}(-2) = \frac{-4}{27} < 0$

$\Longrightarrow$ Thus mode $= 2.5$

(b) **Left as exercise for you.**

(2) A random variable X has its p.d.f given by;

$$f(x) = \begin{cases} A(x^2 - x^3) & ; \quad 0 \leq x \leq 1 \\ 0 & ; \quad \text{elsewhere} \end{cases}$$

Find

(i) Value of constant A,

(ii) Mode of X

(iii) The mode of X

(iv) Variance

Solution

(i) $\int_{-\infty}^{+\infty} f(x)dx = 1 \iff \int_0^1 A(x^2 - x^3)dx = 1$

$\left[A\left(\frac{x^3}{3} - \frac{x^4}{4}\right)\right]_0^1 = 1$

$\left[A\left(\frac{(1)^3}{3} - \frac{(1)^4}{4}\right) - 0\right] = 1$

$\Longrightarrow A = 12$

$\therefore$

$$f(x) = \begin{cases} 12(x^2 - x^3) & ; \quad 0 \leq x \leq 1 \\ 0 & ; \quad \text{elsewhere} \end{cases}$$

(ii) Mode of X

For mode; $\frac{d}{dx}\big(f(x)\big) = 0, \Longrightarrow \frac{d}{dx}\big(12(x^2 - x^3)\big) = 12(2x - 3x^2)$

$12(2x - 3x^2) = 0 \iff$ Either $x = 0$ 0r $x = \dfrac{2}{3}$

Testing for the maximum value of x ,
$\Longrightarrow f''(x) = 12(2 - 6x)$.
When $x = 0, f''(x) = 12(2 - 0) = 24 > 0$, then 0 is not the mode
When $x = 2, f''(x) = 12(2 - (6 \times \dfrac{2}{3})) = -24 < 0$, then $\dfrac{2}{3}$ is the mode
$\Longrightarrow mode = \dfrac{2}{3}$
The remaining parts are left as exercise for you.

7.1.6 Median, Qurtile and other percentiles.

The median (m) is the value of x that gives half of the area under the function $f(x)$. It splits the area under the curve into two halves.
If m is the median, then for $f(x)$ defined for $a \leq x \leq b$,
I.e $\int_a^m f(x)dx = 0.5$ or $\int_m^{+\infty} f(x)dx = 0.5$ or $F(m) = 0.5$
In this text, am going to use
For median $\int_a^m f(x)dx = 0.5$
For lower quartile(q_1), use $\int_a^{q_1} f(x)dx = 0.25$ while upper quartile(q_3), use $\int_a^{q_3} f(x)dx = 0.75$
Remember: Semi - interqurtile range $= \dfrac{(q_3 - q_1)}{2}$, and interqurtile range $= (q_3 - q_1)$.
Examples:

(1) A random variable has it's p.d.f given by;

$$f(x) = \begin{cases} \beta x^3 & ; \quad 0 \leq x \leq 2 \\ 0 & ; \quad \text{elsewhere} \end{cases}$$

(a) Find the value of constant β
(b) Obtain

(i) median

(ii) 9^{th} decile

(iii) Semi - interqurtile range.

(iv) Standard deviation

Solution

(a) $\int_{-\infty}^{+\infty} f(x)dx = 1 \iff \int_0^2 \beta x^3 dx = 1 \iff \left[\frac{\beta}{4}x^4\right]_0^2 = 1 \iff \frac{\beta}{4}(2)^4 - \frac{\beta}{4}(0)^4 = 1$

$4\beta = 1$

$\implies \beta = \frac{1}{4}$

(b)(i) Median (m)

$\int_0^m \frac{1}{4}x^3 dx = 0.5$

$\left[\frac{1}{16}x^4\right]_0^m = 0.5$

$m^4 - 0 = 8$

$m^4 = 8$

$m = 1.6818$

(ii) Let 9^{th} deciles ($\text{d}_9 = a$) such that

$\int_0^a \frac{1}{4}x^3 dx = 0.9 \iff \left[\frac{1}{16}x^4\right]_0^a = 0.9 \iff a^4 - 0 = 14.4 \iff a^4 = 14.4$

$a = 1.9480$

(iii) Semi - interqurtile range $= \dfrac{q_3 - q_1}{2}$

For q_1, we use $\int_0^{q_1} f(x)dx = 0.25 \iff \int_0^{q_1} \frac{1}{4}x^3 dx = 0.25 \iff \left[\frac{1}{16}x^4\right]_0^{q_1} = 0.25 \iff (q_1)^4 - 0 = 4 \iff a^4 = 4$

$a = 1.4142$

For q_3, we use $\int_0^{q_3} f(x)dx = 0.75 \iff \int_0^{q_3} \frac{1}{4}x^3 dx = 0.75 \iff \left[\frac{1}{16}x^4\right]_0^{q_3} = 0.75 \iff (q_1)^4 - 0 = 12 \iff a^4 = 12$

$a = 1.8612$

$\iff$ Semi - interqurtile range $= \dfrac{1.8612 - 1.4142}{2} = 0.2235$

(iv) **The remaining part is left for you.**

NB; In case of more than one sub-function(or intervals) in a given probability function, you must test for the appropriate interval that contain/ accommodate the median (by containing atleast 0.5). This is illustrated as follows.

(2) A random variable X has its p.d.f given by;

$$f(x) = \begin{cases} \frac{2}{7}x & ; \quad 0 \leq x \leq 2 \\ \frac{2}{7}(4-x) & ; \quad 2 \leq x \leq 3 \\ 0 & ; \quad \text{otherwise} \end{cases}$$

Find

(i) median

(ii) eightieth percentile

(iii) inter quartile range

(iv) Variance

Solution:

(i) Median (m)

Testing for the appropriate interval. I.e the one containing the m

For $0 \leq x \leq 2, \int_0^2 \frac{2}{7}xdx = \left[\frac{x^2}{7}\right]_0^2 = \frac{2^2-0}{7} = \frac{4}{7} = 0.5714 > 0.5$ This implies that we have the median in the first interval

$\Longrightarrow \int_0^m \frac{2}{7}xdx = 0.5 \Longleftrightarrow \left[\frac{1}{7}x^2\right]_0^m = 0.5$

$m^2 - 0 = 3.5$

$m^2 = 3.5$

$m = \pm 1.8708$ i.e You take the value in the range.

$\Longrightarrow$ median $= 1.8708$

(ii) Let $P_{80} = n$

Since $\int_0^2 \frac{2}{7}xdx = \left[\frac{x^2}{7}\right]_0^2 = \frac{2^2-0}{7} = \frac{4}{7} = 0.5714 < 0.8$. This automatically implies that we have P_{80} in the second interval. So

$\int_0^2 \frac{2}{7}xdx + \int_2^n \frac{2}{7}(4-x)dx = 0.8 \Longleftrightarrow \frac{4}{7} + \int_2^n \frac{2}{7}(4-x)dx = 0.8 \Longleftrightarrow \frac{2}{7}\left[4x - \frac{x^2}{2}\right]_2^n = \frac{8}{35}$

$5\left[\left(4n - \frac{n^2}{2}\right) - \left(4(2) - \frac{2^2}{2}\right)\right] = 4$

$40n - 5n^2 - 60 = 8$

$5n^2 - 40n + 68 = 0$

Either $n = 2.4508$ or $n = 5.5492$ i.e you take the value in the range.

$\Longrightarrow P_{80} = 2.4508$

(iii) Inter quartile range = upper quartile (Q_3) - lower quartile (Q_1)

For $q_3, \int_0^2 \frac{2}{7}xdx + \int_2^n \frac{2}{7}(4-x)dx = 0.75 \Longleftrightarrow \frac{4}{7} + \int_2^{q_3} \frac{2}{7}(4-x)dx = 0.75 \Longleftrightarrow \frac{2}{7}\left[4x - \frac{x^2}{2}\right]_2^{q_3} = \frac{5}{28}$

$8\left[\left(4(q_3) - \frac{(q_3)^2}{2}\right) - \left(4(2) - \frac{2^2}{2}\right)\right] = 5$

$64(q_3) - 8(q_3)^2 - 96 = 10 \Longleftrightarrow 8(q_3)^2 - 64(q_3) + 106 = 0$

Either $q_3 = 2.3417$ or $q_3 \neq 5.6583$ i.e Should be in the range

$\Longrightarrow Q_3 = 2.3417$

lower quartile (Q_1)

Since $\int_0^2 \frac{2}{7}xdx = \left[\frac{x^2}{7}\right]_0^2 = \frac{2^2-0}{7} = \frac{4}{7} = 0.5714 > 0.25$, then $\int_0^{Q_1} \frac{2}{7}xdx = 0.25$

$\left[\frac{1}{7}x^2\right]_0^{Q_1} = 0.25$

$(Q_1)^2 - 0 = 1.75$

$(Q_1)^2 = 1.75$

$a = \pm 1.3229$

Either $Q_1 = 1.3229$ or $Q_1 \neq -1.3229$

$\Longrightarrow Q_1 = 1.3229$

$\therefore$ Inter quartile range$= (2.3417 - 1.3229) = 1.0188$

Do the last part.

(3) A continuous random variable has the p.d.f given by

$$f(x) = \begin{cases} \frac{1}{4} & ; \quad 0 \leq x < 1 \\ \frac{x^3}{5} & ; \quad 1 \leq x \leq 2 \\ 0 & ; \quad \text{otherwise} \end{cases}$$

Find

(i) median

(ii) inter quartile range

(iii) middle 55%

(iv) Variance

Solution

(i) Median (m)

Testing for the appropriate interval. I.e the one containing the m

For $0 \leq x \leq 1, \int_0^1 \frac{1}{4}dx = \left[\frac{1}{4}x\right]_0^1 = 0.25 < 0.5$ This implies that we have the median in the second interval

$\int_0^1 \frac{1}{4}dx + \int_1^m \frac{x^3}{5} = 0.5 \Longleftrightarrow 0.25 + [\frac{x^4}{20}]_1^m = 0.5$

$(\frac{m^4}{20}) - (\frac{1^4}{20}) = 0.5 - 0.25 \Longleftrightarrow (\frac{m^4}{20}) = 0.3$

1.5651.

$\Longrightarrow$ Median $= 1.5651$

Complete the remaining parts

(4) An example for three interval to get median.

7.1.7 Cumulative Probability Density Function, $F(x)$.

As seen earlier, the cumulative probability density function of a continuous random variable is abbreviated as $F(X) = P(X \leq x)$ and defined as:
For any t in the range $a \leq x \leq b$ of the function $f(x)$, we have
$F(x) = \int_a^x f(t)dt$, for$F(a) = 0$ and$F(b) = 1$
Where:
$a =$ Lower limit or $-\infty$
$b =$ Upper limit or $+\infty$
Properties of $F(X)$: These include;

- $F(-\infty) = 0$
- $F(+\infty) = 1$
- $F(X)$ is a non decreasing function

How to use $F(X)$ to obtain probabilities. This gives probabilities from the lower limit to the mentioned values. The cases to consider include:

(a) $P(X < a) \ = \ F(a)$

(b) $P(a \leq x \leq b) \ = \ F(b) - F(a)$

(c) $P(X > b) \ = \ 1 - F(b)$

How to use $F(X)$ to obtain median,quartile percentile and deciles. Since the integration process has been done during the process of obtaining the $F(X)$, then:

(i) Median(m), we apply $F(m) = 0.5$ There after we simplify to obtain m

(ii) Lower quartile(q_1), we apply $F(q_1) = 0.25$ There after we simplify to obtain q_1

(iii) Upper quartile(q_3), we apply $F(q_3) = 0.75$ There after we simplify to obtain q_3

(iv) Percentile(P_n), we apply $F(P_n) = \dfrac{n}{100}$ There after we simplify to obtain P_n

(v) Decile(D_n), we apply $F(D_n) = \dfrac{n}{10}$ There after we simplify to obtain D_n
Also here, the appropriate interval must be identified as in 7.1.6

Examples

(1) A random variable has its p.d.f given by;

$$f(x) = \begin{cases} \alpha x^2 & ; \quad 0 \leq x \leq 3 \\ 0 & ; \quad \text{elsewhere} \end{cases}$$

(a) Find the value of constant α
(b) Obtain the cumulative probability function and hence find

(i) $P(X < 2)$

(ii) $P(1 < x < 2.5)$

(iii) $P(X \geq 2.5)$

(iv) $P(X \leq 1.8)$

(v) median

(vi) middle 80%

(vii) Semi interquartile range

Solution

Since $\int_{-\infty}^{+\infty} f(x)dx = 1$, then

$\int_0^3 \alpha x^2 dx = 1 \iff \left[\frac{\alpha}{3}x^3\right]_0^3 = 1 \iff \frac{\alpha}{3}(3)^3 - \frac{\alpha}{3}(0)^3 = 1$

$\therefore 9\alpha = 1$

$\alpha = \frac{1}{9}$

$$\implies f(x) = \begin{cases} \frac{1}{9}x^2 & ; \quad 0 \leq x \leq 3 \\ 0 & ; \quad \text{elsewhere} \end{cases}$$

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

For $x \leq 0, f(x) = 0, \implies F(0) = 0$

For $0 \leq x < 3, f(x) = \frac{1}{9}x^2$

$F(X) = F(0) + \int_0^x \frac{1}{9}t^2 dt = 0 + \left[\frac{t^3}{27}\right]_0^x = \left[\frac{x^3}{27} - 0\right] = \frac{x^3}{27}$

When $x = 3, \quad F(3) = \frac{3^3}{27} = 1$

$$\implies F(x) = \begin{cases} 0 & ; \quad x \leq 0 \\ \frac{x^3}{27} & ; \quad 0 \leq x < 3 \\ 1 & ; \quad x \geq 3 \end{cases}$$

(i)$P(x < 2) = P(0 < X < 2) = F(2) - F(0) = \frac{2^3}{27} - \frac{0^3}{27} = \frac{8}{27}$

(ii) $P(1 < x < 2.5) = [F(2.5) - F(1)] = [\frac{(2.5)^3}{27} - \frac{(1)^3}{27}] = 0.5417$

(iii) $P(X \geq 2.5) = P(2.5 \leq x \leq 3) = [\frac{(3)^3}{27} - \frac{(2.5)^3}{27}] = 0.4219$

(iv)$P(x \geq 1.8) = P(0 \leq X \leq 2) = [F(1.8) - F(0)] = [\frac{(1.8)^3}{27} - \frac{0^3}{27}] = 0.216$

(ii) Median (m)

$F(m) = \frac{m^3}{27} = 0.5$

$m^3 = 13.5$

$\therefore m = 2.3811$

(ii) middle 80% $= P_{90} - P_{10}$

Let $\mathrm{P}_{10} = a$and $\mathrm{P}_{90} = b$

For a, Using $F(a) = \frac{a^3}{27} = 0.1$

$a^3 = 2.7$

$a = 1.3925$

For b, Using $F(b) = \frac{b^3}{27} = 0.9$

$b^3 = 24.3$

$b = 2.8965$

$\Longrightarrow$ middle 80% $= 2.8965 - 1.3925 = 1.504$

(vii) Semi interquartile range $= \dfrac{q_3 - q_1}{2}$

For q_3, we use $F(q_3) = 0.75 \Longleftrightarrow \frac{(q_3)^3}{27} = 0.75$
$(q_3)^3 = 20.25 \Longleftrightarrow q_3 = 2.7259$

For q_1, we use $F(q_1) = 0.25 \Longleftrightarrow \frac{(q_1)^3}{27} = 0.25$
$(q_1)^3 = 6.750 \Longleftrightarrow q_1 = 1.8899$
$\Longrightarrow$ Semi interquartile range $= \dfrac{2.7259 - 1.8899}{2} = 0.4180.$

(2) A random variable X has its probability density function f(x) given by;

$$f(x) = \begin{cases} \frac{3}{32}x^2 & ; \quad 0 \leq x \leq 2 \\ \frac{3}{32}(6-x) & ; \quad 2 \leq x \leq 6 \\ 0 & ; \quad \text{otherwise} \end{cases}$$

(a) Obtain the expression of $F(x)$
(b) Find

(i) median,

(ii) 2nd to 9th semi inter-deciles range,

(iii) $P(1 < x < 4)$,

(iii) $P(X \geq 3)$,

(iv) $P(X \leq 5)$,

(vi) $P(X \leq 5/X > 1)$,

(vii) Interquartile range.

Solution (a) Using, $F(x) = \int_{-\infty}^{x} f(t)dt$,

For $x < 0, f(x) = 0, \Longleftrightarrow F(0) = 0$

For $0 \leq x < 2, f(x) = \frac{3}{32}x^2$
Then $F(X) = F(0) + \int_0^x \frac{3}{32}t^2dt = 0 + \left[\frac{t^3}{32}\right]_0^x = \left[\frac{x^3}{32} - 0\right] = \frac{x^3}{32}$

When $x = 2, F(2) = \frac{2^3}{32} = \frac{1}{4}$

For $2 \leq$ x < 6, F(2) $+ \int_2^x \frac{3}{32}(6-t)dt = \frac{1}{4} + \frac{3}{32}\left[6t - \frac{t^2}{2}\right]_2^x$

$$= \frac{1}{4} + \frac{3}{32}\left[\left(6x - \frac{x^2}{2}\right) - \left(6(2) - \frac{2^2}{2}\right)\right] = \frac{3}{32}\left(6x - \frac{x^2}{2}\right) - \frac{11}{16}$$

When $x = 6, \quad F(6) = \frac{3}{32}\left(6(6) - \frac{6^2}{2}\right) - \frac{11}{16} = 1$

$$\implies F(x) = \begin{cases} 0 & ; \quad x \le 0 \\ \frac{x^3}{32} & ; \quad 0 \le x < 2 \\ \frac{3}{32}\left(6x - \frac{x^2}{2}\right) - \frac{11}{16} & ; \quad 2 \le x < 6 \\ 1 & ; \quad x \ge 6 \end{cases}$$

(b)

(ii) Median (m)

Since for $x = 2, F(2) = \frac{2^3}{32} = \frac{1}{4} = 0.25 < 0.5$, then m is in the second interval.Therefore using: $F(m) = \frac{3}{32}\left(6m - \frac{m^2}{2}\right) - \frac{11}{16} = 0.5$

$18m - \frac{3m^2}{2} - 22 = 16$

$36m - 3m^2 - 76 = 0$

$3m^2 - 36m + 76 = 0$

Either $m = 2.7340$ or $mneq9.2660$

$\implies$ median $(m) = 2.7340$

(ii) 2^{nd} to 9^{th} semi-interdecile range $= \frac{(d_9 - d_2)}{2}$ using F(x)

Let $d_2 = n$ and $d_9 = w$

Since for $x = 2, F(2) = \frac{2^3}{32} = \frac{1}{4} = 0.25 > 0.2$, then n is in the first interval.Therefore:
For $n,\ F(n) = \frac{n^3}{32} = 0.2$

$n^3 = 6.4$

$n = 1.8566$

$\implies D_2 = 1.8566$

Since for $x = 2, F(2) = \frac{2^3}{32} = \frac{1}{4} = 0.25 < 0.9$, then w is in the second interval.Therefore:
For $w,\ F(w) = \frac{3}{32}\left(6w - \frac{w^2}{2}\right) - \frac{11}{16} = 0.9$

$18w - \frac{3w^2}{2} - 22 = 28.8$

$36w - 3w^2 - 101.6 = 0$

$3w^2 - 36w + 101.6 = 0$

Either $w = 4.5394$ or $w \neq 7.4606$(discard)

$\implies D_9 = 4.5394$

2^{nd} to 9^{th} semi-interdecile range $= \frac{(d_9 - d_2)}{2} = \frac{4.5394 - 1.8566}{2} = 1.3414$.

(iii) P(1 < X < 4) using F(x)

P(1 < X < 4) = F(4) - F(1) $= \left[\frac{3}{32}\left(6(4) - \frac{4^2}{2}\right) - \frac{11}{16}\right] - \left[\frac{1^3}{32}\right] = \frac{25}{32}$

Complete the remaining parts.

7.1.8 Graph of $f(x)$.

This is plotted as $f(x)$ against x values or range. The resulting graph can be a line(s) only, or curve(s) only, or both (i.e line(s) and curve(s)).
The does not need the graph paper and you need to formute the table for better sketching.

7.1.9 Graph of $F(X)$.

This is plotted as $F(x)$ against x values or range. The resulting graph ever must be like Orgive like figure/shape I.e Graph

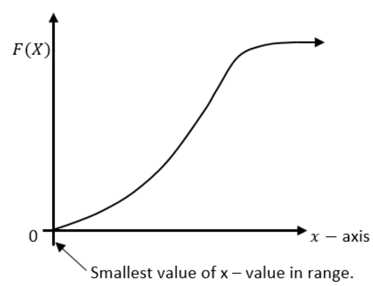


Figure 7.5

Example.

(1) A random variable has a p.d.f given by;

$$f(x) = \begin{cases} kx & ; \quad 0 \leq x \leq 3 \\ 0 & ; \quad \text{elsewhere} \end{cases}$$

(a) Sketch the graph of $f(x)$ and use it to find the value of k
(b) Find $F(x)$ hence sketch it.

(a) $f(x) = kx$ (This is line)

x	0	3
$f(x)$	0	$3k$

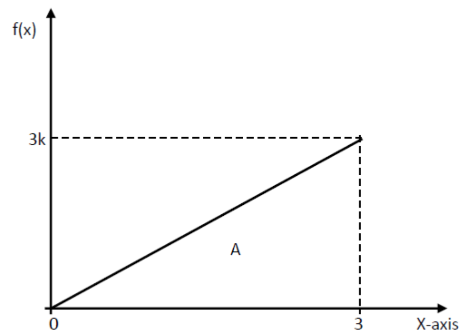


Figure 7.6

(i) **NB:** Total area(A) under the curve but above $x - axis = 1$
$\Longrightarrow A = \frac{1}{2} \times b \times h = 1 \Longleftrightarrow \frac{1}{2} \times 3 \times 3k = 1$
$\therefore k = \frac{2}{9}$
(b) Form

$$\therefore f(x) = \begin{cases} \frac{2}{9}x & ; \quad 0 \leq x \leq 3 \\ 0 & ; \quad \text{elsewhere} \end{cases}$$

For when $x = 0, \; f(x) = 0 \Longleftrightarrow F(X) = 0, \; F(0) = 0$
For $0 \leq x \leq 3, \; f(x) = \frac{2}{9}x$
$F(X) = F(0) + \int_0^x \frac{2}{9}tdt = 0 + [\dfrac{2}{18}t^2]_0^x = \dfrac{2}{18}x^2 - 0 = \dfrac{2}{18}x^2$
When $x = 3, \; F(3) = \dfrac{2}{18}(3)^2 = 1$

$$\Longrightarrow F(x) = \begin{cases} 0 & ; \quad x < 0 \\ \dfrac{2}{18}x^2 & ; \quad 0 \leq x \leq 3 \\ 1 & ; \quad x \geq 3 \end{cases}$$

Sketch of $F(x)$,
Extract the table:

x	0	2	3
$F(x)$	0	0.444	1

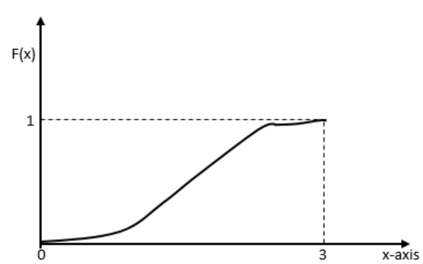


Figure 7.7

(2) A random variable X has its probability density $f(x)$ given by;

$$f(x) = \begin{cases} kx & ; \quad 0 \leq x \leq 2 \\ k(4-x) & ; \quad 2 \leq x \leq 4 \\ 0 & ; \quad \text{otherwise} \end{cases}$$

(a) Sketch the graph of $f(x)$ hence find k
(b) Sketch the graph of $F(X)$ hence sketch it
Solution.

For $0 \leq x \leq 2$, $f(x) = kx$ Using:

x	0	2
$f(x) = kx$	0	$2k$

For $2 \leq x \leq 4$, $f(x) = k(4-x)$ Using:

x	0	2
$f(x) = k(4-x)$	0	$2k$

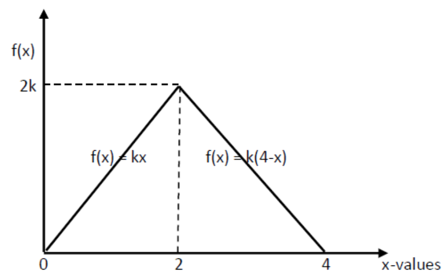


Figure 7.8

From the diagram:

$$\text{Total area} = 1$$
$$\frac{1}{2}bh = 1$$
$$\frac{1}{2} \times 4 \times 2k = 1$$
$$4k = 1$$
$$\Longrightarrow k = \frac{1}{4}$$

$$f(x) = \begin{cases} \frac{1}{4}x & ; \quad 0 \leq x \leq 2 \\ \frac{1}{4}(4-x) & ; \quad 2 \leq x \leq 4 \\ 0 & ; \quad \text{otherwise} \end{cases}$$

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

For $x < 0, f(x) = 0, \Longleftrightarrow F(0) = 0$

For $0 \leq x < 2, f(x) = \frac{1}{4}x$

Then $F(X) = F(0) + \int_{0}^{x} \frac{1}{4}tdt = 0 + \left[\frac{t^2}{8}\right]_0^x = \left[\frac{x^2}{8} - 0\right] = \frac{x^2}{8}$

When $x = 2, F(2) = \frac{2^2}{8} = \frac{1}{2} = 0.5$

For $2 \leq$ x < 4, $\quad$ F(2) + $\int_2^x \frac{1}{4}(4-t)dt = \frac{1}{2} + \frac{1}{4}\Big[4t - \frac{t^2}{2}\Big]_2^x$

$$= \frac{1}{2} + \frac{1}{4}\Big[\Big(4x - \frac{x^2}{2}\Big) - \Big(4(2) - \frac{2^2}{2}\Big)\Big] = \frac{1}{4}\Big(4x - \frac{x^2}{2}\Big) - 1$$

When $x = 4, \quad F(4) = \frac{1}{4}\Big(4(4) - \frac{4^2}{2}\Big) - 1 = 1$

$$\implies F(x) = \begin{cases} 0 & ; \quad x \leq 0 \\ \frac{x^2}{4} & ; \quad 0 \leq x < 2 \\ \frac{1}{4}\Big(4x - \frac{x^2}{2}\Big) - 1 & ; \quad 2 \leq x < 4 \\ 1 & ; \quad x \geq 4 \end{cases}$$

Sketch of ketch

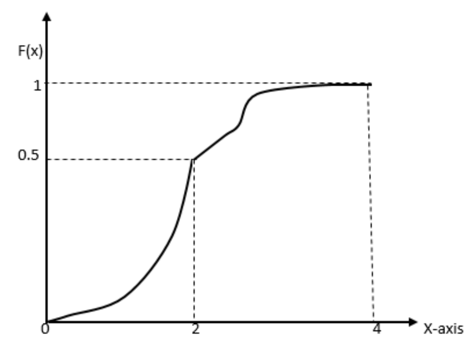


Figure 7.9

(3) A random variable X has its probability density f(x) given by;

$$f(x) = \begin{cases} kx^2 & ; \quad -1 \leq x \leq 2 \\ k(6-x) & ; \quad 2 \leq x \leq 6 \\ 0 & ; \quad \text{otherwise} \end{cases}$$

(a) Show that f(x) is continuous
(b) Sketch the graph of $f(x)$ hence find k
(c) Sketch the graph of $F(X)$ hence sketch it
(c) Find

(i) median $\qquad$ (ii) E(X)

Solution
NB: For $f(x)$ to be continuous, then $\boxed{f(x_1) = f(x_1)}$, for the range $a \leq x_1 \leq b$ of the

function
Now since $x = 2$ is common in the intervals of $f(x)$, Then
(a) $f(2) = k(2)^2 = f(2)$
$k(6-2) = (2)^2k$
$\Longrightarrow 4k = 4k$

Since the two sides are equal, then f(x) is continuous.

(b) For $^-1 \leq x \leq 2, f(x) = kx^2,$

x	-1	0	1	2
f(x)	k	0	k	$4k$

For $2 \leq x \leq 6, f(x) = k(6-x),$

x	2	6
f(x)	$4k$	0

Sketch of $f(x)$.

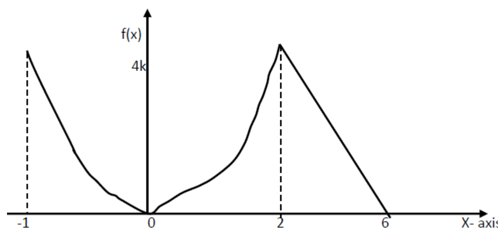


Figure 7.10

Hence, from the graph, $\int_{-1}^{2} kx^2dx + \dfrac{bh}{2} = 1 \Longleftrightarrow [\dfrac{kx^3}{3}]_{-1}^{2} + \dfrac{(6-2)\times(4k)}{2} = 1$
$[\dfrac{k(2)^3}{3}] - [\dfrac{k(-1)^3}{3}] + 8k = 1 \Longleftrightarrow 11k = 1$
$\therefore k = \frac{1}{11}$

$$\Longrightarrow f(x) = \begin{cases} \dfrac{1}{11}x^2 & ; \quad -1 \leq x \leq 2 \\ \dfrac{1}{11}(6-x) & ; \quad 2 \leq x \leq 6 \\ 0 & ; \quad \text{otherwise} \end{cases}$$

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

For $x \leq -1, f(x) = 0, \Longrightarrow F(-1) = 0$

For $-1 \leq x < 2, f(x) = \frac{1}{11}x^2$
$F(X) = F(-1) + \int_{-1}^{x} \frac{1}{11}t^2dt = 0 + \left[\frac{t^3}{33}\right]_{-1}^{x} = \left[\frac{x^3}{33} - \frac{(-1)^3}{33}\right] = \frac{(x^3+1)}{33}$

When $x = 2, \quad F(2) = \frac{(2^3+1)}{33} = \frac{9}{33} = \frac{3}{11}$
For $2 \leq x < 6, f(x) = \frac{1}{33}(6-x)$
$F(X) = F(2) + \int_{2}^{x} \frac{1}{11}(6-t)dt = \frac{3}{11} + \left[\frac{1}{11}(6t - \frac{t^2}{2})\right]_{2}^{x} = \frac{3}{11} + \left[\frac{1}{11}(6x - \frac{x^2}{2})\right] - \left[\frac{1}{11}(6(2) - \frac{2^2}{2})\right] =$
$\frac{3}{11} + \frac{1}{11}(6x - \frac{x^2}{2} - 10) = \frac{1}{11}(6x - \frac{x^2}{2} - 7)$

When $x = 6$, $\quad F(6) = \frac{1}{11}(6(6) - \frac{(6)^2}{2} - 7) = 1$

$$\Longrightarrow F(x) = \begin{cases} 0 & ; \quad x \leq -1 \\ \frac{(x^3+1)}{33} & ; \quad -1 \leq x < 2 \\ \frac{1}{11}(6x - \frac{x^2}{2} - 7) & ; \quad 2 \leq x < 6 \\ 1 & ; \quad x \geq 6 \end{cases}$$

(3) A probability density function of a random variable X is defined by

$$f(x) = \begin{cases} x^2 & ; \quad 0 \leq x \leq 1 \\ \frac{1}{2} & ; \quad 1 \leq x \leq k \\ 0 & ; \quad \text{otherwise} \end{cases}$$

where k is constant.

(i) Sketch the graph of $f(x)$,

(ii) find the value of k and hence the mean of the distribution,

(ii) Calculate the median of the distribution.

Solution.

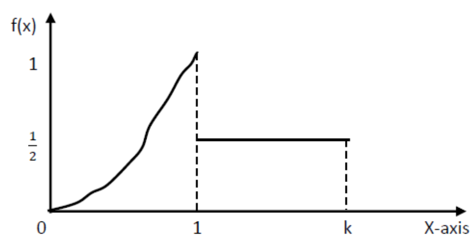


Figure 7.11

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

$\int_0^1 x^2 dx + \int_1^k \frac{1}{2} dx = 1$

$\left[\frac{x^3}{3}\right]_0^1 + \left[\frac{1}{2}x\right]_1^k = 1$

$\left[\frac{1}{3} - 0\right] + \left[\frac{k}{2} - \frac{1}{2}\right] = 1$

$\Longrightarrow k = \dfrac{7}{3}$

$$\Longrightarrow f(x) = \begin{cases} x^2 & ; \quad 0 \leq x \leq 1 \\ \frac{1}{2} & ; \quad 1 \leq x \leq \frac{7}{3} \\ 0 & ; \quad \text{otherwise} \end{cases}$$

Or You use the diagram and apply the area under the curve = 1
Please do the remaining parts.

(4) Sketch the continuous random variable

$$\Longrightarrow f(x) = \begin{cases} \frac{2}{3} & ; \quad 0 \leq x \leq 1 \\ \frac{1}{3} & ; \quad 1 \leq x \leq 2 \\ 0 & ; \quad \text{otherwise} \end{cases}$$

Solution.

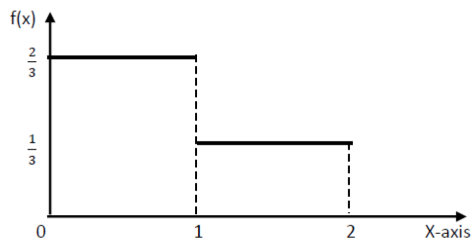


Figure 7.12

(5) The continuous random variable X has a p.d.f ($f(x)$) where

$$f(x) = \begin{cases} k(x+2)^2 & ; \quad -2 \leq x \leq 0 \\ 4k & ; \quad 0 \leq x \leq 1\frac{1}{3} \\ 0 & ; \quad \text{otherwise} \end{cases}$$

(a) Find the value of the constant k,
(b) Find $F(X)$
(c) Find $P(-1 \leq x \leq 1)$
(d) Sketch $f(x)$ and $F(X)$. **Solution:**
$\int_{-\infty}^{+\infty} f(x)dx = 1$

$\Longrightarrow \int_{-2}^{0}(x+2)^2dx + \int_{0}^{1\frac{1}{3}} 4kdx = 1 \Longleftrightarrow \frac{k}{3}\Big[(x+2)^3\Big]_{-2}^{0} + 4k\Big[x\Big]_{0}^{1\frac{1}{3}} = 1$

$\Longleftrightarrow 8k = 1$

$\Longrightarrow k = \dfrac{1}{8}$

$$f(x) = \begin{cases} \dfrac{1}{8}(x+2)^2 & ; \quad -2 \leq x \leq 0 \\ \dfrac{1}{2} & ; \quad 0 \leq x \leq 1\frac{1}{3} \\ 0 & ; \quad \text{otherwise} \end{cases}$$

(b)later (page 318)

(c) $P(-1 \leq x \leq 1) = \int_{-1}^{0}(x+2)^2 dx + \int_0^1 4k dx = \frac{7}{24} + \frac{1}{2} = \frac{19}{24}$.

(d) later

7.1.10 How to find $f(x)$ from it's sketch

This is done be finding the gradients of the sketch at different ranges. It's illustrated with the following examples below:

Examples.

(1) A random variable X has its p.d.f given by the sketch below.

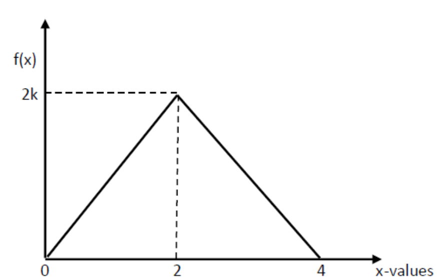


Figure 7.13

(a) Find the mode and value of the constant k.

(b) Obtain the equations of the p.d.f

(c) Find P(X $>$ 2/ X$\leq$ 3)

Solution

(a) Mode = 2

For k, Using Area = 1

$\implies A = \frac{1}{2} \times b \times h = \frac{1}{2} \times 4 \times 2k = 1$

$4k = 1$

$\therefore k = \frac{1}{4}$

(b)

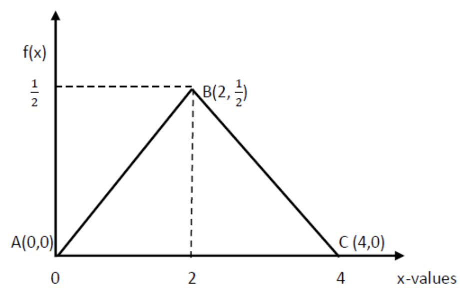


Figure 7.14

For $0 \leq$ x ≤ 2, Using $(0,0), (2, \frac{1}{4})$ From Gradient $= \frac{\Delta y}{\Delta x} = \frac{\frac{1}{4}-0}{2-0} = \frac{1}{4}$

Using at $(0,0)$, Grad= $\frac{1}{4}$ and (x, y)

$\frac{y-0}{x-0} = \frac{1}{4} \implies y = \frac{1}{4}x$

$\implies y = f(x) = \frac{1}{4}x$ with $0 \leq x \leq 2$

For $2 \leq$ x ≤ 4, Using $(2, \frac{1}{4}), (4,0)$ From Gradient $= \frac{\Delta y}{\Delta x} = \frac{0-\frac{1}{4}}{0-2} = \frac{-1}{4}$

Using at $(4,0)$, Grad= $\frac{-1}{4}$ and (x, y)

$\frac{y-0}{x-4} = \frac{-1}{4} \implies y = \frac{1}{4}(4-x)$

$\implies y = f(x) = \frac{1}{4}(4-x)$ with $2 \leq x \leq 4$

$$\implies f(x) = \begin{cases} \frac{1}{4}x & ; \quad 0 \leq x \leq 2 \\ \frac{1}{4}(4-x) & ; \quad 2 \leq x \leq 4 \\ 0 & ; \quad \text{otherwise} \end{cases}$$

P(x $>$ 2/ x $\leq$ 3) $= \frac{P(x>2 \ n \ x\leq 3)}{P(X\leq 3)} = \frac{P(2<x<3)}{P(0<x<3)} = \frac{3}{7}$ **Please show the working.**

(2) Given the illustration below.

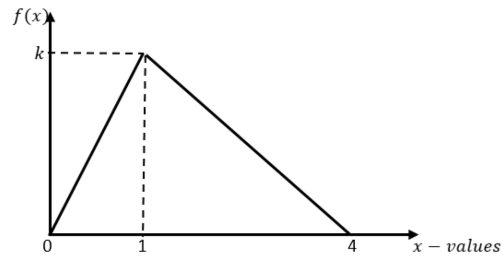


Figure 7.15

(a) Show that $K = \frac{1}{2}$,
(b) Determine $f(x)$ hence find mean and $Var(X)$.
Solution.
(a) Using Total area = 1

$\Longrightarrow A = \frac{1}{2} \times b \times h = 1 \Longleftrightarrow \frac{1}{2} \times 4 \times k = 1 \Longleftrightarrow 2k = 1$

$\therefore k = \frac{1}{4}$

(b)

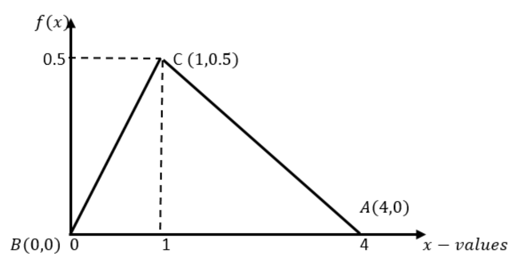


Figure 7.16

For $0 \leq$ x ≤ 1, Using $B(0,0), C(1,0.5)$ From Gradient $= \frac{\Delta y}{\Delta x} = \frac{0.5-0}{1-0} = \frac{1}{2}$

Using at $B(0,0)$, Grad= $\frac{1}{2}$ and (x,y)

$\frac{y-0}{x-0} = \frac{1}{2} \Longrightarrow y = \frac{1}{2}x$

$\Longrightarrow y = f(x) = \frac{1}{2}x$ with $0 \leq x \leq 1$

For $1 \leq$ x ≤ 4, Using $C(1,0.5), B(4,0)$ From Gradient $= \frac{\Delta y}{\Delta x} = \frac{0-0.5}{4-1} = \frac{-1}{6}$

Using at $B(4,0)$, Grad= $\frac{-1}{6}$ and (x,y)

$\frac{y-0}{x-4} = \frac{-1}{6} \Longrightarrow y = \frac{1}{6}(4-x)$

$\Longrightarrow y = f(x) = \frac{1}{4}(4-x)$ with $1 \leq x \leq 4$

$$\Longrightarrow f(x) = \begin{cases} \frac{1}{2}x & ; \quad 0 \leq x \leq 1 \\ \frac{1}{6}(4-x) & ; \quad 1 \leq x \leq 4 \\ 0 & ; \quad \text{otherwise} \end{cases}$$

Please complete the qn.

(3) Given the illustration below,

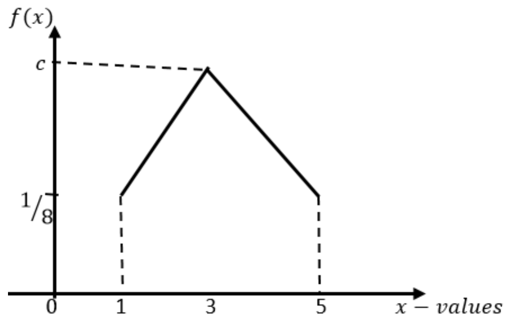


Figure 7.17

(a) Show that $C = \frac{3}{8}$,
(b) Determine $f(x)$
Solution.
(a) Using Total area $= 1 \iff$ Area of triangle + Area of rectangle $= 1$

Let $h = (c - \frac{1}{8})$ I.e from the diagram.
$\implies (\frac{1}{2} \times 4 \times (c - \frac{1}{8})) + (4 \times \frac{1}{8}) = 1 \iff 2(c - \frac{1}{8}) + \frac{1}{2} = 1 \iff 2c - \frac{1}{4} + \frac{1}{2} = 1$

$\therefore C = \frac{3}{8}$

(b)

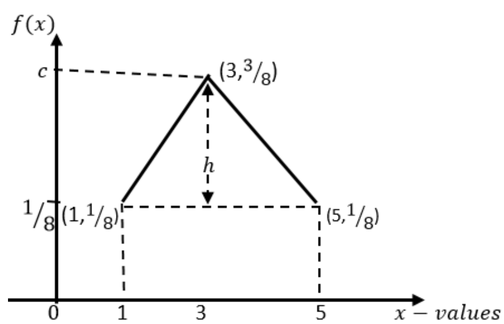


Figure 7.18

For $1 \leq$ x ≤ 3, Using $(1, \frac{1}{8}), (3, \frac{3}{8})$ From Gradient $= \frac{\Delta y}{\Delta x} = \frac{\frac{3}{8} - \frac{1}{8}}{3-1} = \frac{1}{8}$

Using at $(1, \frac{1}{8})$, Grad$= \frac{1}{8}$ and (x, y)

$\frac{y - \frac{1}{8}}{x-3} = \frac{1}{8} \implies y = \frac{1}{8}(x - 3) + \frac{1}{8} = \frac{1}{8}(x - 2)$

$\implies y = f(x) = \frac{1}{8}(x - 2)$ with $1 \leq x \leq 3$

For $3 \leq$ x ≤ 5, Using $(3, \frac{3}{8}), (5, \frac{1}{8})$ From Gradient $= \frac{\Delta y}{\Delta x} = \frac{\frac{1}{8} - \frac{3}{8}}{5-3} = \frac{-1}{8}$

Using at $(5, \frac{1}{8})$, Grad= $\frac{-1}{8}$ and (x, y)

$\frac{y-\frac{1}{8}}{x-5} = \frac{-1}{8} \Longrightarrow y = \frac{1}{8}(6-x)$

$\Longrightarrow y = f(x) = \frac{1}{8}(6-x)$ with $3 \leq x \leq 5$

$$\Longrightarrow f(x) = \begin{cases} \frac{1}{8}(x-2) & ; \quad 1 \leq x \leq 3 \\ \frac{1}{8}(6-x) & ; \quad 3 \leq x \leq 5 \\ 0 & ; \quad \text{otherwise} \end{cases}$$

(4) Given the illustration below,

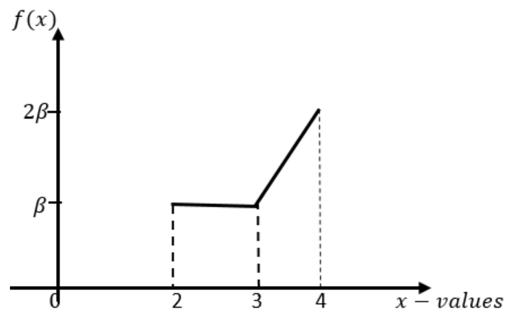


Figure 7.19

(a) Find the value of β
(b) Determine $f(x)$
Solution.
(a) Using Total area = 1 $\Longleftrightarrow$ Area of rectangle + Area of trapezium = 1

$\Longrightarrow ((3-2) \times \beta) + (\frac{1}{2} \times (4-3) \times (\beta + 2\beta) = 1 \Longleftrightarrow \frac{5}{2}\beta = 1$

$\therefore \beta = \frac{2}{5}$ (b)

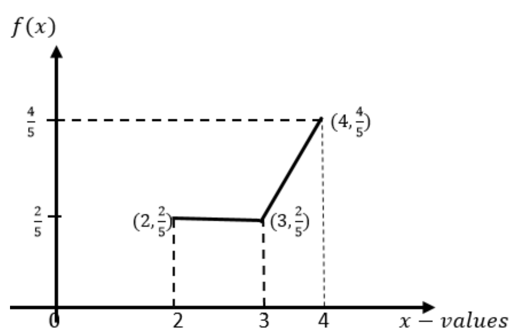


Figure 7.20

For $2 \leq$ x ≤ 3, Using $(2, \frac{2}{5}), (3, \frac{2}{5})$ From Gradient $= \frac{\Delta y}{\Delta x} = \frac{\frac{2}{5}-\frac{2}{5}}{3-2} = 0$

Using at $(2, \frac{2}{5})$, Grad= 0 and (x, y)

$\frac{y-\frac{2}{5}}{x-2} = 0 \Longrightarrow y = \frac{2}{5}$

$\Longrightarrow y = f(x) = \frac{2}{5}$ with $2 \leq x \leq 3$

For $3 \leq$ x ≤ 4, Using $(3, \frac{2}{5}), (4, \frac{4}{5})$ From Gradient $= \frac{\Delta y}{\Delta x} = \frac{\frac{4}{5}-\frac{2}{5}}{4-5} = \frac{2}{5}$

Using at $(3, \frac{2}{5})$, Grad= $\frac{2}{5}$ and (x, y)

$\frac{y-\frac{2}{5}}{x-3} = \frac{2}{5} \Longrightarrow y = \frac{2}{5}(x-2)$

$\Longrightarrow y = f(x) = \frac{2}{5}(x-2)$ with $3 \leq x \leq 4$

$$\Longrightarrow f(x) = \begin{cases} \frac{2}{5} & ; \quad 2 \leq x \leq 3 \\ \frac{2}{5}(x-2) & ; \quad 3 \leq x \leq 4 \\ 0 & ; \quad \text{otherwise} \end{cases}$$

7.1.11 How to find the $f(x)$ from $F(X)$

This is done by differentiating the cumulative distribution function $F(X)$ basing on the given ranges.I.e

$$\boxed{f(x) = \frac{d}{dx}[F(X)]},$$

Examples

1. Given a cumulative distribution function $F(X)$ such that

$$F(X) = \begin{cases} 0 & ; \quad x \leq 0 \\ \frac{x^2}{16} & ; \quad 0 \leq x \leq 4 \\ 1 & ; \quad x \geq 4 \end{cases}$$

(a) Find the p.d.f $f(x)$ hence use it to find $var(x)$
(b) Sketch $F(X)$and$f(x)$.
Solution:

From

$$F(X) = \begin{cases} 0 & ; \quad x \leq 0 \\ \frac{x^2}{16} & ; \quad 0 \leq x \leq 4 \\ 1 & ; \quad x \geq 4 \end{cases}$$

Using

$$f(x) = \frac{d}{dx}F(X) = \begin{cases} \frac{d}{dx}(0) & ; \quad x \leq 0 \\ \frac{d}{dx}[\frac{x^2}{16}] & ; \quad 0 \leq x \leq 4 \\ \frac{d}{dx}(1) & ; \quad x \geq 4 \end{cases}$$

$$f(x) = \begin{cases} \frac{x}{8} & ; \quad 0 \le x \le 4 \\ 0 & ; \quad otherwise \end{cases}$$

(b) **The remaining parts are left as an exercise**

2. Given a cumulative distribution function $F(X)$ such that

$$F(X) = \begin{cases} 0 & ; \quad x \le 0 \\ \frac{x^2}{6} & ; \quad 0 \le x \le 2 \\ \frac{-x^2}{3} + 2x - 2 & ; \quad 2 \le x \le 3 \\ 1 & ; \quad x \ge 3 \end{cases}$$

(a) Find the p.d.f $f(x)$ hence use it to find standard deviation
(b) Sketch $F(X)$and$f(x)$.
Solution:

From

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

$$\text{Then } f(x) = \begin{cases} \frac{x}{3} & ; \quad 0 \le x \le 2 \\ \frac{-2x}{3} + 2 & ; \quad 2 \le x \le 3 \\ 0 & ; \quad otherwise \end{cases}$$

(b) **The remaining parts are left as an exercise**

3. The continuous random variable X has cumulative distribution function $F(X)$ such that

$$F(X) = \begin{cases} 0 & ; \quad x \le -2 \\ \frac{1}{12}(2+x) & ; \quad -2 \le x < 0 \\ \frac{1}{6}(1+x) & ; \quad 0 \le x < 4 \\ \frac{1}{12}(6+x) & ; \quad 4 \le x < 6 \\ 1 & ; \quad x \ge 6 \end{cases}$$

(a) Find the p.d.f of X $f(x)$ hence use it to find standard deviation
(b) Sketch $F(X)$and$f(x)$.
(c) Find $E(X)$
Solution
Let $f(x) = \dfrac{d}{dx}F(X)$
For $-2 \le x \le 0,\ f(x) = \dfrac{d}{dx}\frac{1}{12}(2+x) = \dfrac{1}{12}$
For $0 \le x \le 4,\ f(x) = \frac{1}{6}(1+x) = \dfrac{1}{6}$

For $4 \leq x \leq 6,\ f(x) = \dfrac{d}{dx}\frac{1}{12}(6+x) = \dfrac{1}{12}$

$$\text{Then } f(x) = \begin{cases} \frac{1}{12} & ; \quad -2 \leq x \leq 0 \\ \frac{1}{6} & ; \quad 0 \leq x \leq 4 \\ \frac{1}{12} & ; \quad 4 \leq x \leq 6 \\ 0 & ; \quad otherwise \end{cases}$$

Other parts are left as exercise

4. The continuous random variable T has cumulative distribution function F(t) where;

$$F(t) = \begin{cases} 0 & ; \quad t \leq 1 \\ k(t-1)^2 & ; \quad 1 \leq t \leq 3 \\ \frac{(14ht - t^2 - 25)}{24} & ; \quad 3 \leq t \leq 7 \\ 1 & ; \quad t \leq 7 \end{cases}$$

(a) Determine the values of k and h
(b) Use F(t) to find;

(i) P(2.8 ≤ T ≤ 5.2) (ii) median of t

Solution:

From continuity of $F(t)$
$F(3) = F(3) \Longleftrightarrow k(t-1)^2 = \frac{(14ht - t^2 - 25)}{24}$
At $t = 3$,

$$k(3-1)^2 = \frac{(14h(3) - 3^2 - 25)}{24} \Longleftrightarrow 4k = \frac{42h - 34}{24} 96k - 42h + 34 = 0 \ldots\ldots (i)$$

Also $F(7) = F(7) \Longleftrightarrow \frac{(14ht - t^2 - 25)}{24} = 1$
At $t = 7$, $\frac{(14h(7) - (7)^2 - 25)}{24} = 1 \Longleftrightarrow h = 1$
From (i), with $h = 1$, we get $96k - 42(1) + 34 = 0$
$k = \dfrac{1}{12}$
$\therefore k = \dfrac{1}{12},\ h = 1$
The remaining parts are left for you.

7.1.12 Uniform / Rectangular distribution

This is a special continuous probability function whose random variable X in the given interval $[a, b]$ has p.d.f given by;

$$f(x) = \begin{cases} \frac{1}{b-a} & ; \quad a \leq x \leq b \\ 0 & ; \quad \text{elsewhere} \end{cases}$$

It's $X \sim R[a, b]$
Its graph can be sketched as below;

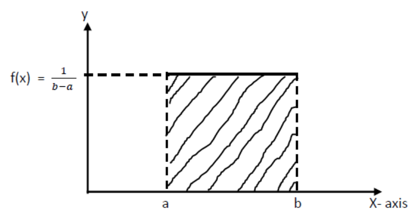


Figure 7.21

$\Longrightarrow$ The total are of the shaded region is 1. I.e $\int_a^b f(x)dx = 1$

Mean and Variance

The expectation and variance of a rectangular distribution $f(x)$ over the interval $[a, b]$ is defined as:

1. Mean $(E(X)) = \dfrac{(b+a)}{2}$
2. Variance$[Var(X)] = \dfrac{(b-a)^2}{12}$

QN: If X is following rectangular distribution $f(x)$ over the interval $[a, b]$. prove that

1. Mean $(E(X)) = \dfrac{(b+a)}{2}$
2. Variance$[Var(X)] = \dfrac{(b-a)^2}{12}$

Solution

(1)Mean= $E(X)$ such that
$E(X) = \int_a^b xf(x)dx = \frac{1}{b-a}\int_a^b xdx$

$$= \frac{1}{2(b-a)}\left[x^2\right]_a^b = \frac{b^2-a^2}{2(b-a)} = \frac{(b-a)(b+a)}{2(b-a)} = \frac{b+a}{2}$$

(2) Variance $= E(X^2) - \left[\text{E(X)}\right]^2$
But $E(X^2) = \int_a^b x^2 f(x)dx = \frac{1}{b-a}\int_a^b x^2 dx$

$$= \frac{1}{3(b-a)}\left[x^3\right]_a^b = \frac{b^3-a^3}{3(b-a)} = \frac{(b-a)^3 + 3ab(b-a)}{3(b-a)} = \frac{b^2+a^2+ab}{3}$$

Also $\left[\mathrm{E(X)}\right]^2 = \left(\frac{b+a}{2}\right)^2 = \frac{b^2+a^2+2ab}{4}$

$\Longrightarrow Var(X) = \left[\frac{b^2+a^2+ab}{3}\right] - \left[\frac{b^2+a^2+2ab}{4}\right]$

$$\frac{4b^2+4a^2+4ab-3b^2-3a^2-6ab}{12} = \frac{b^2+a^2-2ab}{12} = \frac{(b-a)^2}{12}$$

Examples

(1) The number of people crossing the bright follow a uniform distribution with the number between 12 to 20 people every hour. Let the number be x,find

(a) the p.d.f of x,

(b) the mean and standard deviation,

(c) $P(X \geq 18)$,

(d) $P(X \leq 16)$,

(e) $P(13 \leq x \leq 18)$,

Solution
(a) the p.d.f of x,

$$\text{From } f(x) = \begin{cases} \frac{1}{20-12} & ; \quad 12 \leq x \leq 20 \\ 0 & ; \quad otherwise \end{cases}$$

$$\text{Then } f(x) = \begin{cases} \frac{1}{8} & ; \quad 12 \leq x \leq 20 \\ 0 & ; \quad otherwise \end{cases}$$

(b) Mean $= \dfrac{a+b}{2} = \dfrac{12+20}{2} = 16.00$

Standard deviation $= \sqrt{var(X)} = \dfrac{(20-12)}{\sqrt{12}} = 2.3094.$

(c)$P(X \geq 18) = \int_{18}^{20} \dfrac{1}{8} dx = [\dfrac{x}{8}]_{18}^{20} = 0.250$

(d) $P(X \leq 16) = \int_{12}^{16} \dfrac{1}{8} dx = [\dfrac{x}{8}]_{12}^{16} = 0.50.$

(e) $P(13 \leq x \leq 18) = \int_{13}^{18} \dfrac{1}{8} dx = [\dfrac{x}{8}]_{13}^{18} = 0.6250$

(2) The arrival time by schools for a maths seminar at KAPROSS is a uniform distribution with minimum arrival time of 10 minutes and maximum arrival time of 20 minutes after 8am the start of the seminar. Find

(i) mean arrival time

(ii) variance

(iii) the probability that a school chosen at random arrived after 16 minutes.

Solution

X $\sim$ R[10, 20]

$$f(x) = \begin{cases} \frac{1}{10} & ; \quad 10 \le x \le 20 \\ 0 & ; \quad \text{elsewhere} \end{cases}$$

(i) E(X) $= \frac{b+a}{2} = \frac{20+10}{2} = 15$ minutes

(ii) Var(X) $= \frac{(b-a)^2}{12} = \frac{(20-10)^2}{12} = 8.3333$ minutes

P(X > 16) = P($16 <$ X < 20) =

$$\int_{16}^{20} \frac{1}{10} dx = \left[\frac{x}{10}\right]_{16}^{20} = \frac{20-16}{10} = \frac{4}{10} = 0.4$$

(3) During the L.C 1 elections in a certain village, voters lined up to vote for their candidate and their heights followed a uniform distribution in the interval $[\alpha, \beta]$. Given that their mean height was 3ft and variance $\frac{4}{3}$.
(a) Find the
(i) value of α and β,
(ii) probability that a voter selected at random from the line will be more than 2 and at most 4,
(b) Obtain the c.d.f of X and hence sketch it.

Solution

(a)(i) X $\sim$ R[α, β]

$$f(x) = \begin{cases} \frac{1}{\beta-\alpha} & ; \quad \alpha \le x \le \beta \\ 0 & ; \quad \text{elsewhere} \end{cases}$$

From, E(X) $= \frac{(\beta+\alpha)}{2} = 3$

$\beta + \alpha = 6 \cdots\cdots\cdots\cdots\cdots\cdots\cdots\cdots\cdots\cdots\cdots\cdots$ (i)

Similarly from Var(X)$= \frac{(\beta-\alpha)^2}{12} = \frac{4}{3}$

$(\beta - \alpha)^2 = 16$

$(\beta - \alpha)^2 = 4 \cdots\cdots\cdots\cdots\cdots\cdots\cdots\cdots\cdots\cdots\cdots\cdots\cdots$ (ii)

Solving eqn(1) and eqn(2)

$\Longrightarrow \alpha = 1$ and $\beta = 5$

$$\therefore f(x) = \begin{cases} \frac{1}{4} & ; \quad 1 \le x \le 5 \\ 0 & ; \quad \text{elsewhere} \end{cases}$$

P($2 <$ x ≤ 4.5) $= \int_2^{4.5} \frac{1}{4} dx = \left[\frac{x}{4}\right]_2^{4.5} = \frac{4.5-2}{4} = \frac{2.5}{4} = \frac{5}{8}$

(b) F(x) $= \int_{-a}^{x} f(t)dt$,
When $x < 1$, $f(x) = 0$, $F(x) = 0 \Longrightarrow F(1) = 0$

For $1 \leq x < 5$, $\quad F(1) + \int_1^x \frac{1}{4}dt = 0 + \left[\frac{t}{4}\right]_1^x = \left[\frac{x}{4} - \frac{1}{4}\right] = \frac{x-1}{4}$

When $x = 5 \quad F(5) = \frac{5-1}{4} = 1$

$$\Longrightarrow F(x) = \begin{cases} 0 & ; \quad x \leq 1 \\ \frac{x-1}{4} & ; \quad 1 \leq x \leq 5 \\ 1 & ; \quad x \geq 5 \end{cases}$$

The graph of F(x)

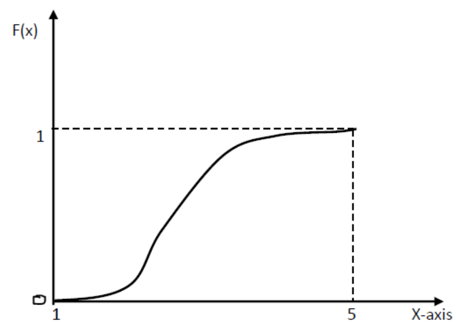


Figure 7.22

7.1.13 Exercise 7

(1) A continuous random variable X has a probapility density function $f(x)$ given by

$$f(x) = \begin{cases} \dfrac{k}{x(4-x)} & ; \quad 1 \leq x \leq 3 \\ 0 & ; \quad \text{otherwise} \end{cases}$$

(a) Show that $k = \dfrac{2}{lnk}$
(b) Calculate the mean and variance of X
Ans; (b) 2, 0.3590

(2) (a) The response time of students to assembly bells is uniformly distributed with maximum respomse time of 20 minutes and mean time of 13 minutes, find
(i) Minimum response time
(ii) probability that a student chosen at random will respond between 8 and 14 minutes after this bell is rung.
Ans; (i) a = 6 (ii) 3/7

(3) The p.d.f of a random variable X is given graphically as shown below.

GRAPH

(i) Express the pdf algebraically and hence find k
(ii) Find Var(X)
(iii) Derive the distribution function for X and hence find;

(a) P(X $\geq$ 6)

(b) the inter-quartile range

$$\textbf{Ans; (i)} f(x) = \begin{cases} \frac{k}{4}x & ; \ 0 \leq x \leq 4 \\ k & ; \ 4 \leq x \leq 10 \\ \frac{k}{4}\Big(14 - x) & ; \ 10 \leq x \leq 14 \\ 0 & ; \ \text{otherwise} \end{cases}, \quad \mathbf{k = \frac{1}{10}, (ii)(a)\ 0.6\ (b)\ 5}$$

(4) A random variable Y is uniformly distributed over interval (α, β). Given that P(Y < 3) = 0.25 and P(Y $\leq$ 7) = 0.7 .
(i) Find the values of α and β
(ii) Obtain and sketch the distribution function of Y.
(iii) Compute the 10 - 90 percentile range.

$$\textbf{Ans;(i)}\alpha = \frac{7}{9},\ \beta = \frac{29}{3} \quad (y) = \begin{cases} \textbf{(ii)}\ \frac{9}{80} & ; \ \frac{7}{9} \leq x \leq \frac{29}{3} \\ 0 & ; \ \text{otherwise} \end{cases}, \ \textbf{(iii)}\frac{64}{9}$$

(5) The continuous random variable T has cumulative distribution function F(T) where;

$$F(t) = \begin{cases} 0 & ; \ t \leq 1 \\ \alpha(t-1)^2 & ; \ 1 \leq t \leq 3 \\ \frac{(14t - t^2 - 25)}{24} & ; \ 3 \leq t \leq \beta \\ 1 & ; \ t \leq \beta \end{cases}$$

(a) Determine the values of α and β
(b) Use F(t) to find;

(i) P(2.8 $\leq$ T $\leq$ 5.2)

(ii) median of t

Ans; α = 1/12, β = 7 (b)(i) 0.595 (ii) 3.536

(6) The mass X kg of a particular substance produced per hour in a chemical process by the p.d.f below;

$$f(x) = \begin{cases} Kx^2 & ; \ 0 \leq x \leq 2 \\ K(6-x) & ; \ 2 \leq x \leq 6 \\ 0 & ; \ \text{otherwise} \end{cases}$$

(i) Sketch f(x) (ii) Find the value of k (iii) Determine E(X)

Ans; (ii) $k = 3/32$ **(iii)** 23/8

(a) The substance produced is sold at shs 9000 per kilogram and the total running cost is shs 4500 per hour. Find the
(i) expected profit per hour
(ii) probability that in an hour, the profit will exceed shs 31500
Ans; (a)(i) 21375 shs (ii) 3/16

(7) A random variable X has p.d.f given by f(x) where;

$$f(x) = \begin{cases} kx & ; \ 0 < x < 3 \\ 3k(4-x) & ; \ 3 < x < 4 \\ 0 & ; \ x < 0 \text{ and } x > 4 \end{cases}$$

(a) Sketch f(x)
(b) Find the

(i) value of k (ii) P(1 < X < 3) (iii) Mean of X

(c) Obtain the cummulative probability function for x

Ans; (b)(i) $\frac{1}{6}$ **(ii)** $\frac{2}{3}$ **(iii)** $\frac{7}{3}$ **(c)** $F(X) = \begin{cases} 0 & ; \ x \leq 0 \\ \frac{x^2}{12} & ; \ 0 \leq x \leq 3 \\ 2x - \frac{x^2}{4} - 3 & ; \ 3 \leq x \leq 4 \\ 1 & ; \ x \geq 4 \end{cases}$

(8) The continuous random variable X has a cumulative distribution function

$$F(x) = \begin{cases} 0 & ; \ x \leq 0 \\ k_1 x & ; \ 0 \leq x \leq 1 \\ \frac{x}{3} + k_2 & ; \ 0 \leq x \leq 2 \\ 1 & ; \ x \geq 2 \end{cases}$$

Find

(i) the values of k_1 and k_2 (ii) P(X < 1.5/ X > 1)

Ans; (i) $k_1 =$ **and** $k_2 =$

(9) A random variable X is uniformly distributes with mean 3. If P(2 < X < 3) = 0.25, find

(i) P(2 < X < 3.7) (ii) the variance of X

Ans; (i) 0.425 (ii) 4/3

(9) A random variable is uniformly distributed over the interval [10, b]. Given that P(x > 25) = 0.5 . Find the value b and mean.
Ans; b = 40, E(X) = 25

(11) The continuous random variable x has a probability density function given by

$$f(x) = \begin{cases} \lambda x^2 & ; \quad 0 < x < 2 \\ \lambda(6 - x) & ; \quad 2 < x < 4 \\ 0 & ; \quad x < 0 \text{ and } x > 4 \end{cases}$$

(a) Sketch f(x)
(b) Find the

(i) value of constant λ (ii) E(3X - 2) (iii) 90^{th} percentile

Ans; (i) 3/26 (ii) 5.3845 (iii) 3.6056

(12) The life span of a battery is known to be uniformly distributed with mean 4 and variance 4/3 issued with a three years' guarantee. If two batteries are picked at random, what is the probability that both will be replaced for failing to beat the guarantee?
Ans; 1/4

(13) The conditions random variable X has a distribution function F(x) given by

$$F(x) = \begin{cases} 0 & ; \quad x \leq 0 \\ Kx^2 & ; \quad 0 \leq x \leq 6 \\ 1 & ; \quad x > 6 \end{cases}$$

(i) Determine the value of the constant k and the median X
(ii) Sketch f(x) and hence find mode.
Ans; 1/36, 4.2426 (ii) 6

(14) The variable X is the distance in metres, that an inexperienced tight rope walker has moved along a given rope before falling off is given as;

P(X > x) = 1 - $\frac{x^3}{64}$; $0 < x < 4$
(a) Show that E(X) = 3
(b) Find the standard deviation, δ of X
(c) Show that P($|x - 3| < \delta$) = $\frac{69}{80}\sqrt{\left(\frac{3}{5}\right)}$
Ans; 9.6

(15) A continuous random variable X has a p.d.f f(x) defined by;

$$f(x) = \begin{cases} kx^2(3-x) & ; \quad 0 \leq x \leq 3 \\ 0 & ; \quad \text{elsewhere} \end{cases}$$

Determine
(i) the value of the constant k
(ii) the mean μ and the variance σ^2 of f(x) and verify that $\mu + 2\sigma = 3$
(iii) the probability that X differs from its mean by more than 2σ.
Ans; (i) 4/27 (ii) μ = 9/5 , δ^2 = 9/25 (iii) 0.0272

(16) The error x made in determining a certain parameter in an experiment is a random variable having a uniform distribution over the interval [a, b], given that the mean error is zero and the variance is o.48, find the value of a and b and the P(X > 0.8)
Ans; a = -1.2, b= 1.2, P = 1/6

(17) A random variable X has the cumulative distribution function;

$$F(x) = \begin{cases} 0 & ; \quad x \leq 0 \\ \frac{x^2}{4} & ; \quad 0 \leq x \leq 1 \\ \alpha x + \beta & ; \quad 1 \leq x \leq 2 \\ \frac{1}{4}(5-x)(x-1) & ; \quad 2 \leq x \leq 3 \\ 1 & ; \quad x \geq 3 \end{cases}$$

(a) Find

(i) the values of the constants α and β (iii) The p.d.f, f(x)

(ii) P(0.5 ≤ x ≤ 2.5/$x > 1$)

(b) Sketch the graph of y = f(x) and hence or otherwise, deduce the mean of X.

Ans; (i) $\alpha = \frac{1}{2}$, $\beta = -\frac{1}{4}$ **(ii)** $\frac{11}{12}$ **(iii)** $f(x) = \begin{cases} \frac{1}{2}x & ; \quad 0 \leq x \leq 1 \\ \frac{1}{2} & ; \quad 1 \leq x \leq 2 \\ \frac{1}{2}(3-x) & ; \quad 2 \leq x \leq 3 \\ 0 & ; \quad \text{otherwise} \end{cases}$ **E(x) = 1.5**

(18) A continuous a random variable X has a cumulative distribution function;

$$P(X \leq x) = \begin{cases} \frac{1}{64}x^3 & ; \quad 0 \leq x \leq \beta \\ 1 & ; \quad x \geq \beta \end{cases}$$

Find the;

(a) value of the constant β

(b) probability density function

$$\textbf{Ans; (a)}\beta = \textbf{4 (b)} \ f(x) = \begin{cases} \frac{3}{64}x^2 & ; \ 0 \leq x \leq 4 \\ 0 & ; \ \text{otherwise} \end{cases}$$

(19) The time taken to perform a particular task, t hours has probability density function

$$f(t) = \begin{cases} 10kt^2 & ; \ 0 \leq t \leq 0.6 \\ 9k(1-t) & ; \ 0.6 \leq t \leq 1.0 \\ 0 & ; \ \text{otherwise} \end{cases}$$

where k is a constant.
(a) Sketch f(t) and hence write down the most likely time
(b) Determine the

(i) value of constant k

(ii) expected time

(iii) probability that the time will be between 24 and 48 minutes.
Ans; (a) 0.6 (b)(i) k = 25/36 (ii) 35.5 (iii) 0.7269

(20) The error in grammes made by a green grocer's scale follows a uniform distribution over the interval [a, 7]. If the expected error in the scale is 2. Determine the;

(i) value of a

(iii) c.d.f, hence sketch it

(ii) probability that the error made is positive

Ans; (i) -3 (ii) 0.7

(21) A random variable has its p.d.f given by;

$$f(y) = \begin{cases} k(y+3) & ; \ -1 \leq y \leq 1 \\ k(5-y) & ; \ 1 \leq y \leq 3 \\ 0 & ; \ \text{otherwise} \end{cases}$$

(a) Sketch the graph of y and hence find the

(i) value of k

(ii) P(X > 1)

(iii) most likely value

Ans; (i) k = 1/12 (ii) 0.5 (iii) 1

(22) The random variable X has a probability function given by;

$$f(x) = \begin{cases} ax & ; \ 1 \leq x \leq 3 \\ c(4-x) & ; \ 3 \leq x \leq 4 \\ 0 & ; \ \text{otherwise} \end{cases}$$

(a) show that c = 3a
(a) Sketch f(x) and hence find

(i) values of a and c (ii) most likely value (iii) 85^{th} percentile

Ans; (i) a = 2/11, c = 6/11 (ii) 3 (iii) 3.26

(23) A random variable W has a probability density function given by;

$$f(w) = \begin{cases} \frac{1}{2c} & ; \quad 0 \leq w \leq 2c \\ 0 & ; \quad \text{otherwise} \end{cases}$$

Show that Var(W) = $\frac{c^2}{3}$, c is a constant

(24) A random variable R has its p.d.f given by

$$f(r) = \begin{cases} ar^2 + b & ; \quad 1 \leq r \leq 3 \\ 0 & ; \quad \text{otherwise} \end{cases}$$

Given that E(R) = $\frac{7}{3}$ and Var(R) = $\frac{11}{45}$, find

(i) the values of a and b (ii) Sketch the graph of f(r)

Ans; a = 1/8, b = 1/24

(25) The p.d.f of random variable X is given by

$$f(x) = \begin{cases} cx \sin \pi x & ; \quad 0 \leq x \leq 1 \\ 0 & ; \quad \text{otherwise} \end{cases}$$

(a) show that $c = \pi$

(26) A continuous random variable X has a probability distribution defined by

$$f(x) = \begin{cases} kx & ; \quad 0 \leq x \leq 8 \\ 8k & ; \quad 8 \leq x \leq 9 \\ 0 & ; \quad \text{otherwise} \end{cases}$$

(i) Sketch the graph of f(x)
(ii) Show that k = 0.025
(iii) Determine for all x, the distribution function f(x)
(iv) Calculate the probability then as observed value of x exceeds 6
(v) Determine the median and the mode of the function of f(x)
Ans; (iv) 0.55 (v) median = 6.325 , mode = 8

(27) The random variable T is the time in minutes that a drunkard women has taken from a night club to her home before falling down on the road side and its given by;

P(T > t) = $\frac{10-t}{t}; 5 \leq t \leq 10$
(a) Obtain an expression for the cumulative distribution function of T.

(b) Derive the p.d.f of T and sketch its graph
(c) Find the mean and variance of the distribution

Ans;(a)$F(t) = \begin{cases} 0 & ; \quad t \leq 0 \\ 2 - \frac{10}{t} & ; \quad 5 \leq t \leq 10 \\ 1 & ; \quad t \geq 10 \end{cases}$, **(b)**$f(t) = \begin{cases} \frac{10}{t^2} & ; \quad 5 \leq t \leq 10 \\ 0 & ; \quad \text{otherwise} \end{cases}$, **(c) E(T) = 6.9315, V**

(28) The distance in kilometres that a faulty car moves a long Masaka road from total petrol station kyengera before stopping can be modelled by the graph below;

GRAPH

(a) Find the equations from the graph between 0 and 3km and hence or otherwise;

(i) Write down the p.d.f of x, f(x)

(ii) Value of constant k

(iii) Most likely distance travelled

(iv) P(X > 1.5 $/1 \leq x < 3$)

$$\textbf{Ans; (i) } f(x) = \begin{cases} kx & ; \quad 0 \leq x \leq 2 \\ 2k & ; \quad 2 \leq x \leq 3 \\ \frac{2}{9}kx^2 & ; \quad 3 \leq x \leq 6 \\ 0 & ; \quad \text{otherwise} \end{cases} \quad \text{(ii) k} = \frac{1}{18} \text{ (iii) 6 (iv) } \frac{5}{7}$$

(29) The arrival time for the wedding party by guests whose reception was estimated to be at 5pm was modelled in to a probability function whose graph is given below;

GRAPH

(a) Find the value of constant k and the most likely time of arrival
(b) Obtain the equations of p.d.f
(c) Obtain the c.d.f of T and hence or otherwise find

(i) P(X > 2)

(ii) median

(iii) 9$^{\text{th}}$ deciles

(d) Sketch the graph of F(t)

Ans;(a) k $= \frac{2}{5}$**, 5:00pm (b)**$f(t) = \begin{cases} \frac{2}{5}(0.5t + 1) & ; \quad -2 \leq t \leq 0 \\ \frac{2}{5}(1 - \frac{t}{3}) & ; \quad 0 \leq t \leq 3 \\ 0 & ; \quad \text{otherwise} \end{cases}$, **(c)**$F(t) = \begin{cases} 0 & ; \\ \frac{2}{5}(t + \frac{t^2}{4}) + \frac{2}{5} & ; \\ \frac{2}{5}(1 - \frac{t^2}{6}) + \frac{2}{5} & ; \\ 1 & ; \end{cases}$

GRAPH

Chapter 8

NORMAL DISTRIBUTION

8.1 NORMAL PROBABILITY DISTRIBUTION

8.1.1 Introduction

It's the most important continuous probability distribution in statistics. Many measured quantities in the nature sciences follow a normal distribution, under certain circumstances it is also a useful approximation to the binomial distribution.
If X is a normally distributed variable with mean μ and variance σ^2, the it's can be abbreviated as $X \sim N(\mu, \sigma^2)$,
It's equation or function $f(x)$ is given as

$f(x) = \frac{1}{\sqrt{2\pi\delta}} e^{-\frac{(X-\mu)^2}{2\delta^2}}; \ -\infty \leq x \leq +\infty$

Since the equation is challenging to be used, then we insteady use the normal tables I.e (either <u>individual</u> normal tables or <u>cumulative</u> normal tables.)
The continuous normal probability distribution can be sketched as

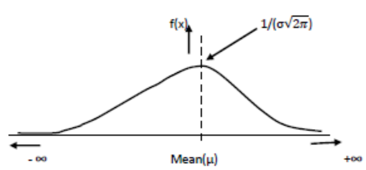


Figure 8.1: The normal curve is maximum at $\dfrac{1}{\sigma\sqrt{2\pi}}$.

8.1.2 Properties of a Normal curve

It has the following properties;

- It is bell-shaped
- Total area under the curve but above $x-$ axis is 1
- The curve is asymptotic about the x-axis.
- It extends from $-\infty$ to $+\infty$
- The maximum value of $f(x)$ is $\dfrac{1}{\sigma\sqrt{2\pi}}$
- The curve is symmetrical about the mean μ
- The mean, mode and median coincides at the maximum value of the function, where x $= \mu$

The Standard Normal Variable, Z: This is obtained by standardizing X where $X \sim N(\mu, \sigma^2)$. It adjusts the normal distribution $X \sim N(\mu, \sigma^2)$ into $Z \sim N(0.1)$ where
$0 =$ Mean of Z
$1 =$ Variance of Z.
$\Longrightarrow$ If $X \sim N(\mu, \sigma^2)$, then $Z = \frac{X-\mu}{\sigma}$ where $Z \sim N(0, 1)$
The standardized variable Z is sketched as below.

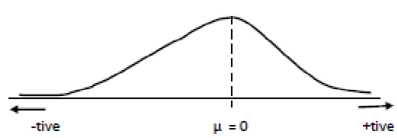


Figure 8.2: The are in this curve is $P(Z < z) = \phi(x)$.

$\therefore$ By definition, standardization is done by using

$$Z = \frac{X-\mu}{\sigma} \text{ where } Z \sim N(0, 1)$$

, Where
$\mu =$ mean of x
$\sigma =$ standard deviation of x
NB: During the process, Z remains continuous thus assuming the same inequality sign as that x had. And z can be obtained using any of these expressions;

- Normal distribution to normal. I.e If $X \sim N(\mu, \sigma^2)$, then $Z = \frac{X-\mu}{\delta}$.
- Distribution of Sample Mean of size n from a Normal Population .
  I.e If $X \sim N(\mu, \sigma^2)$, then $\overline{X} \sim N(\mu, \dfrac{\sigma^2}{n}) \Longleftrightarrow Z = \frac{X-\mu}{\delta/\sqrt{n}}$.
- Normal approximation to binomial. I.e Z $= \frac{X \pm 0.5-\mu}{\delta}$.

8.1.3 Probabilities in normal distribution

The probability that X lies between a and b is written as $P(a < x < b)$.To find the probability, you need to find **the area under the normal curve** between a and b. But since integrating the normal function is complicated and so difficult to integrate, we instead use **cumulative normal table** or **calculator** after obtaining Z mainly at three decimal places

When calculating the probabilities, remember that the total area under the standardized normal curve is 1.

Interpretation of Normal Diagram. This involves how to find the area of shaded portion in the normal diagram.
They are generalized as follows and the the shaded region in each case is the required probability;

- ♦ $P(0 < z < z_1)$

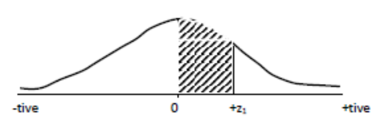


Figure 8.3

- ♦ $P(x > \pm a) = P(z > z_1) = 0.5 - P(0 < z < z_1)$

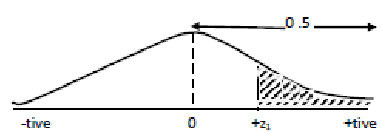


Figure 8.4

- ♦ $P(x > \pm x_2) = P(z > -z_2) = 0.5 + P(0 < z < z_2)$

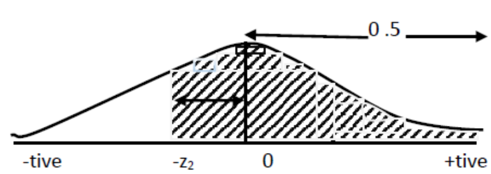


Figure 8.5

♦ $P(x < \pm x_3) = P(z < -z_3) = 0.5 - P(0 < z < z_3)$

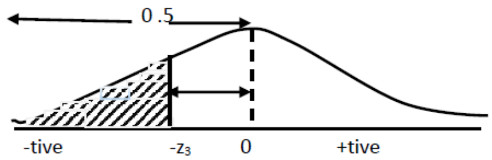


Figure 8.6

♦ $P(x < \pm x_4) = P(z < +z_4) = 0.5 + P(0 < z < z_4)$

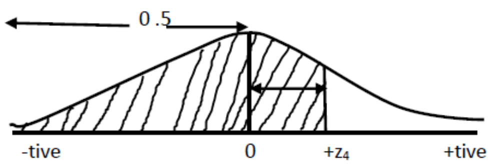


Figure 8.7

♦ $P(\pm x_5 < x < \pm x_6) = p(\pm z_6 < z < \pm z_7) = P(0 < z < \pm z_7) - P(0 < z < \pm z_6)$ I.e If you have the same signs of z values.

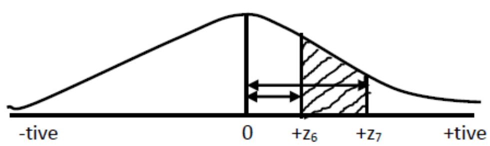


Figure 8.8

Or

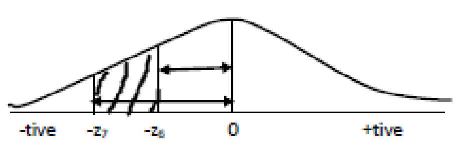


Figure 8.9: I.e Shading on same side, you subtract the two probabilities.

♦ $P(\pm x_8 < x < \pm x_9) = p(-z_8 < z < +z_9) = P(0 < z < \pm z_8) + P(0 < z < \pm z_9)$

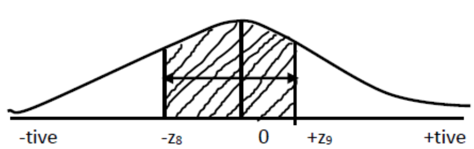


Figure 8.10: I.e Shading on different sides, you add the two probabilities.

NB: It's advisable to clearly indicate the source of your answer i.e either from the calculator as (cal) or table as (tab). E.g $P(X < x_1) = 0.2342(tab)$ or $P(X < x_1) = 0.23423(cal)$. **NB:** Since **normal distribution** is a special form of continuous random variables, then the following are indistinguishable:
$P(a \leq x \leq b) = P(a < x \leq b) = P(a \leq x < b) = P(a < x < b)$
NB: You are advised to draw sketches to illustrate your answers. This helps you to avoid simple mistakes.

8.1.4 Normal distribution to Normal distribution

Remember to standardize X, where $X \sim N(\mu, \sigma^2)$, then , z $= \frac{x-\mu}{\delta}$ where $Z \sim N(0, 1)$
Examples

(1) The mass in kilograms of people on a certain village is normally distributed with mean $68kg$ and standard deviation $9kg$.
(a) Find the probability that a person chosen at random from this village will have

(i) more than $75kg$,
(ii) more than $60kg$,
(iii) less $50kg$,
(vi) less than $71kg$,
(v) less than $45kg$,
(vi) between 55 and $75kg$,
(vii) between 52 and $58kg$,
(viii) between 70 and $79kg$.

(b) Find the percentage of people whose mass lies between $48kg$ and $54kg$
(c) Find how people whose mass lies between $70kg$ and $80kg$ if there were 180 people on the village
Solution
Let mass in kg of people be X

(a) (i) $P(X > 75) = P(Z > \frac{x-\mu}{\sigma}) = P(Z > \frac{75-68}{9}) = P(Z > 0.778)$

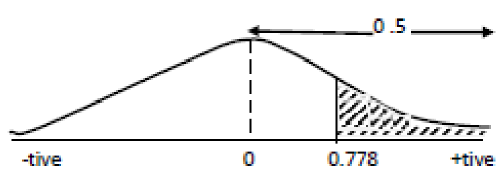


Figure 8.11

$\Longrightarrow P(X > 75) = 0.5 - P(0 < Z < 0.778) = 0.5 - 0.2818 = 0.2182(tab)$
(ii) $P(X > 60) = P(Z > \frac{60-68}{9}) = P(Z > -0.889)$

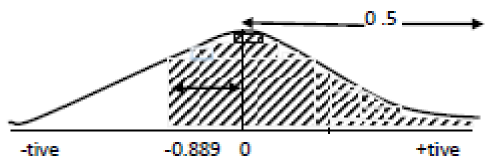


Figure 8.12

$\Longrightarrow P(X > 60) = 0.5 - P(0 < Z < -0.889) = 0.5 + 0.3130 = 0.8130(tab)$

(iii) $P(X < 50) = P(Z < \frac{50-68}{9}) = P(Z < -2.000)$

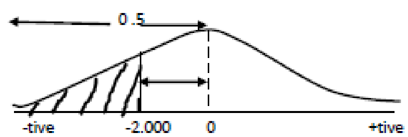


Figure 8.13

$\Longrightarrow P(X < 60) = 0.5 - P(0 < Z < 2.000) = 0.5 - 0.47725 = 0.02275(cal)$

(iv) $P(X < 71) = P(Z < \frac{71-68}{9}) = P(Z < 0.333)$

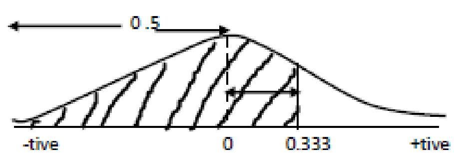


Figure 8.14

$\Longrightarrow P(X < 71) = 0.5 + P(0 < Z < 0.333) = 0.5 + 0.13043 = 0.63043(cal)$

(v) $P(X < 45) = P(Z < \frac{45-68}{9}) = P(Z < -2.556)$

$\Longrightarrow P(X < 45) = 0.5 - P(0 < Z < 2.556) = 0.5 - 0.49471 = 0.00529(cal)$

(vi) $P(55 < x < 75) = P(\frac{55-68}{9} < z < \frac{75-68}{9}) = P(-1.444 < z < 0.778)$

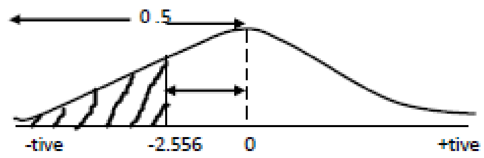


Figure 8.15

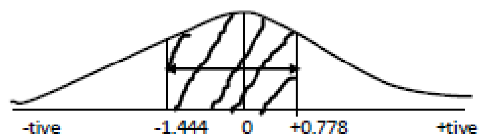


Figure 8.16

$\Longrightarrow P(55 < x < 75) = P(0 < Z < 0.778) + P(0 < Z < 1.444) = 0.4257 + 0.2819 = 0.7076(tab)$

(vii) $P(52 < x < 58) = P(\frac{52-68}{9} < z < \frac{58-68}{9}) = P(-1.778 < z < -1.111)$

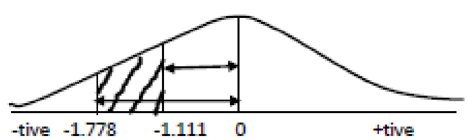


Figure 8.17

$\Longrightarrow P(55 < x < 75) = P(0 < Z < 1.778) - P(0 < Z < 1.111) = 0.46230 - 0.36672 = 0.09558(cal)$

(viii) $P(70 < x < 80) = P(\frac{70-68}{9} < z < \frac{80-68}{9}) = P(0.222 < z < 1.333)$

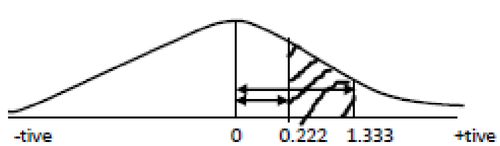


Figure 8.18

$\Longrightarrow P(70 < x < 80) = P(0 < Z < 1.333) - P(0 < Z < 0.222) = 0.40873 - 0.08784 =$

$0.32089(cal)$

(b) $P(48 < x < 54) = P(\frac{48-68}{9} < z < \frac{54-68}{9}) = P(-2.222 < z < -1.556)$

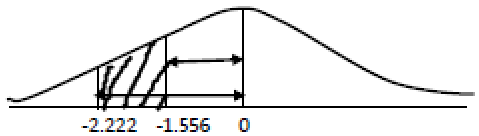


Figure 8.19

$\implies P(48 < x < 54) = P(0 < Z < 2.222) - P(0 < Z < 1.556) = 0.4869 - 0.4411 = 0.0458(tab)$

$\therefore$ 4.58% of the people will have their mass between 48 and $54kg$.

(c) Let $n = 180$ and
$P(70 < x < 79) = P(\frac{70-68}{9} < z < \frac{79-68}{9}) = P(0.222 < z < 1.222)$

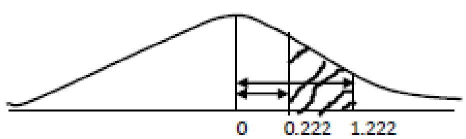


Figure 8.20

$\implies P(70 < x < 79) = P(0 < Z < 1.222) - P(0 < Z < 0.222) = 0.38915 - 0.08784 = 0.30131(cal)$
$\implies$ There are $0.30131\times = 55$ people in the village.

(2) A manufacturing company produces two types of iron sheets, X and Y. Their lengths are normally distributed with type X having an average length of $250cm$ and standard deviation $3cm$. The type Y has an average length of $300cm$ and standard deviation of $5cm$. Find the percentage of type
(i) X iron sheets that have a length of more than $255cm$.
(ii) Y iron sheets that have a length of more than $290cm$

Solution

(i) $\mu = 250, \sigma = 3$;
So $P(X > 255) = P(Z > \frac{x-\mu}{\delta} = \frac{255-250}{3} = P(Z > 1.667)$

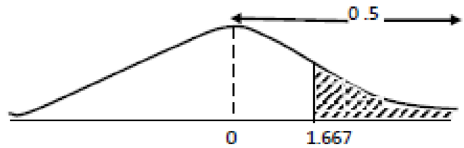


Figure 8.21

$\Longrightarrow P(X > 255) = 0.5 - P(0 < z < 1.667) = 0.5 - 0.4522 = 0.0478(tab)$
$\therefore \%age = 4.780$
(b) Do this part please ($Ans : 97.72\%$)

(3) Given that $X \sim N(-7, 12)$. Find

(a) $P(X < -9.8)$

(b) $P(X > -7.4)$

(c) $P(X < -6.3)$

(d) $P(-6.2 < x < -9.5)$

Solution.
From $X \sim N(-7, 12), \Longleftrightarrow \mu = -7$ and $\sigma = \sqrt{12}$
(a) $P(X < -9.8) = P(Z > \dfrac{^-9.8 -^- 7}{\sqrt{12}}) = P(Z > -0.808)$

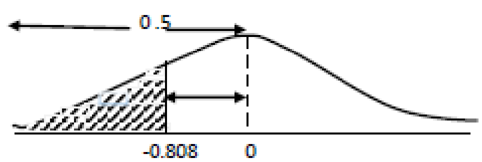


Figure 8.22

Please do the remaining parts.

8.1.5 Finding the Value of μ or σ or Both

This involves how to obtain $Z-$ value(s) or unknown given the corresponding probabilities. This is also referred to as **using the table reverse for any normal variables** X
Given the $z-$ diagram below, I.e

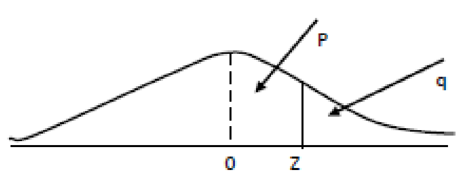


Figure 8.23

From the diagram, we extract the $Z-$ value using the table of the form

p	q	z
		$\pm Z-$value

NB For normal distribution, $p+q=0.5$
The $Z-$ value(s) is/are read from the critical values and if the value is not there, then you should look for the value from the cumulative normal distribution table.
The sign of Z is determined by the point where you have stoped shading on the normal diagram. If you have stoped on the positive side, then the value must be positive otherwise it is negative.

NB: Here we under the a backward move / reverse.

8.1.6 Quartiles, Percentiles and Deciles

Here we can obtain the quartiles, percentiles and deciles as below

1. For Lower quartile q_1, we use $P(X < q_1) = 0.25$
2. For upper quartile q_3, we use $P(X < q_3) = 0.75$
3. For percentile P_n, we use $P(X < P_n) = \dfrac{n}{100}$
4. For decile D_n, we use $P(X < D_n) = \dfrac{n}{10}$

Examples.

(1) Given that $X \sim N(\mu, 6^2)$ and $P(X > 55) = 0.2$. Find:

(a) The mean,

(b) Lower qurtile,

(c) Upper qurtile,

(d) Semi intequrtile range,

(e) 6^{th} decile,

(f) 35% percntile.

Solution.
(a) From $X \sim N(\mu, 6^2) \Longleftrightarrow \mu = \mu, \sigma = 6$ and
$P(X > 55) = 0.2$.
$P(X > 55) = P(Z > \dfrac{55-\mu}{6}) = 0.2$.

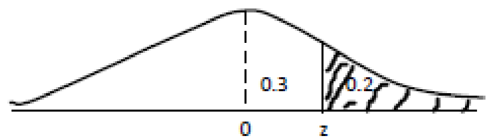


Figure 8.24

From the diagram,

p	q	z
0.30	0.20	$^{+}0.842$

$\therefore$ Using

$$Z =^{+} 0.842$$
$$\frac{55-\mu}{6} = 0.842$$
$$\Longrightarrow \mu = 60.052$$

(b) $P(X \leq q_1) = 0.25$
$P(X \leq q_1) = P(Z \leq \dfrac{q_1 - 60.0520}{6}) = 0.25.$

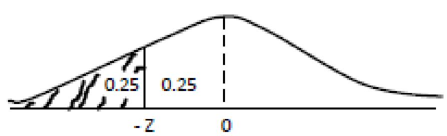


Figure 8.25

From the diagram,

p	q	z
0.25	0.25	$^{-}0.674$

$\therefore$ Using

$$Z =^{-} 0.674$$
$$\frac{q_1 - 60.0520}{6} =^{-} 0.674$$
$$\Longrightarrow q_1 = 56.0080$$

(c) $P(X \leq q_3) = 0.75$
$P(X \leq q_1) = P(Z \leq \dfrac{q_3 - 60.0520}{6}) = 0.75.$

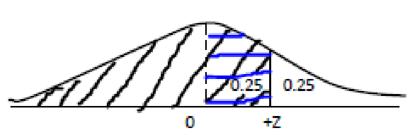


Figure 8.26

From the diagram,

p	q	z
0.25	0.25	$^{+}0.674$

$\therefore$ Using

$$Z =^{+} 0.674$$
$$\frac{q_3 - 60.0520}{6} =^{+} 0.674$$
$$\implies q_3 = 64.0960$$

(d) Semi interqurtile range $= \dfrac{q_3 - q_1}{2} = \dfrac{64.0960 - 56.0080}{2} = 4.0440$

(e) Let 6^{th} decile $= D_6$

$P(X \leq q_3) = 0.6$

$P(X \leq q_3) = P(Z \leq \dfrac{D_6 - 60.0520}{6}) = 0.6$

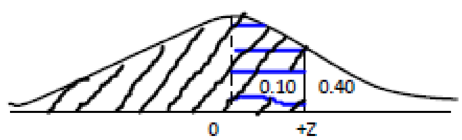


Figure 8.27

From the diagram,

p	q	z
0.10	0.40	$^{+}0.253$

$\therefore$ Using

$$Z =^{+} 0.253$$
$$\frac{D_6 - 60.0520}{6} =^{+} 0.253$$
$$\implies D_6 = 61.570$$

(f) Let 35% percentile $= P_{35}$

$P(X \leq P_{35}) = 0.35$

$P(X \leq P_{35}) = P(Z \leq \dfrac{P_{35} - 60.0520}{6}) = 0.35$

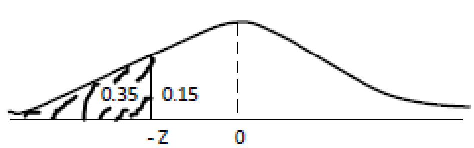


Figure 8.28

From the diagram,

p	q	z
0.35	0.15	$^{-}1.036$

$\therefore$ Using

$$Z =^{-} 1.036$$
$$\frac{P_{35} - 60.0520}{6} =^{-} 1.036$$
$$\Longrightarrow P_{35} = 53.8360$$

(2) Given that $X \sim N(100, \sigma^2)$ and $P(X < 106) = 0.8849$. Find:
(a) the value of the standard deviation σ,
(b) $P(X > 98)$
(c) Semi quatile range.
Solution:
Since $X \sim N(100, \sigma^2)$ compare it with $X \sim N(\mu, \sigma^2)$
$\Longrightarrow \mu = 100, \;\; \sigma = \sigma$
Let $P(X < 106) = P(Z < \dfrac{106 - 100}{\sigma}) = 0.8849$

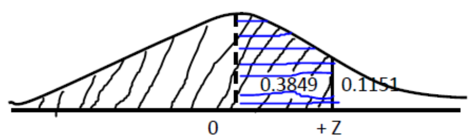


Figure 8.29

From the diagram,

p	q	z
0.3849	0.1151	$^{+}1.200$

$\therefore$ Using

$$Z =^{+} 1.200$$
$$\frac{106 - 100}{\sigma} =^{+} 1.200$$
$$\Longrightarrow \sigma = 5.0$$

(b) and (c) **Please left as the exercise for you.**

(3) The height of football students at KAPROSS is normally distributed with mean $168cm$ and standard deviation $12cm$. Find the

(i) upper quartile (ii) 4^{th} decile (iii) limits of the middle 70%

Please do the question
Answers: (i) Upper quartile $= 176.088cm$
(ii) $4^{\text{th}} = 164.964$
(iii) Middle 70% is P_{15} and P_{85} are Lower limit $= 155.568cm$, Upper limit $= 180.432cm$

(4) Given that $X \sim N(\mu, \sigma^2)$ such that $P(X < 10) = 0.0567$ and $P(X > 30) = 0.01151$
(a) Find the value of μ and σ.
(b) find interqurtile range.
Solution
From $P(X < 10) = P(Z < \dfrac{10-\mu}{\sigma}) = 0.0567$

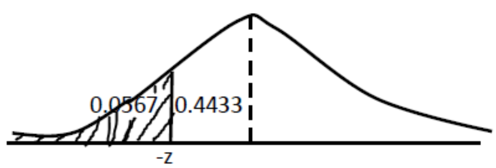


Figure 8.30

From the diagram,

p	q	z
0.4433	0.0567	$^{-}1.583$

$\therefore$ Using

$$Z =^{-} 1.583$$
$$\frac{10-\mu}{\sigma} =^{-} 1.583$$
$$\Longrightarrow \mu - 1.583\sigma = 10 \cdots\cdots\cdots\cdots\cdots (1)$$

From $P(X > 30) = P(Z > \dfrac{30-\mu}{\sigma}) = 0.01151$

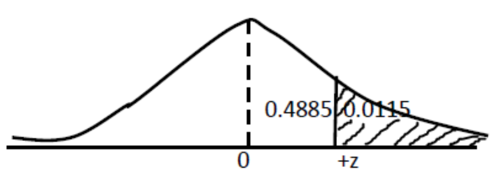


Figure 8.31

From the diagram,

p	q	z
0.4885	0.0115	$^{+}2.272$

$\therefore$ Using

$$Z =^{+} 2.272$$
$$\frac{30-\mu}{\sigma} =^{-} 2.272$$
$$\Longrightarrow \mu + 2.272\sigma = 30 \cdots\cdots\cdots\cdots\cdots (2)$$

From (1) and (2), we have $\mu = 18.2127$ and $\sigma = 5.1881$
(b) **Please do it as exercise.**

(5) Given that $X \sim N(\mu, \sigma^2)$ such that $P(X > 4.00) = 0.3$ and $P(X > 4.53) = 0.2$
(a) Find the value of μ and σ.
(b) $P(|X-2| < 3)$.
Solution
From $P(X > 4.00) = P(Z > \dfrac{4.00-\mu}{\sigma}) = 0.3$

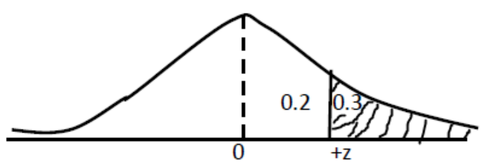


Figure 8.32

From the diagram,

p	q	z
0.2	0.3	$^{+}0.524$

$\therefore$ Using

$$Z =^{+} 0.524$$
$$\frac{4.00-\mu}{\sigma} =^{+} 0.524$$
$$\Longrightarrow \mu + 0.524\sigma = 4.00 \cdots\cdots\cdots\cdots\cdots (1)$$

From $P(X > 4.53) = P(Z > \dfrac{4.53-\mu}{\sigma}) = 0.2$

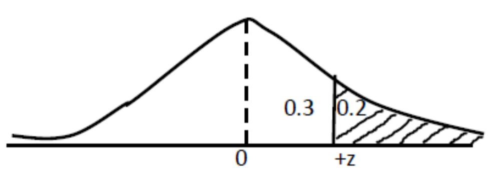


Figure 8.33

From the diagram,

p	q	z
0.3	0.2	$^{+}0.842$

$\therefore$ Using

$$Z =^{+} 0.842$$
$$\frac{4.53-\mu}{\sigma} =^{+} 0.842$$
$$\Longrightarrow \mu + 0.842\sigma = 4.53 \cdots\cdots\cdots\cdots\cdots (2)$$

From (1) and (2), we have $\mu = \dfrac{469}{150} = 3.1267$ and $\sigma = \dfrac{5}{3} = 1.6667$
(b) $P(|X-2| < 3) = P(-3+2 < x < 3+2) = P(-1 < x < 5)$
Please complete the question.

(5) A random variable X is normally distributed such that $X \sim N(\mu, \sigma^2)$. Given that $P(X < 188.440) = 0.9846$, and $P(X > 159.280) = 0.8399$. Find

(a) values of μ and σ

(b) $P(x > 92)$

(c) Find X_0 such that $P(X < X_0) = 0.8$

(d) Find X_1 such that $P(X > X_1) = 0.6$

Solution: From $P(X < 188.440) = P(Z < \dfrac{188.440-\mu}{\sigma}) = 0.9846$

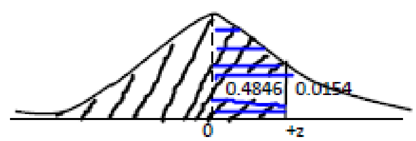


Figure 8.34

From the diagram,

p	q	z
0.4846	0.0154	$^{+}2.160$

$\therefore$ Using

$$Z =^{+} 2.160$$
$$\frac{188.440-\mu}{\sigma} =^{+} 2.160$$
$$\Longrightarrow \mu + 2.160\sigma = 188.440 \cdots\cdots\cdots\cdots\cdots (1)$$

From $P(X > 159.280) = P(Z > \dfrac{159.280-\mu}{\sigma}) = 0.8399$

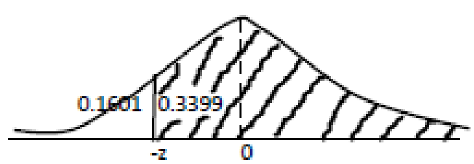


Figure 8.35

From the diagram,

p	q	z
0.3399	0.1601	$^{-}1.080$

$\therefore$ Using

$$Z =^{-} 1.080$$
$$\frac{159.280 - \mu}{\sigma} =^{-} 1.080$$
$$\Longrightarrow \mu - 1.080\sigma = 159.280 \cdots\cdots\cdots\cdots\cdots (2)$$

From (1) and (2), we have $\mu = 169.0$ and $\sigma = 9.0$

(b) From, $P(X < X_0) = P(Z < \dfrac{X_0 - 169.0}{9}) = 0.8$

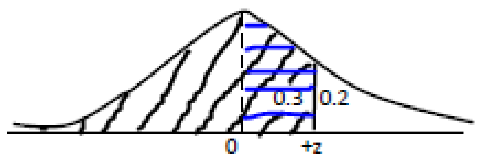


Figure 8.36

From the diagram,

p	q	z
0.3	0.2	$^{+}0.842$

$\therefore$ Using

$$Z =^{+} 0.842$$
$$\frac{X_0 - 169.0}{9} =^{+} 0.842$$
$$\Longrightarrow X_0 = 176.578$$

(c) **Please do the last part.**

(6) A random variable X is normally distributed such that X $\sim N(\mu, \sigma^2)$. Given that $P(X < 55) = 0.44$, and $P(55 < x < 84) = 0.27$. Find

(a) Values of μ and σ

(b) $P(X > 92)$

Solution

From $P(X < 55) = P(X < \dfrac{55 - \mu}{\sigma}) = 0.44$,

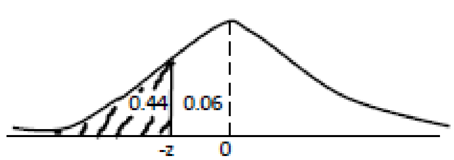


Figure 8.37

From the diagram,

p	q	z
0.06	0.44	$^-0.1510$

$\therefore$ Using

$$Z =^- 0.1510$$
$$\frac{55-\mu}{\sigma} =^- 0.1510$$
$$\Longrightarrow \mu - 0.1510\sigma = 55 \cdots\cdots\cdots\cdots\cdots (1)$$

Also from $P(55 < x < 84) = P(\frac{55-\mu}{\sigma} < Z < \frac{84-\mu}{\sigma}) = 0.27$

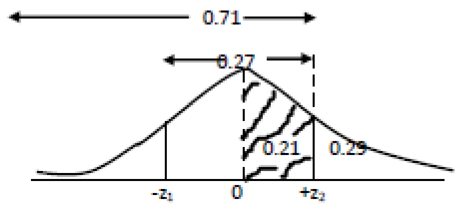


Figure 8.38

From the diagram,

p	q	z_2
0.21	0.29	$^+0.553$

$\therefore$ Using

$$Z =^+ 0.553$$
$$\frac{84-\mu}{\sigma} =^+ 0.553$$
$$\Longrightarrow \mu + 0.553\sigma = 84 \cdots\cdots\cdots\cdots\cdots (2)$$

From (1) and (2), we have $\mu = 61.2202$ and $\sigma = 41.193$
(b) $P(X > 92) = P(Z > \frac{92-61.2202}{41.193}) = P(Z > 0.747)$

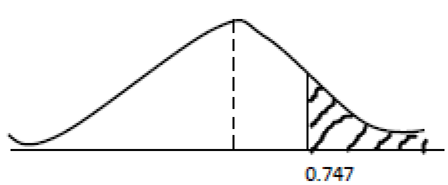


Figure 8.39

$\Longrightarrow P(X > 92) = 0.5 - P(0 < z < 0.747) = 0.5 - 0.27247 = 0.22753.$

(7) A total population fo 400 students sat an examination for which the pass mark was 43%. Their marks were normally distributed such that 16 students scored below 40 marks, while 20 scored above 60

(a) Find the mean μ and standard deviation σ.
(b) suppose the pass mark is increased by 4, how many more students passed?
Solution.
Let the marks be X
$P(X < 40) = \dfrac{16}{400} = 0.04$
$\implies P(X < 40) = P(X < \dfrac{40-\mu}{\sigma}) = 0.04$

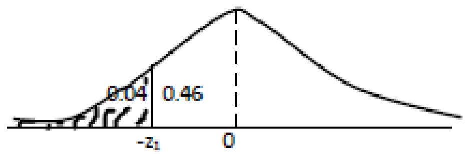


Figure 8.40

From the diagram,

p	q	z_2
0.46	0.04	$^{-}1.751$

$\therefore$ Using

$$Z =^{-} 1.751$$
$$\frac{40-\mu}{\sigma} =^{-} 1.751$$
$$\implies \mu - 1.751\sigma = 40 \cdots\cdots\cdots\cdots\cdots (1)$$

Similarly, $P(X > 60) = \dfrac{20}{400} = 0.05$
$\implies P(X < 60) = P(X < \dfrac{60-\mu}{\sigma}) = 0.05$

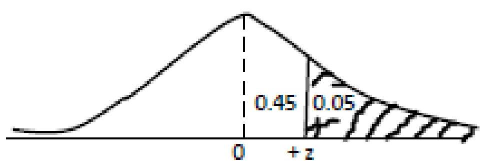


Figure 8.41

From the diagram,

p	q	z_2
0.45	0.05	$^{+}1.645$

$\therefore$ Using

$$Z =^{+} 1.645$$
$$\frac{60-\mu}{\sigma} =^{+} 1.645$$
$$\implies \mu + 1.645\sigma = 60 \cdots\cdots\cdots\cdots\cdots (2)$$

From (1) and (2), we have $\mu = 50.3121$ and $\sigma = 5.8893$
(b) **Please do it.**

8.1.7 Distribution of Sample Mean from a Normal Population.

Let $x_1, x_2, x_3, \cdots x_n$, be a random sample of size n taken from the normal distribution with mean μ and variance σ^2 such that $X \sim N(\mu, \sigma^2)$, then the sample mean $\overline{X}$ is also normally distributed with mean μ and variance $\dfrac{\sigma^2}{n}$ such that $\overline{X} \sim N(\mu, \dfrac{\sigma^2}{n}) \iff Z = \frac{X-\mu}{\sigma/\sqrt{n}}$. This is used when finding the probability that the sample mean(average) sample of size n can occur .

NB: In order to consider a normal distribution as a sample, the word(s) sample mean (average) sample must be mentioned in the sentence and there fore we use $Z = \frac{X-\mu}{\sigma/\sqrt{n}}$ instead of $Z = \frac{X-\mu}{\sigma}$.

Examples

(1) The height in centimeters of student in a certain school were normally distributed with mean $100cm$ and variance $81cm^2$. A sample of 16 students is taken, find the probability that a sample mean will have a height;

(i) more than $104cm$

(ii) Atleast $95cm$

(iii) Atleast $97cm$ and less than $103cm$

Solution

We have $\overline{X} \sim N(100, \dfrac{81}{16}) \iff \sigma = \sqrt{81} = 9, n = 16, \mu = 100$

(i) $P(\overline{\mathrm{X}} > 104) = P(Z > \frac{x-\mu}{\frac{\sigma}{\sqrt{n}}}) = P(Z > \frac{104-100}{\frac{9}{\sqrt{16}}}) = P(Z > 1.778)$

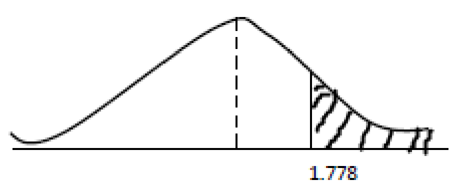


Figure 8.42

$\implies P(\overline{\mathrm{X}} > 104) = 0.5 - P(0 < Z < 1.778) = 0.5 - 0.4623 = 0.0377(cal)$.

(ii) $P(\overline{\mathrm{X}} \geq 95) = P(Z \geq \frac{95-\mu}{\frac{\sigma}{\sqrt{n}}}) = P(Z \geq \frac{95-100}{\frac{9}{\sqrt{16}}}) = P(Z > -2.222)$

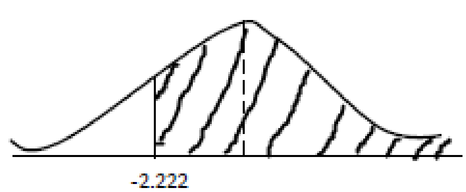


Figure 8.43

$\Longrightarrow P(\overline{\mathrm{X}} \geq 95) = 0.5 - P(0 \geq z \geq 2.222) = 0.5 + 0.48686 = 0.98686(cal).$
(iii) $P(97 \leq \overline{\mathrm{X}} < 103) = P(\frac{97-100}{\frac{9}{\sqrt{16}}} \leq \overline{\mathrm{Z}} < \frac{103-100}{\frac{9}{\sqrt{16}}}) = P(-1.333 \leq \overline{\mathrm{X}} < 1.333)$

Figure 8.44

$\Longrightarrow P(97 \leq \overline{\mathrm{X}} < 103) = 2P(0 < Z < 1.333) = -2 \times 0.40873 = 0.81746(cal).$

(2) The day's of production of milk sold in the packet of the certain firm is normally distributed with mean of 0.5 liters and standard deviation of 0.08 liters.Suppose a sample of 16 packets is drawn from the days of production, find the probability that the mean is between 0.46 and 0.47 liters.
Solution.
Let the litres of the milk produced be X
$\Longrightarrow \overline{X} \sim N(0.5, \dfrac{0.08^2}{16})$.Then
$P(0.46 < \overline{X} < 0.47) = P(\frac{0.46-0.5}{\frac{0.08}{\sqrt{16}}} < Z < \frac{0.47-0.5}{\frac{0.08}{\sqrt{16}}}) = P(-2.000 < Z < -1.500)$

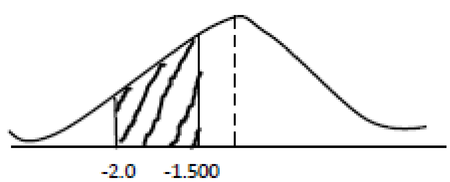


Figure 8.45

$$\begin{aligned}\Longrightarrow P(0.46 < \overline{X} < 0.47) &= P(0 < Z < 2.000) + P(0 < Z < 1.500)\\ &= 0.47725 + 0.43319\\ &= 0.91044(cal)\end{aligned}$$

Task: The distribution of the random variable X is $N(25, 340)$. The mean of the random of size n drawn from the distribution is $\overline{X}$. Find the value of n such that $P(\overline{X} > 28) \approx 0.005$.
$Ans: \ n = 250$

8.1.8 Normal Approximation to Binomial Distribution.

If $X \sim B(n,p)$ and n and p are such that $np > 5$ and $nq > 5$ where $q = 1-p$ then $X \sim N(np, npq)$ approximately.
Howerever, Basing on the target group, the number of trials (n) must be above 20. I.e $n \geq 21$
The normal distribution is used to approximate the binomial distribution when:
(a) the number of trials of the binomial experiment is large. I.e $n \geq 21$
(b) the probability of success p is constant and should be close to 0.5

Continuity Correction.

This is done by ensuring the equality of inequality signs i.e $[\leq$ or $\geq]$ there after you look for the class boundaries of the obtained interval.
NB: Class boundaries are obtained by taking $\pm 0.5 \times 10^{-n}$ where n = number of decimal places. Some books just say ± 0.5
The above adjustment is done to accommodate the fact that a discrete distribution is being approximated by a continuous distribution. **NB:** Remember if $X \sim B(n,p)$, then mean $\mu = np$ and variance $\sigma^2 = npq$,
NB: In this case, we must take into consideration the meaning of the words like at most, at least, more, less, inclusive, and others
All the above key points can be illustrated as below:

- $P(5 \leq X \leq 8) = P(4.5 \leq X \leq 8.5)$
- $P(5.4 \leq X \leq 6.7) = P(5.35 \leq X \leq 6.75)$ ie apply $\pm 0.5 \times 10^{-n}$
- $P(5 < X \leq 8) = P(6 \leq X \leq 8) = P(5.5 \leq X \leq 8.5)$
- $P(5 \leq X < 8) = P(5 \leq X \leq 7) = P(4.5 \leq X \leq 7.5)$
- $P(5 < X < 8) = P(6 \leq X \leq 7) = P(5.5 \leq X \leq 7.5)$
- $P(X < 4) = P(X \leq 3) = P(X \leq 3.5)$
- $P(X \leq 4) = P(X \leq 4.5)$
- $P(X > 4) = P(X \geq 5) = P(X \leq 4.5)$
- $P(X \geq 4) = P(X \leq 3.5)$
- $P(X = 4) = P(3.5 \leq X \leq 4.5)$
- $P(X = x_1) = P((x_1 - 0.5) \leq X \leq (x_1 + 0.5))$
- $P(X = 0) = P(-0.5 \leq X \leq 0.5)$

NB: You should define X as a binomial, then check that the conditions are suitable before defining the approximate normal distribution.

Examples.

(1) Given that $X \sim B(200, 0.7)$, find:

(a) $P(X \leq 130)$

(b) $P(136 \leq X < 148)$

(c) $P(X < 142)$

(d) $P(X = 152)$

(e) $P(X > 158)$

Solution

(a) $P(X \leq 130)$
From $X \sim B(200, 0.7), \Longleftrightarrow n = 200 > 20, p = 0.7$ and $q = 0.3$, then, $X \sim N(np, npq)$ such that:
$\mu = np = 200 \times 0.7 = 140$
$\sigma = \sqrt{npq} = \sqrt{200 \times 0.7 \times 0.3} = \sqrt{42}$
$\Longrightarrow P(X \leq 130) = P(X \leq 130.5) = P(Z \leq \dfrac{130.5 - 140}{\sqrt{42}}) = P(Z \leq^{-} 1.466)$

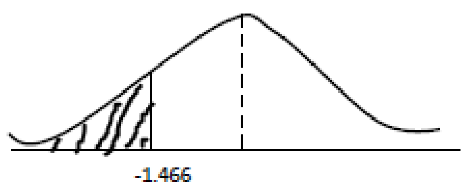


Figure 8.46

$\Longrightarrow P(X \leq 130) = 0.5 - P(0 < Z < 1.466) = 0.5 - 0.42868 = 0.07132(cal)$.
(b)

$$\begin{aligned} P(136 \leq X < 148) &= P(136 \leq X \leq 147) \\ &= P(135.5 \leq X \leq 147.5) \\ &= P(\frac{135.5 - 140}{\sqrt{42}} \leq Z \leq \frac{147.5 - 140}{\sqrt{42}}) \\ &= P(-0.694 \leq Z \leq 1.157) \end{aligned}$$

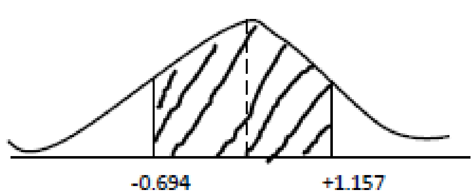


Figure 8.47

$$\begin{aligned}\Longrightarrow P(136 \leq X < 148) &= P(-0.694 \leq Z \leq 1.157)\\ &= P(0 \leq Z \leq 0.694) + P(0 \leq Z \leq 1.157)\\ &= 0.25616 + 0.37636\\ &= 0.63252(cal).\end{aligned}$$

(c)

$$\begin{aligned}P(X < 142) &= P(X \leq 141)\\ &= P(X \leq 141.5)\\ &= P(Z \leq \frac{141.5 - 140}{\sqrt{42}})\\ &= P(Z \leq 0.231)\end{aligned}$$

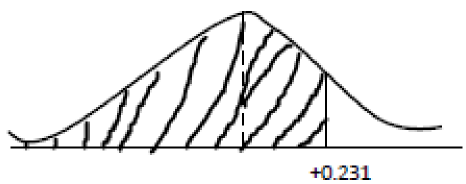


Figure 8.48

$\Longrightarrow P(X < 142) = 0.5 + P(0 \leq Z \leq 0.231) = 0.5 + 0.09134 = 0.59134(cal).$
(d)

$$\begin{aligned}P(X = 152) &= P(151.5 \leq X \leq 152.5)\\ &= P(\frac{151.5 - 140}{\sqrt{42}} \leq Z \leq \frac{152.5 - 140}{\sqrt{42}})\\ &= P(1.774 \leq Z \leq 1.929)\end{aligned}$$

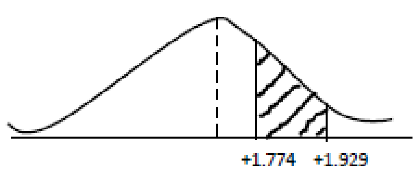


Figure 8.49

$$\begin{aligned}\Longrightarrow P(X = 152) &= P(0 \leq Z \leq 1.929) - P(0 \leq Z \leq 1.774)\\ &= 0.47313 - 0.46197\\ &= 0.01116(cal)\end{aligned}$$

(e)

$$
\begin{aligned}
P(X > 158) &= P(X \geq 159) \\
&= P(X \geq 158.5) \\
&= P(Z \geq \frac{158.5 - 140}{\sqrt{42}}) = P(Z \geq 2.855)
\end{aligned}
$$

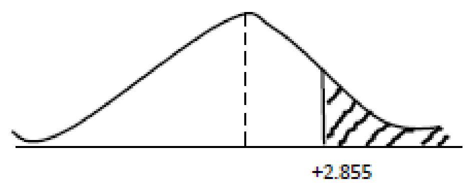


Figure 8.50

$\Longrightarrow P(X > 158) = 0.5 - P(0 \leq Z \leq 2.855) = 0.5 - 0.49785 = 0.00215(cal)$

(2) A coin is biased such that the probability of obtaining a head is three times that of getting a tail. Suppose the coin is tossed 120 times, find the probability of obtaining:

(i) less than 80 heads,

(ii) More than 86 heads,

(iii) at least 98 heads,

(iv) at most 82 heads,

(v) exactly 100 heads,

(vi) between 81 and 87 heads inclusive,

(vii) between 92 and 95 heads,

(viii) between 84 to 94 heads.

(a)

Solution.
Since $P(H) = 3P(T), n = 120 > 20$
Let $P(H) + P(T) = 1 \Longleftrightarrow 3P(T) + P(T) = 1 \Longleftrightarrow P(T) = \dfrac{1}{4}$
and $P(H) = 3P(T) = 3 \times \dfrac{1}{4} = \dfrac{3}{4}$
$\therefore p = \dfrac{3}{4}$, while $q = \dfrac{1}{4}$ This is because the qn is interested in heads. Then
Mean$(\mu) = np = 120 \times \dfrac{3}{4} = 90$
Similary, Standard deviation$(\sigma) = \sqrt{npq} = \sqrt{120 \times \frac{3}{4} \times \frac{1}{4}} = 4.7434$

(i)

$$\begin{aligned} P(X < 80) &= P(X \leq 79) \\ &= P(X \leq 79.5) \\ &= P(Z \leq \frac{79.5 - 90}{4.7434}) \\ &= P(Z \leq -2.214) \end{aligned}$$

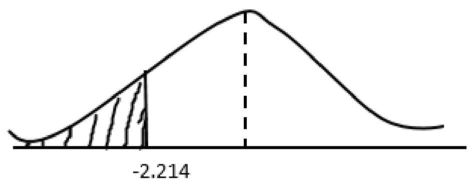


Figure 8.51

$\Longrightarrow P(X < 80) = 0.5 - P(0 \leq Z \leq 2.214) = 0.5 - 0.48659 = 0.01341(cal)$.
(ii)

$$\begin{aligned} P(X > 86) &= P(X \geq 87) \\ &= P(X \geq 86.5) \\ &= P(Z \geq \frac{86.5 - 90}{4.7434}) = P(Z \geq -0.738) \end{aligned}$$

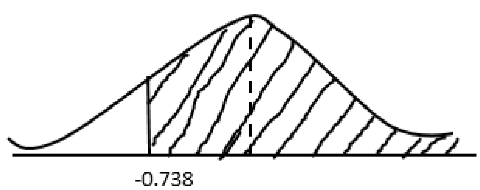


Figure 8.52

$\Longrightarrow P(X > 86) = 0.5 + P(0 \leq Z \leq 0.738) = 0.5 + 0.26974 = 0.76974(cal)$

(iii)

$$\begin{aligned} P(X \geq 98) &= P(X \geq 97.5) \\ &= P(Z \geq \frac{97.5 - 90}{4.7434}) = P(Z \geq 1.581) \end{aligned}$$

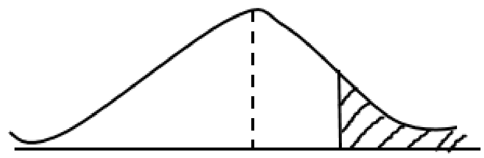


Figure 8.53

$$\implies P(X \geq 98) = 0.5 - P(0 \leq Z \leq 1.581) = 0.5 - 0.44306 = 0.05694(cal)$$

(iv)

$$\begin{aligned} P(X \leq 82) &= P(X \leq 82.5) \\ &= P(Z \leq \frac{82.5 - 90}{4.7434}) = P(Z \leq -1.581) \end{aligned}$$

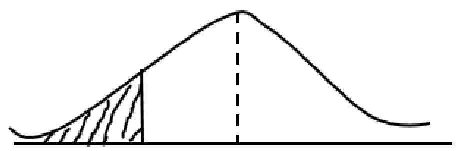


Figure 8.54

$$\implies P(X \leq 82) = 0.5 - P(0 \leq Z \leq 1.581) = 0.5 - 0.44306 = 0.05694(cal)$$

(v)

$$\begin{aligned} P(X = 100) &= P(99.5 \leq X \leq 100.5) \\ &= P(\frac{99.5 - 90}{4.7434} \leq Z \leq \frac{100.5 - 90}{4.7434}) \\ &= P(2.003 \leq Z \leq 2.214) \end{aligned}$$

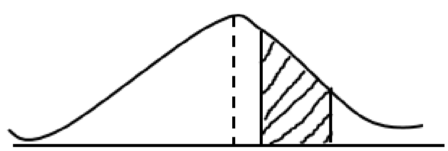


Figure 8.55

$$\begin{aligned} \implies P(X = 100) &= P(0 \leq Z \leq 2.214) - P(0 \leq Z \leq 2.003) \\ &= 0.48659 - 0.47741 \\ &= 0.00918(cal) \end{aligned}$$

(vi)

$$\begin{aligned} P(81 \le X \le 87) &= P(80.5 \le X \le 87.5) \\ &= P(\frac{80.5-90}{4.7434} \le Z \le \frac{87.5-90}{4.7434}) \\ &= P(-2.003 \le Z \le -0.527) \end{aligned}$$

Figure 8.56

$$\begin{aligned} \Longrightarrow P(81 \le X < 87) &= P(-2.003 \le Z \le -0.527) \\ &= P(0 \le Z \le 2.003) - P(0 \le Z \le 0.527) \\ &= 0.47741 - 0.2009 \\ &= 0.27651(cal). \end{aligned}$$

(vii)

$$\begin{aligned} P(92 < X < 95) &= P(93 \le X \le 94) \qquad = P(92.5 \le X \le 94.5) \\ &= P(\frac{92.5-90}{4.7434} \le Z \le \frac{94.5-90}{4.7434}) \\ &= P(0.527 \le Z \le 0.949) \end{aligned}$$

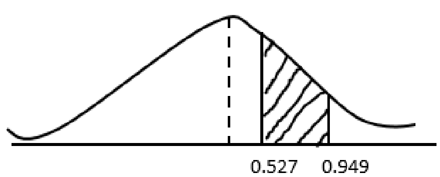


Figure 8.57

$$\begin{aligned} \Longrightarrow P(92 < X < 95) &= P(0.527 \le Z \le 0.949) \\ &= P(0 \le Z \le 0.949) - P(0 \le Z \le 0.527) \\ &= 0.32869 - 0.2009 \\ &= 0.12779(cal). \end{aligned}$$

(viii)

$$
\begin{aligned}
P(84 < X \leq 94) &= P(85 \leq X \leq 94) \qquad\qquad = P(84.5 \leq X \leq 94.5) \\
&= P(\frac{84.5 - 90}{4.7434} \leq Z \leq \frac{94.5 - 90}{4.7434}) \\
&= P(-1.160 \leq Z \leq 0.949)
\end{aligned}
$$

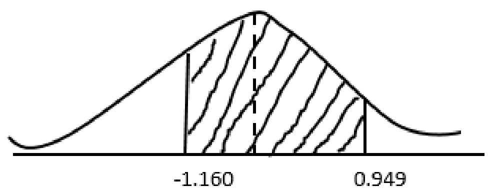


Figure 8.58

$$
\begin{aligned}
\Longrightarrow P(84 < X \leq 94) &= P(-1.160 \leq Z \leq 0.949) \\
&= P(0 \leq Z \leq 0.949) + P(0 \leq Z \leq 1.160) \\
&= 0.32869 + 0.37698 \\
&= 0.70567(cal).
\end{aligned}
$$

(3) Of the students in a school, 10% travel to school by taxi. If 400 students are chosen at random, find the probability that;

(i) more than 50 students travel by taxi,

(ii) at most 54 students travel by taxi,

(iii) exactly 38 students travel by taxi,

(iv) from 35 but less than 47 students travel by taxi.

Solution
Let $X \sim B(n,p), p = 0.1$, and $q = 0.9$ such that

$\mu = np = 400 \times 0.1 = 40$ and

$\sigma^2 = npq = 400 \times 0.1 \times 0.9 = 36 \Longleftrightarrow \sigma = 6$

Please complete the question.
Ans: 0.0401, 0.9921, 0.0629 and **0.6810**

8.1.9 Estimation

Introduction:

Statistical estimation is a statistical procedure used to describe the unknown parameters of the population using the sample characteristics.
A sample is the representation of a population.
Population **parameters** refers to constants that can be used to describe the population. E.g population mean μ, population variance σ^2 and others.
Sample **characteristics** refers to constants that can be used to describe the sample. E.g sample mean $\overline{X}$, sample variance $\hat{\sigma}^2$ and others.
Types of Estimation. There are two types estimations i.e
(a) Point estimations and
(b) Interval estimations.

(a) Point estimation.

This refers to a single point estimation of the parameters.
If a random sample of size n is taken from the large population,then

1. The unbiased estimate of population mean μ is $\hat{\mu}$ such that
   $\hat{\mu} = \overline{x} = \dfrac{\sum x}{n}$ where $\overline{x}$ = sample mean.
2. The unbiased estimate of population variance σ is $\hat{\sigma}^2$ such that
   $\hat{\sigma}^2 = \dfrac{n}{n-1} \times s^2$ where s^2 = sample variance.
   Since s^2 can be expressed in two ways, then
   When $s^2 = \dfrac{\sum(x-\overline{x})}{n}$,
   $\Longrightarrow \hat{\sigma}^2 = \dfrac{n}{n-1} \times s^2 = \dfrac{n}{n-1} \times \dfrac{\sum(x-\overline{x})}{n} = \dfrac{\sum(x-\overline{x})}{n-1}$
   While When $s^2 = \dfrac{\sum x^2}{n} - \left(\dfrac{\sum x}{n}\right)^2$,
   $\Longrightarrow \hat{\sigma}^2 = \dfrac{n}{n-1} \times s^2 = \dfrac{n}{n-1} \times \left(\dfrac{\sum x^2}{n} - \left(\dfrac{\sum x}{n}\right)^2\right) = \dfrac{1}{n-1} \times \left(\sum x^2 - \dfrac{(\sum x)^2}{n}\right)$.

Example.
Given the random sample below,
(a) $46, 48, 51, 50, 45, 53, 50, 48.$
(b) $1.684, 1.691, 1.687, 1.688, 1.688, 1.690, 1.693, 1.685, 1.689$
(c)

x	20	21	22	23	24	25
f	4	14	17	26	20	9

(d) $\sum x = 100, \sum X^2 = 1028, n = 10$
(e) $\sum fx = 738, \sum fX^2 = 16526, \sum f = 50$
Calculate the unbiased estimate for
(i) Population mean,

(ii) Population variance.

Solution.

(a) Using

x	f	fx	fx^2
45	1	45	2025
46	1	46	2116
48	2	96	4608
50	2	100	5000
51	1	51	2601
53	1	53	2809
	$\sum f = 8$	$\sum fx = 391$	$\sum fx^2 = 19159$

(i) Let $\hat{\mu} = \overline{x} = \dfrac{\sum fx}{n} = \dfrac{391}{8} = 48.875$

(ii)

$$\begin{aligned}
\text{Unbiased estimate of population variance } (\hat{\sigma}^2) &= \frac{n}{n-1} \times s^2 \\
&= \frac{n}{n-1} \times \left(\frac{\sum fx^2}{n} - \left(\frac{\sum fx}{n}\right)^2\right) \\
&= \frac{8}{8-1} \times \left(\frac{19159}{8} - \left(\frac{391}{8}\right)^2\right) \\
&= 6.982
\end{aligned}$$

(b)Using

x	f	fx	fx^2
1.684	1	1.684	2.835856
1.685	1	1.685	2.839225
1.687	1	1.687	2.845969
1.688	2	3.376	5.698688
1.689	1	1.689	2.852721
1.690	1	1.690	2.856100
1.691	1	1.691	2.859481
1.693	1	1.693	2.866249
	$\sum f = 9$	$\sum fx = 15.195$	$\sum fx^2 = 25.654289$

(i) Let $\hat{\mu} = \overline{x} = \dfrac{\sum fx}{n} = \dfrac{15.195}{9} = 1.688$

(ii)

$$\begin{aligned}
\text{Unbiased estimate of population variance } (\hat{\sigma}^2) &= \frac{n}{n-1} \times s^2 \\
&= \frac{n}{n-1} \times \left(\frac{\sum fx^2}{n} - \left(\frac{\sum fx}{n}\right)^2\right) \\
&= \frac{8}{8-1} \times \left(\frac{25.654289}{9} - \left(\frac{15.195}{9}\right)^2\right) \\
&= 8 \times 10^{-6} \\
&\approx 0
\end{aligned}$$

(c) Using

x	f	fx	fx^2
20	4	80	1600
21	14	294	6174
22	17	374	8228
23	26	598	13754
24	20	480	11520
25	9	225	5625
	$\sum f = 90$	$\sum fx = 2051$	$\sum fx^2 = 46901$

(i) Let $\hat{\mu} = \overline{x} = \dfrac{\sum fx}{n} = \dfrac{2051}{90} = 22.789$

(ii)

$$\begin{aligned}\text{Unbiased estimate of population variance } (\hat{\sigma}^2) &= \frac{n}{n-1} \times s^2 \\ &= \frac{n}{n-1} \times \left(\frac{\sum fx^2}{n} - \left(\frac{\sum fx}{n}\right)^2\right) \\ &= \frac{90}{90-1} \times \left(\frac{46901}{90} - \left(\frac{2051}{90}\right)^2\right) \\ &= 1.809\end{aligned}$$

(d) $\sum x = 100, \sum X^2 = 1028, n = 10$

(i) $\hat{\mu} = \overline{x} = \dfrac{\sum x}{n} = \dfrac{100}{10} = 10.000$

(ii)

$$\begin{aligned}\text{Unbiased estimate of population variance } (\hat{\sigma}^2) &= \frac{n}{n-1} \times \left(\frac{\sum x^2}{n} - \left(\frac{\sum x}{n}\right)^2\right) \\ &= \frac{10}{10-1} \times \left(\frac{1028}{10} - \left(\frac{100}{10}\right)^2\right) \\ &= 3.111\end{aligned}$$

(d) $\sum fx = 738, \sum fX^2 = 16526, \sum f = 50$

(i) $\hat{\mu} = \overline{x} = \dfrac{\sum fx}{n} = \dfrac{738}{50} = 14.760$

(ii)

$$\begin{aligned}\text{Unbiased estimate of population variance } (\hat{\sigma}^2) &= \frac{n}{n-1} \times \left(\frac{\sum fx^2}{n} - \left(\frac{\sum fx}{n}\right)^2\right) \\ &= \frac{10}{10-1} \times \left(\frac{16526}{50} - \left(\frac{738}{50}\right)^2\right) \\ &= 114.961\end{aligned}$$

(b) Interval estimation.

This involves estimating the population parameter by using interval of sample characteristics known as a confidence interval.
If the population parameter lies within $[a, b]$, then a and b are the confidence limits where a

is the lower confidence limit and b is the upper confidence limit.
NB: The interval has an associated confidence level that the true parameter is in the proposed range.

8.1.10 The Confidence interval / limit

If any parameter w lies within a and b such that $a \leq w \leq b = [a, b]$ is referred to as the confidence interval for w, and $a < w < b = (a, b)$ is referred to as the confidence limit for w. The probability that the true population parameter will lie within the stated interval is denoted by $(1 - \alpha)$ and this is referred to as the confidence coefficient or the degree of the confidence.

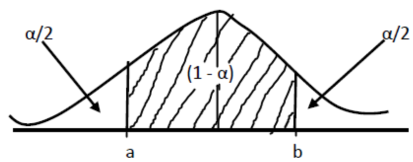


Figure 8.59

NB: $(1-\alpha)100$ is called the confidence interval. From the diagram it can be concluded that $(1 - \alpha)100 =$**Required confidence interval**
NB: The wider the confidence interval, the more confident we can be that the given interval contains the unknown parameter.
Under this text, we are only going to discuss the confidence interval/limits for **population mean.**μ Before constructing your interval for μ, it is essential to ask the following questions:

- Is the distribution of the population normal or not?
- Do you know the population variance σ^2?
- Is the sample size large or small?

NB: The above questions are referred to as the cases considered when obtaining the confidence interval/limits.

Confidence Interval for Population Mean μ of a Normal Population with known Population Variance σ^2

Under this case even if the sample size n is big or small, the sample mean $\overline{X}$ will be used as a point estimate of the population mean, μ
From sampling distribution theory,
$s^2 = \dfrac{\sigma^2}{n} \Longleftrightarrow s = \dfrac{\sigma}{\sqrt{n}}$

By standardizing the distribution, we have $Z = \dfrac{\overline{X} - \mu}{\frac{\sigma}{\sqrt{n}}}$ under the standard normal curve leaving an area $\dfrac{\sigma}{2}$ to the right of the standard normal curve.

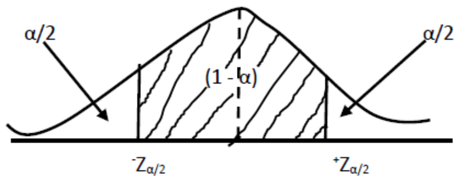


Figure 8.60

From the diagram,

$$P(-Z_{\alpha/2} < Z < Z_{\alpha/2}) = 1 - \alpha$$

$$textBut\ Z = \frac{\overline{X} - \mu}{\frac{\sigma}{\sqrt{n}}}$$

$$P\Big(-Z_{\alpha/2} < \frac{\overline{X} - \mu}{\frac{\sigma}{\sqrt{n}}} < Z_{\alpha/2}\Big) = 1 - \alpha$$

$$P\Big(-Z_{\alpha/2}.\frac{\sigma}{\sqrt{n}} < (\overline{X} - \mu) < Z_{\alpha/2}.\frac{\sigma}{\sqrt{n}})\Big) = 1 - \alpha$$

$$\implies P\Big(-\overline{X} - Z_{\alpha/2}.\frac{\sigma}{\sqrt{n}} < -\mu < Z_{\alpha/2}.\frac{\sigma}{\sqrt{n}} + \overline{X}\Big) = 1 - \alpha$$

Multiply both sides of the inquality -1

$$P\Big(\overline{X} - Z_{\alpha/2}.\frac{\sigma}{\sqrt{n}} < \mu < \overline{X} + Z_{\alpha/2}.\frac{\sigma}{\sqrt{n}}\Big) = 1 - \alpha$$

Hence if $\overline{X}$ is the mean of a random sample of any size n taken from the normal population with known σ^2, then the $(1-\alpha)100$ confidence interval for μ is given by
$\Big[\overline{X} - Z_{\alpha/2}.\frac{\sigma}{\sqrt{n}}, \overline{X} + Z_{\alpha/2}.\frac{\sigma}{\sqrt{n}}\Big]$
Or
$\mu = \overline{X} \pm +Z_{\alpha/2}.\frac{\sigma}{\sqrt{n}}$
NB: $\pm Z_{\alpha/2}\frac{\sigma}{\sqrt{n}}$ is an error (E) and the width $2Z_{\alpha/2}\frac{\sigma}{\sqrt{n}}$

Confidence Interval for Population Mean μ of a Normal Population with unknown Population Variance σ^2

Under this case we are going to consider only a large sample size n .The sample is said to be large if $n \geq 30$.
Here we obtain the confidence for the population mean using $\hat{\sigma}^2$ instead of σ^2 such that $\mu = \overline{X} \pm Z_{\alpha/2}.\frac{\hat{\sigma}}{\sqrt{n}}$ or $\Big[\overline{X} - Z_{\alpha/2}.\frac{\hat{\sigma}}{\sqrt{n}}, \overline{X} + Z_{\alpha/2}.\frac{\hat{\sigma}}{\sqrt{n}}\Big]$
Remember that $\hat{\sigma} = \dfrac{n}{n-1}s^2$

Examples.

(1) The marks obtained are normally distributed with standard deviation of 8.2% marks. The random sample of 30 students had an average mark of 40% marks. find

(i) 95%,

(ii) 90%,

(iii) 99%,

(iv) 89% confidence interval for mean.

Solution

(i) Using $\mu = \overline{X} \pm Z_{\alpha/2}.\frac{\sigma}{\sqrt{n}}, \ \overline{X} = 40, \ \sigma = 8.2, \ n = 30$

For $Z_{\alpha/2}$, from $(1-\alpha)100 = 95$ Making α the subject, then

$\alpha = 1 - 0.95 \Longleftrightarrow \alpha = 0.05$

$\frac{\alpha}{2} = \frac{0.05}{2} \Longleftrightarrow \alpha/2 = 0.25$

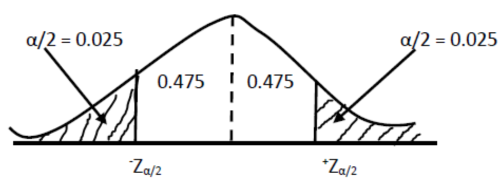


Figure 8.61

p	q	z
0.475	0.025	1.960

$$\Longrightarrow \mu = 40 \ \pm \ 1.960 \times \frac{8.2}{\sqrt{30}}$$
$$= 40 \ \pm 2.9343$$
$$\text{Lower limit} = 40 \ - 2.9343 = 37.0657$$
$$\text{Upper limit} = 40 \ + 2.9343 = 42.9343$$
$$\therefore 37.0657 \leq \mu \leq 42.9343 \text{ or } [37.0657, 42.9343]$$

(ii) Using $\mu = \overline{X} \pm Z_{\alpha/2}.\frac{\sigma}{\sqrt{n}}, \ \overline{X} = 40, \ \sigma = 8.2, \ n = 30$

For $Z_{\alpha/2}$, from $(1-\alpha)100 = 90$ Making α the subject, then

$\alpha = 1 - 0.90 \Longleftrightarrow \alpha = 0.1$

$\frac{\alpha}{2} = \frac{0.1}{2} \Longleftrightarrow \alpha/2 = 0.05$

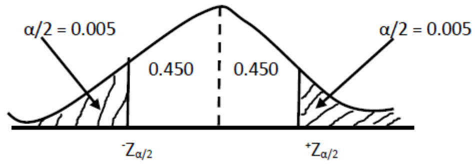


Figure 8.62

p	q	z
0.450	0.005	1.645

$$\begin{aligned}
\Longrightarrow \mu &= 40 \ \pm \ 1.645 \times \frac{8.2}{\sqrt{30}} \\
&= 40 \ \pm 2.4627 \\
\text{Lower limit} &= 40 \ - 2.4627 = 37.5373 \\
\text{Upper limit} &= 40 \ + 2.4627 = 42.4627
\end{aligned}$$

$$\therefore 37.5373 \leq \mu \leq 42.4627 \text{ or } [37.5373, 42.4627]$$

Please complete the remaining parts.

(2) The chance that a cow recovers from a certain mouth disease when treated is 0.72 . If 100 cows are treated by the same vaccine, find the 95% confidence limits for the mean number of cows that recover.

Solution

$p = 0.72, q = 0.28, n = 100$

$\overline{\mathrm{X}} = np = 100 \times 0.72 = 72$

Sample variance $\sigma^2 = npq = 100 \times 0.72 \times 0.28 = 20.16$

$\hat{\sigma}^2 = \frac{n}{n-1}(\sigma^2) = \frac{100}{99}(20.16) = 20.3636$ (I.e estimating the variance for the whole population

For $Z_{\alpha/2}$, from $(1-\alpha)100 = 95$ Making α the subject, then

$\alpha = 1 - 0.95 \Longleftrightarrow \alpha = 0.1$

$\dfrac{\alpha}{2} = \dfrac{0.05}{2} \Longleftrightarrow \alpha/2 = 0.025$

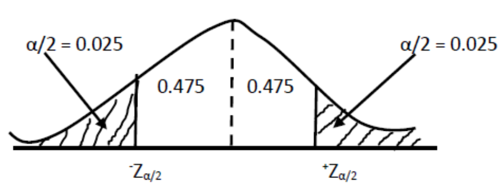


Figure 8.63

p	q	z
0.475	0.025	1.960

$$\Longrightarrow \mu = 72 \ \pm \ 1.960 \times \frac{\sqrt{20.3636}}{\sqrt{100}}$$
$$= 72 \ \pm 0.8845$$
$$\text{Lower limit} = 72 \ - 0.8845 = 71.1155$$
$$\text{Upper limit} = 72 \ + 0.8845 = 72.8845$$

(3) To determine the mean monthly leisure time of Manchester united fans, the sports analyst decide to study a sample of Man-U fans from Uganda. The mean monthly leisure time is to be estimated within 1.22hours of the true mean monthly time with 98% confidence level. Its assumed that the leisure time is of the Man-U fans is normally distributed with standard deviation 6.5 hours. Find how many people must be chosen for the study.

Solution

$E = Z_{\alpha/2}.\frac{\sigma}{\sqrt{n}}, E = 1.22, \sigma = 6.5$

For $Z_{\alpha/2}$, from $(1-\alpha)100 = 98$ Making α the subject, then

$\alpha = 1 - 0.98 \Longleftrightarrow \alpha = 0.1$

$\frac{\alpha}{2} = \frac{0.02}{2} \Longleftrightarrow \alpha/2 = 0.01$

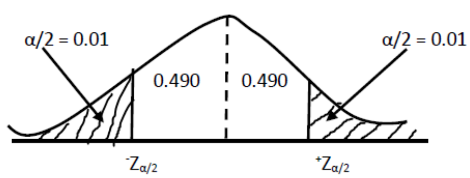


Figure 8.64

p	q	z
0.490	0.01	2.326

$\Longrightarrow 1.22 = 2.326 \times \frac{6.5}{\sqrt{n}}$

$\sqrt{n} = 12.3926$

$n = 153.5765 \approx 154$

154 people must be taken for the study.

(4) The height x cm of each man in a random sample of 200 men living in Mbale district was measured and results are as follows $\sum x = 35,000$ and $\sum x^2 = 6,200,000$.
(a) Calculate the unbiased estimate of mean and variance of the heights of all men living in Mbale district,
(ii) Determine 56% confidence interval of the heights of all men living in Mbale district.

Solution

$\sum x = 35,000, \sum X^2 = 6,200,000, \sum n = 200$

(a) $\hat{\mu} = \overline{x} = \dfrac{\sum x}{n} = \dfrac{35,000}{200} = 175.0$

$$\text{Unbiased estimate of population variance } (\hat{\sigma}^2) = \frac{n}{n-1} \times \left(\frac{\sum x^2}{n} - \left(\frac{\sum x}{n}\right)^2\right)$$
$$= \frac{200}{200-1} \times \left(\frac{6,200,000}{200} - \left(\frac{35,000}{200}\right)^2\right)$$
$$= 376.884$$

(b) Using $\mu = \overline{X} \pm Z_{\alpha/2}.\frac{\hat{\sigma}}{\sqrt{n}}$, $\overline{X} = 175.0$, $\hat{\sigma} = \sqrt{376.884}$, $n = 200$

For $Z_{\alpha/2}$, from $(1-\alpha)100 = 56$ Making α the subject, then

$\alpha = 1 - 0.56 \Longleftrightarrow \alpha = 0.1$

$\dfrac{\alpha}{2} = \dfrac{0.44}{2} \Longleftrightarrow \alpha/2 = 0.22$

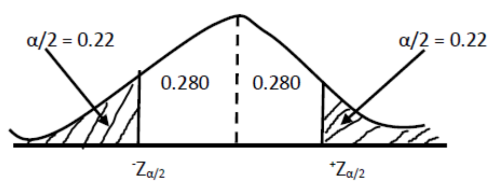


Figure 8.65

p	q	z
0.280	0.220	0.772

$$\Longrightarrow \mu = 175.0 \ \pm \ 0.772 \times \frac{\sqrt{376.884}}{\sqrt{30}}$$
$$= 175.0 \ \pm 1.05976$$
$$\text{Lower limit} = 175.0 \ - 1.05976 = 173.94024$$
$$\text{Upper limit} = 175.0 \ + 1.05976 = 176.05976$$
$$\therefore 173.94024 \leq \mu \leq 176.05976 \text{ or } [173.94024, 176.05976]$$

(5) The table below shows the marks obtained in a biology test by 50 students.

Marks	0 – 9	10 – 19	20 – 29	30 – 39	40 – 49	50 – 59	60 – 69
Number of students	4	10	5	15	8	2	6

(a) Calculate the mean
(b) Obtain the unbiased estimate for the variance.
(c) Find the 87.5% confidence limit for the mean

Solution

Class boundaries	f	X	fX	fX^2
-0.5 – 9.5	4	4.5	18	81
9.5 – 19.5	10	14.5	145	2102.5
19.5 – 29.5	5	24.5	122.5	3001.25
29.5 – 39.5	15	34.5	517.5	17853.75
39.5 – 49.5	8	44.5	356	15842
49.5 – 59.5	2	54.5	109	5940.5
59.5 – 69.5	6	64.5	387	24961.5
	$\sum f = 50$		$\sum fX = 1655$	$\sum fX^2 = 69782.5$

(a) mean$(\overline{\mathrm{X}})$ = $\frac{\sum f\mathrm{X}}{\sum f} = \frac{1655}{50} = 33.100$

(b) Variance , $\sigma^2 = \frac{\sum f\mathrm{X}^2}{\sum f} - (\overline{\mathrm{X}})^2 = \frac{69782.5}{50} - (33.1)^2 = 300.04$

$\hat{\sigma}^2 = \frac{n}{n-1}(\sigma^2) = \frac{50}{49}(300.04) = 306.1633$

(c) Using $\mu = \overline{X} \pm Z_{\alpha/2}.\frac{\hat{\sigma}}{\sqrt{n}}$, $\overline{X} = 33.100$, $\hat{\sigma} = \sqrt{306.1633}$, $n = 50$

For $Z_{\alpha/2}$, from $(1-\alpha)100 = 87.5$ Making α the subject, then

$\alpha = 1 - 0.875 \Longleftrightarrow \alpha = 0.1$

$\frac{\alpha}{2} = \frac{0.125}{2} \Longleftrightarrow \alpha/2 = 0.0.0625$

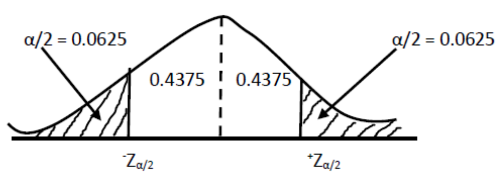


Figure 8.66

p	q	z
0.4375	0.0625	1.534

$$\begin{aligned}
\Longrightarrow \mu &= 33.100 \ \pm \ 0.772 \times \frac{\sqrt{306.1633}}{\sqrt{50}} \\
&= 33.100 \ \pm 3.7959 \\
\text{Lower limit} &= 33.100 \ - 3.7959 = 29.3041 \\
\text{Upper limit} &= 33.100 \ + 3.7959 = 36.8959
\end{aligned}$$

8.1.11 Exercise 8

(1) Given that $X \sim N(400, 64)$, find:

(i) the limits within which the central 95% of the distribution lies.

(ii) the interqurtile range of the distribution.

(ii) $P(X > 390)$.

(iv) $P(X \leq 443)$.

Ans: (i) 384.32, 415.68, (ii) 10.784

(2) The height of boys at $KAPROSS$ at a particular age follow a normal distribution with mean $150.3cm$ and variance $25cm$. Find the probability that a boy picked at random from this age group has height:

(a) less than $153cm$,

(b) more than $158cm$,

(c) between $150cm$ and $158cm$,

(d) More than $10cm$ difference from the mean height,

(e) at least $140cm$.

Ans: (a) 0.7054, (b) 0.0618, (c) 0.4621, (d) 0.00456

(3) Given that $X \sim N(100, 25)$, find:

(a) $P(X > 108)$

(b) show that $P(|X - 100|) < 6.80 = 0.8262$

(c) $P(X > 91)$.

(d) $P(X \leq 110)$.

Ans: (a) 0.0548,

(4) Given that $X \sim N(-3, 49)$, find:

(i) the probability of obtaining a positive value,

(ii) $P(X > -5)$,

(iii) $P(X \leq -1)$,

(iv) $P(X \leq -6)$,

(v) $P(|X + 2| < 3.50)$,

(vi) $P(|X + 1| > 3.20)$,

(5) The time taken by the president to brief his press secretary is normally distributed with mean 12 minutes and standard deviation 2 minutes. The president briefs the secretary every day. Estimate the number of days during the year when he takes

(i) longer than 17 minutes

(ii) less than 10 minutes

(iii) between 9 and 13 minutes

Ans; (i) 0.0062 (ii) 0.1587 (iii) 0.6247

(6) A company packs salt in bags which are normally distributed with mean weight 50kg and variance 9. A random sample of 20 bags was taken.

(i) Find the probability that the sample mean lies between 49 and 50.5kg.

(ii) Find the probability that the sample mean is more that 52.35kg.

(iii) Find the probability that the sample mean is less 53.5kg.

Ans; 0.7039

(7) 40% of the students in a certain school are from Northern Uganda. If 60 students are randomly chosen, find the probability that:

(i) more than half of them are from the north.

(ii) at least 26 of them are from the north.

(iii) at most 28 of them are from the north.

(iv) between 20 to 25 of them are from the north.

Ans; 0.0433

(8) The 400 marks obtained by students in a mathematics paper were normally distributed with mean 50 and standard deviation 16. Given that 8.5% failed the paper and 12.5% got distinctions. Calculate the (a)

(i) pass mark, (ii) the lowest mark for a distinction,

(b) If 5.3% of the students obtain a distribution by scoring x_1 marks or more, estimate the value of x_1.

Ans; (i) 28.048 (ii) 68.176

(9) Given that $X \sim N(\mu, \sigma^2)$ such that $P(X < 10) = 0.0578$ and $P(X > 30) = 0.02252$

(a) Find the value of μ and σ,

(b) find interqurtile range,

(c) $P(X > 20$.

Ans:

(10) Given that $X \sim N(\mu, \sigma^2)$ such that $P(X > 4.00) = 0.35$ and $P(X > 4.53) = 0.2291$

(a) Find the value of μ and σ.

(a) $P(|X - 3| < 1$.

Ans:

(11) A random variable X is normally distributed such that $X \sim N(\mu, \sigma^2)$. Given that $P(X < 28) = 0.9846$, and $P(X > 25) = 0.9429$. Find:

(a) values of μ and σ

(b) $P(X > 23)$

(c) Find X_0 such that $P(X < X_0) = 0.9$

(d) Find X_1 such that $P(X > X_1) = 0.6$

(12) A random variable X is normally distributed such that X $\sim N(\mu, \sigma^2)$. Given that $P(X < 55) = 0.25$, and $P(55 < x < 84) = 0.65$. Find:

(a) Values of μ and σ

(b) $P(X > 92)$

(13) The height of the female students at a particular collage are normally distributed with a mean of $169cm$ and a standard deviation $9cm$.

(i) Given that 88% of these female students have a height less than bcm, find the value of b.

(ii) Given that 58% of these female students have a height less than scm, find the value of s.

(14) The masses of boxes 0f oranges are normally distributed such that 30% of them are greater than $4.00kg$ and 20% are greater than $4.53kg$. Estimate the mean and standard deviation.
Ans: $\mu =$**3.13,** $\sigma = 1.67$

(15) The weight X in grams of chocolate chips in packet sold by a certain supermarket is given by the distribution $X \sim N(80, 21)$

(a) Find:

(i) $P(X > 70)$

(ii) $P(|X - 79| < 6.8)$

(b) Four packets are selected at random from the packets chocolate chips on the supermarket sheet, find the probability that exactly three of them will have weights in the range $|X - 75| < 9$
Ans; (a)(i) (ii) (b)

(16) The speed of cars on Masaka - Kampala high way have a normal distribution. If 10% of the cars are traveling faster than $120kmh^{-1}$ while 30% of the slower than $85kmh^{-1}$. Find the:

(i) mean speed and standard deviation,

(ii) quartile deviation of the speeds,

(iii) maximum safe speed if 6% of the cars were stopped by police for speeding,

(iv) limits within which central 70% of the speed lies.

Ans; (i) $\mu =$ **95.1550,** $\sigma = 19.3798$ **(ii)** $26.124kmh^{-1}$ **(iii) 125.2906**kmh^{-1} **(iv) 75.0775, 115.2325**

(17) A dog can jump and clear a height of $1.88m$ four in five attempts and a height of $1.05m$ one time out of ten attempts. Given that the height cleared are normally distributed,

(i) Determine the mean (μ) and standard deviation (σ) of the height.

(ii) The percentage of dogs clearing the height greater than $1.7m$

(iii) Assuming the number of cats that follow distribution are 12, calculate the expected number of dogs that can fail to clear $1.70m$

(18) The random variable X is $N(\mu, \sigma^2).P(X < 40) = 0.3$ and $P(40 < X < 70) = 0.55$. find μ and σ. Hence, determine percentage of obtaining less than 80%.
Ans: 50.0769, 19.2308, 94.015.

(19) The drying time of a newly manufactured paint is normally distributed with mean 110.5 minutes and standard deviation 12 minutes.

(a) Find the probability that the paint dries for less than 104 minutes.

(b) If a random sample of 20 tins of the paint was taken, find the probability that the mean drying time of the sample is more than 112 minutes.

(c) A random sample of 16 tins taken from a different type of paint of standard deviation 15 minutes is found to have a mean time of 105.5 minutes, determine the 90% confidence limits the mean time of this type of paint.
Ans; (a) 0.2939 (b) 0.2881 (c) (99.3313, 111.6688)

(20) The random variable X is distributed $B(200, 0.7)$. find;

(i) $P(X \geq 130)$,

(ii) $P(136 \leq X < 148)$,

(iii) $P(X < 142)$,

(iv) $P(X = 152)$
Ans: (a) 0.9474 (b) 0.8345 (c) 0.5914 (d) 0.0111

(21) An experiment consists of removing 2 pens. One at a time with replacement from a box containing 3 red and 4 blue pens. If A is the event that both pens are of the same colour.

(i) Find $P(A)$,

(ii) If the experiment is repeated 70 times, find the probability that event A occurred at least 25 times.
Ans; (i) 25/49 (ii) 0.9936

(22) The germination time for a certain type of seeds is known to be normally distributed. If for a batch of 10,000 seeds, 1643 take more than 6 days to germinate and 1112 take less than 4 days to germinate.

(ii) Determine the mean and variance of the germination time.

(ii) Find the 95% confidence limits of the germination time.
Ans; (i) $\mu = 5.1106$, $\sigma^2 = 0.8286$ (iii) (3.3264, 6.8948)

(23) The masses of soap powder in certain packets are normally distributed with mean 842 grams and variance 225(grams)^2. Find the probability that a random sample of 120 packets has sample mean with mass:

(i) between 844 grams and 846 grams,

(ii) less than 843 grams,

(iii) at least 839 grams,

(iv) at most 840 grams,

(v) more than 847 grams,

(vi) at least 850 grams.

Ans; (i) 0.07062 (ii) 0.7673

(24) A random sample of size 76 electrical components produced by a certain manufacturer have resistances $r_1, r_2, \cdots, r_{76}$ ohms where $\sum r_i = 740$ and $\sum r_i^2 = 8,216$
Calculate the:

(i) Unbiased estimate for the population variance,

(ii) 91.86% confidence interval for the mean resistance of the electrical components produced, (give answers correct to 3 decimal places)

(iii) 56.3% confidence interval for the mean resistance of the electrical components produced.
Ans; (i) $\hat{\sigma}^2$ = 3.6711 (ii) [9.003, 10.470], (iii)

(25) A random sample of 100 observations from a normal population with mean μ gave the following data; $\sum x = 82,00$, and $\sum x^2 = 686,000$.

(i) Find a 98% confidence interval for μ.

(ii) Find a 99% confidence limit for μ.

(iii) Find a 72% confidence limit for μ.
Ans: (i) [79.19, 84.81] ,(ii) 78.89, 85.11

(26) Sixty students sat for a mathematics contest whose pass mark was 40 marks. Their scores in the test were approximately normally distributed. 9 students scored less than 20 marks while 3 scored more than 70 marks. Find the:

(i) mean score and standard deviation of the contest,

(ii) The probability that a student chosen at random passed the contest.
Ans; (i) μ = 39.3208 , σ = 18.6500 (ii) 0.4856

(27) The times a machine takes to print each of the 10 documents were recorded in minutes as given below; 16.5, 18.3, 18.5, 16.6, 19.4, 16.8, 18.6, 16.0, 20.1, 18.2. If the times of printing documents are approximately normally distributed with variance of 2.56 minutes,

(i) find the 80% confidence interval for the mean time of printing the documents,

(ii) find the 59% confidence interval for the mean time of printing the documents,

(iii) find the 88.5% confidence interval for the mean time of printing the documents.
Ans;

(28) Given the sample of the data: 15.7, 18.3, 18.5, 16.6, 19.4, 16.8, 18.6, 16.0, 20.1, 18.2.

(a) find the 76% confidence interval for μ,

(b) find the 89% confidence interval for μ,

(c) find the 98.5% confidence interval for μ.

Ans;

(29) Metal rods produced by a machine have lengths that are normally distributed. 2% of the rods are rejected as being too short and 5% are rejected as being too long. Given that the least and greatest acceptable lengths of the rods are $6.32cm$ and $7.52cm$.

(a) calculate the mean and variance of the lengths of the rods.

(b) If the rods are chosen at random from a batch produced by the machine. Find the probability that exactly 3 of them are rejected as being too short.

Ans; (a) $\mu = $ **6.9863 ,** $\sigma^2 = $ **0.1052 (b) 0.00083**

(30) Cartons of oranges are filled by a machine. A sample of 10 cartons selected at random from the production contained the following quantities (in milliliters)
201.2, 205.0, 209.1, 202.3, 204.6, 206.4, 210.1, 201.9, 203.7, 207.3
(i) Calculate the unbiased estimate of the mean and variance of the population from which the sample was taken.
(ii) Obtain the 98% confidence interval for the population mean.
Ans; $\overline{x} = $ **205.16,** $\hat{\sigma} = $ **3.0369 (ii) [202.9261, 207.3939]**

(31) At a particular hospital, records show that each day, on average, only 80% of people keep their appointments at the out patient's clinic. Find the probability that on a day when 200 appointments have been booked,
(i) More than 170 patients keep their appointments
(ii) at least 155 patients keep their appointments.
Ans; (i) 0.0318 (ii) 0.8345

(32) The probability that a man aims and hits a target with a single shot is 0.4. If he is given 25 bullets, estimate the probability that he hits the target with;

(i) Exactly 8 shots

(ii) Between 9 and 15 shots

(iii) atleast 12 shots

Ans; 0.11658 (ii) 0.547723 (iii) 0.27027

(33) A secondary school in Masaka district has 700 students whose weights are normally distributed with mean 52kg and standard deviation 5. Find the number of students weighing over 61kg
Ans; 25.13

(34) The mean weight of tilapia in a fish farm is 980g and standard deviation 100g. What is the probability that a catch of 10 tilapia will have a mean weight per fish between

910g and 1050g
Ans; 0.97318

(35) In a certain town, 75% of its people are university graduates, if a random sample of 400 people is selected, determine the probability that between 280 t0 320 people are graduates.
Ans; 0.9790

(36) The life time of batteries for a transitor radio is normally distributed with mean of 160 hours and standard deviation 30 hours. Calculate;
(a) the percentage of batteries which have a life between 150 hours and 180 hours inclusive.
(b) the range, symmetrical about the mean with in which 75% of the battery lives lies.
(c) If a radio takes four of these batteries and requires all of them to be working. Calculate the probability that, the radio will ran for atleast 135 hours
Ans; (a) 37.81% (b) from 125.47 hrs to 194.53 hrs (c)

(37) A machine produces rubber balls whose diameters are normally distributed with mean 5.50cm and standard deviation 0.08cm.
(a) What proportion of balls will have diameters

(i) less than 5.65cm
(ii) between 5.44 and 5.54cm

Ans; (a)(i) 96.96% (ii) 46.48%

(38) The balls are packed in cylindrical tubes whose internal diameters are normally distributed with mean 5.70cm and standard deviation 0.12cm. If a ball, selected at random is placed in a tube, selected at random, what is the distribution of the clearance? (The clearance is the internal diameter of the tube minus the diameter of the ball.) What is the probability that the clearance is between 0.08cm and 0.22cm?
Ans; 0.6902

(39) The development engineer of a company making razors records the time it takes him to shave, on seven mornings, using a standard razor made by the company. The times, in seconds, were 217, 210, 254, 237, 232, 228 and 243. Assuming that this may be regarded as a random sample from a normal distributio. Calculate a 95% confidence interval for the mean.
Ans; [220.4436, 242.6992]

(40) The mean weight of a trout fish in a fish farm is 980g and the standard deviation is 100g. What is the probability that a catch of 10 trout will have a mean

(i) greater than 1050g?
(ii) between 1030g and 1060g

Ans; (i) 0.24196 (ii) 0.09668

(41) The times taken to complete the test are normally distributed with mean 40.5 and standard deviation 75minutes. Applicants who complete the testing less than 30 minutes are immediately accepted for training. Those who take between 30 to 36 minutes are required to take a further test, and other applicants are rejected.
(a) For a randomly chosen applicant, calculate the probability of

(i) immediate acceptance for training (ii) requirement to take a further test

(b) Given that a random chosen applicant was not rejected after this first test, calculate to 3 decimal places the probability that the applicant was immediately accepted for training
(c) On a certain occasion there were 100 applicants. Use a suitable distribution approximation to calculate the probability that more than 25 applicants were required to take a further test. **Ans; (a)(i) 0.0808 (ii) 0.1935 (b) (c) 0.0588**

(42) (a) A certain variety of seeds is solid in packets containing 1000 seeds. It's claimed on the packet that 40% will bloom white and 60% will bloom red. 5 seeds are planted. Find the probability that

(i) exactly 3 will bloom white (ii) at least one will bloom red

(b) 100 seeds are planted. Use the normal approximation to estimate the probability of obtaining between 30 and 45, (inclusive) white flowers.
Ans; (a)(i) 0.2304 (ii) 0.98976 (b) 0.8531

(43) A coin is three times as likely to show a head as it is to show a tail. If it is tossed 48 times, what is the probability that between 30 and 40 heads will show?
Ans; 0.7925

(44) A class of 100 pupils has 60 boys and 40 girls. A pupil is selected at random; find the
(a) 99% confidence interval for mean number of boys in a class.
(b) 86% confidence interval for mean number of girls in a class.
Ans;

(45) A coin is biased such that it is thrice as likely to show head as tail. Find the probability that in 90 tosses of a coin,

(i) exactly 50 heads are obtained,

(ii) between 45 to 70 heads are obtained,

(iii) between 55 and 60 heads inclusive are obtained,

(iv) at least 69 heads are obtained.

Ans; (i) 0.0043

(46) A population of 200 students sat for a maths test for which the pass mark was 60. Their marks were normally distributed, 30 students scored below 40 marks while 50 scored above 50 marks.

(i) If a sample of 100 students is taken, find the mean mark and standard deviation.

(ii)If a student who scores at least 70 marks gets a prize, determine the number of students who got prizes.

(iii) What is the probability that a student chosen at random failed the test?
Ans; (i) $\mu = 46.0585$, $\delta = 58.4795$ (ii) 68.28 students (iii) 0.5941

(47) In a large consignment of tomatoes, 4 in every 10 are rotten. If a random sample of 100 tomatoes is taken, calculate the

(i) mean and variance of the rotten tomatoes

(ii) probability that between 30 and 45 inclusive is rotten
Ans; (i) $\mu = 40$, $\delta^2 = 24$ (ii) 0.1771

(48) Marks of candidates in an examination are normally distributed with mean of 45 marks and a standard deviation of 20 marks.

(i) Find the probability that a candidate selected at random scored; between 41 and 49 marks.

(ii) If 310 candidates passed the examination with a pass mark of 40, estimate the total number of students who sat the examination.

(iii) Find the 50^{th} percentile of the distribution of marks.
Ans; (i) 0.1586 (ii) 518 students (iii) 45

(49) The height of the population in a certain village is approximately normal distributed. Given that 10% have height above $2m$ while 18% of the population have heights below $1.75m$, estimate:

(i) the mean,

(ii) the standard deviation of heights of the whole population,

(iii) semi inter qurtile range,

(iv) probability of having the hieght above $1.9m$.
Ans; (i) $\mu = 1.8541$, $\sigma = 0.1138$ (ii) 68.28 students (iii) 0.0767.

(50) A pair of fair dice is rolled 72 times. What is the probability that a total of 7 occurs between six and twelve times inclusive.
Ans; 0.5428

(51) An extinct species of animals was considered to have an average life of 25 years with standard deviation of 2 years. Of the family of 50 members of the species born on the same time, how many would be expected to have lived (assuming a normal distribution)

(i) between 22 and 27 years

(ii) less than 19 years

(iii) Calculate the age range within which 95% of the family would have been expected to live
Ans; (i) 38.725 animals (ii) 0.065 animals (iii) [24.4456, 25.5544]

(52) The probability that a chameleon changes its colour after walking a distance of 50m is 0.45, if it walks a total distance of 5km. Determine the probability that the chameleon changes its colour:

(i) exactly 40 times, (ii) at most 42 times.

Ans; (i) 0.0484 (ii) 0.3075

(53) The marks in an examination were normally distributed with mean μ and standard deviation δ, 10% of the candidates scored more than 75 marks and 20% scored less than 40 marks.

(i) 25 students 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.
Ans; (i) μ = 53.8748, δ = 16.4783, 0.0316 (iii) 0.4648

(54) A sack contains bean seeds where only 65% are capable of germinating. If a random sample of 400 seeds are drawn from the sack and planted, find the probability that less than 120 seeds will germinate.
Ans; 0.0158

(55) If X $\sim$ N(100, 81), find P($|X - \mu| < 2\delta$) where μ is mean and δ is standard deviation for the population.
Ans; 0.9544

(56) 150 bags of beans are taken from the store and found to have a mean mass of 748kg and standard deviation of 3.6kg.
(i) Calculate the unbiased estimate of the standard deviation of a bag of beans.
(ii) Calculate a 98% confidence interval for the mean mass of a bag of beans at the store.
Ans; (i) 3.6121 (ii) [747.7051, 748.2949]

(57) The admission aggregates for MEG combination in S.5 at Buddo secondary school were normally distributed. If a sample of 36 applicants has a mean of 40 aggregates variance 5.76 aggregates, dtermine the 97.5% confidence limits for the mean admission aggregates for the combination.
Ans; lower limit = 40.909, lower limit = 39.091

(58) A certain point on motor way indicates that the speeds of the cars are normally distributed. 15% of the cars have speeds of less than 35kmh^{-1} , 58% of the buses have speeds 35kmh^{-1} and 45kmh^{-1}.
(a) Find the mean and standard deviation of the speeds of the cars.
(b) Find the probability that the speed of a randomly selected car is more that 12ms^{-1}.
Ans; μ = 41.2827, δ = 6.0640 (b) 0.3761

(59) The chance of a cow being infected on a farm is 2/5 . If there are 150 cows on the farm, find the probability that atmost 60 cows will be infected.
Ans; 0.5331

(60) In Kajjansi town council 70% of the people have been immunized against Hepatitis B. If a sample of 120 people is taken, find the;
(i) 94% confidence interval for mean number of people immunized.
(ii) 98%confidence limits for mean number of people immunized.

Chapter 9

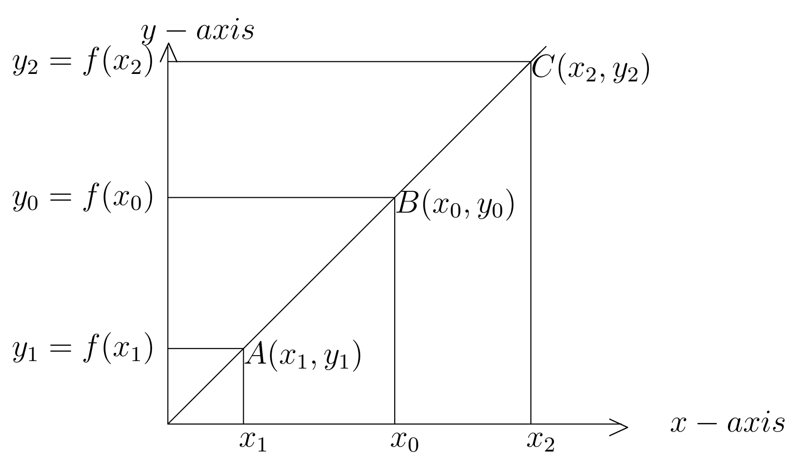
Linear Interpolation and Linear Extrapolation.

9.1 Linear Interpolation

Linear Interpolation: Refers to the process in which the non-tabulated values of the function are estimated on the assumption that the function behaves sufficiently smooth between the tabular points.
We apply the idea of **gradient/slope** to obtain the required value(s).I.e

$$\frac{f(x_0) - f(x_1)}{x_0 - x_1} = \frac{f(x_2) - f(x_1)}{x_2 - x_1}$$

Where we are looking for either $f(x_0)$ or x_0
Proof
Given the illustration below;



Take $A(x_1, y_1)$ and $B(x_0, y_0)$;

$\Longrightarrow$ Gradient $AB = \dfrac{f(x_0) - f(x_1)}{x_0 - x_1}$ .... (i)

Similarly, take $A(x_1, y_1)$ and $C(x_2, y_2)$;

$\Longrightarrow$ Gradient $AB = \dfrac{f(x_2) - f(x_1)}{x_2 - x_1}$ .... (ii)

But for a straight line, gradient is the same and there fore,

$$\text{(i)} = \text{(ii)}$$

$$\Longrightarrow \frac{f(x_0) - f(x_1)}{x_0 - x_1} = \frac{f(x_2) - f(x_1)}{x_2 - x_1} \qquad \square$$

, Hence proved.

Examples.

1. Given the tabular information below;

x	1	2	3
$y = f(x)$	2	8	11

Use linear interpolation to find:

(i) y when $x = 1.4$

(ii) y when $x = 2.2$

(iii) x when $y = 4.0$

(iv) x when $y = 7.7$

Solution.

(i) Extract:

1	1.40	2
2	y	8

, Using $\dfrac{y - 2}{1.40 - 1} = \dfrac{8 - 2}{2 - 1} \Longrightarrow y = 5.40$

(ii) Extract:

2	2.2	3
8	y	11

, Using $\dfrac{y - 8}{2.2 - 2} = \dfrac{11 - 8}{3 - 2} \Longrightarrow y = 8.60$

(iii) Extract:

1	x	2
2	4.0	8

, Using $\dfrac{x - 1}{4.0 - 2} = \dfrac{2 - 1}{8 - 2} \Longrightarrow x = 1.3333$

(iv) Extract:

1	x	2
2	7.7	8

, Using $\dfrac{x - 1}{7.7 - 2} = \dfrac{2 - 1}{8 - 2} \Longrightarrow x = 1.950$

NB: To reduce on time wastage during the calculation, please start with the unknown value in the extracted table.

9.2 Linear Extrapolation

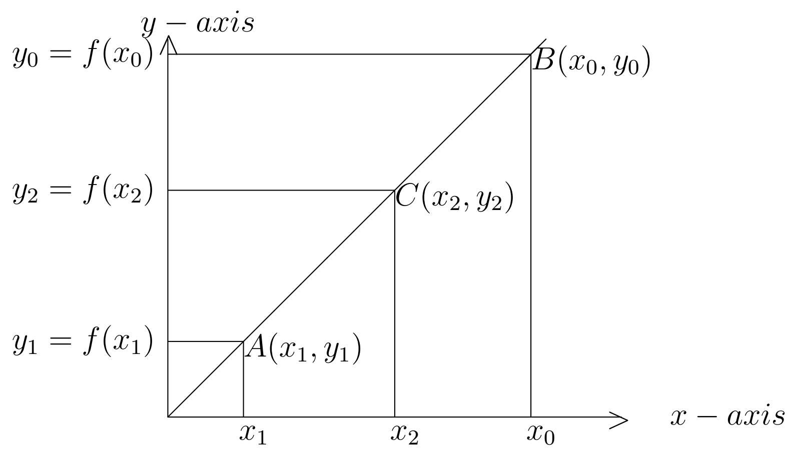
This involves approximating the values lying outside the given data.
Also we apply the idea of gradient to find the unknown values. It's given as;

$$\frac{f(x_0) - f(x_1)}{x_0 - x_1} = \frac{f(x_2) - f(x_1)}{x_2 - x_1}$$

Where we are looking for either $f(x_0)$ or x_0

Proof

Given the illustration below;



Take $A(x_1, y_1)$ and $B(x_0, y_0)$;

$\Longrightarrow$ Gradient $AB = \dfrac{f(x_0) - f(x_1)}{x_0 - x_1}$ .... (i)

Similarly, take $A(x_1, y_1)$ and $C(x_2, y_2)$;

$\Longrightarrow$ Gradient $AB = \dfrac{f(x_2) - f(x_1)}{x_2 - x_1}$ .... (ii)

But for a straight line, gradient is the same and there fore,

$$\text{(i)} = \text{(ii)}$$

$$\Longrightarrow \frac{f(x_0) - f(x_1)}{x_0 - x_1} = \frac{f(x_2) - f(x_1)}{x_2 - x_1} \qquad \square$$

, Hence proved.

Examples.

1. Given the tabular information below;

x^0	10	12	15
$y = Cos(x^0)$	0.9848	0.9781	0.9659

Use linear interpolation/extrapolation to find;

(i) y when $x = 17^0$

(ii) y when $x = 8^0$

(iii) x when $y = 0.9955$

(iv) x when $y = 0.8322$

(v) $\text{Cos}^{-1}(0.9452)$

(vi) $\text{Cos}^{-1}(0.9899)$

(vii) $\text{Cos}13^0$

(viii) $\text{Cos}9.5^0$

Solution.

(i) Extract:

12	15	17
0.9781	0.9659	y

, Using $\dfrac{y - 0.9659}{17 - 15} = \dfrac{0.9659 - 0.9781}{15 - 12} \Longrightarrow y = 0.9578$

(ii) Left for You.
(iii) Extract:

x	10	12
0.9955	0.9848	0.9781

, Using $\dfrac{x-10}{0.9955-0.9848}=\dfrac{10-12}{0.9848-0.9781}\Longrightarrow y=6.806^0$

(iv): Left for You.
(v) Let $x=\text{Cos}^{-1}(0.9452),\Longleftrightarrow \text{Cos}x=0.9452$. Then, we need x. Extract

12	15	x
0.9781	0.9659	0.9452

, Using $\dfrac{x-15}{0.9452-0.9659}=\dfrac{15-12}{0.9659-0.9781}\Longrightarrow y=20.0902^0$

(vi) Left for You.
(vii) Extract:

12	13	15
0.9781	y	0.9659

, Using $\dfrac{y-0.9781}{13-12}=\dfrac{0.9659-0.9781}{15-12}\Longrightarrow y=0.9740$

(viii) Left for You.

2. Given that the function $y=f(x)$ passes through $(-0.10,1.1800)$ and $(-0.09,1.1781)$.
   (a) Find the value of x such that $f(x)=1.1764$.
   (b) Find the value of y such that $f(-0.11)=y$

Solution.
(a) Using;

-0.10	-0.09	x
1.1800	1.1781	1.1764

, Using $\dfrac{x--0.09}{1.1764-1.1781}=\dfrac{-0.09--0.10}{1.1781-1.1800}\Longrightarrow y=-0.0811$

(b) Here, we are given $x=-0.11$. So

-0.11	-0.10	-0.09
y	1.1800	1.1781

, Using $\dfrac{y-1.1800}{-0.11--0.10}=\dfrac{1.1800-1.1781}{-0.10--0.09}\Longrightarrow y=1.1819$

3. In the table below is an extract of part of $y=log_e{}^{x}$ with the given x values.

x	60.00	60.26	60.42	60.78
y	4.0943	4.0987	4.1013	4.1073

Using linear interpolation/Extrapolation, estimate the unknown values below:

(a) $y=log_e{}^{60.54}$
(b) $log_e{}^{60.80}$
(c) $4.4019=log_e{}^{x}$
(d) $4.0091=log_e{}^{q}$

Solution. NB: $log_e{}^{x}$ implies $ln(x)$.
(a) Using;

60.42	60.54	60.78
4.1013	y	4.1073

, $\Longrightarrow \dfrac{y-4.1013}{60.54-60.42}=\dfrac{4.1073-4.1013}{60.78-60.42}\Rightarrow y=4.1033$

(b). Let $k=log_e{}^{60.80}$. Complete this part using the above knowledge.
(c) (a) Using;

60.42	60.78	x
4.1013	4.1073	4.4019

, $\Longrightarrow \dfrac{x-60.78}{4.4019-4.1073} = \dfrac{60.78-60.42}{4.1073-4.1013} \Rightarrow y = 78.456$

(c) Left as exercise.

4. The below is showing an extract of a part of secx^0

$x = 60'$	$0'$	$12'$	$24'$	$36'$	$48'$
sec$x$	2.0000	2.0122	2.0245	2.0371	2.0498

Using linear interpolation,
(i) value of sec$60^0 15'$ (ii) angle whose secant is 2.0436

5. The below is showing an extract of a part of secx^0

x^0	0.85^0	0.86^0	1.85^0	1.86^0	1.88^0
sinx^0	0.7513	0.7578	0.9612	0.9585	0.9526

Using linear interpolation,
(i) value of sin0.857^0 (ii) value of sin1.857^0
(ii) value of sin1.84^0 (iv) value of sin1.980^0

Solution.

NB: Since we are dealing with very small angles in degree (I.e less than 10^0), then the calculator must be in Radians.

(i)

0.85^0	0.857^0	0.86^0
0.7513	y	0.7578

, Take $\dfrac{y-0.7513}{0.857-0.850} = \dfrac{0.7578-0.7513}{0.860-0.85} \Longrightarrow y = 0.7559$

(ii),(iii) and(iv) are left as exercise.

9.3 Exercise 9

(1) Given that the function $y = f(x)$ passes through $(0.19, 1.1878)$ and $(0.09, 1.7581)$.
  (a) Find the value of x_0 such that $f(x_0) = 1.1764$.
  (b) Find the value of y_1 such that $f(0.14) = y_1$
  (c) Find $f^{-1}(1.8581)$

Ans: $(a)x_0 = 0.192, (b)y_1 = 1,4730$,(c) 0.0725.

(2) Given that the function $y = f(x)$ passes through $(-0.13, 1.1880)$ and $(-0.06, 1.1781)$.
  (a) Find the value of x_0 such that $f(x_0) = 1.2964$.

(b) Find the value of m such that $f(-0.16) = m_0$

(c) Find the value of y_0 such that $P(-0.19, y_0)$ is a point on $f(x)$.

Ans: (a) -0.896, (b) 1.192 and (c) 1.197

(3) In the table below is an extract of part of $y = log_e{}^{x}$ with the given x values.

x	30.00	30.22	30.42	30.68
y	3.4012	3.4085	3.4151	3.4236

Using linear interpolation/Extrapolation, estimate the unknown values below:

(a) $y = log_e{}^{30.50}$

(b) $log_e{}^{30.25}$

(c) $3.4019 = log_e{}^{x}$

(d) $3.3991 = log_e{}^{m}$

Ans:(a) 3.418,(b) 3.410 (c) 30.020 and (d) 29.940

(4) The distance covered by the a certain body and the time in seconds are as shown in the table below:

Distance(m)	0	5	10	15	20
Time(s)	0	12	25	39	54

Using linear interpolation/extrapolation, estimate;

(i) The time when the distance $x = 12m$,

(ii) The distance covered when $t = 45$

(iii) The time when the distance $x = 22m$ **Ans:** (i) 30.60, (ii) 17.0 and (iii) 60.0

(5) A body freely moving from rest is acted on by a variable force(N) as shown in the table below.

Distance(m)	0	4	10	15	20	25	31
Force(N)	5	8	11	12	13.6	10.5	5

Using linear interpolation/extrapolation, determine;

(a) the force when the body has traveled a distance of $22m$

(b) the distance when a force $F = 12.8N$

(c) the force when the body has traveled a distance of $34.7m$.

Ans: (a) 12.360, (b) 17.50 and (c) 1.608

(6) The below is showing an extract of a part of secx^0

$60^0x'$	$0'$	$12'$	$24'$	$36'$	$48'$
sec$60^0x'$	2.0000	2.0122	2.0245	2.0371	2.0498

Using linear interpolation,

(i) value of sec$60^015'$

(ii) angle whose secant is 2.0436

Ans: (i) 2.0307 (ii) $60^042'$.

(7) The amount of fuel, (in litres) remaining in a car tank recorded after every 10 minutes was as follows: 40, 37, 34.6, 30.3, 28.6. If the initial amount of fuel was 42.5 litres, estimate using linear interpolation/ extrapolation:

(i) amount of fuel left after 17.5 minutes,

(ii) time taken when 26.4 litres remained.

Ans: (i) 40.625 (ii) 62.941

(8) The charges of sending parcels by a certain distributing company depends on the weights, $500g, 1kg, 1.5kg, 2kg and 5kg$ the charges are 750/=, 1000/=, 2000/=, 3500/= and 5500/
respectively. Estimate:

(i) What would the distributor charge for a parcel of weight 2.5kg?

(ii) If the sender pays 6200/=, what is the weight of the parcel.

Ans: (i) 3833.30/= (ii) $6.05kg$

(9) Given the table below,

x	0.1	0.2	0.3	0.4
$\sqrt{x}$	0.3162	0.4472	0.5477	0.6325

Use linear interpolation/ extrapolation to estimate:

(i) $\sqrt{0.25}$

(ii) $\sqrt{x} = 0.75$:

(iii) 0.6677^2

(iv) $p^2 = 0.5$, find p.

Ans: (i) 0.4975 (ii) 0.5386 (iii) 0.4415 (iv) 0.7487

(10) In an experiment the following observations were made;

$T:$	0	12	20	30
$Q:$	0.6	2.9	-0.1	0.6325

Use linear interpolation to find:

(i) Q when $T = 16$

(ii) T when $Q = -1$

Ans: (i) 1.40 (ii) 23.210

(11) Given the following values.

x^0	0.85^0	0.86^0	1.85^0	1.86^0
$Sinx^0$	0.7513	0.7578	0.9612	0.9585

Use linear interpolation to estimate:

(i) $sin 0.857^0$,

(ii) $sin 1.857^0$.

Ans: (i) 0.7559 (ii) 0.9593

(12) The table below shows the cost y shillings for hiring a motor cycle for a distance x kilometers.

Distance$(x)km$	10	20	35	47
Cost$(y)shs.$	3800	4600	5600	6200

Use linear interpolation or extrapolation to calculate the;

(i) Cost of hiring the motorcycle for a distance of $45km$

(ii) Distance Mukasa travelled if he paid $shs.4000$.

Ans: (i) 6100 (ii) $12.50Kg$.

(13) The table below shows the time and the corresponding velocity for a particle projected vertically upwards with other factors affecting its motion.

Time in minutes	3.66	3.87	4.04	4.66
Velocity(ms^{-1})	2.794	1.633	0.810	-2.80

Using linear interpolation and extrapolation find;

(i) The time taken for the particle to reach the maximum height

(ii) Initial velocity of projection

(iii) Velocity of the particle when time is 186 seconds after projections.

Hint: Change from minutes to seconds first. Ans: (i) 4.1729 (ii) 1.633 (iii) 5.890.

(14) The public mean cost from Lukaya to Kampala city is $shs.10000$ for a total distance of $120km$.

(i) If the distance from Kyengera town to Kampala is $20km$, how much should be paid from Lukaya to Kyengera.

(iii) Peter was driving from Lukaya to Kampala and his car got a puncture after $55km$, how much is he supposed to pay for the remaining distance to Kampala.
Ans: (i)

Chapter 10

ROOTS OF THE EQUATION.

10.1 A root

A **root** :An x_0 is called the root of the polynomial or transcendental function $f(x) = 0$ if $f(x_0) = 0$. i.e A root refers to the value of x that must be substituted in $f(x) = 0$ in order to obtain a zero(0).
The function $f(x) = 0$ can be of higher powers that can be so hard to look for the root(s), this calls for **Numeriacl Methods**

10.1.1 Location of the Real Roots of Equation $f(x) = 0$

This involves identifying where the root (x_0) of the equation is situated. This can be in form of interval or a starting value that is so close to the x_0 called the **initial approximation.** This can be done using the following methods:

1. Graphical Method.
   With this approach, either **one** or **two** graphs can be plotted in order to locate the root(s) of the equation.This is normally done in two ways:

   (i) Single Graph Approach/plot: This involves drawing the graph $f(x) = 0$ as one graph for values of x close to zero or in the interval of $\pm(|k| + 1)$ where k is the largest coefficient in $f(x)$ by magnitude.
   In this case, the root(x_0) exists at point where the function $f(x) = 0$ cuts the $x-$ axis.I.e

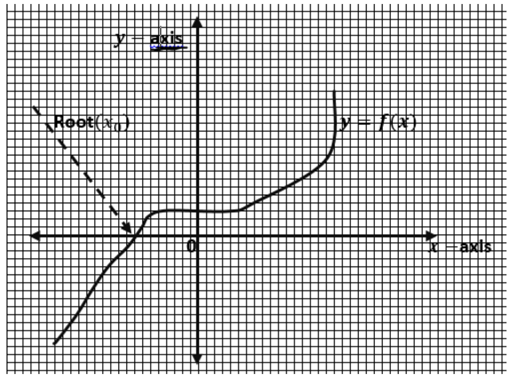


Figure 10.1: Here you plot a single graph and you identify where it cuts the $x-$ from which becomes the root.

(ii) Two Graphs Method/Plot. This is also known as the **Double plot approach**. Here the function $y = f(x)$ is split into two and the equations are plotted separately on the same axes. I.e

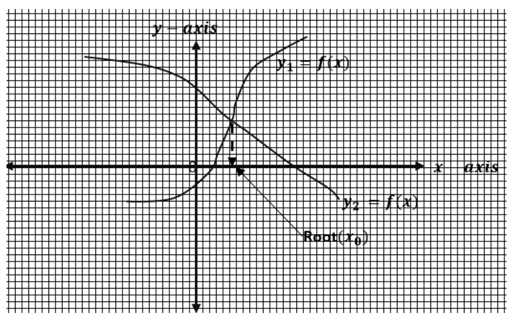


Figure 10.2: Combining y_1 and y_2 gives the original $f(x)$. I.e $y_1 = y_2 \iff f(x) = y_1 - y_2 = 0$.

The root(s) is/are the value(s) of x corresponding to the point of intersection of the two plotted graphs.
NB

(i) In case your are given the interval of x values, make sure that your plot at least three points within the given interval. This is done in order to increase on the level of accuracy of the value your to obtain.

(ii) In case your are **not** given the interval of x values, if a **polynomial** is given, apply $\underline{\pm(|k| + 1)}$ where $k =$ Highest coefficient of x in the given equation. But most people recommend $\pm$ which is some times <u>not true.</u> While if a **transcendental** is given, apply a calculator to identify the interval.

(iii) If your required to find the negative root, please only consider the negative the x valves in order to serve time including zero and vice verse.

(iv) The over all importance of any graph plotted is to get the **initial approximation of the root.** However you may be requested to read it as an interval or a range.

2. Sign Change Approach/ Analytical Approach: This involves applying the pure mathematics approach.
If the function $y = f(x)$ has a root between $[a, b]$, then there must be a sign change say from positive to negative or negative to positive.
The sign change is identified from $f(x)$ values but the conclusion is made from the $x-$ values.
So in this case $\boxed{f(a) \times f(b) < 0}$
Conclusion: Since the $f(a) \times f(b) < 0$, then there is a root of $f(x)$ between $x = a$ and $x = b$
Examples:

(1) Show that the positive real root of $2x^2 - 6x - 3 = 0$ lies between 3.2 and 3.98.
Solution:
Let $f(x) = 2x^2 - 6x - 3$.
When $x = 3.2$, $\quad f(3.2) = 2(3.2)^2 - 6(3.2) - 3 = -1.720$
When $x = 3.98$, $\quad f(3.98) = 2(3.98)^2 - 6(3.98) - 3 = 4.801$
Since $f(3.2).f(3.98) < 0$, then there is a root of the equations $f(x) = 2x^2 - 6x - 3$. between 3.2 and 3.98.

(2) Show that the negative real root of $x^3 - 6x + 1 = 0$ lies between -3 and -2.2.
Solution:
Let $f(x) = x^3 - 6x + 1$.
When $x = -3$, $\quad f(-3) = (-3)^3 - 6(-3) - 3 = -8.000$
When $x = -2.2$, $\quad f(-2.2) = (-2.2)^3 - 6(-2.2) - 3 = 3.552$
Since $f(-3).f(-2.2) < 0$, then there is a root of the equations $x^3 - 6x + 1 = 0$ between -3 and -2.2.

(3) Show that the real root of $Xe^{-X} = 2X - 5$ lies between 2 and 3.
Solution:
NB: In order to check for the existence of the root, you must make sure that $f(x) = 0$
So from $Xe^{-X} = 2X - 5, \Longleftrightarrow Xe^{-X} - 2X + 5 = 0$.
Let $f(x) = Xe^{-X} - 2X + 5$
When $x = 2$, $\quad f(2) = 2e^2 - 2(2) + 5 = 1.2706$
When $x = 3$, $\quad f(3) = 3e^3 - 2(3) + 5 = -0.8506$
Since $f(2).f(3) < 0$, then there is a root of the equations $Xe^{-X} = 2X - 5$ between 2 and 3.

(5) Show that the real root of $4x = tan(x)$ lies between 1.0 and 1.2.
Solution:
NB: In case any trig ratio (i.e sine, cosine tangent and others) are given when the angles are small, then the **calculator** must be converted to RADIANS.
So from $4x = tan(x), \Longleftrightarrow 4x - tan(x) = 0$.
Let $f(x) = 4x - tan(x)$
When $x = 1.0$, $\quad f(1.0) = 4(1.0) - tan(1.0) = 2.4426$
When $x = 1.2$, $\quad f(1.2) = 4(1.2) - tan(1.2) = -0.8506$

Since $f(2).f(3) < 0$, then there is a root of the equations $Xe^{-X} = 2X - 5$ between 2 and 3.

3. Using single graph method, Show that $2x^2 - 6x - 3 = 0$ has two positive roots.
Solution:

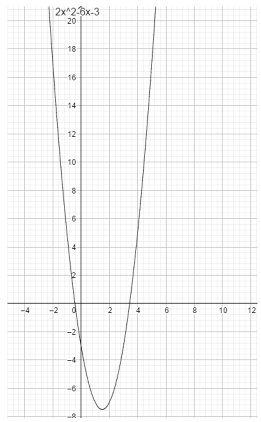


Figure 10.3

4. Using single graph method, Show that $Xe^{-X} = 2X - 5$ has one real roots.

Solution:

5. By plotting graphs, Show that $Xe^{-X} = 2X - 5$ has one real roots.

Solution:

6. Using a suitable graph, locate the interval over which the roots of the equation $2x^2 + 3x - 3 = 0$ lie.
Solution:
Table of values
let $y = 2x^2 + 3x - 3$

x	-3	-2	-1	0	1	2	
y	6	-1	-4	-3	2	11	

7. Using suitable graphs, locate the interval over which the roots of the equation $2x^2 + 3x - 3 = 0$ lie.
Solution:
let $y = 2x^2 + 3x - 3 = 0 \iff 2x^2 = 3 - 3x$
Then $y_1 = 2x^2$ and $y_2 = 3 - 3x$.
Table of values

x	-3	-2	-1	0	1	2	
$y_1 = 2x^2$	18	8	2	0	2	8	
$y_2 = 3 - 3x.$	12			3			

8. UNEB 2018 & 2019

Chapter 11

ERRORS

11.1 ERRORS

11.1.1 Introduction

Error refers to the difference between the exact value and inexact(approximated) value. Errors arise from many calculations and approximations used when working out numbers.

11.1.2 Types of Errors

These are also referred to as the **sources of errors.** and these include;

1. Random Error: This refers to the error that arises due to human or machine failure. They result from poor reading or a mistake in the calculations. E.g A person reading 62 as 26.
   Random error cannot be treated numerically.

2. Rounding Off Error.This refers to the type of error arising from approximating the true value by a different value of lower decimal places.
   NB: Rounding off is done in a single step.

3. Truncation Error : This occurs when the infinite process is terminated at some point guided by decimal places or significant figures. In this case the higher terms are neglected.
   Examples. Study the table below;

Value	Required Accuracy	Rounding off	Truncation
2.02871	$3dps$	2.029	2.028
0.3333333	$3sf$	0.333	0.333
0.006789	$4dps$	0.0068	0.0068
0.0076589	$4sf$	0.007659	0.007658
1.0034562	$5sf$	1.0035	1.0034
0.6666666	$3dps$	0.667	0.666
52,457,829	$3sf$	52,500,000	52,400,00

11.1.3 Common Terms Used in Studying Errors.

(1) Error: This is also known as the **Absolute error**. It refers to the magnitude of error.
$\Longrightarrow$ Error$(\Delta x) = |Exact\ Value(X) - Approximated\ Value(x)|$
(2) Relative error: This refers to the ratio of absolute error to the working value. It tells us how the big error is.
Working value is exact value or the approximated value.
$\Longrightarrow$ Relative Error $= \dfrac{Absolute\ Error}{Working\ value}$
NB: If you have the Exact Value, take it as the working value other wise take the approximated value.
(3) percentage Error: This refers to the relative error expressed as a percentage. This is done by multiplying relative error by 100%.
$\Longrightarrow$ %age error $= \dfrac{Absolute\ Error}{Working\ value} \times 100\%$
NB: The percentage sign(%) can be included on percentage error or omitted.
Example:
(1) Given the number $X = 2.6666668$. Truncate it to 2 decimal places and take it as x hence calculate

(i) Absolute error (ii) Relative error (iii) percentage error.

Solution
Let Exact value$(X) = 2.6666668 \Longrightarrow$ approximated value$(x) = 2.66(2dps)$ Hence:
(i) Absolute error$(\Delta x) = |X - x| = |2.6666668 - 2.66| = 0.0066668$
(ii) Relative error $= \dfrac{Absolute\ Error}{Working\ value} = \dfrac{0.0066668}{2.6666668} = 0.0025(4dps)$
(iii) %age error $= \dfrac{Absolute\ Error}{Working\ value} \times 100 = \dfrac{0.0066668}{2.6666668} \times 100 = 0.250\%$
(2) The length of a rectangle was measured as $5.72m$ instead of $5.5m$. Find the;

(i) Absolute error (ii) Relative error (ii) percentage error.

Solution
Let Exact value$(L) = 5.5 \Longrightarrow$ approximated value$(l) = 5.72$ Then:
(i) Absolute error$(\Delta x) = |L - l| = |5.5 - 5.72| = 0.22$
(ii) Relative error $= \dfrac{Absolute\ Error}{Working\ value} = \dfrac{0.22}{5.5} = 0.040$
(iii) %age error = Relative error $\times 100 = 0.040 \times 100 = 4.0\%$
Absolute Error can also be obtained based on the level of accuracy.I.e Based on the

number of decimal places or Significant figures used in the calculation. Here absolute error

can be obtained as $\boxed{\text{Absolute error} = 0.5 \times 10^{-n}}$, where $n =$ Number of decimal places.
NB; Absolute Error in a value is also known as **Maximum Absolute Error in the value.**

Limits for accuracy:

This is done by starting the Upper limit(maximum value) and the lower limit (minimum value).
The Maximum value/ Upper limit = Approximated value + Maximum Absolute Error in the value while
The Minimum value/ Lower limit = Approximated value − Maximum Absolute Error in the value

Given the Maximum value and the Minimum value in each term, then the absolute error in the working value is give by:

$$\textbf{Absolute Error in the working value} = \frac{1}{2}|(Max.\ value - Min.\ value)|$$

Range of values within which the exact value is lying: This is given by starting the minimum value and the maximum value as below;

$$\text{minimum value} \leq \text{Working Value} \leq \text{maximum value}$$
or
$$[minimum\ value,\ maximum\ value]$$

NB: The range of values is the same as interval

Error Bounds in the main Binary Operations.

This shows how the errors are developed in Addition, Subtraction, Multiplication and Division.

(i) Addition: let $Z = X + Y$, then
$\diamond Z_{max} = X_{max} + Y_{max}$
$\diamond\ Z_{min} = X_{min} + Y_{min}$

(ii) Subtraction: let $Z = X - Y$, then
$\diamond Z_{max} = X_{max} - Y_{min}$
$\diamond\ Z_{min} = X_{min} - Y_{max}$

(iii) Multiplication: let $Z = X \times Y$, then
$\diamond Z_{max} = X_{max} \times Y_{max}$
$\diamond\ Z_{min} = X_{min} \times Y_{min}$

(iv) Division: Let $Z = \dfrac{X}{Y}$, then
$\diamond Z_{max} = \dfrac{X_{max}}{Y_{min}}$
$\diamond\ Z_{min} = \dfrac{X_{min}}{Y_{max}}$

Examples.

1. Given that $x = 4.96$ and $y = 2.234$ each rounded off to the given number of decimal places. Calculate the maximum possible error in . Suppose

(i) $P = x + y$ (ii) $Z = x - y$ (a) $M = xy$ (iv) $T = \dfrac{x}{(x+y)}$.

(a) Find the maximum absolute error in each of x and y,
(b) Find the limits within which each of P, Z, M and T lie.
(c) Find the absolute error in each of P, Z, M and T .
(d) Find the percentage error in each of P, Z, M and T .

Solution.
For P,
(a) From $x = 4.96$ and $y = 2.234$
Let maximum absolute error in $x = \Delta x = 0.5 \times 10^{-2} = 0.005$
Let maximum absolute error in $y = \Delta y = 0.5 \times 10^{-3} = 0.0005$

(b) From $P = x + y$,
$P_{max} = x_{max} + y_{max} = (4.96 + 0.005) + (2.234 + 0.0005) = 7.1995$
$P_{main} = x_{min} + y_{min} = (4.96 - 0.005) + (2.234 - 0.0005) = 7.1885$

So lower Limit $= P_{main} = 7.1885$,
Upper limit $= P_{max} = 7.1995$

(c) Absolute error in $P = \dfrac{1}{2}|P_{max} - P_{main}| = \dfrac{|7.1995 - 7.1885|}{2} = 0.0055$
(d) %age error of $P = \dfrac{Absolute\ Error}{Working\ value} \times 100$
But Working value $P = (x + y) = (4.96 + 2.234) = 7.174$

Then %age error of $P = \dfrac{0.0055}{7.174} \times 100 = 0.0765\%(4dps)$

For Z,
(a) From $x = 4.96$ and $y = 2.234$
Let maximum absolute error in $x = \Delta x = 0.5 \times 10^{-2} = 0.005$
Let maximum absolute error in $y = \Delta y = 0.5 \times 10^{-3} = 0.0005$

(b) From $Z = x - y$,
$Z_{max} = x_{max} - y_{min} = (4.96 + 0.005) - (2.234 - 0.0005) = 2.7315$
$Z_{main} = x_{min} - y_{max} = (4.96 - 0.005) - (2.234 + 0.0005) = 2.7205$

So lower Limit $= Z_{main} = 2.7205$,
Upper limit $= Z_{max} = 2.7315$

(c) Absolute error in $Z = \dfrac{1}{2}|Z_{max} - Z_{main}| = \dfrac{|2.7315 - 2.7205|}{2} = 0.0055$

(d) %age error of $Z = \frac{Absolute\ Error}{Working\ value} \times 100$
But Working value $Z = (x - y) = (4.96 - 2.234) = 2.7260$

Then %age error of $Z = \frac{0.0055}{2.7260} \times 100 = 0.2018\%(4dps)$

For M,
(a) From $x = 4.96$ and $y = 2.234$
Let maximum absolute error in $x = \Delta x = 0.5 \times 10^{-2} = 0.005$
Let maximum absolute error in $y = \Delta y = 0.5 \times 10^{-3} = 0.0005$

(b) From $M = x \times y$,
$M_{max} = x_{max} \times y_{max} = (4.96 + 0.005) \times (2.234 + 0.0005) = 11.0943$
$M_{main} = x_{min} \times y_{min} = (4.96 - 0.005) \times (2.234 - 0.0005) = 11.0670$

So lower Limit $= M_{main} = 11.0670$,
Upper limit $= M_{max} = 11.0943$

(c) Absolute error in $M = \frac{1}{2}|M_{max} - M_{main}| = \frac{|11.0943 - 11.0670|}{2} = 0.01365$
(d) %age error of $M = \frac{Absolute\ Error}{Working\ value} \times 100$
But Working value $M = (x \times y) = (4.96 \times 2.234) = 11.0806$

Then %age error of $M = \frac{0.01365}{11.0806} \times 100 = 0.1232\%(4dps)$

For T,
(a) From $x = 4.96$ and $y = 2.234$
Let maximum absolute error in $x = \Delta x = 0.5 \times 10^{-2} = 0.005$
Let maximum absolute error in $y = \Delta y = 0.5 \times 10^{-3} = 0.0005$
(b) Let $T = \frac{x}{y}$, then
$T_{max} = \frac{X_{max}}{Y_{min}} = \frac{4.96 + 0.005}{2.234 - 0.0005} = 2.2230$
$T_{min} = \frac{X_{min}}{Y_{max}} = \frac{4.96 - 0.005}{2.234 + 0.0005} = 2.2175$ So lower Limit $= T_{main} = 2.2175$,
Upper limit $= T_{max} = 2.2230$

(c) Absolute error in $T = \frac{1}{2}|T_{max} - T_{main}| = \frac{|2.2230 - 2.2175|}{2} = 0.0055$
(d) %age error of $T = \frac{Absolute\ Error}{Working\ value} \times 100$
But Working value $T = (\frac{x}{y}) = (\frac{4.96}{2.234}) = 2.2202$

Then %age error of $T = \dfrac{0.0055}{2.2202} \times 100 = 0.2477\%(4dps)$

2. Obtain the range of the values within which the exact value of $\dfrac{15.36 + 27.1 - 1.672}{2.36 \times 1.043}$ lies.
Solution:
Let $Z = \dfrac{15.36 + 27.1 - 1.672}{2.36 \times 1.043}$
Absolute Error in $15.36 = 0.005$,
Absolute Error in $27.1 = 0.05$,
Absolute Error in $1.672 = 0.0005$,
Absolute Error in $2.36 = 0.005$,
Absolute Error in $1.043 = 0.0005$,
Then;
$Z_{max} = \dfrac{15.365 + 27.15 - 1.6715}{2.355 \times 1.0425} = 16.9203(4dps)$
$Z_{min} = \dfrac{15.355 + 27.05 - 1.6725}{2.365 \times 1.0435} = 16.5051(4dps)$
Range:

$$16.5051 \leq Z \leq 16.9203 \text{ or } [16.5051\ ,\ 16.9203]$$

3. Obtain the limits within which the exact value of $2.674\Big(\dfrac{15.175 - 4.8006}{0.82}\Big)$ lies. Each of the values were rounded to the given number of decimal places
Solution:
Let $Z = 3.674\Big(\dfrac{10.175 - 4.8006}{0.82}\Big)$
Absolute Error in $3.674 = 0.0005$,
Absolute Error in $10.175 = 0.0005$,
Absolute Error in $4.8006 = 0.00005$,
Absolute Error in $0.82 = 0.005$,
Then;
$Z_{max} = 3.6745\Big(\dfrac{10.1755 - 4.80055}{0.815}\Big) = 24.2334(4dps)$
$Z_{min} = 3.6735\Big(\dfrac{10.1745 - 4.80065}{0.825}\Big) = 23.9305(4dps)$

$\Longrightarrow$ Lower limit $= 23.9305$, Upper limit $= 24.2334$

4. (a) Obtain the limits within which the exact value of $P = 2.874\Big(\dfrac{6.175 + 4.8006}{3.92 \times 1.0}\Big)$ lies. Each of the values were rounded to the given number of decimal places.
(b) Find the percentage error in approximating P .
Solution:
(a) From $P = 2.874\Big(\dfrac{6.175 + 4.8006}{3.92 \times 1.0}\Big)$
Absolute Error in $2.874 = 0.0005$,

Absolute Error in 6.175 = 0.0005,
Absolute Error in 4.8006 = 0.00005,
Absolute Error in 3.92 = 0.005,
Absolute Error in 1.0 = 0.05,
Then;

$$P_{max} = 2.8745\Big(\frac{6.1755 + 4.80065}{3.915 \times 0.95}\Big) = 8.4831(4dps)$$

$$P_{min} = 2.8735\Big(\frac{6.1745 + 4.80055}{3.925 \times 1.05}\Big) = 7.6549(4dps)$$

$\Longrightarrow$ Lower limit = 7.6549 , Upper limit = 8.4831

(b) percentage error in approximating $P = \dfrac{Absolute\ Error}{Working\ value} \times 100$
But
Absolute error in $P = \dfrac{1}{2}|P_{max} - P_{main}| = \dfrac{|8.4831 - 7.6549|}{2} = 0.4141$

But Working value $P = 2.874\Big(\dfrac{6.175 + 4.8006}{3.92 \times 1.0}\Big) = 8.0469$
Then
%age error of $P = \dfrac{0.4141}{8.0469} \times 100 = 5.146\%(4dps)$

5. Determine the maximum percentage absolute error in $\dfrac{\sqrt{z}}{(x^2y^3)}$, given that $x = 3.41, y = 5.0$ and $z = 2.82$ all numbers rounded off to the given number of decimal places.
Solution. Let $M = \dfrac{\sqrt{z}}{(x^2y^3)}$ and
Error in $x = 0.005$,
Error in $y = 0.05$,
Error in $z = 0.005$,
Then

$$M_{max} = \frac{\sqrt{z_{max}}}{((x_{min})^2(y_{min})^3)} = \frac{\sqrt{2.825}}{(3.405)^2(4.95)^3} = 0.001195$$

$$M_{min} = \frac{\sqrt{z_{min}}}{((x_{max})^2(y_{max})^3)} = \frac{\sqrt{2.815}}{(3.415)^2(5.05)^3} = 0.001117$$

So Absolute error in $M = \dfrac{|0.001195 - 0.001117|}{2} = 0.00004$ and

Working value $= \dfrac{\sqrt{2.82}}{(3.41^2 5.0^3)} = 0.001155$

$\Longrightarrow$ %age error of $M = \dfrac{0.00004}{0.001155} \times 100 = 3.4632(4dps)$

6. Given numbers $x = 2.24$ and $y = 4.8$, with percentage errors 4.0% and 6.1% respectively. Find the limits in;

(a) $Z = x^3y$,

(b) $Z = \dfrac{2}{x+y}$

(c)$Z = \dfrac{1}{x} + \dfrac{1}{y} - \dfrac{1}{xy}$

Solution.

(a) From $Z = x^3y$,

Since percentage error $= \dfrac{Absolute\ Error}{Working\ value} \times 100$, Then

For $x,\ 4.0 = \dfrac{\Delta x}{2.24} \times 100 \Longleftrightarrow \Delta x = 0.0896$

For $y,\ 6.1 = \dfrac{\Delta y}{4.8} \times 100 \Longleftrightarrow \Delta y = 0.2928$

$\Longrightarrow Z_{max} = (x_{max})^3 y_{max} = (2.24 + 0.0896)^3 \times (4.8 + 0.2928) = 64.3874$ and

Also $Z_{min} = (2.24 - 0.0896)^3 \times (4.8 - 0.2928) = 44.8192$

Lower limit = 44.8192, Upper limit= 64.3874

(b) From $Z = \dfrac{2}{x+y}$,

$$Z_{max} = \frac{2}{x+y} = \frac{2}{x_{min} + y_{min}} = \frac{2}{(2.24 - 0.0896) + (4.8 - 0.2928)} = 0.3004$$

$$Z_{min} = \frac{2}{x+y} = \frac{2}{x_{max} + y_{max}} = \frac{2}{(2.24 + 0.0896) + (4.8 + 0.2928)} = 0.2695$$

Lower limit = 0.2695, Upper limit= 0.3004

(c)$Z = \dfrac{1}{x} + \dfrac{1}{y} - \dfrac{1}{xy}$

$$Z_{max} = \frac{1}{x_{min}} + \frac{1}{y_{min}} - \frac{1}{x_{max} \times y_{max}} = \frac{1}{(2.24 - 0.0896)} + \frac{1}{(4.8 - 0.2928)} - \frac{1}{(2.24 + 0.0896) \times (4.8 + 0.292}$$

0.6026

$$Z_{min} = \frac{1}{x_{max}} + \frac{1}{y_{max}} - \frac{1}{x_{min} \times y_{min}} = \frac{1}{(2.24 + 0.0896)} + \frac{1}{(4.8 + 0.2928)} - \frac{1}{(2.24 - 0.0896) \times (4.8 - 0.292}$$

0.5224

So Lower limit = 0.5224, Upper limit= 0.6026

Applications of Error Bound.

7. A trader in books and pens makes annual profits in books of $sh.60$ million with a margin of error of $sh.5$ million and annual profits in pens of $sh.20$ million with a margin error of $sh.1.53$ million.

(i) Find the interval of values corresponding to her gross income.

(ii) Find the percentage error in the gross income to three decimal places.

Solution.

NB: Since she is getting profits from each department, then we should the profits.

(i) Let Total profit(P) = profits from books(B) + profits from pens(A) $\Longleftrightarrow P = B + A$
$\Delta B = 5$ million and $\Delta A = 1.53$ million
From, $P = B + A$
$P_{max} = B_{max} + A_{max} = (60 + 5) + (20 + 1.53) = 86.530$ million.
$P_{min} = B_{min} + A_{min} = (60 - 5) + (20 - 1.53) = 73.470$ million
For interval $[73.470\ ,\ 86.530]$
(ii) Relative error $= \dfrac{Absolute error in profits}{gross\ income} \times 100 = \dfrac{0.5|86.530 - 73.470|}{(60 + 20)} \times 100 =$ 8.163

8. A company dealing in clothes and shoes had a capital of $sh.700$ million. The profit in a certain year was $sh.50$ million from clothes and $sh.43.6$ million from shoes. There was possible error of 5.5% from the clothes and 11% error from shoes.

(a) Find the maximum and minimum values of the total profit of the departments as a percentage of the capital.

(b) Find the limits with in which the exact profits from the both departments companies lies.

Solution:

(i) Let Total profit(P) = profits from clothes(B) + profits from shoes(A)
$\Longrightarrow P = B + A$
$\Delta B = \dfrac{5.5}{100} \times 50 = 2.75$ million and $\Delta A = \dfrac{11}{100} \times 43.6 = 4.796$ million
From, $P = B + A$
$P_{max} = B_{max} + A_{max} = (50 + 2.75) + (43.6 + 4.796) = 101.1460$ million.
$P_{min} = B_{min} + A_{min} = (50 - 2.75) + (43.6 - 4.796) = 86.0540$ million
Lower limit $= \dfrac{86.0540}{700} \times 100 = 12.2934.$ and
Upper limit $= \dfrac{101.1460}{700} \times = 14.4494.$
(ii) Lower limit = 86.0540 and Upper limit = 101.1460.

9. A trader in books and pens makes annual profits in books of $sh.60$ million with a margin of error of $sh.600000$ and annual profits in pens of $sh.20$ million with a margin error of $sh.350000$.
(i) Find the range of values corresponding to her gross income.
(ii) Find the relative error in the gross income to three decimal places.

Solution.

NB: Please makes sure that you have the same units of measure. Ie Either millions or

(i) Let Total profit(P) = profits from books(B) + profits from pens(A) $\Longleftrightarrow P = B + A$
$\Delta B = 0.60$ million and $\Delta A = 0.35$ million
From, $P = B + A$
$P_{max} = B_{max} + A_{max} = (60 + 0.60) + (20 + 0.35) = 80.950$ million.
$P_{min} = B_{min} + A_{min} = (60 - 0.60) + (20 - 0.35) = 79.050$ million

For range: $79.050 \leq$ gross income ≤ 80.950

(ii) Relative error $= \dfrac{Absolute error in profits}{gross\ income} = \dfrac{0.5|80.950 - 79.050|}{(60+20)} = 0.012$

10. A trader in soap and sanitizers makes annual profits in soap of $sh.120$ million with a margin of error of $sh.25$ millions and annual loss in sanitizers of $sh.20$ million with a margin error of $sh.0.998$ millions
(i) Find the range of values corresponding to his gross income.
(ii) Find the relative error in the gross income to three decimal places.
Solution.
NB: Since he is getting profits and losses, then we should subtract off the losses.
(i) Let Gross income(P) = profits from soap(N) − profits from sanitizers(M) $\Longleftrightarrow$ $P = N - M$
$\Delta N = 0.60$ million and $\Delta M = 0.35$ million
From, $P = N - M$
$P_{max} = N_{max} - M_{min} = (120+25) - (20-0.998) = 125.9980$ million.
$P_{min} = N_{min} - M_{max} = (120-25) - (20+0.998) = 74.0020$ million
For range
$74.002 \leq$ gross income ≤ 125.9980
(ii) Relative error $= \dfrac{Absolute error in profits}{gross\ income} = \dfrac{0.5|125.9980 - 74.0020|}{(120+20)} = 0.1857$

11. The radius of the circle is measured as $7.48m$ to the nearest cm. calculate the upper bound of its area correct to four significant figures. Hence, the percentage error in approximating the area.
Solution:
$r = 7.48 \Longleftrightarrow \Delta r = 0.5 \times 10^{-2} = 0.005$
Let area(A) $= \pi r^2$
$\Longrightarrow A_{max} = \pi(7.48 + 0.005) = 23.5148.$
Hence %age error $= \dfrac{\text{Absolute error in area.}}{\text{Working value}}$
But $A_{min} = \pi(7.48 - 0.005) = 23.4834.$
%age error $= \dfrac{\text{Absolute error in area.}}{\text{Working value}} = \dfrac{0.5|23.5148 - 23.4834|}{(\pi(7.48)^2)} = 0.1857$

12. The sides of a rectangle are measured as $10.026m$ and $4.5m$, to the nearest centimeters.
(a) Calculate the maximum possible error in each side
(b) Calculate the maximum and minimum values of the perimeter. Hence, determine the absolute error in the perimeter.
(c) Calculate the relative error in the area.
Solution.
$l = 10.026 \Longleftrightarrow \Delta l = 0.5 \times 10^{-3} = 0.0005.$ and $w = 4.5 \Longleftrightarrow \Delta l = 0.5 \times 10^{-1} = 0.05.$
(a) $\Delta l = 0.0005.$ and $\Delta l = 0.05.$
(b) Let Perimeter(P) $= 2(l + w)$
$\Longrightarrow P_{max} = 2((10.026 + 0.0005) + (4.5 + 0.05)) = 29.1530$and
$P_{min} = 2((10.026 - 0.0005) + (4.5 - 0.05)) = 28.9510$
Hence Absolute error $= \dfrac{|29.1530 - 28.9510|}{2 \times (10.026 + 4.50)} = 0.0035.$

Please do the remaining part.

11.1.4 Error Propagation

This involves deriving formulae for error propagation in Addition, Subtraction, Multiplication and Division.
Here we must apply **triangular inequality** I.e $|a+b| \leq |a|+|b|$
Here we also consider two assumptions:
(1) We are dealing with fixed numbers and there we don't say that they tend to zero.
(2) The errors from each value is so small and there for at least their products tend to zero.

QN: Given that x and y are approximations of X and Y with respective errors Δx and Δy. Derive the simplest formula for percentage error in;

1. Addition,
2. Subtraction,
3. Multiplication,
4. Division.

Solution:
Hint: Here the reader must first read and interpret the question in order to identify the Exact terms given.

1. Addition,
   Using $X = x + \Delta x$ and $Y = y + \Delta y$
   Let $Z = X + Y$ and $z = x + y$

$$\begin{aligned}
\Longrightarrow \Delta z &= Exact\ Value(Z) - Approximated\ Value(z) \\
&= X + Y - (x + y) \\
&= (x + \Delta x) + (y + \Delta y) - x - y \\
&= \Delta x + \Delta y \\
\therefore Absolute\ error &= |\Delta z| \\
&= |\Delta x + \Delta y| \leq |\Delta x| + |\Delta y| \\
Maximum\ absolute\ error &= |\Delta x| + |\Delta y| \\
Relative\ Error &= \frac{(|\Delta x| + |\Delta y|)}{X + Y} \\
\Longrightarrow \%age\ error &= \frac{(|\Delta x| + |\Delta y|)}{X + Y} \times 100.
\end{aligned}$$

2. Subtraction,
   Using $X = x + \Delta x$ and $Y = y + \Delta y$
   Let $Z = X - Y$ and $z = x - y$

$$
\begin{aligned}
\implies \Delta z &= Exact\ Value(Z) - Approximated\ Value(z) \\
&= X - Y - (x - y) \\
&= (x + \Delta x) - (y + \Delta y) - x + y \\
&= \Delta x + \Delta y \\
\therefore Absolute\ error &= |\Delta z| \\
&= |\Delta x + \Delta y| \leq |\Delta x| + |\Delta y| \\
Maximum\ absolute\ error &= |\Delta x| + |\Delta y| \\
Relative\ Error &= \frac{(|\Delta x| + |\Delta y|)}{X - Y} \\
\implies \%age\ error &= \frac{(|\Delta x| + |\Delta y|)}{X - Y} \times 100.
\end{aligned}
$$

3. Multiplication,
   Using $X = x + \Delta x$ and $Y = y + \Delta y$
   Let $Z = XY$ and $z = xy$

$$
\begin{aligned}
\implies \Delta z &= Exact\ Value(Z) - Approximated\ Value(z) \\
&= (XY) - (xy) \\
&= (x + \Delta x)(y + \Delta y) - (xy) \\
&= xy + y\Delta x + x\Delta y + \Delta x\Delta y - xy \\
&= \Delta x + \Delta y + \Delta x\Delta y.
\end{aligned}
$$

$Since\ \Delta x\ and\ \Delta y$ are so small, then $\Delta x\Delta y \approx 0$

$$
\begin{aligned}
\therefore Absolute\ error &= |\Delta z| \\
&= |y\Delta x + x\Delta y| \leq y|\Delta x| + x|\Delta y| \\
Maximum\ absolute\ error &= y|\Delta x| + x|\Delta y| \\
Relative\ Error &= \frac{(y|\Delta x| + x|\Delta y|)}{xy} \\
&= \frac{y|\Delta x|}{xy} + \frac{x|\Delta y|}{xy} \\
&= \frac{|\Delta x|}{x} + \frac{|\Delta y|}{y} \\
\implies \%age\ error &= \left(\left|\frac{\Delta x}{x}\right| + \left|\frac{\Delta y}{y}\right|\right) \times 100.
\end{aligned}
$$

4. Division.
   Using $X = x + \Delta x$ and $Y = y + \Delta y$
   Let $Z = \dfrac{X}{Y}$ and $z = \dfrac{x}{y}$

$$
\begin{aligned}
\implies \Delta z &= Exact\ Value(Z) - Approximated\ Value(z)\\
&= \frac{X}{Y} - \frac{x}{y}\\
&= \frac{(x+\Delta x)}{(y+\Delta y)} - \frac{x}{y}\\
&= \frac{(x+\Delta x)(y-\Delta y)}{(y+\Delta y)(y-\Delta y)} - \frac{x}{y}
\end{aligned}
$$

Since Δx *and* Δy are so small, then $\Delta x \Delta y \approx 0$. $(\Delta y)^2 \approx 0$

$$
\begin{aligned}
\implies \Delta z &= \frac{(xy - x\Delta y + y\Delta x - xy)}{y^2}\\
&= \frac{(y\Delta x - x\Delta y)}{y^2}\\
\therefore Absolute\ error &= |\Delta z|\\
&= |\frac{y\Delta x}{y^2} + \frac{x\Delta y}{y^2}| \leq y|\frac{\Delta x}{y^2}| + x|\frac{\Delta y}{y^2}|\\
Maximum\ absolute\ error &= y|\frac{\Delta x}{y^2}| + x|\frac{\Delta y}{y^2}|\\
Relative\ Error &= y|\frac{\Delta x}{y^2}| + x|\frac{\Delta y}{y^2}| \div \frac{x}{y}\\
&= \frac{|\Delta x|}{x} + \frac{|\Delta y|}{y}\\
&= \frac{|\Delta x|}{x} + \frac{|\Delta y|}{y}\\
\implies \%age\ error &= \left(\left|\frac{\Delta x}{x}\right| + \left|\frac{\Delta y}{y}\right|\right) \times 100.
\end{aligned}
$$

(1) The numbers x and y were rounded off to give X and Y with errors E_1 and E_2 respectively. Obtain the expressions for maximum possible absolute error in the following expressions.

(a) $x + y$ (b) xy (c) $\frac{x^2}{y}$ (d) $x\sqrt{y}$

Solution

(a) $x + y$
Using $x = X + \Delta x$ and $y = Y + \Delta y$
Error = exact - approximate value

$$Let\ \text{exact}(n) = x + y, \text{approximate}(N) = X + Y$$

$$\Longrightarrow E_3 = [(X + E_1) + (Y + E_2)] - N$$
$$E_3 = (X + E_1) + (Y + E_2) - (X + Y)$$
$$= (X +_1) + (Y + E_2) - X - Y$$
$$= E_1 + E_2$$
$$\Longrightarrow absoluteerror = |E_3| = |E_1 + E_2| \leq |E_1| + |E_2|$$
$$\Longrightarrow MaximumAbsoluteerror|E_3| = |E_1| + |E_2|$$

(b) xy
Using $x = X + \Delta x$ and $y = Y + \Delta y$
Error = exact - approximate value

$$Let\ m = xy, M = XY$$
$$E_3 = (X + E_1)(Y + E_2) - M$$
$$= (X + E_1)(Y + E_2) - M$$
$$= XY + YE_1 + XE_2 + E_1E_2 - XY$$

Since E_1 and E_2 are very small numbers, thus $E_1E_2 \approx 0$

$$\therefore Absolute\ error = |E_3| = |YE_1 + XE_2| \leq |YE_1| + |XE_2|$$
$$\Longrightarrow Maximum\ bsolute\ error\ |E_3| = |YE_1| + |XE_2|$$

(c) $\frac{x^2}{y}$ Using $x = X + \Delta x$ and $y = Y + \Delta y$

$$Error = exact - approximatevalue$$
$$\text{Let exact}(n) = \frac{x^2}{y}, \text{approximate}(N) = \frac{X^2}{Y}$$
$$\Longrightarrow E_3 = \frac{(X + E_1)^2}{(Y + E_2)} - N$$
$$= \frac{(X^2 + 2XE_1 + E_1^2)(Y - E_2)}{(Y^2 - E_2^2)} - \frac{X^2}{Y}$$

Since E_1, E_2 and E_3 are very small numbers, thus $E_1^2 \approx 0$, $E_2^2 \approx 0$

$$\Longrightarrow E_3 = \frac{(X^2 + 2XE_1)(Y - E_2)}{Y^2} - \frac{X^2}{Y}$$
$$= \frac{(X^2 + 2XE_1)(Y - E_2)}{Y^2} - \frac{X^2}{Y}$$
$$= \frac{(X^2Y - X^2E_2 + 2XE_1Y - 2XE_1E_2)}{Y^2} - \frac{X^2}{Y}$$

Since E_1, E_2 and E_3 are very small numbers, then $\mathrm{E_1E_2} \approx 0$

$$\implies E_3 = \frac{2XE_1Y - X^2E_2}{Y^2}$$

$$\implies Absolute error|E_3| = \Big|\frac{2XE_1Y - X^2E_2}{Y^2}\Big| \leq \Big|\frac{2XE_1Y}{Y^2}\Big| + \Big|\frac{X^2E_2}{Y^2}\Big|$$

$$\implies \text{Maximum absolute error}|E_3| = \Big|\frac{2XE_1}{Y}\Big| + \Big|\frac{X^2E_2}{Y^2}\Big|.$$

(d) $x\sqrt{y}$
Using $x = X + \Delta x$ and $y = Y + \Delta y$
NB: For easy derivation, you should first eliminate the radicals(i.e square root, cube root etc)

$$\text{Let exact}(z) = x\sqrt{y}, \text{approximate}(Z) = X\sqrt{Y}$$

$$\text{From } Exact(z) = approximate(Z) + error(\Delta z)$$

$$(Z + \Delta z) = (X + \Delta x)\sqrt{(Y + \Delta y)},$$

Squaring both sides to eliminate the radical.

$$(Z + \Delta)^2 = (X + \Delta x)^2(Y + \Delta y)$$

$$2Z\Delta z + (\Delta z)^2 = (X^2 + 2X\Delta x + (\Delta x)^2)(Y + \Delta y)$$

$$Z^2 + 2Z\Delta z + (\Delta z)^2 = (X^2 + 2X\Delta x + (\Delta x)^2)(Y + \Delta y) - Z^2$$

$$2Z\Delta z + (\Delta z)^2 = X^2Y + 2XY\Delta x + Y(\Delta x)^2 + X^2\Delta y + 2X\Delta x\Delta y + (\Delta x)^2\Delta y - X^2Y$$

Since Δx and Δy are so small, then $(\Delta x)^2\Delta y \approx 0, 2X\Delta x\Delta y \approx 0, Y(\Delta x)^2 \approx 0$, $(\Delta z)^2 \approx 0$

$$\implies 2Z\Delta z = 2XY\Delta x + X^2\Delta y$$

$$\Delta z = \frac{2XY\Delta x + X^2\Delta y}{2Z}, \text{But } Z = X\sqrt{Y}$$

$$\Delta z = \frac{2XY\Delta x + X^2\Delta y}{2X\sqrt{Y}}$$

$$Absolute\ error|\Delta z| = \Big|\frac{2XY\Delta x + X^2\Delta y}{2X\sqrt{Y}}\Big| \leq \Big|\frac{2XY\Delta x}{2X\sqrt{Y}}\Big| + \Big|\frac{X^2\Delta y}{2X\sqrt{Y}}\Big|$$

$$\implies Max.\ absolute\ error|\Delta z| = \Big|\frac{2XY\Delta x}{2X\sqrt{Y}}\Big| + \Big|\frac{X^2\Delta y}{2X\sqrt{Y}}\Big|$$

Your free to reduce the above expression !

$$\text{From, } Max\ absolute\ error|\Delta z| = \Big|\frac{2XY\Delta x}{2X\sqrt{Y}}\Big| + \Big|\frac{X^2\Delta y}{2X\sqrt{Y}}\Big|$$

It can be deduced that relative error is given by

$$
\begin{aligned}
Relative Max.\ absolute\ error &= \frac{|\Delta z|}{Z} \\
&= \left(\left|\frac{2XY\Delta x}{2X\sqrt{Y}}\right| + \left|\frac{X^2\Delta y}{2X\sqrt{Y}}\right|\right) \div 2X\sqrt{Y} \\
&= \left|\frac{\Delta x}{X}\right| + \frac{1}{2}\left|\frac{\Delta y}{Y}\right|
\end{aligned}
$$

(2) The numbers A and B were rounded off to give a and b errors e_1 and e_2 respectively. Obtain the expressions for maximum possible relative error in the following expressions.

(i) $A - B$ (ii) $\frac{A}{B}$ (iii) A^2 B (iv) $\sqrt{AB}$

Solution Left for you.

(3) The height (h) and radius (r) of a cone are measured as e_h and e_r respectively. Show that the maximum percentage error in its volume (V) is given by $\left[2\left|\frac{e_r}{r}\right| + \left|\frac{e_h}{h}\right|\right] \times 100$

Solution

$$\begin{aligned}
\text{Let Volume}(V) &= \frac{1}{3}\pi r^2 h \\
V + e_v &= \frac{1}{3}\pi(r+e_r)^2(h+e_h) \\
e_v &= \frac{1}{3}\pi(r^2+2re_r+e_r^2)(h+e_h) - V
\end{aligned}$$

Since e_r, e_h and e_v are very small numbers, thus $e_r^2 \approx 0$

$$\begin{aligned}
\implies e_v &= \frac{1}{3}\pi(r^2+2re_r)(h+e_h) - \frac{1}{3}\pi r^2 h \\
e_v &= \frac{1}{3}\pi(r^2h + 2rhe_r + r^2e_h + 2re_re_h - r^2h)
\end{aligned}$$

Since e_r, e_h and e_v are very small numbers, then $e_re_h \approx 0$

$$\begin{aligned}
\implies |e_v| &= \frac{1}{3}\pi|2rhe_r + r^2e_h| \leq \frac{2}{3}\pi rh|e_r| + \frac{1}{3}\pi r^2|e_h| \\
\text{Maximum absolute error}|e_v| &= \frac{2}{3}\pi rh|e_r| + \frac{1}{3}\pi r^2|e_h| \\
\text{Relative error} &= \left|\frac{e_V}{V}\right| \\
&= \left(\frac{2}{3}\pi rh|e_r| + \frac{1}{3}\pi r^2|e_h|\right) \div \frac{1}{3}\pi r^2 h \\
&= 2\left|\frac{e_r}{r}\right| + \left|\frac{e_h}{h}\right| \\
\implies \%ge\ error &= \left(2\left|\frac{e_r}{r}\right| + \left|\frac{e_h}{h}\right|\right) \times 100 \\
&= \left(2\left|\frac{e_r}{r}\right| + \left|\frac{e_h}{h}\right|\right) \times 100\%
\end{aligned}$$

as required.

11.1.5 Error in a function

If the expression whose error function is required contains trigonometrical functions, exponential functions or logarithmic functions, then error derivations of such an expression is got by differentiation but maintaining the small changes approach assumption that your dealing with fixed numbers. The 1st derivative gives the absolute error. I.e

$$\begin{aligned}
\text{If} y = f(x) &\Longleftrightarrow \frac{dy}{dx} = f'(x)\\
\text{For small chages } \frac{dy}{dx} &\approx \frac{\Delta y}{\Delta x}\\
\therefore \frac{\Delta y}{\Delta x} &\approx f'(x)\\
\Longrightarrow \Delta y \approx f'(x)\Delta x &\Longleftrightarrow |\Delta y| = f'(x)|\Delta x|
\end{aligned}$$

E.g Given $y = sinx$,
Then $\Delta y = cosx.\Delta x \Longleftrightarrow$ Absolute error $= |\Delta y| = cosx|\Delta x|$
$\therefore$ Relative error $= \dfrac{|\Delta y|}{y} = \dfrac{(cosx)|\Delta x|}{sinx} = |\Delta x|cotx$

$\Longrightarrow$ %ge error $=$ Relative error $\times 100 = |\Delta x|cotx$.
NB: If your given two or more variables in the function, use partial differentiation as illustrated in the example below.
Examples:

(1) Quantity P is given by $P = x\tan\theta$. If the errors in x and θ are e_x and e_θ respectively. Obtain an expression for maximum possible absolute error.
Solution

$$\begin{aligned}
\text{From } P &= x\tan\theta\\
\Delta P &= \Delta x\tan\theta + \Delta\theta x\sec^2\theta\\
\text{Absolute error} = \left|\Delta P\right| &= \left|e_x\tan\theta + e_\theta x\sec^2\theta\right| \le \left|e_x\tan\theta\right| + \left|e_\theta x\sec^2\theta\right|\\
\text{Maximum absolute error} = \left|\Delta P\right| &= \left|e_x\right|\tan\theta + \left|e_\theta\right|x\sec^2\theta
\end{aligned}$$

(2) The area of a triangle (A) is given by $A = \frac{1}{2}xy\sin\theta$, show that the maximum possible relative error in A is given by $\left|\frac{\Delta x}{x}\right| + \left|\frac{\Delta y}{y}\right| + \cot\theta\left|\Delta\theta\right|$. where $\Delta x, \Delta y$ and $\Delta\theta$ are small compared to x, y and θ respectively.

Solution

$$
\begin{aligned}
\text{From } A &= \frac{1}{2}xy\sin\theta \\
\Delta A &= \frac{1}{2}y\sin\theta\Delta x + \frac{1}{2}x\sin\theta\Delta y + \frac{1}{2}xy\cos\theta\Delta\theta \\
\implies \big|\Delta A\big| &= \Big|\frac{1}{2}y\sin\theta\Delta x + \frac{1}{2}x\sin\theta\Delta y + \frac{1}{2}xy\cos\theta\Delta\theta\Big| \\
\Big|\frac{1}{2}y\sin\theta\Delta x + \frac{1}{2}x\sin\theta\Delta y + \frac{1}{2}xy\cos\theta\Delta\theta\Big| &\leq \Big|\frac{1}{2}y\sin\theta\Delta x\Big| + \Big|\frac{1}{2}x\sin\theta\Delta y\Big| + \Big|\frac{1}{2}xy\cos\theta\Delta\theta\Big| \\
\therefore \text{Max. absolute error} &= \Big|\frac{1}{2}y\sin\theta\Delta x\Big| + \Big|\frac{1}{2}x\sin\theta\Delta y\Big| + \Big|\frac{1}{2}xy\cos\theta\Delta\theta\Big| \\
\text{Ralative absolute error}\Big(\Big|\frac{\Delta A}{A}\Big|\Big) &= \Bigg|\frac{\Big|\frac{1}{2}y\sin\theta\Delta x\Big| + \Big|\frac{1}{2}x\sin\theta\Delta y\Big| + \Big|\frac{1}{2}xy\cos\theta\Delta\theta\Big|}{\frac{1}{2}xy\sin\theta}\Bigg| \\
\Big|\frac{\Delta A}{A}\Big| &= \Big|\frac{y\sin\theta\Delta x}{xy\sin\theta}\Big| + \Big|\frac{x\sin\theta\Delta y}{xy\sin\theta}\Big| + \Big|\frac{xy\cos\theta\Delta\theta}{xy\sin\theta}\Big| \\
&\text{On simplification,} \\
\text{Ralative absolute error} = \Big|\frac{\Delta A}{A}\Big| &= \Big|\frac{\Delta x}{x}\Big| + \Big|\frac{\Delta y}{y}\Big| + \cot\theta\big|\Delta\theta\big|
\end{aligned}
$$

(3) The quantity z is given by $z = ye^{-2x}$ and the errors in x and y are Δ_1 and Δ_2 respectively. Find the expression for percentage error in z.

Solution

$$
\begin{aligned}
From\ z &= ye^{-2x} \\
\Delta z &= -2ye^{-2x}\Delta_1 + e^{-2x}\Delta_2 \\
\implies \big|\Delta z\big| &= \Big| - 2ye^{-2x}\Delta_1 + e^{-2x}\Delta_2\Big| \leq \Big| - 2ye^{-2x}\Delta_1\Big| + \Big|e^{-2x}\Delta_2\Big| \\
\therefore \text{Max. absolute error} &= 2ye^{-2x}\Big|\Delta_1\Big| + e^{-2x}\Big|\Delta_2\Big| \\
\text{Ralative absolute error} = \Big|\frac{\Delta z}{z}\Big| &= \Bigg|\frac{2ye^{-2x}\Big|\Delta_1\Big| + e^{-2x}\Big|\Delta_2\Big|}{ye^{-2x}}\Bigg| \\
\Big|\frac{\Delta z}{z}\Big| &= \big|2\Delta_1\big| + \Big|\frac{\Delta_2}{y}\Big| \\
&\text{On simplification,} \\
\Big|\frac{\Delta z}{z}\Big| &= 2\big|\Delta_1\big| + \Big|\frac{\Delta_2}{y}\Big| \\
\implies \%\text{age error} &= \Big[2\big|\Delta_1\big| + \Big|\frac{\Delta_2}{y}\Big|\Big] \times 100
\end{aligned}
$$

11.1.6 Special Techniques/More examples

Drawing conclusions from error propagation,

In general, we can deduce the expression of Relative and %age error given the function.

This however works when the experession of the function is in either division or multiplication.
I.e

- If $p = \frac{x^n\sqrt{z}}{y^m} \Longrightarrow$ Ralative absolute error$\left(\dfrac{\Delta p}{p}\right) = \dfrac{1}{n}\left|\dfrac{\Delta x}{x}\right| + \dfrac{1}{2}\left|\dfrac{\Delta z}{z}\right| + \dfrac{1}{m}\left|\dfrac{\Delta y}{y}\right|$
- If $Z = \pi\frac{x^n\sqrt[k]{y}}{t^m} \Longrightarrow$ Ralative absolute error$\left(\dfrac{\Delta Z}{Z}\right) = \dfrac{1}{n}\left|\dfrac{\Delta x}{x}\right| + \dfrac{1}{k}\left|\dfrac{\Delta y}{y}\right| + \dfrac{1}{m}\left|\dfrac{\Delta t}{t}\right|$

NB: Any constant in the given formula vanishes in the the expression of relative error or in the %age error. Like π in the above formula and also if $V = \frac{1}{3}\pi r^2 h$, then basing on this formula, $\frac{1}{3}$ and π are constants indicating that they must vanish and therefore Relative error is given by $2\left|\frac{\Delta r}{r}\right| + \left|\frac{\Delta h}{h}\right|$
Examples

1. The numbers $x = 1.65, y = 2.6$ and $z = 4.3$ were each rounded to the given decimal places. Find the percentage error in $\frac{x^3\sqrt{z}}{y^4}$

   Solution

   $x = 1.65, e_1 = 0.005$
   $y = 2.6, e_2 = 0.05$
   $z = 4.3, e_3 = 0.05$

   Method 1 (Deducing from standard error expression proof)

   $\%\text{ error} = \left[3\left|\frac{e_1}{x}\right| + \frac{1}{2}\left|\frac{e_3}{z}\right| + 4\left|\frac{e_2}{y}\right|\right] \times 100$

   $\%\text{ error} = \left[3\left|\frac{0.005}{1.65}\right| + \frac{1}{2}\left|\frac{0.05}{4.3}\right| + 4\left|\frac{0.05}{2.6}\right|\right] \times 100$

   $\%\text{ error} = 9.183$

   Method 2 (Deducing from maximum and minimum)

   $\%\text{age error} = \frac{\text{Absolute error}}{\text{working value}} \times 100$
   But working value $W = \frac{x^3\sqrt{z}}{y^4} = \frac{(1.65)^3\sqrt{4.3}}{(2.6)^4} = 0.203842$

   $W_{Max} = \frac{(1.655)^3\sqrt{4.35}}{(2.55)^4} = 0.223603$

   $W_{Min} = \frac{(1.645)^3\sqrt{4.25}}{(2.65)^4} = 0.186084$

   Absolute error $= \frac{|W_{Max} - W_{Min}|}{2} = \frac{|0.223603 - 0.186084|}{2} = 0.01876$

$\Longrightarrow$ %ge error $= \frac{0.018760}{0.203842} \times 100 = 9.203$

2. Given that $z = mn\cos\theta$ is such that $m = 2.5 \pm 0.04, n = 3.34 \pm 0.001$ and $\theta = 30^0$. Find the limit within which the exact value of the z lies.

   Solution

   Method 1 Deducing absolute error from Relative error.
   $m = 2.5 \Longrightarrow \Delta m = 0.04$
   $n = 3.34 \Longrightarrow \Delta n = 0.001$
   $\theta = 30^0 \Longrightarrow \Delta\theta = 0.5^0 = \frac{\pi}{360}$ (to radians)
   Let the working value $z = mn\cos\theta = 2.5 \times 3.34 \times \cos 30 = 7.23131$
   For absolute error, from $z = mn\cos\theta \Longleftrightarrow \left|\frac{\Delta z}{z}\right| = \left|\frac{\Delta m}{m}\right| + \left|\frac{\Delta n}{n}\right| + \left|\frac{-\sin\theta\Delta\theta}{\cos\theta}\right|$

$$\Longrightarrow |\Delta z| = \left(\left|\frac{\Delta m}{m}\right| + \left|\frac{\Delta n}{n}\right| + \left|\frac{\sin\theta\Delta\theta}{\cos\theta}\right|\right) \times z$$

$$\text{Absolute error}|\Delta z| = \left(\left|\frac{0.04}{2.5}\right| + \left|\frac{0.001}{3.34}\right| + \left|\frac{\frac{\pi}{360}\sin(30)}{\cos 30}\right|\right) \times 7.23131$$

$$= 0.1543.$$

$$\text{Lower limit} = 7.23131 - 0.15430 = 7.0770$$

$$\text{Upper limit} = 7.23131 + 0.15430 = 7.3856.$$

   Method 2 I.e From maximum and minimum.
   $m = 2.5 \Longrightarrow \Delta m = 0.04$
   $n = 3.34 \Longrightarrow \Delta n = 0.001$
   $\theta = 30^0 \Longrightarrow \Delta\theta = 0.5^0 = \frac{\pi}{360}$ (to radians)
   Let the working value $z = mn\cos\theta = 2.5 \times 3.34 \times \cos 30 = 7.23131$
   $z_{\min} = (2.5 - 0.04) \times (3.34 - 0.001) \times \cos(30 + 0.5) = 7.07737$
   $z_{\max} = (2.5 + 0.04) \times (3.34 + 0.001) \times \cos(30 - 0.5) = 7.38600$
   Absolute error $= \frac{|z_{\max} - z_{\min}|}{2} = \frac{|7.38600 - 7.07737|}{2} = 0.1543$
   Lower limit $= 7.23131 - 0.1543 = 7.0770$
   Upper limit $= 7.23131 + 0.1543 = 7.3856.$

3. Given that $A = xy\sin\theta$ is such that $x = 2.5, y = 3.4$ and $\theta = 30^0$ each rounded off to the given number of decimal places. Find the limit within which the exact value of the A lies.

   Solution

   Method 1 Deducing absolute error from Relative error.
   $x = 2.5 \Longrightarrow \Delta x = 0.05$
   $y = 3.4 \Longrightarrow \Delta y = 0.05$
   $\theta = 30^0 \Longrightarrow \Delta\theta = 0.5^0 = \frac{\pi}{360}$ (to radians)
   Let the working value $A = xy\sin\theta = 2.5 \times 3.4 \times \sin 30 = 4.250$

For absolute error, from $A = xy\sin\theta \Longleftrightarrow \left|\frac{\Delta A}{A}\right| = \left|\frac{\Delta x}{x}\right| + \left|\frac{\Delta y}{y}\right| + \left|\frac{\cos\theta\Delta\theta}{\sin\theta}\right|$

$$\Longrightarrow |\Delta A| = \left(\left|\frac{\Delta x}{x}\right| + \left|\frac{\Delta y}{y}\right| + \left|\frac{\cos\theta\Delta\theta}{\sin\theta}\right|\right) \times A$$

$$\text{Absolute error}|\Delta A| = \left(\left|\frac{0.05}{2.5}\right| + \left|\frac{0.05}{3.4}\right| + \left|\frac{\frac{\pi}{360}\cos(30)}{\sin 30}\right|\right) \times 4.250$$

$$= 0.2117.$$

$$\text{Lower limit} = 4.250 - 0.2117 = 4.038$$

$$\text{Upper limit} = 4.250 + 0.2117 = 4.462.$$

Method 2 I.e From maximum and minimum

$x = 2.5 \Longrightarrow \Delta x = 0.05$

$y = 3.4 \Longrightarrow \Delta y = 0.05$

$\theta = 30^0 \Longrightarrow \Delta\theta = 0.5^0 = \frac{\pi}{360}$ (to radians)

Let the working value $A = xy\sin\theta = 2.5 \times 3.4 \times \sin 30 = 4.250$

$A_{\min} = (2.5 - 0.05) \times (3.4 - 0.05) \times \sin 29.5 = 4.0416$

$A_{\max} = (2.5 + 0.05) \times (3.4 + 0.05) \times \sin 30.5 = 4.4651$

$\Longrightarrow$Absolute error $= \frac{|A_{\max} - A_{\min}|}{2} = \frac{|4.4651 - 4.0416|}{2} = 0.21175$

Lower limit $= 4.250 - 0.21175 = 4.038$

Upper limit $= 4.250 + 0.21175 = 4.462.$

4. The cylindrical pipe has a height of $3.6m$ measured to the nearest units. If the relative absolute error made in calculating its volume is 0.213. Find the relative absolute error made in measuring its radius.

   Solution.

   let $h = 3.6,\ \Delta h = 0.05,$ and $\left|\frac{\Delta V}{V}\right| = 0.213$

   Let Volume$(V) = \pi r^2 h$

   $\Longrightarrow \left|\frac{\Delta V}{V}\right| = 2\left|\frac{\Delta r}{r}\right| + \left|\frac{\Delta h}{h}\right|$

   $0.213 = 2\left|\frac{\Delta r}{r}\right| + \frac{0.05}{3.6} \Longleftrightarrow \left|\frac{\Delta r}{r}\right| = \frac{(0.213 - \frac{1}{72})}{2} = \frac{112}{1125} = 0.09956.$

5. Given that $a = 2.5, b = 5.55$ are rounded off to the given number of decimal places, determine the interval within which the exact value of;

   $(i)\frac{(b-a)}{(a+b)}$,

   $(ii)\frac{(a-b)}{(a+b)}$,

   $(iii)\frac{(a-b)}{(ab)}$,

   $(iv)\frac{(a+b)}{(a-b)}$ *Lies.* [*Reference question UNEB* 2010 *Number* 14$b(ii)$]

   Solution:

   $a = 2.5 \Longrightarrow \Delta a = 0.05$ and $b = 5.55 \Longrightarrow \Delta a = 0.005$

$(i)\dfrac{(b-a)}{(a+b)}$ =?

Let $P = \dfrac{(b-a)}{(a+b)}$

$\implies P_{min} = \dfrac{(b-a)_{min}}{(a+b)_{max}} = \dfrac{(5.545-2.55)}{(2.55+5.555)} = 0.36952.$

Similarly, $P_{max} = \dfrac{(b-a)_{max}}{(a+b)_{min}} = \dfrac{(5.555-2.45)}{(2.45+5.545)} = 0.38837.$

Thus [0.36952, 0.38837].

$(ii)\dfrac{(a-b)}{(a+b)}$ =?

Let $Z = \dfrac{(a-b)}{(a+b)}$

NB: We adjust the rule minimizing and maximizing whenever a negative value is developed within the expression when the given values are not negative.

$\implies Z_{min} = \dfrac{(a-b)_{min}}{(a+b)_{max}} = \dfrac{(a-b)_{min}}{(a+b)_{min}} = \dfrac{(2.45-5.555)}{(2.45+5.545)} = -0.3884.$

Similarly, $Z_{max} = \dfrac{(a-b)_{max}}{(a+b)_{min}} = \dfrac{(a-b)_{max}}{(a+b)_{max}}\dfrac{(2.55-5.545)}{(2.55+5.555)} = -0.3695.$

Thus [-0.3884, -0.3695].

$(iii)\dfrac{(a-b)}{(ab)}$ =?

Let $P = \dfrac{(a-b)}{(a\times b)}$

$\implies P_{min} = \dfrac{(a-b)_{min}}{(a\times b)_{max}} = \dfrac{(a-b)_{min}}{(a\times b)_{min}} = \dfrac{(2.45-5.555)}{(2.45\times 5.545)} = -0.2286.$

Similarly, $P_{max} = \dfrac{(a-b)_{max}}{(a\times b)_{min}} = \dfrac{(a-b)_{max}}{(a\times b)_{max}}\dfrac{(2.55-5.545)}{(2.55\times 5.555)} = -0.2114.$

Thus [-0.2114, -0.2286].

Please the last part.

11.1.7 Exercise 9

(1) The cylindrical pipe has a height of $7.76m$ measured to the nearest units. If the relative absolute error made in calculating its volume is 0.3. Find the absolute relative error made in measuring its radius.
Ans: 0.1497.

(2) The length of a rectangle was measured as $2.22m$ instead of $2.61m$. Find the;

(i) Absolute error, (ii) Relative error, (iii) percentage error.

Ans: (i) 0.390 (ii) 0.1494 and (iii) 14.94%

(3) The radius of a circle was measured as $10.42m$ instead of $10.00m$. Find the;

(i) Absolute error, (ii) Relative error, (iii) percentage error.

Ans: (i) 0.420 (ii) 0.0420 and (iii) 4.20%

(4) Obtain the interval within which the exact value of $4.53 + \dfrac{(1.102 \times 1.53)}{0.9772}$ lies. Hence, obtain the percentage error in made in approximating the value.
Ans: [6.232, 6.5260] hence 3.64%

(5) The value of $L = 100.23m$ was obtained when measuring the length of the football pitch. Given that, the percentage error in this value was 0.04%. Find the range within which the value of L lies.
Ans: $100.1899 \leq L \leq 100.2701$

(6) (a) Obtain the limits within which the exact value of $P = 1.467\Big(\dfrac{7.225 + 4.8006}{3.92 \times 2.0}\Big)$ lies. Each of the values were rounded to the given number of decimal places.

(b) Find the percentage error in approximating P.
Ans: (a) Lower limit 2.9610, Upper limit 3.0428, (b) 1.3629

(7) Determine the maximum percentage absolute error in $\dfrac{\sqrt{z}}{(x^2y^3)}$, given that $x = 2.41, y = 5.011$ and $z = 2.45$ all numbers rounded off to the given number of decimal places.
Ans: 0.0%

(8) Given numbers $x = 4.24$ and $y = 3.80$, with percentage errors 0.4% and 0.6% respectively. Find the percentage error in;

(a) $Z = x^3y$,

(b) $Z = \dfrac{3}{x-y}$

(c) $Z = \dfrac{1}{x} - \dfrac{1}{x-y} - \dfrac{1}{xy}$

(d) $\frac{x^2}{y-x}$, correct to four decimal places.
Ans: (a) 0.4854%, (b) 2.273%, (c) 2.2482% and (d) 2.5100%

(9) The numbers $A = 10.4, B = 18.433$ and $C = 4.25$ are each rounded off with percentage errors $3.5\%, 1.05\%$ and 1.0% respectively. Find the;

(i) limits within which the exact value of $\dfrac{\sqrt{A}}{(B-C)^2}$ lies,

(ii) percentage error in $\dfrac{\sqrt{A}}{(B-C)^2}$.

(10) The sides of a rectangle are measured as $6.026m$ and $7.5m$, to the nearest centimeters.

(a) Calculate the maximum possible error in each side,

(b) Calculate the maximum and minimum values of the perimeter. Hence, determine the absolute error in the perimeter,

(c) Calculate the percentage error in the area.

(11) Real numbers R and T are approximated to r and t with maximum possible errors of Δr and Δt respectively.

(i) Show that the maximum possible percent error made in computing $\frac{T}{\sqrt{R}}$ is $\left(\left|\frac{\Delta t}{t}\right| + \frac{1}{2}\left|\frac{\Delta r}{r}\right|\right) \times 100$. State any assumptions made.

(ii) Deduce an expression for the relative error in $\frac{y^{\frac{1}{5}}\sqrt{R^3}}{\sqrt{T}}$. Given that Δy is the maximum possible error in y.

(12) Given that the error in each of the values of e^x and ee^{-x} is ± 0.0005, taking $x = 0.04$, find the minimum and the maximum value of the quotient $\dfrac{e^x}{e^{-x}}$.
Ans: 1.082, 1.084.

(13) (a) Decimal numbers x, y and z are approximated to X, Y and Z with maximum possible errors Δx, Δy and Δz. Show that the maximum percentage relative error made in computing $\frac{xy}{z}$ is $\left(\left|\frac{\Delta x}{X}\right| + \left|\frac{\Delta y}{Y}\right| + \left|\frac{\Delta z}{Z}\right|\right) \times 100$.

(b) If $X = 2.4, Y = 10.40$ and $Z = 0.008$, deduce an expression for the percentage relative error in, $\frac{x^2 z^3}{\sqrt{y}}$, hence determine this percentage error.
Ans: 22.9407

(14) A right angled triangle has height measured as 4.2cm and its bases as 3.0cm . Determine the percentage error made in calculating its area
Ans: error = 2.857%

(15) Given that $X = 2.504, Y = 4.41 \pm 0.02$ and $Z = 1.8 \pm 20\%$, for X rounded off to the nearest number of decimal places. Find the:

(i) maximum possible error in the expression $\frac{X\sqrt{Y}}{Z^2}$,

(ii) Limits within which the exact value of $\frac{X\sqrt{Y}}{Z^2}$ lies.
Ans: (i) 0.36466 (ii) (1.2583, 1.98762)

(16) The numbers 2.6954, 4.6006, 16.175 and 0.82 have been rounded to the given number of decimal places. Find the range of values within which the exact values of $2.6954\left(4.6006 - \frac{16.175}{0.82}\right)$ lies.
Ans: [-41.09669, -40.44320]

(17) Given that $A = |a||b|\sin\theta$ where θ is the angle between the vectors. Given that $|a| = 7.5cm, |b| = 6.64cm$ and $\theta = 30^0$, all rounded off to the given number of decimal places.
find

(i) error made in A,

(ii) values within which A lies.
Ans: (i) 0.21174 (ii) (4.03826, 4.46174)

(18) Given that $a = 6.150, b = -5.26$ and the relative errors for a and b are 0.01% and 0.02% respectively. State the range within which $\frac{a+b}{ab}$ lies.
Ans: [-0.02758, -0.02745]

(19) If $x = 4.0, y = 2.50$ are rounded off to the given number of decimal points, find the interval within which the exact value of $x + \frac{xy}{x-y}$ lie.
Ans: [10.2877, 11.07093]

(20) The numbers A and B are rounded off to a and b with errors e_1 and e_2, respectively.

(i) Show that the maximum relative error made in the approximation of $\frac{A}{B}$ by $\frac{a}{b}$ is $\left|\frac{e_1}{a}\right| + \left|\frac{e_2}{b}\right|$,

(ii) If also the number C is rounded off to c with error e_3 deduce the expression for the maximum relative error in taking the approximation of $\frac{A+C}{B+C}$ as $\frac{a+c}{b+c}$ in terms of e_1, e_2, e_3, a, b, and c.

(iii) 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.
Ans: [2.95157, 2.95576]

(21) A school canteen started its operation with a capital of 2 million shillings. At the end of a certain term its profits in soft drinks and foodstuff were 0.2 million and 0.45 million respectively. There were possible errors of 6.5% and 8.2% in the sales respectively. Find the maximum and minimum values of the sales as a percentage of the capital.
Ans:

(22) A trader in face masks and sanitizers makes annual profits from face masks of $sh.80$ million with a margin of error of $sh.700,000$ and annual profits from sanitizers of $sh.45$ million with a margin error of $sh.600,000$.

(i) Find the range of values corresponding to her gross income.

(ii) Find the relative error in the gross income to three decimal places.

(23) The height of a triangle is given $h = |a| \sin\theta$ were $|a| = 2.54$ and $\theta = 30^0$.

(i) Find the limits within where the exact value of the height lies.

(ii) Determine the percentage error in h.
Ans: (i) (1.2483, 1.2917), (ii)

(24) If $P = \frac{15.36+27.1-1.672}{2.36\times1.043}$ all numbers corrected to given decimal places. Find a and b such that $a \leq P \leq b$. Hence determine the absolute error in P.
Ans: a = 16.5051, b = 16.6363 and absolute error = 0.0656

(25) The y coordinate of a hyperbola is given by $b \sec \phi$,

(a) Find an expression for the percentage error in y,

(b) Given $\phi = 30^0 30'$ and b $= 4.2 \pm 0.3$, find the range with in which the exact value of y lies.

Ans: [4.52371, 5.22527]

(26) The base and height of a triangle were measured as $5.1m$ and $6.4m$. The relative error in the base was 0.04 while the percentage error in the height was 5%. Find the:

(i) absolute error in each dimension,

(ii) percentage error in the area,

(iii) limits within which the value of its area lies.

Ans: (i) 0.204 , 0.32 (ii) 9 (iii) 17.1328

Chapter 12

APPROXIMATE NUMERICAL METHODS.

12.1 Methods of Finding the Root(s) of the Equation

When the root has been located we can use the following methods to obtain a better approximation based on either a given number of iterations or number of decimal places(or significant figures.)
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 that can be used in finding the real root(s) include the following:

- Repeated bisection,
- Linear interpolation,
- Newton Rampson method(NRM),
- General iterative method.

Under this case, we must test for the convergence of root(s) basing on 0.5×10^{-n} where $n =$ Number of decimal places or the number of iterations.

12.1.1 Repeated bisection

Suppose we want to find the solution to the equation $f(x) = 0$, where f is continuous. Given a function $f(x)$ continuous on an interval $[a_0, b_0]$ and satisfying $f(a) \times f(b) < 0$. For $n = 0, 1, 2, \cdots$ until convergent. Where $n =$ (Number of iterations + 1)

Compute $x_r = \dfrac{1}{2}(a_0 + b_0)$. Accept x_r after the required number of iterations or if

$|x_{(r)} - x_{(r-1)}| \leq \epsilon$ where $\epsilon = 0.5 \times 10^{-n}$ for $n =$ number of dps

Examples.

(1) (a) Show that $f(x) = x^3 - x - 1$ has areal root of between 1 and 1.6.

(b)Find a real root of the equation $f(x) = x^3 - x - 1 = 0$. using bisection method to 1 decimal places.

Solution.

(a) let $f(x) = x^3 - x - 1$

When $x = 1,\ f(1) = (1)^3 - (1) - 1 = -1$

When $x = 1.6,\ f(1.6) = (1.6)^3 - (1.6) - 1 = 1.490$

Since $f(1) \times f(1.6) < 0$, then there is a root of the equation $x^3 - x - 1$.

(b) let $f(x) = x^3 - x - 1$ $\epsilon = 0.5 \times 10^{-1} = 0.05$ and $x_0 = \dfrac{(1 + 1.6)}{2} = 1.30$

$\Longrightarrow f(x_0) = f(1.30) = (1.30)^3 - (1.30) - 1 = -0.103$

Since $f(1.50)$ is negative,then take $x = 1.6$ and $x_0 = 1.3$

$$\Longrightarrow x_1 = \frac{(1.3 + 1.6)}{2} = 1.450$$

Error $= |x_1 - x_0| = |1.450 - 1.3| = 0.150 > 0.05$

$f(1.450) = (1.450)^3 - (1.450) - 1 = 0.599$

$$x_2 = \frac{(1.3 + 1.450)}{2} = 1.375$$

$|x_2 - x_1| = |1.375 - 1.450| = 0.075 > 0.05$

$f(1.375) = (1.375)^3 - (1.375) - 1 = 0.2246$

$$x_3 = \frac{(1.375 + 1.30)}{2} = 1.338$$

$|x_3 - x_2| = |1.338 - 1.375| = 0.037 < 0.05$

$\Longrightarrow$ The root $= 1.3(1dp)$

NB: Here we are required to give the answer after a given number of iterations.

(2) Find a positive root of the equation $xe^x = 1$, which lies between 0.4 and 0.7 using the bisection method and give the answer to after three iterations.

Solution

let $f(x) = xe^x - 1$

When $x = 0.4,\ f(0.4) = (0.4)e^{0.4} - 1 = -0.403$

When $x = 0.7,\ f(0.7) = (0.7)e^{0.7} - 1 = 0.410$

$$x_0 = \frac{(0.4 + 0.7)}{2} = 0.5500$$

$f(0.550) = (0.550)e^{0.550} - 1 = -0.047$

$$x_1 = \frac{(0.550 + 0.7)}{2} = 0.6250$$

$f(0.6250) = (0.6250)e^{0.6250} - 1 = 0.1677$

$x_2 = \dfrac{(0.550 + 0.6250)}{2} = 0.5875$
$\implies$ The root $= 0.588(3dp)$
NB: Here we are required to give the answer after a given number of iterations.

12.1.2 Linear Interpolation

Since the root of the function $f(x)$ is the value of x say x_0 such that $f(x_0) = 0$, then to find x_0 by linear interpolation, we employ the following table.
I.e Let $f(x)$ be an equation with the root within x_1 and x_2 such that $f(x_1) \times f(x_2) < 0$,

x_1	x_0	x_2
$f(x_1)$	0	$f(x_2)$

then by linear interpolation, we can obtain x_r for $r = 0, 1, 2, \cdots$ until convergent. Accept x_r after the required number of iterations or if $|x_{(r)} - x_{(r-1)}| \le \epsilon$ where $\epsilon = 0.5 \times 10^{-n}$ for $n =$ number of dps.
Examples:

(1) (a) Show that the equation $x^3 - 2x^2 + 4 = 0$ has a root between -2 and -1,
(b) Use linear interpolation to find the root to 2 decimal places.

Solution.
(a) $f(x) = x^3 - 2x^2 + 4$
$f(-2) = (-2)^3 - 2(-2)^2 + 4 = -12$
$f(-1) = (-1)^3 - 2(-1)^2 + 4 = 1$
Since $f(-2) \times f(-1) < 0$, thus there is a root of the equation $x^3 - 2x^2 + 4 = 0$.

(b) Using

-2	x_0	-1
-12	0	1

$\therefore \dfrac{x_0 -^- 2}{0 -^- 12} = \dfrac{^-1 -^- 2}{1 -^- 12} \implies x_0 =^- 1.077$
$f(^-1.077) = (^-1.077)^3 - 2(^-1.077)^2 + 4 = 0.431$

From

-2	x$_1$	-1.077
-12	0	0.431

Using $\dfrac{x_1 -^- 2}{0 -^- 12} = \dfrac{^-1.077 -^- 2}{0.431 -^- 12} \implies x_1 =^- 1.109$
Error bound $= \left|x_1 - x_0\right| = \left|^-1.109 -^- 1.077\right| = 0.032 > 0.005$
$f(^-1.109) = (^-1.109)^3 - 2(^-1.109)^2 + 4 = 0.176$

From

-2	x_2	-1.109
-12	0	0.176

, Using $\dfrac{x_2 -^- 2}{0 -^- 12} = \dfrac{^-1.109 -^- 2}{0.176 -^- 12} \implies x_1 =^- 1.122$
$\left|x_2 - x_1\right| = \left|^-1.122 -^- 1.109\right| = 0.013 > 0.005$
$f(^-1.122) = (^-1.122)^3 - 2(^-1.122)^2 + 4 = 0.070$

From

-2	x_3	-1.122
-12	0	0.070

, Using $\dfrac{x_3 -{}^{-}2}{0 -{}^{-}12} = \dfrac{{}^{-}1.122 -{}^{-}2}{0.070 -{}^{-}12} \Longrightarrow x_3 ={}^{-}1.127$

$\left|x_3 - x_2\right| = \left|{}^{-}1.127 -{}^{-}1.122\right| = 0.005 \leq 0.005$

$\Longrightarrow$Thus the root is ${}^{-}1.13$ to 2 decimal places

(2) (a) Locate the root of the equation $e^x - \cos x - 2 = 0$ in the interval $-2 \leq x < 2$

(b) Use linear interpolation twice to estimate the root to three decimal place.

Solution

(a) Let $f(x) = e^x - \cos x - 2$

x	-2	-1	0	1
$f(x)$	5.805	0.178	-2.0	-2.172

Since $f({}^{-}1) \times f(0) < 0$, then there is a root of the equation $f(x) = e^x - \cos x - 2$ between $x ={}^{-}1$ and $x = 0$.

(b) Take

-1	x0	0
0.178	0	-2.0

, Using $\dfrac{x_0 -{}^{-}1}{0 - 0.178} = \dfrac{0 -{}^{-}1}{{}^{-}2.0 - 0.178}$

$\Longrightarrow x_0 ={}^{-}0.9200$

$f({}^{-}0.9200) = e^{(-0.9200)} - \cos({}^{-}0.9200) - 2 = -2.2073.$

From

-1	x1	${}^{-}0.9200$
0.178	0	-2.2073

, Using $\dfrac{x_0 -{}^{-}1}{0 - 0.178} = \dfrac{{}^{-}0.9200 -{}^{-}1}{-2.2073 - 0.178}$

$\Longrightarrow x_0 ={}^{-}0.994$

$\Longrightarrow$ The root is ${}^{-}0.994$.

(3) (a) Show that the equation $3x^2 + x - 5 = 0$ has areal between $x = 1$ and $x = 1.5$.

(b) Use linear interpolation to calculate the root to $2dps$.

Solution:

(a) Let $f(x) = 3x^2 + x - 5$

When $x = 1,\ f(1) = 3(1)^2 + 1 - 5 = -1$

When $x = 1.5,\ f(1.5) = 3(1.5)^2 + 1.5 - 5 = 3.25$

Since $f(1) \times f(1.5) < 0$, then there is a root of the equation $3x^2 + x - 5 = 0$.

(b) Using

1	x_0	1.5
-1	0	3.250

$\therefore \dfrac{x_0 - 1}{0 -{}^{-}1} = \dfrac{1.5 - 1}{3.250 -{}^{-}1} \Longrightarrow x_0 = 1.1176$

$f(1.1176) = 3(1.1176)^2 + 1.1176 - 5 = -0.1353$

From

1.1176	x1	1.5
-0.1353	0	3.250

Using $\dfrac{x_1 - 1.1176}{0 -{}^{-}0.1353} = \dfrac{1.50 - 1.1176}{3.250 -{}^{-}0.1353} \Longrightarrow x_1 = 1.1329$

Error bound $= \left|x_1 - x_0\right| = \left|1.1329 - 1.1176\right| = 0.0153 > 0.005$

$f(1.1329) = 3(1.1329)^2 + 1.1329 - 5 = -0.0167$

From

1.1329	x$_2$	1.5
-0.0167	0	3.250

Using $\dfrac{x_2 - 1.1329}{0 -^{-} 0.0167} = \dfrac{1.50 - 1.1329}{3.250 -^{-} 0.0167} \implies x_1 =$ 1.1348

Error bound $= \left|x_2 - x_1\right| = \left|1.1348 - 1.1329\right| = 0.0019 \leq 0.005$

$\implies$ The root is 1.13.

12.1.3 The Newton Raphson's Method (N.R.M)

For any function $f(x)$ where $f'(x)$ is it's first derivative and exist, then the Newton Raphson's Method is given by:

$$x_{n+1} = x_n - \frac{f(x_n)}{f'(x_n)},$$

for $\quad n = 0, 1, 2, 3, \cdots$ until convergent and $f'(x_n) \neq 0$

Proof.

Let $(x_n, f(x_n))$ and $(x_{n+1}, f(x_{n+1}))$ be two points on the curve $y = f(x)$ as illustrated below.

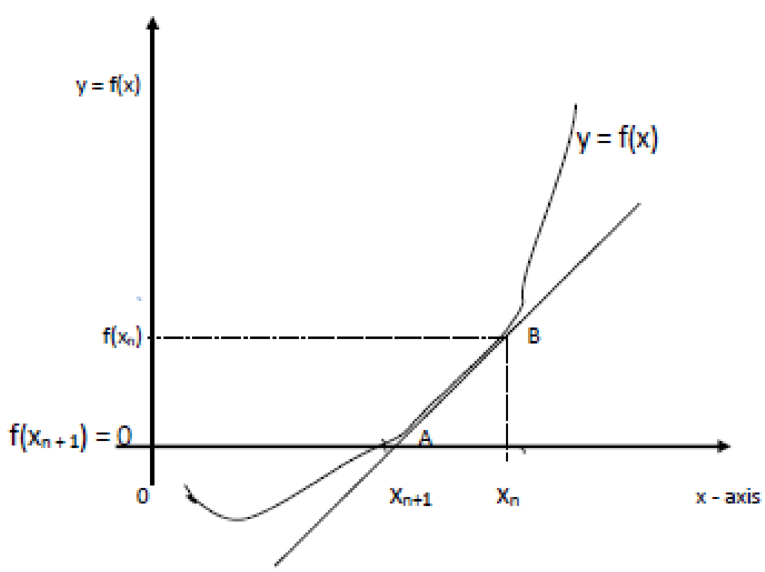


Figure 12.1

$$
\begin{aligned}
\text{Gradiant of AB} &= \frac{\Delta y}{\Delta x} \\
&= \frac{f(x_n) - f(x_{n+1})}{x_n - x_{n+1}} \\
&= \frac{f(x_n) - 0}{x_n - x_{n+1}} \\
&= \frac{f(x_n)}{x_n - x_{n+1}}
\end{aligned}
$$

Similary, the gradient of $y = f(x)$ is $\dfrac{dy}{dx} = f'(x)$ while at any point x_n is given as $\dfrac{dy}{dx} = f'(x_n)$
Because we are dealing with the same curve $y = f(x)$, then we have the same gradients. I.e

$$
\begin{aligned}
\text{Gradiant of AB} &= \frac{dy}{dx} \\
\frac{f(x_n)}{x_n - x_{n+1}} &= f'(x_n) \\
\Longrightarrow \frac{f(x_n)}{f'(x_n)} &= x_n - x_{n+1} \\
\therefore x_{n+1} &= x_n - \frac{f(x_n)}{f'(x_n)} \quad \text{hence shown.}
\end{aligned}
$$

Examples

(1) (a) Find the iterative based on Newton Raphson's formula for finding the root of the equation $x^2 - 3x + 2 = 0$.

(b) Taking the initial approximation as 0.7, find the root of the equation to two decimal places.

Solution

Let $x^2 - 3x + 2 = 0$

$$
\begin{gathered}
f(x) = x^2 - 3x + 2 \Longrightarrow f(x_n) = x_n^2 - 3x_n + 2 \\
f'(x) = 2x - 3 \Longrightarrow f'(x_n) = 2x_n - 3
\end{gathered}
$$

Using $x_{n+1} = x_n - \frac{f(x_n)}{f'(x_n)}$

$x_{n+1} = x_n - \left(\frac{x_n^2 - 3x_n + 2}{2x_n - 3}\right)$

$x_{n+1} = \frac{2x_n^2 - 3x_n - x_n^2 + 3x_n - 2}{2x_n - 3}$

$x_{n+1} = \frac{x_n^2-2}{2x_n-3}, \quad n = 0, 1, 2, 3, \cdots$

(b) Using $x_0 = 0.7$, and Error bound $= 0.5 \times 10^{-2} = 0.005$

From $x_{n+1} = \frac{x_n^2-2}{2x_n-3}$

$x_1 = \frac{(0.7)^2-2}{2(0.7)-3} = 0.9436$

$\left|x_1 - x_0\right| = \left|0.9436 - 0.7\right| = 0.2436 > 0.005$

$x_2 = \frac{(0.9436)^2-2}{2(0.9436)-3} = 0.9971$

$\left|x_2 - x_1\right| = \left|0.9971 - 0.9436\right| = 0.0535 > 0.005$

$x_3 = \frac{(0.9971)^2-2}{2(0.9971)-3} = 1.0000$

$\left|x_3 - x_2\right| = \left|1.0000 - 0.9971\right| = 0.0029 \leq 0.005$

$\Longrightarrow$ The root $= 1.00$

(2) (a) Find the iterative based on Newton Raphson's formula for finding the root of the equation $2x^2 = 6x + 1$.

(b) Taking $x_0 = -0.1$, find the root of the equation to three decimal places.

Solution

Hint: Here you should ensure that your equation is equated to zero.

Since $2x^2 = 6x + 1 \Longleftrightarrow 2x^2 - 6x - 1 = 0$

$$\text{Let } f(x) = 2x^2 - 6x - 1 \Longrightarrow f(x_n) = 2x_n^2 - 6x_n - 1$$
$$f'(x) = 4x - 6 \Longrightarrow f'(x_n) = 4x_n - 6$$

Using $x_{n+1} = x_n - \dfrac{f(x_n)}{f'(x_n)}$

$x_{n+1} = x_n - \left(\dfrac{x_n^2 - 3x_n + 2}{4x_n - 6}\right)$

$x_{n+1} = \frac{2x_n^2+1}{4x_n-6}, \quad n = 0, 1, 2, \cdots$

(b) Using $x_0 = -0.1$, and Error bound $= 0.5 \times 10^{-3} = 0.0005$

From $x_{n+1} = \frac{2x_n^2+2}{4x_n-6}$

$x_1 = \dfrac{2(-0.1)^2 + 1}{4(-0.1)_n - 6} = -0.1594$

$\left|x_1 - x_0\right| = \left|-0.1594 -^{-} 0.1\right| = 0.0594 > 0.005$

$x_2 = \dfrac{(-0.1594)^2 - 2}{2(-0.1594) - 3} = -0.1583$

$\left|x_2 - x_1\right| = \left|-0.1583 -^{-} 0.1594\right| = 0.0011 > 0.005$

$x_3 = \dfrac{(-0.1583)^2 - 2}{2(-0.1583) - 3} = -0.1583$

$\left|x_3 - x_2\right| = \left|-0.1583 -^{-} 0.1583\right| = 0.0000 \leq 0.005$

$\Longrightarrow$ The root $= -0.158$

(3) Show that the iterative formula based on $N.R.M$ for solving the root of the equation $e^x = \cos 2x$ is given by $x_{n+1} = \dfrac{(x_n - 1)e^{x_n} + 2x_n \sin 2x_n + \cos 2x_n}{e^{x_n} + 2\sin 2x_n}$, $n = 0, 1, 2, \cdots$. Hence taking $x_0 = 0.1$, and use the formula once, find the approximate root of the equation to three decimal places.

Solution
NB: Remember that the calculator must be in radians.
From, $e^x = \cos 2x \iff e^x - \cos 2x = 0$

Let $f(x) = e^x - \cos 2x$

$$
\begin{aligned}
f(x) = e^x - \cos 2x &\Longrightarrow f(x_n) = e_n^x - \cos 2x_n \\
f'(x) = e^x + 2\sin 2x &\Longrightarrow f'(x_n) = e_n^x + 2\sin 2x_n
\end{aligned}
$$

$$
\begin{aligned}
\text{Using, } x_{n+1} &= x_n - \frac{f(x_n)}{f'(x_n)} \\
&= x_n - \left(\frac{e^{x_n} - \cos 2x_n}{e^{x_n} + 2\sin 2x_n}\right) \\
&= \frac{x_n(e^{x_n} + 2\sin 2x_n) - e^{x_n} + \cos 2x_n}{e^{x_n} + 2\sin 2x_n} \\
\Longrightarrow x_{n+1} &= \frac{(x_n - 1)e^{x_n} + 2x_n \sin 2x_n + \cos 2x_n}{e^{x_n} + 2\sin 2x_n}, \quad n = 0, 1, 2, 3, \cdots
\end{aligned}
$$

Hence using $x_0 = 0.1$ with error bound $= 0.5 \times 10^{-3} = 0.0005$

From $x_{n+1} = \dfrac{(x_n - 1)e^{x_n} + 2x_n \sin 2x_n + \cos 2x_n}{e^{x_n} + 2\sin 2x_n}$

$$x_1 = \frac{((0.1) - 1)e^{(0.1)} + 2(0.1)\sin(2 \times 0.1) + \cos(2 \times 0.1)}{e^{(0.1)} + 2\sin(2 \times 0.1)} = 0.083$$

$\Longrightarrow$ The root $= 0.083$

(3) (i) Show that the N.R.M formula for finding the k$^{\text{th}}$ root of number N is

$$x_{n+1} = \frac{1}{K}\left[(k-1)x_n + \frac{N}{{x_n}^{k-1}}\right], \quad n = 0, 1, 2, 3, \cdots$$

(ii) Use your formula to find the positive cube root of 13 to three *dps*.

Solution
Let $x = \sqrt[k]{N} \iff x^k = N$

$x^k - N = 0$

$f(x) = x^k - N$

$$
\begin{aligned}
f(x) &= x^k - N \Longrightarrow f(x_n) = x_n^k - N \\
f'(x) &= kx^{k-1} \Longrightarrow f'(x_n) = kx_n^{k-1} \\
\text{Using } x_{n+1} = x_n &- \frac{f(x_n)}{f'(x_n)} \\
&= x_n - \Big(\frac{x_n^k - N}{kx_n^{k-1}}\Big) \\
&= \frac{x_n(kx_n^{k-1}) - x_n^k + N}{kx_n^{k-1}} \\
&= \frac{(kx_n^k) - x_n^k + N}{kx_n^{k-1}} \\
&= \frac{(k-1)x_n^k + N}{kx_n^{k-1}}, \\
\Longrightarrow x_{n+1} &= \frac{1}{K}\Big[(k-1)x_n + \frac{N}{{x_n}^{k-1}}\Big] \quad n = 0, 1, 2, 3, \cdots,
\end{aligned}
$$

hence shown.
By comparing $x = \sqrt[k]{N}$ with $\sqrt[3]{13} \Longrightarrow N = 13$ and $k = 3$.
$\therefore x_{n+1} = \frac{1}{K}\Big[(k-1)x_n + \frac{N}{{x_n}^{k-1}}\Big]$ becomes $x_{n+1} = \frac{1}{3}\Big[2x_n + \frac{13}{x_n}^2\Big]$ and choose $x_0 = 2.3$

$x_1 = \frac{1}{3}\Big[2(2.3) + \frac{13}{(2.3)^2}\Big] = 2.3525$

$x_2 = \frac{1}{3}\Big[2(2.3525) + \frac{13}{(2.3525)^2}\Big] = 2.3513.$

$|x_1 - x_0| = |2.3513 - 2.3525| = 0.0012 > 0.0005.$

$x_2 = \frac{1}{3}\Big[2(2.3513) + \frac{13}{(2.3513)^2}\Big] = 2.3513.$

$|x_2 - x_1| = |2.3513 - 2.3513| = 0.000 \leq 0.0005.$

$\Longrightarrow$ the root is 2.351

(4) (a) Show that the equation $3tanx + \frac{x}{3} = 0$ has a root between $x = 2$ and $x = 3.2$.

(b) Show that the Newton Raphson's method for finding the root of the equation $3tanx + \frac{x}{3} = 0$ is $\frac{(6x_n - 3sin2x_n)}{(6 + 2cos2x_n)}$. Hence, approximate the root to three decimal places using the answer in (a) above as the initial approximation.

12.1.4 General Iterative Method

This is the method that involves making any of the $x's$ in the function $f(x) = 0$ the subject so as x is expressed as the root of the equation.
If $f(x) = 0$, let $x = g(x)$ then x on the left hand side becomes our better approximation while that/those on the right hand side become(s) the preceding(current) approximation. I.e $x_{n+1} = g(x_n)$.
Example:
Given the function$f(x) = 2x^3 + 3x - 8 = 0$, generate the different iterative equations that can be used to solve $f(x) = 0$.
From $2x^3 + 3x - 8 = 0 \Longleftrightarrow 3x = 8 - 2x^3 \Longleftrightarrow x = \dfrac{(8 - 2x^3)}{3}$
$\therefore x_{n+1} = \dfrac{(8 - 2x_n^3)}{3} \cdots\cdots\cdots\cdots (i)$
$2x^3 + 3x - 8 = 0 \Longleftrightarrow x^3 = \dfrac{(8 - 3x)}{2} \Longleftrightarrow x = \dfrac{(8 - 3x)}{2}$
$\therefore x_{n+1} = \sqrt[3]{\left(4 - \dfrac{3x}{2}\right)} \cdots\cdots\cdots\cdots (ii)$
From $2x^3 + 3x - 8 = 0 \Longleftrightarrow x(2x^2 + 3) = 8 \Longleftrightarrow x = \dfrac{8}{(2x^2 + 3)}$
$\therefore x_{n+1} = \dfrac{8}{(2x^2 + 3)} \cdots\cdots\cdots\cdots (iii)$
Please generate any other two!!

12.1.5 Testing for Convergence.

This involving identifying from all the formula generated, the best(i.e the one that can give you the root of the equation as faster as possible)
This can be done in two forms:

- Analytical approach. This involve differentiating the generated formula.I.e
  From $x_{n+1} = g(x_n) \Longleftrightarrow x = g(x)$ because of dropping $(n + 1)$ and n
  If $|g'(x_0)| < 1$ for x_0 = initial approximation, then $x_{n+1} = g(x_n)$ becomes the better equation for approximating the root of the equation otherwise it tend to diverge.
- In this case, we continuously use $x_{n+1} = g(x_n)$ for $n = 0, 1, 2, 3, \cdots$ If the error $|x_{n+1} - x_n|$ keeps on reducing, then $x_{n+1} = g(x_n)$ becomes becomes the better equation for approximating the root of the equation otherwise it tend to diverge.

Examples

(1) (a) Show that the iterative formula for finding the root of the equation $x^3 - x - 1 = 0$ is given by $x_{n+1} = \sqrt{\frac{(x_n+1)}{x_n}}$ $\quad n = 0, 1, 2, 3, \cdots$

(b) Find the root of the equation in (a) above to $3dps$.

Solution

From $x^3 - x - 1 = 0 \Longleftrightarrow x^3 = x + 1 \Longleftrightarrow x^2 = \dfrac{(x+1)}{x} \Longleftrightarrow x = \sqrt{\dfrac{(x+1)}{x}}$

$\Longrightarrow x_{n+1} = \sqrt{\dfrac{(x_n+1)}{x_n}}$.

hence shown

(b)Take $x_0 = 1.3$ and using $x_{n+1} = \sqrt{\dfrac{(x_n+1)}{x_n}}$

Then $x_1 = \sqrt{\dfrac{((1.3)+1)}{(1.3)}} = 1.3301.$

$x_2 = \sqrt{\dfrac{((1.3301)+1)}{(1.3301)}} = 1.3236.$

$|x_2 - x_1| = |1.3236 - 1.3301| = 0.0065 > 0.0005$

$x_3 = \sqrt{\dfrac{((1.3236)+1)}{(1.3236)}} = 1.3250.$

$|x_3 - x_2| = |1.3250 - 1.3236| = 0.0014 > 0.0005$

$x_4 = \sqrt{\dfrac{((1.3250)+1)}{(1.3250)}} = 1.3247.$

$|x_4 - x_3| = |1.3247 - 1.3250| = 0.0003 \leq 0.0005 \Longrightarrow$ the root $= 1.325.$

(2) Given the iterative formula as $x_{n+1} = \dfrac{1}{2}\Big(x_n + \dfrac{5}{x_n}\Big) \qquad \text{for} n = 0, 1, 2, 3, \cdots$

(a) State the purpose of the formula.

(b) Find the root of the equation above taking the initial approximation as 2.2 to two decimal places.

Solution:

(a) From $x_{n+1} = \dfrac{1}{2}\Big(x_n + \dfrac{5}{x_n}\Big) \Longleftrightarrow x = \dfrac{1}{2}\Big(x + \dfrac{5}{x}\Big)$

$2x = x + \dfrac{5}{x} \Longleftrightarrow x^2 = 5 \Longleftrightarrow x^2 - 5 = 0.$

To compute the root of the equation $x^2 - 5 = 0$.

(b) Take $x_{n+1} = \dfrac{1}{2}\Big(x_n + \dfrac{5}{x_n}\Big)$ and $x_0 = 2.2$ with $E = 0.005$

$x_1 = \dfrac{1}{2}\Big((2.2) + \dfrac{5}{(2.2)}\Big) = 2.2364$

$x_2 = \dfrac{1}{2}\Big((2.2364) + \dfrac{5}{(2.2364)}\Big) = 2.2361$

$|x_2 - x_1| = |2.2361 - 2.2364| = 0.0003 \leq 0.0005$

The root of equation is 2.236.

(3) Given the two iterative formulas for solving the root of the equation $f(x)$ as (F_1); $x_{n+1} = \frac{2x_n^3-1}{3x_n^2+1}$ and (F_2); $x_{n+1} = \sqrt{-1-\frac{1}{x_n}}$, by taking $x_0 = -0.6$ and using each formula once, which formula is suitable?

Solution

Let $g(x) = \frac{2x^3-1}{3x^2+1}$

$\Longrightarrow g'(x) = \frac{(3x^2+1)6x^2-(2x^3-1)3x}{(3x^2+1)^2}$

$$g'(^-0.6) = \frac{\left(3(^-0.6)^2+1\right)6(^-0.6)^2-\left(2(^-0.6)^3-1\right)3(^-0.6)}{\left(3(^-0.6)^2+1\right)^2} = 0.4427$$

$\left|g'(^-0.6)\right| = 0.4427 < 1$

Let $h(x) = \sqrt{-1-\frac{1}{x}}$

$$\Longrightarrow h'(x) = \frac{1}{2}\left(-1-\frac{1}{x}\right)^{\frac{-1}{2}}\left(\frac{1}{x^2}\right)$$

$$h'(^-0.6) = \frac{1}{2}\left(-1-\frac{1}{^-0.6}\right)^{\frac{-1}{2}}\left(\frac{1}{(^-0.6)^2}\right) = 1.7010$$

$\left|h'(^-0.6)\right| = 1.7010 > 1$ Thus formula F_1 is more suitable since it converges.

(4) Given the two iterative formulas for solving the root of the equation $f(x)$ as I; $x_{n+1} = \frac{2x_n^3-1}{3x_n^2+1}$ and II; $x_{n+1} = \sqrt{-1-\frac{1}{x_n}}$, by taking $x_0 = -0.6$ and using each formula twice, state with a reason, which formula is more suitable?

Solution

For I,

Using $x_{n+1} = \dfrac{2x_n^3-1}{3x_n^2+1}$, and $x_0 = {}^-0.6$

Then $x_1 = \dfrac{2(^-0.6)^3-1}{3(^-0.6)+1} = {}^-0.6885$

Also $x_2 = \dfrac{2(^-0.6885)^3-1}{3(^-0.6885)+1} = {}^-0.6824$

$\left|x_2 - x_1\right| = \left|^-0.6824 - {}^-0.6885\right| = 0.0061$

For formula II,

Using $x_{n+1} = \sqrt{-1-\frac{1}{x_n}}$, and $x_0 = {}^-0.6$

Then $x_1 = \sqrt{-1-\frac{1}{^-0.6}} = {}^-0.8165$

Also $x_2 = \sqrt{-1-\frac{1}{^-0.8165}} = {}^-0.4741$

$\left|x_2 - x_1\right| = \left|^-0.4741 - {}^-0.8165\right| = 0.3424$

$\Longrightarrow$ formula II is more suitable since it gives a smaller error showing convergence.

12.1.5.1 Further Examples:

(1) (a) Show that the equation $x^2 = 3x + 3$ has only two roots.

(b) Use Newton Rampson Method the find the negative root correct to three decimal places.

(c) Use linear interpolation to approximate the positive root to two decimal places

(2) (a) Show that $x^3 - 3x^2 + 1 = 0$ has a real root between $x = 2$ and $x = 3$.

(b) Using linear interpolation, find the first approximation for the root

(c) Using the newton – Rapson formula, find the value of the root correct to 4 significant figures.

(3) (a) Given the equation $y = x^2 - cosx$, show by plotting suitable graphs on the same axes that the root lies between $\dfrac{\pi}{8}$ and $\dfrac{\pi}{2}$ inclusive putting your answer to one decimal places.

(b) Use linear interpolation twice to approximate the root to two decimal places,

(c)Using the NRM, find the value of the root correct to 3 decimal places.

12.1.6 Exercise 12

(1) By plotting graphs of x^3 and $3x - 4$ on the same axes, find the approximate root of the equation $x^3 - 3x + 4 = 0$. Hence use the Newton Raphson method to find the root, correct to three decimal places.
Ans: -2.196

(2) Show that the iterative formula based on NRM for finding the fourth root of a number α is given by $x_{n+1} = \frac{1}{4}\left(3x_n + \frac{\alpha}{x_n^3}\right), \quad n = 0, 1, 2, \cdots$ taking $x_0 = 2.5$ find $\sqrt[4]{45.7}$ correct to three significant figures.
Ans: 2.60

(3) Given the equation $e^x + x - 3 = 0$

(a) Obtain graphically the approximate root of the equation.

(b) Derive the simplest iterative formula based on Newton Raphson method that can be used to find a better approximation of the root. Using the value in (a) and the equation in (b), find the root of the equation correct to 3 decimal places.
Ans: (a) 0.8 (b) 0.792

(4) Given the equation $e^x - 1 = 2^x$; show by plotting suitable graphs on the same axes that there is only one root between 1 and 2.
Ans: x_0 = 1.15

(5) By using a suitable table of values show that the equation $x^3 - 3x + 1 = 0$ has three roots in the interval [-2, 2]. Use linear interpolation once to find the negative

root of the equation.
Ans: x_0 =**-1.75**

(6) Find the iterative formula for approximating the root $f(x) = 0$ by the N.R.M for the equation $xe^x + 5x - 10 = 0$, using $x_0 = 1.5$ solve the root correct to 4 significant figures.
Ans: 1.201

(7) Using Newton Raphson formula and $x_0 = \frac{\pi}{2}$ as the initial approximation to the root of the equation $10\cos x - x = 0$. Show that the next approximation is $\frac{5\pi}{11}$.

(8) Show graphically that $\sin x = \frac{1}{2}x$ has a root between 1.8 and 1.9. Hence or otherwise use linear interpolation to find the root of the equation to 3 significant figures.
Ans: 1.89
Find graphically an approximation to the root of the equation $x^3 - 4x + 5 = 0$. Use linear interpolation to give a better approximation of the root to 3 decimal places.
Ans: -2.457

(9) Find graphically the largest root of the equation $x^3 = 2x + 6$. Hence use linear interpolation to find a better approximation of the root.
Ans:

(10) Show that the root of the equation $f(x) = e^x + x^3 - 4x$ has a root between 1 and 2. By using NRM method, find the root to 2 dps.
Ans: 1.12

(11) Show that there is a root of the equation $8\sin x - x = 0$ between x = 2.7 and 2.8. Use linear innterpolation once, find the root of the equation correct to 2 decimal places.
Ans: 2.79

(12) Find graphically the approximate root of the equation $x^3 + 4x - 16 = 0$. Hence use the linear interpolation once to find a better approximation.
Ans:

(13) Show that the equation $e^{-x} = \sin x$ has a root between 0 and 1.5. Use the Newton Raphson method once to find the root taking $x_0 = 0.75$, give your answer to 2 decimal places.
Ans:

(14) (a) Show graphically that the equation $x\sin x - 1 = 0$ has a root between $x = 1$ and $x = 1.5$.
(b) Using the Newton - Raphson method, find this root correct to 2 decimal places
Ans:

(15) (a) Show that the root of the equation $e^{-2x} + 2x - 10 = 0$ lies between -1.2 and -1.4
(b) Use Newton Raphson's method to find the root of the equation in (a) above. Give ypur answer correct to three decimal places.
Ans:

(16) Determine the equation being solved using the iterative formula $x_{n+1} = \frac{x_n^2 e^{x_n}+10}{e^{x_n}(x_n+1)+5}$.
Ans: $xe^x + 5x - 10$

(17) Show that the iterative formula for solving the equation $x^3 = x + 1$ is $x_{n+1} = \sqrt{1 + \frac{1}{x_n}}$, $\quad n = 0, 1, 2, \cdots$. Starting with $x_0 = 1$, find the solution to the equation to 4 significant figures.
Ans: 1.325

(18) The equation $x \ln x + x - 3 = 0$ has root in the interval (1, 2). Using a more suitable formula of the iterative below find the root to 3 significant figures
$x_{n+1} = 3 - x_n \ln x_n$ and $x_{n+1} = \frac{3}{1+\ln x_n}$
Ans: 1.87

(19) The sequence given by iteration foermula $x_{n+1} = \frac{x_n}{\sqrt{\tan x_n}}$; $\quad n = 0, 1, 2, \cdots$ with $x_0 = 1$, converges to the number α, where $0 < \alpha < 1$. State an equation of which α is a root and hence use the approximations to the positive root of the equation $x^2 - 7 = 0$. Hence find $\sqrt{7}$ to 1 decimal place.
Ans:

(20) Show that the root of the equation $xe^x - 3$ lies between x = 1.0 and x = 1.1

(21) Show that one of the roots of the equation $e^x - 2x = 1$ is zero and the other lies between x = 1 and x = 1.5, hence,
(a) Use linear interpolation once to find a better approximation of the root that lies between 1 and 1.5
(b) Use Newton Raphson formula to find the root in (a) above, to two decimal places.
Ans: (a) 1.185 (b) 1.26

(22) (a) Show that the root of the equation $2x - 3\cos\left(\frac{x}{2}\right) = 0$ lies between 1 and 2.
(b) Use the Newton Raphson's method to find the root of the equation in (a) above. Give your answer correct to two decimal places.
Ans: 1.23

(23) 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.
Ans: 1.6

(24) Given the equation $x^3 - 6x^2 + 9x + 2 = 0$;
(a) Find graphically the root of the equation which lies between -1 and 0
(b)(i) Show that the Newton Rapson's formula for approximating the root of the equation is given by $x_{n+1} = \frac{2}{3}\left[\frac{x_n^3 - 3x_n^2 - 1}{x_n^2 - 4x_n + 3}\right]$
(ii) Use the formula in (b)(i) above, with an initial approximation in (a) above to find the root of the given equation correct to two decimal places.
Ans: -0.20

(25) Use Newton Raphson's method to find the approximate value of $^3\sqrt{66}$, taking 4 as the initial approximation. Give your answer correct to 3 d.ps.
Ans: 4.041

(26) Given the formula $x_{n+1} = \sqrt[3]{3 - x_n^{-2}}$. Find the equation whose root is saught. Hence show that the equation has three real roots.
Ans: $x^5 - 3x^2 + 1$

(27) (a) Show that the equation $e^x = 4\sin x$ has a root between 1.0 and 1.5. Hence use linear interpolation to estimate the root to 3 d.ps
Ans: 1.284

(28) Given the two iterative formulea;
$x_{n+1} = \frac{x_n^3 - 1}{5}$ $\qquad x_{n+1} = \sqrt{5 + \frac{1}{x_n}}$
Using $x_0 = 2$, deduce a more suitable formula for solving the equation. Hence use the formula once to find the second approximation of the root, correct to three decimal places.
Ans: 1.345

(29) By plotting graphs of $y = e^x$ and asuitable line on the same axes, show that the equation $e^x + x - 4 = 0$ has a root between 1 and 2. Hence use Newton Raphson's method to find the root of the equation, correct to 3 d.ps
Ans: 1.074

(30) (a) Determine the number of roots of the equation; $x^3 + 2x^2 = 4x + 4$.
(b) Use the linear interpolation to find an approximation of the greatest root of the equation $x^3 + 2x^2 = 4x + 4$
Ans: (a) (-3, -2) (-1,0) (1,2) (b) 1.6

(31) Taking the first approximation as 2.0, use use the Newton Raphson method twice to find $\log_e 10$, correct to three decimal places
Ans: 2.304

(32) Show that the equation: $3x + \cos 2x = 2$ has a real root between x = 0.4 and x = 0.6

(33) (a) By plotting graphs of $y = x^3$ and $y = 12x + 6$ for the interval [-4, 4], find graphically the real roots of the equation $x^3 = 12x + 6$.
(b) Use linear interpolation twice to find a better approximation of the positive root, correct to two decimal places.
Ans:

(34) (a) Show that, if α is the first approximation of the root of the equation $x^3 = 3x + 12$, Newton Raphson method gives the second approximation as $\frac{2\alpha+12}{3\alpha^2-3}$
(b) Show that the equation $x^3 - 3x - 12 = 0$ has only one real root between x = 0 and x = 4. Hence taking $\alpha = 2.5$, find the root correct to four decimal places.
Ans: btn (2, 3) , 2.729

(35) The N.R.M formula for solving the equation $x^k = N$ derivative as $x_{n+1} = \frac{1}{3}\left(2x_n + \frac{N}{x_n^2}\right)$
(i) Determine the value of k and state the purpose of the formula.
(ii) Taking $x_0 = 1.5$, N = 10.5 determine the k$^{\text{th}}$ root of N using the iterative formula above to 3 d.ps
Ans: (i) k = 3, purpose : to find the cube root of N (ii) 2.190

(36) Prove that the iterative formula for finding the reciprocal of a number β by Newton Raphson method is given by $x_{r+1} = \frac{x_r}{3}\left(4 - \beta x_r^3\right); \quad r = 0, 1, 2, \cdots$. Hence find the second approximation of the reciprocal of 1.5, using $x_0 = 0.9$
Ans: 0.87

(37) (a) Use a suitable table of values to locate each of the two real roots of the equation $x^4 - 4x - 4 = 0$.
(b) Use the linear interpolation;
(i) twice to find the largest root, to 1 d.p
(ii) to estimate the least root to 1d.p
Ans: (i) 1.8 (ii) -0.8

(38) By plotting the graphs of $y = \sin x$ and $y = 1 - x$ on the same axes for the interval 0 < x $\leq$ 2.0, show that the two graphs have a common root and hence obtain the root of the equation $\sin x + x - 1 = 0$, correct to 1 decimal place.
Ans: 0.5

(39) (a) Show that the equation $xe^x = x + 1$ has a root between 0 and 1.
(b) Using the Newton Raphson method, find the root of the equation in (a) above correct to 3 decimal places.
Ans: (b) 0.807

(40) Use the graphically method to estimate the positive root of the equation $x = (2x - 5)e^x$, correct to one decimal place.
Ans: x_0 = 2.6

(41) Show that the equation $x = \cos x - 3$ has a root between -4 and -3. Hence use linear interpolation once to estimate the root to two decimal places.
Ans: -3.74

Chapter 13

NUMERICAL INTEGRATION.

13.1 TRAPEZIUM RULE / TRAPEZOIDAL RULE

This is the method used to estimate the area under the curve over a given interval $[a, b]$. The area under the curve is referred to as the integral of the curve over an interval$[a, b]$
Given the curve $y = f(x)$ over the interval$[a, b]$, By trapezium Rule,

$$\int_a^b f(x)dx \approx \frac{1}{2}h[y_0 + y_n + 2(y_2 + y_3 + \cdots + y_{n-1})].$$

Where: $h = \dfrac{Range\ in\ interval}{Number\ of\ sub-intervals} = \dfrac{b-a}{Number\ of\ sub-intervals}$.
NB: Sub - intervals = Strips = Sub - divisions = Ordinates -1
NB. If h is recurring, please express it as a fraction.

Proof.

Given the approximate sketch of the curve $y = f(x)$ between $x_0 = a$ and $x_n = b$

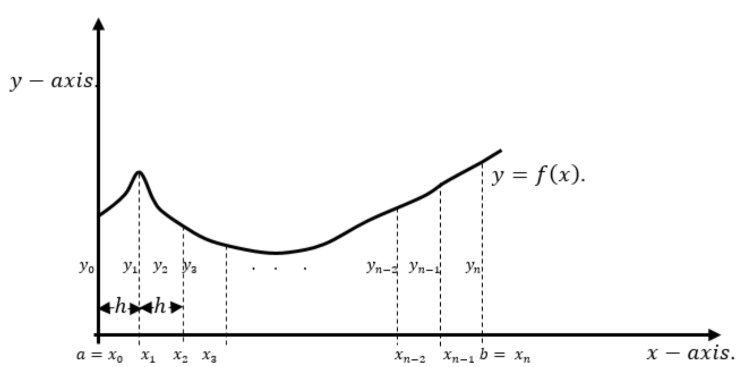


Figure 13.1: Each trapezium is of height h. I.e $h = (x_1 - x_0) = (x_2 - x_1) = \cdots = (x_{n+1} - x_n)$.

From the diagram, the function is sub divided into different trapeziums each of area $\frac{1}{2}h(a+b)$ and the total area is given as:

$$\begin{aligned} Area &= \int_{x_0}^{x_n} f(x)dx \\ &= \frac{1}{2}h(y_0+y_1) + \frac{1}{2}h(y_1+y_2) + \frac{1}{2}h(y_2+y_3) + \cdots + \frac{1}{2}h(y_{n-1}+y_n) \\ &= \frac{1}{2}h[y_0 + 2(y_1+y_2+\cdots+y_{n-1}) + y_n] \end{aligned}$$

$$\int_a^b f(x)dx \approx \frac{1}{2}h[y_0 + y_n + 2(y_2+y_3+\cdots+y_{n-1})].$$

The **exact** value of Area $= \int_{x_0}^{x_n} f(x)dx$ is obtained by using direct pure mathematics approach.
With the **Approximated** value and the **Exact** value, then we can obtain the;

(i) Error, (ii) Relative Error, (iii) %age Error.

The Error, Relative error or percentage error can be reduced by:
- **Increasing the Number of Strips.**

Examples

1. Use trapezium rule with six strips to estimate $\int_0^3 \frac{x}{1+x}dx$, put your answer to 3 decimal places.
   Solution: Let $h = \frac{3-0}{6} = 0.50$ and $f(x) = \frac{x}{1+x}$

x	$y_o + y_n$	$y_1 + y_2 + \cdots + y_{n-1}$
0.0	0.0000	
0.5		0.3333
1.0		0.5000
1.5		0.6000
2.0		0.6667
2.5		0.7143
3.0	0.7500	
Total	0.7500	2.8143

   $\Longrightarrow \int_0^3 \frac{x}{1+x}dx = \frac{1}{2} \times 0.50[0.7500 + 2(2.8143)] \approx 1.595.$

2. (a) Use trapezium rule with 7 ordinates to estimate $\int_1^2 \frac{1}{1+x^2}dx$, put your answer to 3 decimal places.
   (b) Find the exact value of $\int_1^2 \frac{1}{1+x^2}$ hence obtain the;

(i) absolute error, (ii) relative error and (iii) %age error.

(c) State how the percentage error can be reduced.

Solution: Let $h = \dfrac{2-1}{7-1} = \dfrac{1}{6}$ and $f(x) = \dfrac{1}{1+x^2}$

x	$y_o + y_n$	$y_1 + y_2 + \cdots + y_{n-1}$
1	0.5000	
$\frac{7}{6}$		0.4235
$\frac{8}{6}$		0.3500
$\frac{9}{6}$		0.3077
$\frac{10}{6}$		0.2647
$\frac{11}{6}$		0.2293
2.0	0.2000	
Total	0.7000	1.5752

$\Longrightarrow \int_1^2 \dfrac{1}{1+x^2}dx = \dfrac{1}{2} \times \dfrac{1}{6}[0.7000 + 2(1.5752)] \approxeq 0.321.$

(b) Exact value $= \int_1^2 \dfrac{1}{1+x^2}dx = tan^{-1}(x)|_1^2 = [tan^{-1}(2) - tan^{-1}(1)] = 0.322$

hence: (i) Absolute error $= |Exact - Approximate| = |0.322 - 0.321| = 0.001.$

(ii) Relative error $= \dfrac{Absolute\ error}{Exact\ value} = \dfrac{0.001}{0.322} = 0.003$

(iii) percentage error = Relative error $\times 100 = 0.003 \times 100 = 0.30\%$

(c) By increasing the number of ordinates.

3. (a)Use the trapezium rule to estimate the under $(x + tan(x))$ using 5 ordinates from $x = 1.0$ to $x = 1.4$ to 2decimal places.
(b) Find the percentage error in approximating the area under $(x + tan(x))$. Suggest how the error can be reduced.
Solution: Let $h = \dfrac{1.4-1.0}{5-1} = \dfrac{0.4}{4} = 0.1$ and $f(x) = x + tanx$

x	$y_o + y_n$	$y_1 + y_2 + \cdots + y_{n-1}$
1.0	2.5574	
1.1		3.0648
1.2		3.7722
1.3		4.9021
1.4	7.1979	
Total	9.7553	11.7391

$\Longrightarrow \int_{1.0}^{1.4}(x + tanx)dx = \dfrac{1}{2} \times 0.1 \times [9.7553 + 2(11.7391)] \ \approx\ 1.66.$

(ii) %age error $= \dfrac{Absolute\ error}{Exact\ value} \times 100$

But Exact value$= \int_{1.0}^{1.4}(x+tanx)dx = \left[\frac{x^2}{2}+lnsecx\right]_{1.0}^{1.4} = 1.64$
$\Longrightarrow$ %age error $= \frac{|1.64-1.66|}{1.64} \times 100 = 0.01\%$ to $2dps$.
The error can be reduced by increasing the number of ordinates.

(1) Use trapezium rule to evaluate $\int \frac{x}{x+3}dx$ over 1(0.2)2 to three decimal places.
Solution:
Let $h = 0.2$ and $\quad f(x) = \frac{x}{x+3}$

x	1	1.2	1.4	1.6	1.8	2.0
$f(x)$	0.25	0.2857	0.3182	0.3478	0.37	0.4

$$\begin{aligned}\int_1^2 \frac{x}{x+3}dx &\approx \frac{1}{2}(0.2)\big[(0.25+0.4)+2(0.2857+0.3182+0.3478+0.375)\big]\\ &= \frac{1}{2}(0.2)\big[(0.65)+2(1.4569)\big]\\ &= 0.330\end{aligned}$$

(2) Use the trapezium find $\int(x\sin x)dx$ over $0(\frac{\pi}{16})\frac{\pi}{4}$ to 3 significant figures.
Solution:
Let $h = \frac{\pi}{16}$, $\quad$ and $f(x) = x\sin x$

x	$y_o + y_n$	$y_1 + y_2 + \cdots + y_{n-1}$
0	0.0	
$\frac{\pi}{16}$		0.03831
$\frac{2\pi}{16}$		0.1503
$\frac{3\pi}{16}$		0.3273
$\frac{\pi}{4}$	0.5554	
Total	0.5554	0.51591

$$\begin{aligned}\int_0^{\frac{\pi}{4}} x\sin x dx &\approx \frac{1}{2}(\frac{\pi}{16})\big[(0.5554)+2(0.51591)\big]\\ &= \frac{\pi}{32}\big[1.58722\big]\\ \Longrightarrow \int_0^{\frac{\pi}{4}} x\sin x dx &\approx 0.156\end{aligned}$$

(3) (a) Evaluate $\int_{0.2}^{1.0}\cos x dx$ using 4 strips give your answer to four decimal places.
(b) Obtain the exact value of $\int_{0.2}^{1.0}\cos x dx$ to four decimal places.
(c) Find the percentage error in your calculations.

Solution

(a) Let $h = \frac{1.0-0.2}{4} = 0.2$ and , $\quad f(x) = \cos x$

x	$y_o + y_n$	$y_1 + y_2 + \cdots + y_{n-1}$
0.2	0.98007	
0.4		0.92106
0.6		0.82534
0.8		0.69671
1.0	0.54030	
total	1.52037	2.44311

$$\Longrightarrow \int_{0.2}^{1.0} \cos x dx \approx \frac{1}{2}(0.2)\big[(1.52037) + 2(2.44311)\big] = 0.33034$$
$$\approx 0.1\big[6.40659\big] = 0.640659$$
$$\Longrightarrow \int_{0.2}^{1.0} \cos x dx \approx 0.6407$$

(b) Exact value $= \int_{0.2}^{1.0} \cos x dx \; = \; \Big[\sin x\Big]_{0.2}^{1.0} \; = \; \sin(1.0) - \sin(0.2) \; = \; 0.6428$

(c) %age error $= \left|\frac{0.6428-0.6407}{0.6428}\right| \times 100 \; = \; 0.327$

13.2 Exercise 10

(1) (a) Use trapezium rule with 5 sub - intervals to estimate $\int_2^3 \dfrac{x^2}{1+x^2} dx$, put your answer to 4 decimal places.

(b) Find the exact value of $\int_2^3 \dfrac{x^2}{1+x^2} dx$ hence obtain the;

(i) absolute error, (ii) relative error and (iii) $\%age$ error.

(c) State how the percentage error can be reduced.

(2) 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 error in the approximation is less than 0.001
Ans:0.4778.

(3) In estimating the value of $\int_a^b \frac{x}{x^2+3} dx$, five equal sub-intervals of 0.1 were used. If $a = 0.1$, find:

(i) value of b hence,

(ii) the approximate value of $\int_a^b \frac{x}{x^2+3} dx$, correct to four significant figures.
Ans: (i) 1.5 (ii) 0.1357

(4) Use the trapezium rule with strips of width $\frac{\pi}{12}$ to determine the approximate value of $\int_0^{\frac{\pi}{3}} e^x \cos x dx$, correct to four significant figures
Ans: 1.435 $(4SF)$

(5) (a) Use the trapezoidal rule to find the area between the curve $y = \dfrac{1}{x^2-4}$ and the lines $x = 0$ and $x = 1$, using five strips, giving your answer to three significant figures.

(b) Determine the exact value of the area in (a) above, correct to three significant figures. Hence find the percentage in your calculation in (a) above.

Ans:

(6) (a) Use the trapezium rule with seven ordinates to find the approximate value $\int_1^2 x \ln x dx$ correct to 3 decimal places.

(b) Find the percentage error in your calculations in (a) above.

Ans: (a) 0.638 (b) 0.314%

(7) Use the trapezium rule with seven ordinates to estimate $\int_0^\pi x\sqrt[3]{(x+\sin x)}dx$, correct to three significant figures.

Ans:

(8) Given that $f(x) = \frac{x}{1+x}$, determine the;

(i) approximate value of $\int_0^1 f(x)dx$ correct to 3 decimal places using the trapezium rule with 6 ordinates,

(ii) percentage error for the calculation in (i).

(iii) How can this error be reduced?

Ans:

(9) The value of a continuous function $f(t)$ was found experimentally as given below;

t	0	0.3	0.6	0.9	1.2	1.5	1.8
$f(t)$	2.72	3.00	3.32	y	4.48	4.48	4.95

(a) Estimate the value of y.

(b) Use the trapezoidal rule to estimate $\int_0^{1.8} f(t)dt$ correct to 1 dp

Ans: (a) $y = 3.9$ (b) 6.9

(10) (a)Use the trapezium rule to estimate the under $\dfrac{1+x}{x^2+1}$ over 1.0(0.5)3 putting your answer to 3decimal places.

(b) Find the relative error in approximating the area under $\dfrac{1+x}{x^2+1}$. Suggest how the error can be reduced.

Solution

(11) A particle moves in a straight line so that after t seconds its velocity in ms^{-1} is given by $v = (2t^2 - 25)(t - 1)$. Find the displacement of the particle from its starting point after 3 seconds using the trapezium rule with 6 ordinates. State your answer to 2 dps.

(12) Given that $y = \frac{x}{x^2+1}$

(a) Copy and complete the table below using at least 4 decimal places of y.

x	1	1.4	1.8	2.2	2.6	3
y						

(b)

(i) Use the trapezoidal rule to estimate $\int_1^3 \frac{x}{x^2+1} dx$, correct to three decimal places.

(ii) Find the error made in using the method in (b)(i) to $3dps$.

(c) Suggest how the accuracy of y in (b)(i) can be improved and find the interval within which the exact value of y lies.

Ans: (b)(i) 0.804 (ii) 0.001 (c) [0.803, 0.805]

Chapter 14

THE FLOW CHARTS / DIAGRAMS

14.1 THE FLOW CHARTS / DIAGRAMS

14.1.1 Introduction

A flow chart is a graphical representation of an algorithm. Programmers often use it as a program-planning tool to solve a problem. It makes use of symbols which are connected among them to indicate the flow of information and processing.
The process of drawing a flowchart for an algorithm is known as "flowcharting".
NB: Algorithm: This refers to the step by step procedure for solving a problem
A flow chart or diagram is also defined as the logical sequence of steps that must be followed in order to solve the problem.

14.1.2 Basic Symbols used in Flowchart Designs.

(a) **Terminal:** The oval symbol indicates Start, Stop or Halt in a program's logic flow. A pause/halt is generally used in a program logic under some error conditions.
Terminal is the first and last symbols in the flowchart. It tells you to start or stop the program.
It is illustrated as below.

Figure 14.1

(b) **Input/Output**: A parallelogram denotes any function of input/output type. Program instructions that take input from input devices and display output on output devices are indicated with parallelogram in a flowchart.
It is illustrated as below.

Figure 14.2

(c) **Processing Or Assignment box**: A box represents arithmetic instructions. All arithmetic processes such as adding, subtracting, multiplication and division are indicated by action or process symbol. This is where the formula to use in the flow chart is put. It is commonly represented by rectangular boxes.
It is illustrated as below.

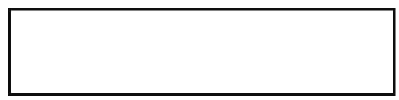


Figure 14.3

(d) **Decision Box:** Here we use a diamond symbol. It represents a decision point. Decision based operations such as yes/no question or true/false are indicated by diamond in flowchart.
Therefore it must contain the question that requires a binary answer.
It also must have one inlet and two outlets.
It is illustrated as below.

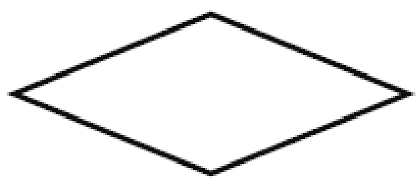


Figure 14.4

NB: Tolerance (TOL) is the difference between the current number(value) and the previous number(value) also known as the error bound, i.e TOL $= |b - a|$. The TOL is evaluated in the decision.

(e) **Flow lines:** Flow lines indicate the exact sequence in which instructions are executed. Arrows represent the direction of flow of control and relationship among different symbols of flowchart.Therefore each line must have an arrow.
They are illustrated as below.

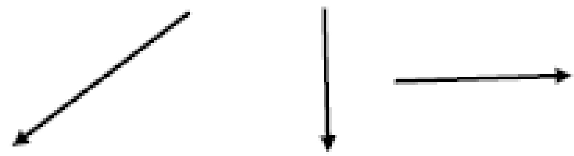


Figure 14.5

(f) **Connectors**: Whenever flowchart becomes complex or it spreads over more than one page, it is useful to use connectors to avoid any confusions. It is represented by a circle. It is illustrated as below.

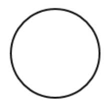


Figure 14.6

Note: Any flow chart constructed, a dry run must be performed.

14.1.3 Dry run

This refers to a check up in the diagram. This is intended to confirm that the program in the flowchart gives the expected outcome.

Examples.

(1) Given the flow chart below, perform a dry run and state its purpose.

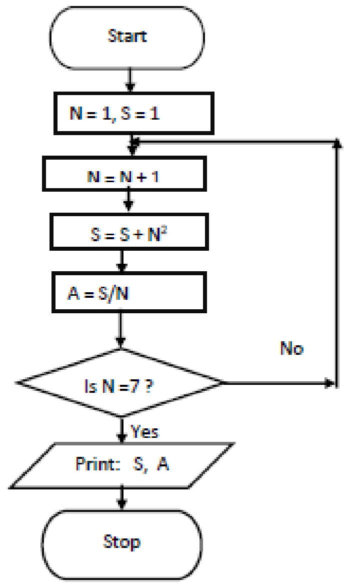


Figure 14.7

Solution
The dry run

n	S	A
1	1	1
2	5	2.5
3	14	4.6667
4	30	7.5
5	55	11
6	91	15.1667
7	140	20

$\Longrightarrow$ output; S = 140, A = 20

Purpose; To compute and print the average of the squares of the first 7 counting numbers.

(2) Study the flow chart below.

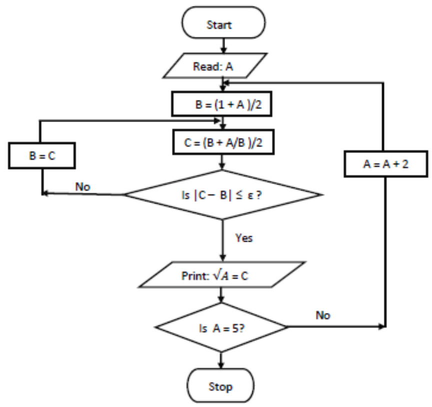


Figure 14.8

(i) Perform a dry run for the flow chart for A = 1, and $\varepsilon = 0.0005$

(ii) State the purpose of the flow chart.

Solution;

The dry run

A	B	C	|B - C|	$\sqrt{\text{A}} = \text{C}$
1	1	1	0	1
3	2	1.75	0.25	
	1.75	1.7321	0.01785	
	1.7321	1.7321	0	1.732
5	3	2.3333	0.6667	
	2.3333	2.2381	0.0952	
	2.2381	2.2361	0.0020	
	2.2361	2.2361	0.000	2.236

Purpose; To compute and print the square root of the first three odd numbers.

(3) Given the flow chart below;

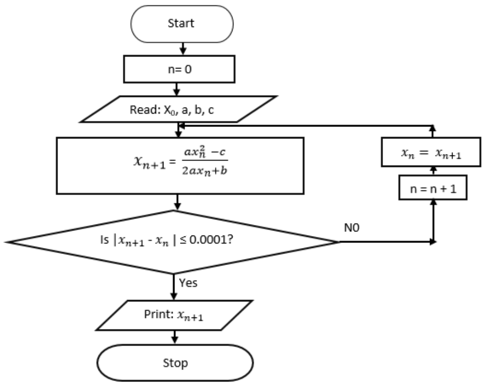


Figure 14.9

(a) Perform a dry run for your flow chart taking $a = 1, b = 0, c = -22.5$ and $x_0 = 4.7$

(b) State the purpose of the flow chart

(a) Solution

(a) Dry run. Let $a = 1, b = 0, c = -22.5$ and $x_0 = 4.7$

n	x_n	x_{n+1}	$\lvert x_{n+1} - x_n \rvert$
0	4.7	4.7457	0.0457
1	4.7457	4.7434	0.0023
2	4.7434	4.7434	0.0000

$\Longrightarrow$ the root is $x_{n+1} = 4.743$ (3 dps)

(b) **Purpose;** To compute and print the positive square root of 22.5

(4) Study the follow below:

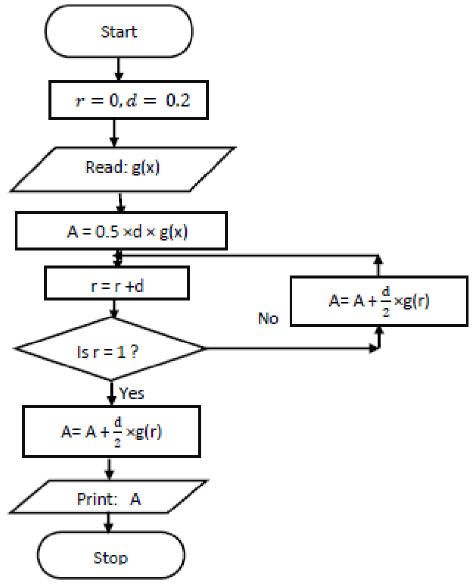


Figure 14.10

(a) Using $f(x) = X^2e^x$, Perform a dry run for the above flow to $3sf$.
(b) State the purpose of the flow chart.
(c) Find the exact value of $\int_0^1 X^2e^x dx$. Hence determine the absolute error.

Solution

(a) (a) Dry run. Let $d = 0.2, r = 0$, and $A = 0$

d	r	A
0.2	0	0
	0.2	0.009771
	0.4	0.057510
	0.6	0.188102
	0.8	0.4735714
	1	0.74539

$$\implies A = 0.7454$$

(b) **Purpose;** To compute and print the area under the curve $g(x) = \int_0^1 x^2e^x dx$ to $4sf$.

(c) Exact value $= \int_0^1 x^2e^x dx = 0.7183$. Please show the working.

$$\text{Hence Relative error} = \frac{|0.7183 - 0.7454|}{0.7183} = 0.0377.$$

14.1.4 Flow charts based on NRM

Here we must initialize the flow chart by setting the first iteration. NRM based flowcharts therefore follow the following basic step in their order and must be put in the basic sharp discussed above.

- Start
- Initialization I.e $n = 0$ or $n = 1$ basing on your interest.
- Input box: Here we read the initial approximation x_0, and any other constant stated/required.
- Assignment box: Here you enter the formula(s) in the process.
- Decision box: Here you test whether $\left|x_{n+1} - x_n\right| \leq TOL$, if NO, vary n by $n = n + 1$ and adjust to proceeding root by using $x_n = x_{n+1}$ then go back to the formula while, if YES, then follow the arrow(s).
- Output box: Here we print the root, x_{n+1} and any thins else required like number of iteration, or constant(s).
- Stop: Here you terminate the process.

Examples:

(1) (a) Show that the iterative formula for finding the cube root of a number β is given by $x_{n+1} = \frac{2x_n^3+\beta}{3x_n^2}, \quad n = 0, 1, 2, \cdots$

(b) Construct a flow chart that:

(i) reads the initial approximation to the root as x_0 and constant β.

(ii) computes and prints the root to 3 decimal places.

(c) Perform a dry run for your flow chart by taking $x_0 = 2.7$ to evaluate $\sqrt[3]{22}$ to 3 decimal places.

Solution

Let $x = \sqrt[3]{\beta} \Longleftrightarrow x^3 = \beta \Longleftrightarrow x^3 - \beta = 0$

$$\begin{aligned}
\text{Let } f(x) = x^3 - \beta &\Longrightarrow f(x_n) = x_n^3 - \beta \\
f'(x) = 3x^2 &\Longrightarrow f'(x_n) = 3x_n^2 \\
\therefore x_{n+1} &= x_n - \frac{f(x_n)}{f'(x_n)} \\
&= x_n - \left(\frac{x_n^3 - \beta}{3x_n^2}\right) \\
&= \frac{x_n(3x_n^2) - x_n^3 + \beta}{3x_n^2} \\
\Longrightarrow x_{n+1} &= \frac{2x_n^3 + \beta}{3x_n^2}, \quad n = 0, 1.2, \cdots
\end{aligned}$$

(b)

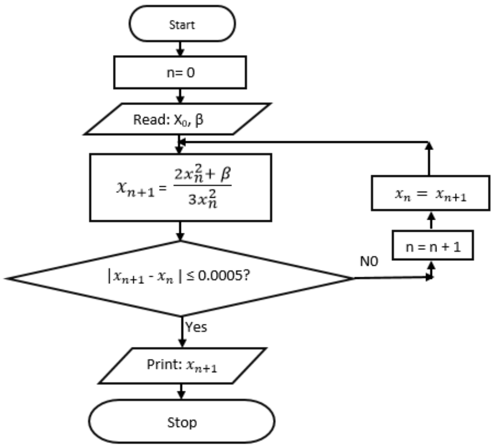


Figure 14.11

(c) Performing a dry run

$x_0 = 2.7$, $\beta = 22$, then $x_{n+1} = \frac{2x_n^3+22}{3x_n^2}$ and $|x_{n+1} - x_n| = 0.0005$

n	x_n	x_{n+1}	$\lvert x_{n+1} - x_n \rvert$
0	2.7	2.8059	0.1059
1	2.8059	2.8020	0.0039
2	2.8020	2.8020	0.0000

$\Longrightarrow$ the root is $x_{n+1} = 2.802$ $(3dps)$

(2) (a)Show that the iterative formula based on Newton -Raphson Method for finding the reciprocal of a number N, is given by $x_{(n+1)} = x_n(2 - x_nN), n = 0, 1, 2, \ldots$ Hence, draw a flow chart that reads the number N and the initial approximation x_0, computes and prints the root to three decimal places and the number of iteration n.

(b)Perform the dry run for the flow chart taking $N = 14$.

Solution:

Let $x = \dfrac{1}{N} \Longleftrightarrow N = \dfrac{1}{x} \Longleftrightarrow N - x^{-1} = 0$

$$\begin{aligned}
\text{Let } f(x) = N - x^{-1} &\Longrightarrow f(x_n) = N - x_n^{-1} \\
f'(x) = x^{-2} &\Longrightarrow f'(x_n) = x_n^{-2} \\
\therefore x_{n+1} &= x_n - \frac{N - x_n^{-1}}{x_n^{-2}} \\
&= \frac{x_n(x_n^{-2}) - N + x_n^{-1}}{x_n^{-2}} \\
\Longrightarrow x_{n+1} &= x_n(2 - x_nN), \quad n = 0, 1.2, \cdots
\end{aligned}$$

(b)

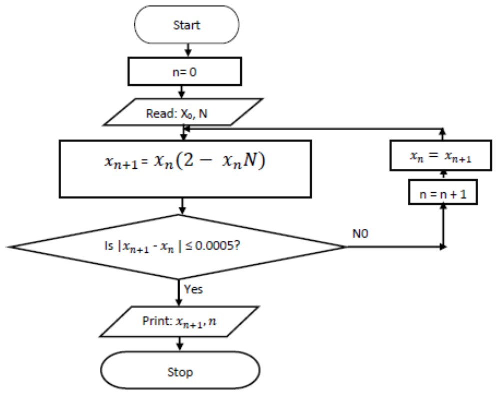


Figure 14.12

(c) Performing a dry run
Take $x_0 = 0.06, N = 14$, then $x_{n+1} = x_n(2 - 14x_n)$ and $|x_{n+1} - x_n| = 0.0005$

n	x_n	x_{n+1}	$\lvert x_{n+1} - x_n \rvert$
0	0.06	0.0696	0.0096
1	0.0696	0.0714	0.0018
2	0.0714	0.0714	0.0000

$\Longrightarrow$ the root is $x_{n+1} = 0.071$, and iterations$n = 3$

(3) (a) Show that the iterative formula for finding arc cot of number N is given by $x_{n+1} = x_n + \frac{1}{2}\sin 2x_n - N\sin^2 x_n, \quad n = 0, 1, 2, \cdots$

(b) Construct a flow chart that;

(i) Reads the initial approximation to the root as P and constant N,

(ii) Computes and prints the root to 4 decimal places or after 4 iterations,

(c) Perform a dry run for your flow chart by taking $x_0 = 0.3$ and $N = 3.8$.

Solution

(a) Let $x = \cot^{-1}(N) \Longleftrightarrow \cot x = N \Longleftrightarrow \cot x - N = 0$

$$\text{Let } f(x) = \cot x - N \implies f(x_n) = \cot x_n - N$$
$$f'(x) = -\text{cosec}^2 x \implies f'(x_n) = -\text{cosec}^2 x_n$$
$$\therefore x_{n+1} = x_n - \frac{f(x_n)}{f'(x_n)}$$
$$= x_n - \left(\frac{\cot x_n - N}{-\text{cosec}^2 x_n}\right)$$
$$= x_n + \cos x_n \sin x_n - N \sin^2 x_n$$
$$\implies x_{n+1} = x_n + \frac{1}{2}\sin 2x_n - N\sin^2 x_n, \quad n = 0, 1.2, \cdots$$

(b)

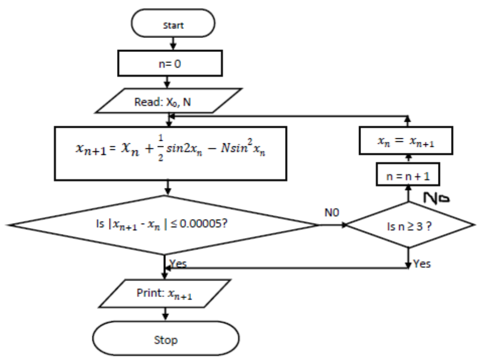


Figure 14.13

(c) Performing a dry run
$x_0 = 0.3, N = 3.8$, then $x_{n+1} = x_n + \frac{1}{2}\sin 2x_n - N\sin^2 x_n$ and Error $= 0.00005$

n	x_n	x_{n+1}	$\lvert x_{n+1} - x_n \rvert$
0	0.3	0.25046	0.04954
1	0.25046	0.25714	0.00668
2	0.25714	0.25732	0.00018
3	0.25732	0.25732	0.00000

$\implies$ the root is $x_{n+1} = 0.2573$.

14.1.5 Flow charts based discounts and taxation.

Examples.

1. Given the information information below showing a tax system (T) calculations for the monthly salary (S), earned by a certain employee.

Salary (S) shillings	Tax (T) in % of S
S $< 50,000$	0.0
50,000 $\leq$ S $<$ 80,000	5.0
80,000 $\geq$ S	8.0

(a) Draw a flow chart that can compute the tax of basing on the above table.

(b) Calculate the Tax for employees earning $70,000,\ 100,000$.monthly.

Solution.

(a)

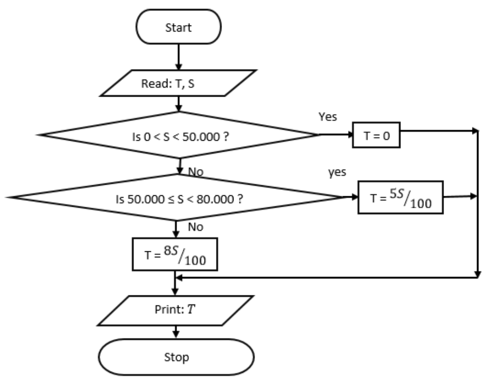


Figure 14.14

(b) Dry run:

A	T
70,000	3500
100,000	8000

2. The taxation on the income of employees of a certain company is given in by,

Taxable income, I (shs)	Taxation rate (%)
$0 - 120,000$	0.0
120,001 - 230,000	6.0
230,001 - 400,000	10.0
400,001 and above	18.0

Construct a flow chart that reads the taxable income I, calculates and prints the tax amount T and the taxable income. Hence perform a dry run for employees earning; $sh.154,000, sh.250,000$ and $sh.558,000$ respectively.

Solution.

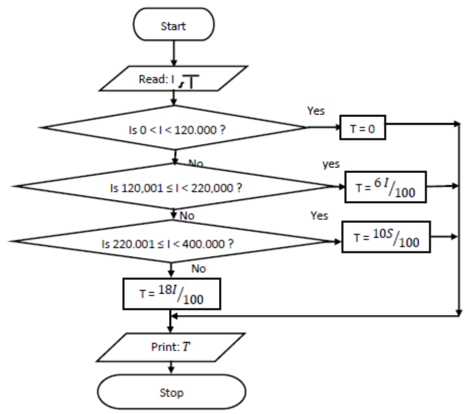


Figure 14.15

(b)Dry run:

I	T
154,000	9240
250,000	25,000
558,000	100,440

3. Draw a flow chart to compute tax for 100 employees as follows.

Monthly salary (A) shillings	Tax (T) in % of A
$0 < A < 100,000$	0.5
$100,001 \leq A < 400,000$	6.0
$400,000 \geq A$	10.0

Solution.

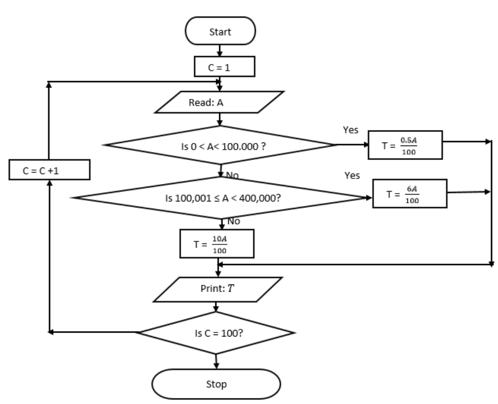


Figure 14.16

4. Construct a flow chart for a salary earner who earns A millions and he invest it in the bank at a rate of 8% compound interest for n years.Hence perform it's dry run for $n = 4$ and $A = 2$ millions.

Solution.

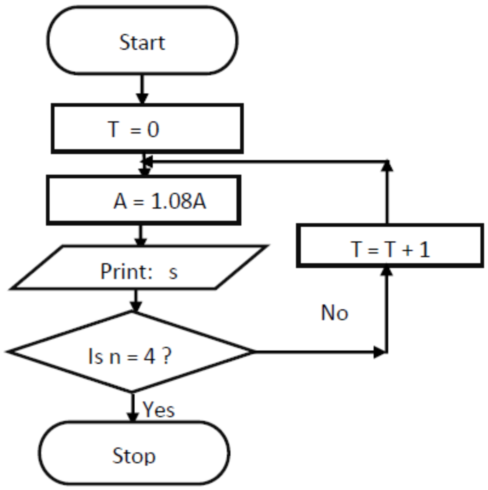


Figure 14.17

Dry run.

T	A
0	2
1	2.60
2	2.3328
3	2.5194
4	2.9390

$\Longrightarrow A = 2.9390$ after four years.

5. Task [Uneb concerning interest rate]

14.1.6 Flow charts based on real life situation.

Here you are requested to draw a flow chart that that be used to perform a certain purpose. This can be illustrated in the following examples.
Example.

(1) (a) Construct a flow chart that can be used to compute and print the sum of the first 8 counting numbers.

(b) Perform a dry run for your flow chart.
Solution:
(a) The flow chart

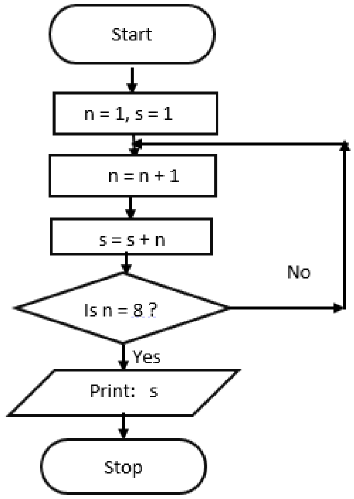


Figure 14.18

(b) The dry run

n	S
1	1
2	3
3	6
4	10
5	15
6	21
7	28
8	36

$\Longrightarrow S = 36$

(2) (a) Construct a flow chart that can be used to compute and print the average of cubes of the first 5 counting numbers.

(b) Perform a dry run for your flow chart

Solution:

(a) The flow chart

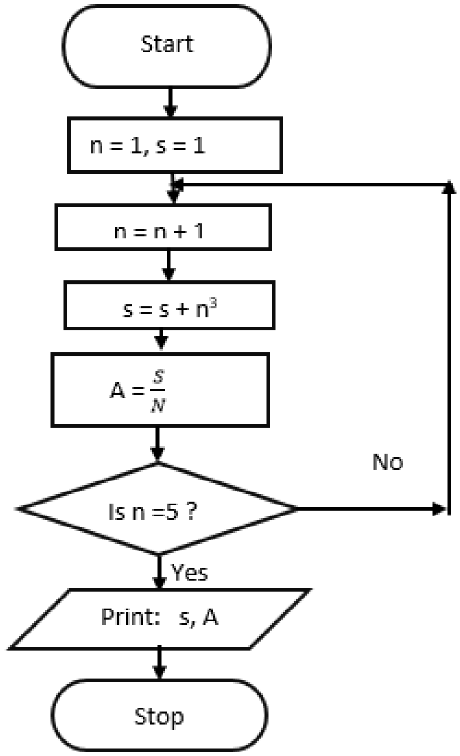


Figure 14.19

(b) The dry run.

n	S	A
1	1	1
2	9	4.5
3	36	12
4	100	25
5	225	45

$\Longrightarrow$ S = 225, A = 45

(3) Construct a flow chart that can be used to compute and print the factorial of six.
(b) Perform a dry run for your flow chart.
Solution.

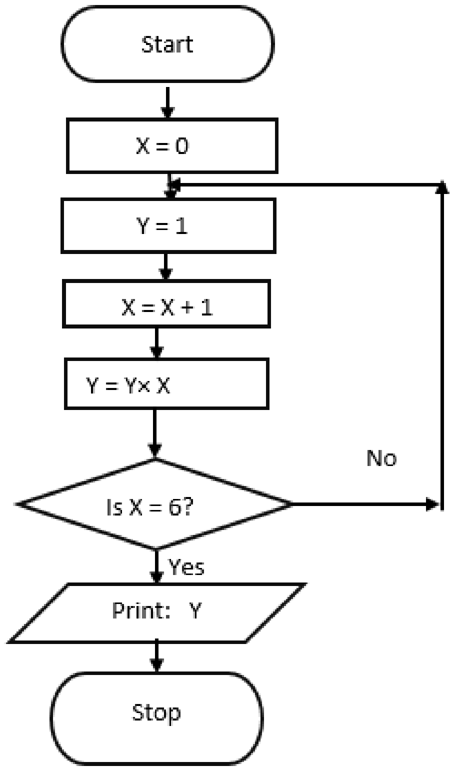


Figure 14.20

The dry run.

X	Y
0	1
1	1
2	2
3	6
4	24
5	120
6	**720**

$\Longrightarrow Y = 720$. Hence relationship $Y = X!$

(4) Construct a flow chart that can be used to compute and print the real root of the equation $ax^2 + bx + c = 0$ for $a \neq 0$.
Solution.
From $ax^2 + bx + c = 0 \Longleftrightarrow x = \dfrac{-b \pm \sqrt{b^2 - 4ac}}{2a}$, for real roots, $b^2 - 4ac \geq 0$
Let $M = b^2 - 4ac \Longrightarrow M \geq 0$

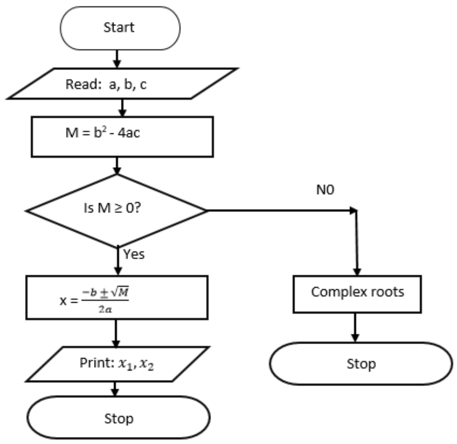


Figure 14.21

(5) Draw a flow chart for the events you perform when reporting to a serious boarding school.
Solution.

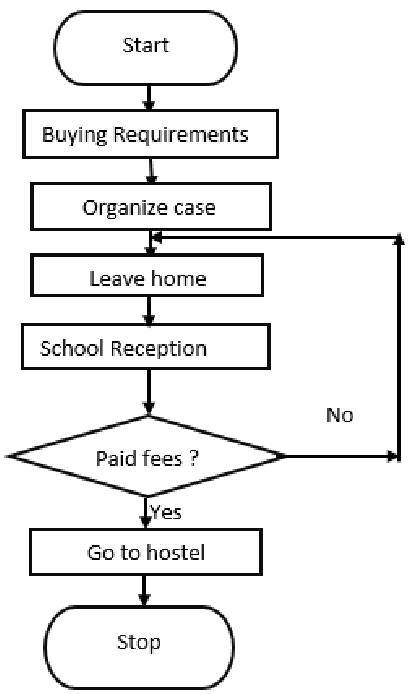


Figure 14.22

(6) Aidah has to cross a busy two lean high way road while going to school every morning. Design a flow chart that she has to follow and cross the road successfully.

Left as exercise for you.

14.1.7 Exercise 13

(1) (a) By plotting graphs $y_1 = \sin x$ and $y_2 = 1 - x$, find the initial approximation, x_0 of the root of the equation $\sin x + x - 1 = 0$

(b)Find the Newton Raphson formula for solving the equation $\sin x + x - 1 = 0$ hence draw a flow chart that reads:

(i) initial approximation x_0,

(ii) Computes and prints the root correct to three decimal places and number of iterations.

(c) Perform a dry run, using x_0 above
Ans: (iii) 0.511

(2) (a) Determine the iterative formulae for finding the fourth root of a given number,N.

(b)(i) Draw a flow chart that reads N and the initial approximation, x_0 computes and prints the fourth root of N correct to 3 decimal places and N.

(ii) Perform a dry run for N = 150.1 and $x_0 = 3.2$
Ans: 3.500

(3) (a) Show that the Newton Raphson formula for finding the natural logarithm of a number N is given by $x_{n+1} = \frac{e^{x_n}(x_n-1)+N}{e^{x_n}}, \quad n = 0, 1, 2, \cdots$

(b) Hence construct a flow chart that;

(i) Records N, the initial approximation ,x_0

(ii) Computes the logarithm to two decimal places.
(iii) Prints out N, and the natural logarithm.
(c) Perform a dry run for $N = 1.76, x_0 = 0..5$
Ans: 0.57

(4) (a) Show that the Newton Raphson iterative formula for finding the natural logarithm of a number N is given by $x_{n+1} = (x_n - 1) + Ne^{-x_n}, \quad n = 0, 1, 2, \cdots$

(b) Draw a flow chart that:

(i) Reads the initial approximations x_0 and N

(ii) Computes and prints the natural logarithm of N after 4 iterations and gives it to 3 decimal places.

(c) Perform a dry run to find $\log_e 10$, taking $x_0 = 2$

Ans: 2.303

(5) Use graphical method to show that the equation $f(x) = e^x + x - 5 = 0$ has only one positive real root. Hence, using the N–R–M, find the root to three decimal places. Illustrate your answer on a flow chart.

(6) (a) Show that the iterative formula for finding the eighth root of a number N, is given by $x_{(n+1)} = \frac{1}{8}(4x_n + \frac{N}{x_n^7}), n = 0, 1, 2, \cdots$ hence, find $80^{\frac{1}{2}}$.

(b) Draw a flow chart that reads the number N and the initial approximation x_0, computes and prints the root to three decimal places and after 3 iteration.

Ans: (ii)

(7) (a) Show that the iterative formula based on $N-R$ process for finding the logarithm to base k of a number N is given by $x_{n+1} = \frac{k^{x_n}(x_n \ln k -) + N}{k^{x_n} \ln k}, \quad n = 0, 1, 2. \cdots$

(b) construct a flow chart that;

(i) reads N, k and initial approximation x_0,

(ii) computes and prints the logarithm of N to three decimal places and after three iterations.

(c) Perform a dry run for your flow chart by taking $x_0 = 2.3$ to evaluate $\log_3 14$ to three decimal places.

Ans:

(8) (a) Draw a flow chart that computes and prints the sum of the first 15 even numbers.

(b) The information below gives a system of tax (T) calculations for the amount of money (A), earned annually by employees working in roofing company.

(A)	(T)
$<$ £2000	2%
£2,000 $\leq$ A $<$ £5,000	3.5% of A
£5,000 $\leq$ A $<$ £8,000	4.2% of A
£8,000 $\geq A$	6.6%

(b)(i) Draw a flow chart using the above data, given that the algorithm stops when 200 counts (N) are made.

(b)(ii) Calculate the Tax for an employee who earns £11,000 annually.

(9) Study the flow chart below

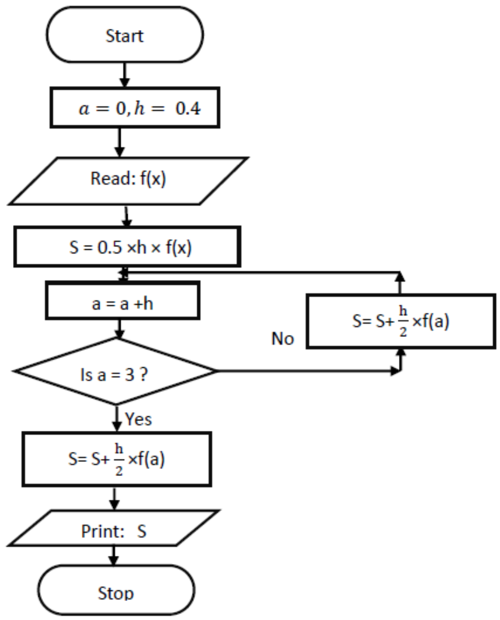


Figure 14.23

(a) Using $f(x) = \sqrt{(1+x)}$, Perform a dry run for the flow chart and you answer correct to $3dps.$,

(b) state the purpose of the flowchart,

(c) Find the exact value of $\int_1^3 \sqrt{(1+)}dx$,

(d) Find the percentage error in approximating $f(x)$.

RELEVANT FORMULAE

DESCRIPTIVE STATISTICS

Mean $\overline{x} = \frac{\sum x}{n}$ or $\overline{x} = \frac{\sum fx}{\sum f}$ Or $\overline{x} =$ A + $\left[\frac{\sum fd}{\sum f}\right] i$ Let $d = \frac{x-A}{i}$

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

Or $Var(x) = \left[\frac{\sum fd^2}{\sum f} - \left(\frac{\sum fd}{\sum f}\right)^2\right]$ Note ;$d = x - A$

Note A is working mean

Standard deviation $(s) = \sqrt{\left[\frac{\sum fx^2}{\sum f} - \left(\frac{\sum fx}{\sum f}\right)^2\right]}$

Median $= L_1 + \left(\frac{N/2 - fb}{fm}\right) i$

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

PROBABILITY THEORY

P(A∪B) = P(A) + P(B) - P(A∩B) (additional rule)

P(A∪B) = P(A) + P(B) (mutually exclusive)

P(A∩B) = P(A) P(B/A) (multiplication rule)

P(A∩B) = P(A) P(B) (for independent events)

conditional probability P(A/B) = $\frac{P(A \cap B)}{P(B)}$

Bayes'theorem P(B_1/A) = $\frac{P(B_1 \cap B)}{P(A)}$

⇒ P(B_1/A)=

$$\frac{P(B_1)P(A/B_1)}{P(B_1)P(A/B_1) + P(B_2)P(A/B_2) + \cdots + P(B_k)P(A/B_k)}$$

DISCRETE PROBABILITY DISTRIBUTION

$\sum(X = x) = 1$

Mean E(x) = $\sum xP(X = x)$

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

$$= \sum x^2 P(X = x) - \Big[\sum xP(X = x)\Big]^2$$

Note for series

$1 + 2 + 3 + 4 + 5 + \cdots + \text{n} = \frac{1}{2}n(n+1)$

$1^2 + 2^2 + 3^2 + 4^2 + 5^2 + \cdots + n^2 = \frac{1}{6}n(n+1)(2n+1)$

If S_n = a + ar + ar^2 + $\cdots$ + ar^{n-1}

$S_n = \frac{a(r^n-1)}{r-1} = \frac{a(1-r^n)}{1-r}$

If $-1 < r < 1$ then $S_n = \frac{a}{1-r}$

BINOMIAL DISTRIBUTION n or $\begin{pmatrix} n \\ n-r \end{pmatrix} = \begin{pmatrix} n \\ r \end{pmatrix}$

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	21	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	0.05	0.10	0.15	0.20	0.25	0.30	0.35	0.40	0.45	0.50
2	0	0.9801	9025	8100	7225	6400	5626	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	2603	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

n	r	0.01	0.05	0.10	0.15	0.20	0.25	0.30	0.35	0.40	0.45	0.50
8	0	0.9227	6634	4305	2725	1678	1001	0576	0319	0168	0084	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	0.0001	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
	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	0036	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

n	r	0.01	0.05	0.10	0.15	0.20	0.25	0.30	0.35	0.40	0.45	0.50
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	0146
	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	0098	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	0146
	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	0.25	0.30	0.35	0.40	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	0.0297	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	1856	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	4716	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

n	r	0.01	0.05	0.10	0.15	0.20	0.25	0.30	0.35	0.40	0.45	0.50
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
	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	6613	8031	8926	9437	9718	9865	9940	9975	9990
	2	0.0043	0961	2639	4557	6242	7560	8507	9140	9536	9787	9893
	3	0.0001	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	6852	8327	9141	9578	9802	9912	9964	9986	9995
	2	0.0052	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	2857	4304	5744	7037	8089	8867
	5		0001	0028	0159	0504	1146	2103	3317	4672	6029	7256

n	r	0.01	0.05	0.10	0.15	0.20	0.25	0.30	0.35	0.40	0.45	0.50
	6			0003	0027	0117	0343	0782	1487	2465	3669	5000
	7				0033	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
	6			0005	0046	0194	0544	1176	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

n	r	0.01	0.05	0.10	0.15	0.20	0.25	0.30	0.35	0.40	0.45	0.50
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

NORMAL DISTRIBUTION $N(0,1)p(Z)$

z	0	1	2	3	4	5	6	7	8	9	SUBTRACT 1	2	3	4	5
0.0	0.3989	3989	3989	3988	3986										
						3984	3982	3980	3977	3973	0	1	1	1	1
0.1	0.3970	3965	3961	3956	3951						0	1	1	2	2
						3945	3939	3932	3925	3918	1	1	2	3	3
0.2	0.3910	3902	3894	3885	3876						1	2	3	4	4
						3867	3857	3847	3836	3825	1	2	3	4	5
0.3	0.3814	3802	3790	3776	3765						1	2	4	5	6
						3752	3739	3725	3712	3697	1	3	4	6	7
0.4	0.3683	3668	3653	3637	3621	3605	3589	3572	3556	3538	2	3	5	6	8
0.5	0.3521	3503	3485	3467	3448	3429	3410	3391	3372	3352	2	4	6	8	10
0.6	0.3332	3312	3292	3271	3251	3230	3209	3187	3166	3144	2	4	6	8	10
0.7	0.3123	3101	3079	3056	3034	3011	2989	2966	2943	2020	2	5	7	9	11
0.8	0.2897	2874	2850	2827	2803	2780	2756	2732	2709	2685	2	5	7	10	12
0.9	0.2661	2637	2613	2589	2565	2541	2516	2492	2468	2444	2	5	7	10	12
1.0	0.2420	2396	2371	2347	2323	2299	2275	2251	2227	2203	2	5	7	10	12
1.1	0.2179	2155	2131	2107	2083	2059	2036	2012	1989	1965	2	5	7	10	12
1.2	0.1942	1919	1895	1872	1849	1826	1804	1781	1758	1736	2	5	7	9	12
1.3	0.1714	1691	1669	1647	1626	1604	1582	1561	1539	1518	2	4	7	9	11
1.4	0.1497	1476	1456	1435	1415	1394	1374	1354	1334	1315	2	4	6	8	10
1.5	0.1295	1276	1257	1238	1219	1200	1182	1183	1145	1127	2	4	6	8	9
1.6	0.1109	1092	1074	1057	1040	1023	1006	0989	0973	0957	2	3	5	7	8
1.7	0.0940	0925	0909	0893	0878	0863	0848	0833	0818	0804	2	3	5	6	8
1.8	0.0790	0775	0761	0748	0734						1	3	4	6	7
						0721	0707	0694	0681	0669	1	3	4	5	7
1.9	0.0656	0644	0632	0620	0608	0596	0584	0773	0561	0551	1	2	4	5	6
2.0	0.0540	0529	0519	0508	0498	0488	0478	0468	0459	0449	1	2	3	4	5
2.1	0.0440	0431	0422	0413	0404	0396	0389	0379	0371	0363	1	2	3	4	4
2.2	0.0355	0347	0339	0332	0325	0317	0310	0303	0297	0290	1	1	2	3	4
2.3	0.0283	0277	0270	0264	0258	0252	0246	0241	0235	0229	1	1	2	2	3
2.4	0.0244	0219	0213	0208	0203	0198	0194	0189	0184	0180	0	1	1	2	2
2.5	0.0175	0171	0167	0163	0158	0154	0151	0147	0143	0139	0	1	1	2	2
2.6	0.0136	0132	0129	0126	0122	0119	0116	0113	0110	0107	0	1	1	1	2
2.7	0.0104	0101	0099	0096	0093	0091	0088	0086	0084	0081	0	1	1	1	1
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	0.0044	0033									1	2	3	4	5
			0024	0017							1	1	2	2	3
					0012	0009	0006	0004	0003	0002					

CUMULATIVE NORMAL DISTRIBUTION ***P(z)*** ADD

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									

If the random variable Z is distributed as the standard normal curve distribution $N(0,1)$ then:

1. $P(0 < Z < z_p) = P$ (shaded area)
2. $P(Z > z_p) = Q = 0.5 - P$
3. $P(Z > |z_p|) = 1 - 2P = 2Q$

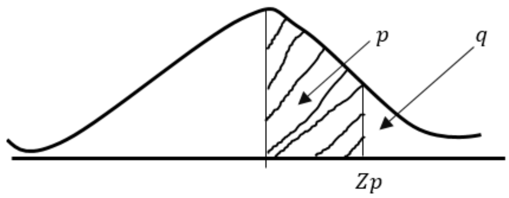


Figure 14.24

CRITICAL POINTS OF THE NORMAL DISTRIBUTION z_P

P	Q	Z	P	Q	Z	P	Q	Z
.00	.50	0.000	.460	.040	1.751	.490	.010	2.326
.05	.45	0.126	.462	.038	1.774	.491	.009	2.366
.10	.40	0.253	.464	.036	1.799	.492	.008	2.409
.15	.35	0.385	.466	.034	1.825	.493	.007	2.457
.20	.30	0.524	.468	.032	1.852	.494	.006	2.512
.25	.25	0.674	.470	.030	1.881	.495	.005	2.575
.30	.20	0.842	.472	.028	1.911	.496	.004	2.652
.35	.15	1.036	.474	.026	1.943	.497	.003	2.748
.40	.10	1.282	.476	.024	1.977	.498	.002	2.878
.45	.05	1.645	.478	.022	2.014	.499	.001	3.090
.450	.050	1.645	.480	.020	2.054	.4995	.0005	3.291
.452	.048	1.665	.482	.018	2.097	.4999	.0001	3.719
.454	.046	1.685	.484	.016	2.144	.49995	.00005	3.891
.456	.044	1.706	.486	.014	2.197	.49999	.00001	4.265
.458	.042	1.728	.488	.012	2.257	499995	.000005	4.417

SIGNIFICANCE LEVELS FOR CORRELATION COEFFICIENTS.

	Product moment coefficient of correlation (r_{xy})		Spearman's correlation coefficient (ρ)		Kendell's rank correlation coefficient (τ)	
No. of	Significance if $\lvert r_{xy}\rvert$ exceeds		Significance if $\lvert\rho\rvert$ exceeds		Significance if $\lvert\tau\rvert$ exceeds	
pairs.	at 5%	at1%	5%	at 1%	5%	at 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	0.60
12	0.58	0.71	0.58	0.71	0.45	0.58
13	0.55	0.68	0.55	0.69		
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	0.47	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	0.20	0.25		

References

14.1.8 References

1. A concise course in Advanced Level Statistics by J. CRAWSHAW and J.CHAMBERS.

2. A comprehensive Approach Advanced Level Statistics and Numerical Methods, Mukose Muhammed.