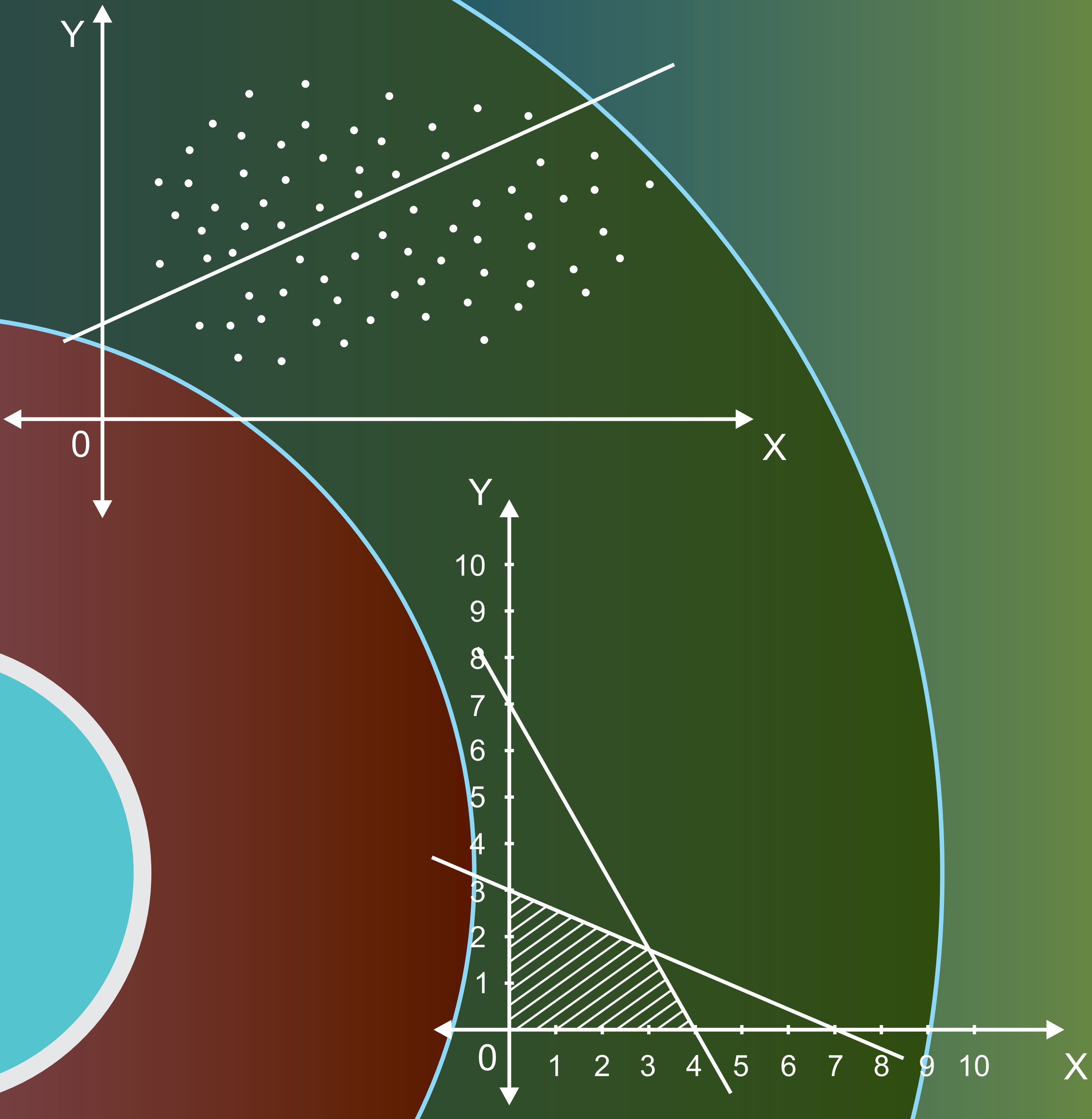
STANDARD XII

Mathematics & Statistics

Commerce Part 2


Y
0
X
Y
10
9
8
7
6
5
4
3
2
1
0
1
2
3
4
5
6
7
8
9
10
X

The Coordination Committee formed by GR No. Abhyas - 2116/(Pra.Kra.43/16) SD - 4
Dated 25.4.2016 has given approval to prescribe this textbook in its meeting held on
30.01.2020 and it has been decided to implement it from the educational year 2020-21.

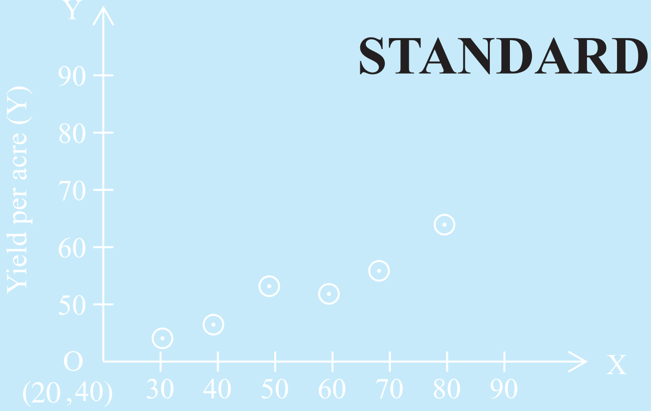
Mathematics and Statistics

(Commerce)

Part - II

Y
L2
10
L1
C(0,6)
5
B(-3,4)
X
D(-9,0)
A(-5,0)
(0,0)
X
O
L1
L2
Y

STANDARD TWELVE



Maharashtra State Bureau of Textbook Production and Curriculum Research,
Pune - 411 004

F3E1A8

Download DIKSHA App on your smartphone. If you scan the Q.R.Code on this page of your textbook, you will be able to access full text and the audio-visual study material relevant to each lesson provided as teaching and learning aids.

First Edition : 2020 © Maharashtra State Bureau of Textbook Production and Curriculum Research, Pune- 411 004.

Maharashtra State Bureau of Textbook Production and Curriculum Research reserves all rights relating to the book. No part of this book should be reproduced without the written permission of the Director, Maharashtra State Bureau of Textbook Production and curriculum Research, Pune.

Mathematics and Statistics (Commerce) Committee Members

Dr. Sharad Gore	(Chairman)
Dr. Mangla Narlikar	(Member)
Shri. Milind Patil	(Member)
Shri. Pradip Bhavsar	(Member)
Smt. Prajkti Gokhale	(Member)
Shri. Prahallad Chippalagatti	(Member)
Shri. Prasad Kunte	(Member)
Shri. Sujit Shinde	(Member)
Shri. Ramakant Sarode	(Member)
Smt. Ujjwala Godbole	(Member-Secretary)

Mathematics and Statistics (Commerce) Study Group Members

- Dr. Ishwar Patil
- Dr. Pradnyankumar Bhojnkar
- Shri. Uday Mahindrakar
- Smt. Nileema Khaldkar
- Shri. Pravin Phalake
- Shri. Pratap Kashid
- Smt. Rajashri Shinde
- Smt. Swati Powar
- Shri. Dhananjay Panhalkar
- Shri. Sharadchandra Walagade
- Shri. Prafullchandra Pawar
- Shri. Amit Totade
- Shri. Pramod Deshpande
- Dr. Pradeep Mugale
- Smt. Pradnya Kashikar
- Shri. Balkrishna Mapari
- Shri. Devanand Bagul
- Shri. Swapnil Shinde
- Shri. Vilas Bagore
- Smt. Gauri Prachand
- Smt. Mahadevi Mane
- Shri. Sachin Batwal
- Smt. Supriya Abhyankar
- Shri. Dilip Panchamia
- Smt. Deepti Kharpas

Cover, Illustrations and Computer Drawings

Shri. Sandip Koli, Artist, Mumbai

Typesetter

Baladev Computers, Mumbai

Co-ordinator

Ujjwala Shrikant Godbole
I/C Special Officer for Mathematics

Production

Sachchitanand Aphale
Chief Production Officer

Sanjay Kamble
Production Officer

Prashant Harne
Asst. Production Officer

Paper

70 GSM Cream wove

Print Order No.

N/PB/2020-21/70,000

Printer

SHAH TRADING CORPORATION, NAGPUR

Publisher

Vivek Uttam Gosavi, Controller
Maharashtra State Textbook Bureau,
Prabhadevi Mumbai- 400 025

The Constitution of India

Preamble

WE, THE PEOPLE OF INDIA, having solemnly resolved to constitute India into a SOVEREIGN SOCIALIST SECULAR DEMOCRATIC REPUBLIC and to secure to all its citizens:

JUSTICE, social, economic and political;

LIBERTY of thought, expression, belief, faith and worship;

EQUALITY of status and of opportunity; and to promote among them all

FRATERNITY assuring the dignity of the individual and the unity and integrity of the Nation;

IN OUR CONSTITUENT ASSEMBLY this twenty-sixth day of November, 1949, do HEREBY ADOPT, ENACT AND GIVE TO OURSELVES THIS CONSTITUTION.

NATIONAL ANTHEM

Jana-gana-mana-adhināyaka jaya hē
Bhārata-bhāgya-vidhātā,

Panjāba-Sindhu-Gujarāta-Marāthā
Drāvida-Utkala-Banga

Vindhya-Himāchala-Yamunā-Gangā
uchchala-jaladhi-taranga

Tava subha nāmē jāgē, tava subha āsisa māgē,
gāhē tava jaya-gāthā,

Jana-gana-mangala-dāyaka jaya hē
Bhārata-bhāgya-vidhātā,

Jaya hē, Jaya hē, Jaya hē,
Jaya jaya jaya, jaya hē.

PLEDGE

India is my country. All Indians are my brothers and sisters.

I love my country, and I am proud of its rich and varied heritage. I shall always strive to be worthy of it.

I shall give my parents, teachers and all elders respect, and treat everyone with courtesy.

To my country and my people, I pledge my devotion. In their well-being and prosperity alone lies my happiness.

PREFACE

Dear Students,

Welcome to Standard XII, an important milestone in your student life.

Standard XII or Higher Secondary Certificate opens the doors of higher education. After successfully completing the higher secondary education, you can pursue higher education for acquiring knowledge and qualification. Alternatively, you can pursue other career paths like joining the workforce. Either way, you will find that mathematics education helps you every day. Learning mathematics enables you to think logically, constistently, and rationally. The curriculum for Standard XII Commerce Mathematics and Statistics has been designed and developed keeping both of these possibilities in mind.

The curriculum of Mathematics and Statistics for Standard XII Commerce is divided in two parts. Part I deals with more theoretical topics like Mathematical Logic, Differentiation and Integration. Part II deals with application oriented topics in finance and management. Random Variables and Probability Distributions are introduced so that you will understand how uncertainty can be handled with the help of probability distributions.

The new text books have three types of exercises for focused and comprehensive practice. First, there are exercises on every important topic. Second, there are comprehensive exercises at end of all chapters. Third, every chapter includes activities that students must attempt after discussion with classmates and teachers. Textbooks cannot provide all the information that the student can find useful. Additional information has been provided on the E-balbharati website (www.ebalbharati.in).

We are living in the age of the Internet. You can make use of the modern technology with help of the Q.R. code given on the title page. The Q.R. code will take you to websites that provide additional useful information. Your learning will be fruitful if you balance between reading the text books and solving exercises. Solving more problems will make you more confident and efficient.

The text books are prepared by a subject committee and a study group. The books are reviewed by experienced teachers and eminent scholars. The Bureau would like to thank all of them for their valuable contribution in the form of creative writing, constructive cristicism and useful suggestions for making the text books valuable. The Bureau is grateful to the members of the subject committee, study group and review committee for sparing their valuable time while preparing these text books. The Bureau hopes and wishes that the text books are appreciated and well received by students, teachers and parents.

Students, you are now ready to study. All the best wishes for a happy learning experience and a well deserved success. Enjoy learning and you will succeed.

(Vivek Gosavi)
Director
Maharashtra State Bureau of Textbook
Production and Curriculum Research, Pune.

Pune
Date : 21 February 2020
Bharatiya Saur : 2 Phalguna 1941

Competency statements

Sr. No.	Area / Topic	Competency Statement
1	Commission, Brokerage, Discount	• understand terms like Agent, Commission Agent, Broker, Auctioneer, Factor, Del credere Agent • identify trade discount and cash discount • know meaning and formula of present worth, true discount, sum due, date of bill, face value, period, nominal due date, discount, banker's gain • solve problems on commission and brokerage
2	Insurance	• understand the terms premium, policy value, types of insurance (fire, marine and accident) • know rules and formulae for claims
	Annuity	• identify types of annuity • know terms related to annuity • understand annuity formulae including abbreviations used in them • solve annuity problems
3	Linear Regression	• understand the meaning of regression • understand types of regression • understand meaning of linear regression • find the regression coefficient • state the equations of regression lines • state interrelations between standard deviations, regression coefficients and correlation coefficient • remember the properties of regression coefficients • solve problems based on regression
4	Time Series	• understand the concept of a time series • identify the components of a time series • use graphical method to find the trend line for a time series • use moving averages to find the trend line for a time series • use least squares method to find the trend line for a time series
5	Index Numbers	• understand the concept of index numbers • identify types of index numbers • understand the terminology of index number • construct different index numbers • solve economic problems involving index numbers

6	Linear Programming	• understand the concept of linear programming • understand the general form and meaning of LPP • formulate a given problems as LPP • draw constraint lines and find the region of feasible solutions • obtain the optimal solution of LPP
7	Assignment	• understand the assignment problem • formulate an assignment problem • solve an assignment problem by Hungarian method • identify the special cases of assignment problem
	Sequencing	• understand the concept of job sequencing • solve problems of processing n jobs through two machines • solve problem of processing n job through three machine
8	Probability Distribution	• understand the meaning of random variables and types of random variables • understand probability mass function and its properties • understand the cumulative distribution function and its properties • find the expected value and variance of a discrete random variable • understand the probability density function and its properties • find the cumulative distribution function, expected value and variance of continous random variables
	Binomial Distribution	• understand Bernauli trial, Bernouli distribution, condition for Binomial distribution and their properties • use Binomial distribution to calculate required probabilities
	Poisson Distribution	• understand the Poission distribution and its properties • use Poisson distribution to calculate required probabilities

INDEX

Sr. No.	Chapter	Page No.
1	Commission, Brokerage and Discount	1
2	Insurance and Annuity	16
3	Linear regression	34
4	Time series	57
5	Index numbers	72
6	Linear programming	95
7	Assignment Problem and Sequencing	108
8	Probability distributions	132
9	Answers	158

1

Commission, Brokerage and Discount

Let's Study

- Commission, Brokerage
- Discount

Let's Discuss…

1.1 Commission and Brokerage Agent:

When transactions like sale, purchase, auction etc. are done through some middlemen, such middlemen are called agents.

(The charges paid to an agent for doing the work on behalf of some other person, called principal, is called commission.)

Principal: Principal refers to an individual party or parties participating in a transaction.

Commission: The charges paid to an agent for doing the work on behalf of principal is called commission. The commission or the remuneration paid to an agent is generally fixed as some percentage of the value of the transaction. For example, suppose an agent charges commission at 5% on the sales. Then in the transaction of Rs.15,000/-, the agent will get a commission of Rs.750 as follows.

$$\text{Rs. } 15000 \times \frac{5}{100} = \text{Rs. } 750/-$$

Following are different types of agents named according to their specialization.

Commission Agents: A commission agent is a person who buys or sells goods on behalf of his principal and gets commission for his service.

Broker: A broker is an agent who brings together the buyer and the seller for the purpose of purchase or sale. This commission is called brokerage and is charged to both the parties.

Auctioneer: An auctioneer is an agent who sells goods by auction. He sells goods to the highest bidder. Many a time name of the principal is not disclosed in the transaction.

Factor: A factor is an agent who is given the possession of goods and enters a contract for sale in his/her own name.

Del Credere Agent: A del credere agent gives guarantee to his principal that the party to whom he/she sells the goods will pay the sale price of goods. If a buyer is unable to pay after the transaction is completed, a del credere agent is liable for the payment.

Agent gets additional commission other than the usual commission for this. This is known as delcredere commission.

SOLVED EXAMPLES

Ex.1: A merchant gives his agent 5% ordinary commission plus 2% del credere commission on sale of goods, worth Rs.55,000/-. How much does the agent receive? How much does the merchant receive?

Solution:

Agent's Commission at 5% $= 55000 \times \frac{5}{100}$

$= \text{Rs.}2750$

Rate of delcredere $= 2\%$

Amount of delcredere $= 55000 \times \frac{2}{100}$

$= \text{Rs.}1100$

$\therefore$ Agent's Total Commission $= 2750 + 1100$

$= \text{Rs.}3850$

Merchant receives $= 55000 - 3850$

$= \text{Rs.}51150$

Ex.2: The price of a refrigerator is Rs.47,000. An agent charges commission at 6% and earns Rs.42,300. Find the number of refrigerators.

Solution: The price of a refrigerator = Rs.47000/-

Rate of commission = 6%

Commission for one refrigerator

$$= 47000 \times \frac{6}{100}$$

$$= \text{Rs. } 2820$$

Total Commission earned = Rs. 42300 (given)

∴ Number of total refrigerators sold

$$= \frac{\text{Total Commission}}{\text{Commission per refrigerator}}$$

$$= \frac{42300}{2820}$$

$$= 15$$

∴ 15 refrigerators were sold.

Ex.3: A house was sold through an agent for Rs.60 lakh. He charged 1% Commission from both buyer and seller. Calculate agent's commission. Calculate the net amount received by the seller and amount paid by the buyer.

Solution: Selling price of the house = Rs. 60 lakh.

Commission on the sale of the house

$$= 60 \times \frac{1}{100}$$

= Rs.0.60 lakh from each party.

∴ Agents Commission is Rs.1.20 lakh from both the parties.

∴ Price of the house paid by the buyer

= Rs. 60 lakh + Rs. 0.60

= Rs. 60.60 lakh

Net amount received by the seller

= 60 − 0.60

= Rs. 59.4 lakh

= Rs. 59,40,000

Ex.4: A sales representative gets fixed monthly salary plus commission based on the sales. In two successive months he received Rs.23,500 and Rs.24,250 on the sale of Rs.70,000 and Rs.85,000 respectively. Find his monthly salary and the rate of commission on sales.

Solution: Income of sales representative

= Salary + Commission on the sales

23500 = Salary+Commission on Rs.70,000 ...(1)

24250 = Salary+Commission on Rs.85,000 ...(2)

subtracting (1) from (2) we get

750 = commission on 15,000

$$\text{Rate of Commission} = \frac{100\times 750}{15000}$$

$$= 5$$

∴ Rate of Commission = 5%

$$\therefore \text{ Commission on Rs.70000} = 70000 \times \frac{5}{100}$$

$$= 3500$$

Substituting in equation (1), we get

23500 = Salary + 3500

∴ Salary = 23500 − 3500

= Rs. 20000

Ex.5: The income of an agent remain unchanged though the rate of commission is increased from 5% to 6.25%. Find the percentage reduction in the value of business.

Solution: Let the initial value of the business be Rs.100

∴ Original income of the agent = Rs.5

Let the new value of the business be Rs.x

$$\therefore \text{ New income of the agent} = x \times \frac{6.25}{100}$$

$$= \frac{x}{100} \times \frac{625}{100}$$

$$= \frac{x}{16}$$

Now Original income = new income (given)

$\therefore \quad 5 = \dfrac{x}{16}$

$\therefore \quad x = 80$

$\therefore$ New value of the business = Rs. 80

$\therefore$ There is 20% reduction in the value of the business.

Ex.6: A salesman receives 8% commission on the total sales. If his sales exceeds Rs. 20,000 he receives an additional commission at 2% on the sales over Rs. 20,000/-. If he receives Rs. 7,600 as commission, find his total sales.

Solution: Let the total sales be Rs. x.

$\therefore$ Commission at 8% on total sales

$$= x \times \frac{8}{100}$$

$$= \frac{8x}{100}$$

Sales exceeding Rs.20000 = $x - 20000$

$\therefore$ Commission at 2% on excess sales

$$= (x - 20000) \times \frac{2}{100}$$

But total commission earned = Rs.7,600

$$\therefore \quad \frac{8x}{100} + \frac{2(x-20000)}{100} = 7,600$$

$$\therefore \quad 8x + 2x - 40000 = 7600 \times 100$$

$$\therefore \quad 10x = 800000$$

$$\therefore \quad x = 80000$$

$\therefore$ His total sales is Rs.80,000

Trade Discount and Cash Discount

Discount is the reduction in the price of an article, allowed by the seller to the purchaser. It is generally expressed in terms of percentage.

There are two types of Discounts :

1) **Trade discount:** Trade discount is allowed by one trader to another. It is given on the catalogue price, list price or market price of the goods.

2) **Cash discount:** Cash discount is allowed in consideration of ready cash payment.

The buyer may be allowed both of these discounts. In such a case the trade discount is first calculated on the catalogue (list) price. The cash discount is then calculated on the price obtained after deducting the trade discount from the list price.

Invoice price = List price (Catalogue price) – Trade discount.

Selling Price / Net Selling Price = Invoice price – Cash discount

Profit = Net selling price – Cost price

Loss = Cost price – Net selling price

SOLVED EXAMPLES

M/s. Saket Electronics is given 15% trade discount and 5% cash discount on purchase of television sets by the distributor. Find the total discount availed if M/s. Saket Electronics purchases TV sets worth Rs. 12,00,000 from the distributor.

Solution: Discount at 15% on Rs 12,00,000

$$= 1200000 \times \frac{15}{100}$$

= Rs.1,80,000

$\therefore$ Invoice price of T.V. Sets

= 12,00,000 – 1,80,000

= Rs.10,20,000

Now cash discount is given on Rs.10,20,000

Cash discount at 5% on Rs.10,20,000

$$= 10,20,000 \times \frac{5}{100}$$

= Rs.51,000

$\therefore$ Total discount availed = 1,80,000 + 51,000

= Rs.2,31,000

Ex.8: Vaishnavi wants to buy an i-phone worth Rs 55,000. A shopkeeper gives 8% trade discount and 8% cash discount. Another shopkeeper gives 10% trade discount and 5% cash discount. Which shopkeeper should be preferable?

Solution: The first shopkeeper gives

$$8\% \text{ on Rs.}55000 = 55000 \times \frac{8}{100}$$

$$= \text{Rs.}4400$$

$\therefore$ Invoice price of the i-phone $= 55000 - 4400$

$$= 50,600$$

Cash discount at 8% on Rs.50,600

$$= 50600 \times \frac{8}{100}$$

$$= \text{Rs. } 4048$$

Net amount payable to the first shopkeeper

$$= 50600 - 4048$$

$$= \text{Rs.}46552$$

Second shopkeeper gives 10% discount on

$$\text{Rs.}55000 = 55000 \times \frac{10}{100}$$

$$= \text{Rs.}5500$$

Net price after deducting trade discount

$$= 55000 - 5500$$

$$= \text{Rs.}49500$$

Cash discount at 5% on Rs.49,500

$$= 49500 \times \frac{5}{100}$$

$$= \text{Rs. } 2475$$

Net amount payable to the second shopkeeper.

$$= 49,500 - 2475$$

$$= \text{Rs. } 47025$$

$\therefore$ The first shopkeeper should be preferred.

Ex.9: A motor bike is marked at Rs 50,000. A retailer allows a discount at 16% and still gains 20% on the cost. Find purchase price of the retailer.

Solution: List price of the motor bike $= \text{Rs } 50,000$

Discount at 16% on Rs 50,000.

$$= 50000 \times \frac{16}{100}$$

$$= \text{Rs.}8,000$$

$\therefore$ Selling price $= 50000 - 8000$

$$= \text{Rs.}42,000$$

In case the purchase price is Rs.100, the selling price is Rs.120

$\therefore$ For selling price $= 42,000$

$$\text{the purchase price} = \frac{100 \times 42000}{120}$$

$$= \text{Rs.}35,000$$

$\therefore$ Purchase Price of the motor bike is Rs.35,000.

Ex.10: Prakash gets a commission at 10% on cash sales and 8% on credit sales. If he receives Rs 4,400 as commission on the total sales of Rs 50,000. Find the sales made by him in cash and on credit.

Solution: Let the cash sales be Rs. x

$\therefore$ Credit sales = Rs. $(50,000 - x)$

Total commission =10% on x + 8% on $(50,000 - x)$

$$\therefore 4400 = x \times \frac{10}{100} + (50000 - x) \times \frac{8}{100}$$

$$= \frac{10x}{100} + \frac{400000 - 8x}{100}$$

$$= \frac{400000 + 2x}{100}$$

$$4400 \times 100 = 400000 + 2x$$

$$\therefore \quad 2x = 40000$$

$$x = 20,000$$

$\therefore$ Prakash's cash sales is Rs.20,000 and his credit sales is

$$50,000 - 20,000 = \text{Rs.}30,000$$

Ex.11: Mr. Anand charges 10% commission on cash sales and 8% commission on credit sales. If his overall commission is 8.8%, Find the ratio of cash sales to the credit sales.

Solution: Let the cash sales be Rs. x

and the credit sales be Rs. y

Commission on cash sales is 10%

$$= x \times \frac{10}{100}$$

$$= \frac{10x}{100}$$

Commission on credit sales is 8%

$$= y \times \frac{8}{100}$$

$$= \frac{8y}{100}$$

Anand's Total Sales = $(x + y)$

∴ Commission at 8.8% on the total sales

$$= \frac{(x+y)\times 8.8}{100}$$

$$= \frac{(x+y)\times 88}{1000}$$

$$\therefore \quad \frac{10x}{100} + \frac{8y}{100} = \frac{88x+88y}{1000}$$

$$\therefore \quad 100x + 80y = 88x + 88y$$

$$\therefore \quad 12x = 8y$$

$$\therefore \quad \frac{x}{y} = \frac{8}{12} = \frac{2}{3}$$

∴ Ratio of cash sales to the credit sales is 2:3

EXERCISE 1.1

1. An agent charges 12% commission on the sales. What does he earn if the total sale amounts to Rs. 48,000? What does the seller get?

2. A salesman receives 3% commission on the sales up to Rs. 50,000 and 4% commission on the sales over Rs. 50,000. Find his total income on the sale of Rs. 2,00,000.

3. Ms. Saraswati was paid Rs. 88,000 as commission on the sale of computers at the rate of 12.5%. If the price of each computer was Rs. 32,000, how many computers did she sell ?

4. Anita is allowed 6.5% commission on the total sales made by her, plus a bonus of $\frac{1}{2}$% on the sale over Rs.20,000. If her total commission amount to Rs. 3400. Find the sales made by her.

5. Priya gets salary of Rs. 15,000 per month and commission at 8% on the sales over Rs.50,000. If she gets Rs. 17,400 in a certain month, Find the sales made by her in that month.

6. The income of a broker remains unchanged though the rate of commission is increased from 4% to 5%. Find the percentage reduction in the value of the business.

7. Mr. Pavan is paid a fixed weekly salary plus commission based on percentage of sales made by him. If on the sale of Rs.68,000 and Rs. 73,000 in two successive weeks, he received in all Rs.9,880 and Rs.10,180, Find his weekly salary and the rate of commission paid to him.

8. Deepak's salary was increased from Rs.4,000 to Rs. 5,000. The sales being the same, due to reduction in the rate of commission from 3% to 2%, his income remained unchanged. Find his sales.

9. An agent is paid a commission of 7% on cash sales and 5% on credit sales made by him. If on the sale of Rs.1,02,000 the agent claims a total commission of Rs.6,420, find his cash sales and credit sales.

10. Three cars were sold through an agent for Rs.2,40,000 , Rs.2,22,000 and Rs.2,25,000 respectively. The rates of commission were 17.5% on the first, 12.5% on the second. If the agent overall received 14% commission on the total sales, find the rate of commission paid on the third car.

11. Swatantra Distributers allows 15% discount on the list price of washing machine. Further 5% discount is given for cash payment. Find the list price of the washing machine if it was sold for the net amount of Rs. 38356.25.

12. A book seller received Rs.1,530 as 15% commission on list price. Find list price of the books.

13. A retailer sold a suit for Rs.8,832 after allowing 8% discount on marked price and further 4% cash discount. If he made 38% profit, find the cost price and the marked price of the suit.

14. An agent charges 10% commission plus 2% delcreder. If he sells goods worth Rs.37,200, find his total earnings.

15. A whole seller allows 25% trade discount and 5% cash discount. What will be the net price of an article marked at Rs. 1600.

Let's discuss

1.2 Discount

a) Present worth, sum due, true discount

When businessmen sell goods on credit, the price quoted for goods includes a sufficient margin of interest for the period of credit allowed.

Suppose the goods are worth Rs.100, if the payment is made on the spot. However if a credit of 4 months is allowed, then the businessmen will quote the price by adding interest for 4 months to Rs.100. If the rate of interest is 12% per annum then the interest for 4 months will be Rs.4. Therefore the customer has to pay Rs.104 after 4 months.

In other words Rs.104 due after 4 months at 12% p.a. are equivalent to Rs.100 today. Hence Rs.100 are known as present value (P.V.) of Rs.104 due after 4 months. Hence at 12% per annum Rs.104 is known as sum due (S.D.) and Rs.4 is known as the true discount (T.D) on the sum due.

The true discount is the interest on the present worth at the given rate of interest for the given period.

We have,

Present worth + True discount = Sum due

i.e. P.W. + T.D. = S.D.

$$\text{T.D.} = \frac{\text{P.W.} \times n \times r}{100},$$

where P.W. is the principal or the present worth, n is period of the bill in years, r is the rate of interest per annum.

Drawer and Drawee:

A person who draws the bill is called the drawer. A person on whom the bill is drawn is called as Drawee.

Date of bill and Face value:

The date on which the bill is drawn is called as 'date of bill'. The amount for which the bill is drawn is called face value (F.V.) of the bill. It is the sum due on the present worth.

Period of the bill is the time after completion of which the drawee receives the payment.

Nominal Due Date and Legal Due Date:

The date on which the period of bill expires is called the nominal due date. The buyer has to make the payment to the seller on this date.

However, the buyer is allowed to pay the amount even 3 days later. These 3 days are called the days of grace. The date obtained after adding 3 days of grace to the nominal due date is known as the legal due date.

Discounting a Bill:

If the drawee of the bill wants money before the legal due date, then there is a facility available at the bank or with an agent who can discount a bill and pay the amount to the drawer (after deducting some amount from face value of the bill).This is called discounting the bill.

Banker's Discount, Cash Value, Banker's Gain:

When a bill is discounted in a bank, the banker will deduct the amount from the face value of the bill at the given rate of interest for the period from the date of discounting to the legal due date and pay the balance to the drawer. This amount is known as Banker's Discount(B.D).

The amount paid to the holder of the bill after deducting banker's discount is known as Cash Value (C.V) of the bill paid on the date of discounting.

The banker's discount is called commercial discount.

Thus, true discount is calculated on present worth and the banker's discount is calculated on the face value. Hence the banker's discount is always higher than the true discount.

The difference between the banker's discount and the true discount is called Banker's Gain (B.G). It is equal to the interest on true discount.

Abbreviations:

Present Worth : P.W. or P
Sum Due/Face Value : S.D. or F.V.
True Discount : T.D.
Banker's Gain : B.G.
Banker's Discount : B.D.
Cash Value : C.V.

Notation

Period (in Years) : n
Rate of Interest (p.a.) : r

List of Formula:

(1) S.D. $= \text{P.W.} + \text{T.D.}$

(2) T.D. $= \dfrac{\text{P.W.}\times n\times r}{100}$

(3) B.D. $= \dfrac{\text{S.D.}\times n\times r}{100}$

(4) B.G. $= \text{B.D.} - \text{T.D.}$

(5) B.G. $= \dfrac{\text{T.D.}\times n\times r}{100}$

(6) Cash value $= \text{S.D.} - \text{B.D.}$

SOLVED EXAMPLES

Ex.1: If the present worth of a bill due six months hence is Rs.23,000 at 8% p.a., What is sum due?

Solution: P.W. = 23,000, $r = 8\%$

$$n = 6 \text{ months} = \frac{1}{2} \text{ year}$$

$$\text{T.D.} = \frac{\text{P.W.}\times n\times r}{100}$$

$$= \frac{23000\times\frac{1}{2}\times 8}{100}$$

$$= \text{Rs.}920$$

$$\text{Now S.D.} = \text{P.W.} + \text{T.D.}$$

$$= 23{,}000 + 920$$

$$= \text{Rs.}23{,}920$$

$\therefore$ The sum due is Rs.23,920

Ex.2: What is the true discount on a sum of Rs.12,720 due 9 months hence at 8% p.a. simple interest?

Solution: S.D. $= \text{Rs.}12{,}720,\ n = \dfrac{9}{12}$ years,

$$r = 8\%$$

$$\text{Now S.D.} = \text{P.W.} + \text{T.D.}$$

$$= \text{P.W.} + \frac{\text{P.W.}\times n\times r}{100}$$

$$= \text{P.W.}\left[1+\frac{n\times r}{100}\right]$$

$$\therefore\ 12720 = \text{P.W.}\left[1+\frac{\frac{9}{12}\times 8}{100}\right]$$

$$= \text{P.W.}\left[1+\frac{6}{100}\right]$$

$$= \text{P.W.}\ \frac{106}{100}$$

$$\therefore\ \text{P.W.} = \frac{12720\times 100}{106}$$

$$= \text{Rs.}12{,}000$$

$$\therefore \quad \text{T.D.} = \text{S.D.} - \text{P.W.}$$

$$= 12{,}720 - 12{,}000$$

$$\therefore \quad \text{T.D.} = \text{Rs.}720$$

Ex.3: The present worth of sum of Rs.8,268 due 8 months hence is Rs.7,800. Find the rate of interest.

Solution: S.D. = Rs.8,268, P.W. = 7,800,

$$n = \frac{8}{12} \text{ years}$$

$$\text{Now T.D.} = \text{S.D.} - \text{P.W.}$$

$$= 8{,}268 - 7{,}800$$

$$= 468$$

$$\text{T.D.} = \frac{\text{P.W.} \times n \times r}{100}$$

$$468 = \frac{7800 \times \frac{8}{12} \times r}{100}$$

$$468 = 78 \times \frac{2}{3} \times r$$

$$r = \frac{468 \times 3}{78 \times 2}$$

$$= 9$$

$\therefore$ Rate of interest is 9%

Ex.4: A bill of Rs.15,000 drawn on 15th February 2015 for 10 months was discounted on 13th May 2015 at $3\frac{3}{4}\%$ p.a. Calculate banker's discount.

Solution: F.V. of the bill= Rs 15,000

$$r = 3\frac{3}{4}\,\% = \frac{15}{4}\,\%$$

Date of drawing = 15th February 2015

Period of bill = 10 months

Nominal due date = 15th December 2015

Legal due date = 18th December 2015

Date of discounting = 13th May 2015

$\therefore$ Number of days from the date of discounting to the legal due date.

May	June	July	Aug.	Sept.	Oct.	Nov.	Dec.	Total
18	30	31	31	30	31	30	18	219

$$\text{Period} = n = \frac{219}{365} = \frac{3}{5} \text{ years}$$

B.D.= interest on F.V. for 219 days at $3\,{}^{3}\!/_{4}\,\%$ p.a.

$$= \frac{15000 \times \frac{3}{5} \times \frac{15}{4}}{100}$$

$$= 337.5$$

$\therefore$ Banker's discount is Rs.337.5

Ex.5: A bill of Rs.10,100 drawn on 14th January for 5 months was discounted on 26th March. The customer was paid Rs.9,939.25. Calculate the rate of interest.

Solution: F.V = Rs.10,100, C.V = Rs.9,939.25

Banker's Discount (B.D) = F.V- C.V

$$= 10{,}100 - 9{,}939.25$$

$$= 160.75$$

Date of drawing = 14th January

Period = 5 months

Nominal due date = 14th June

Legal due date = 17th June

Date of discounting = 26th March

Number of days from the date of discounting to legal due date.

March	April	May	June	Total
05	30	31	17	83

$\therefore$ B.D. = interest on F.V. for 83 days at r %

$$160.75 = 10{,}100 \times \frac{83}{365} \times \frac{r}{100}$$

$$\therefore \quad r = \frac{16075 \times 365}{10100 \times 83}$$

$$= 6.99$$

$\therefore$ Rate of interest is 6.99% ≈ 7%

Ex.6: A bill of Rs.18,000 was discounted for Rs.17,568 at a bank on 25th October 2017. If the rate of interest was 12% p.a., what is the legal date?

Solution : S.D. = 18,000 C.V. = 17,568

$r = 12\%$ p.a.

Now, B.D. = S.D. - C.V.

= 18,000 - 17,568

= Rs 432

$$\text{Also B.D.} = \frac{\text{S.D.} \times n \times r}{100}$$

$$432 = \frac{18000 \times n \times 12}{100}$$

$$\therefore n = \frac{432 \times 100}{18000 \times 12}$$

$$\therefore n = \frac{1}{5}$$

$$\therefore n = \frac{1}{5} \text{ years}$$

$$\frac{365}{5} = 73 \text{ days}$$

The period for which the discount is deducted is 73 days, which is counted from the date of discounting i.e. 25th October 2017.

Oct.	Nov.	Dec.	Jan.	Total
06	30	31	6	73

Hence legal due date is 6th January 2017.

Ex.7: A bill of Rs.29,200 drawn on 15th June for 6 months, was discounted for Rs.28,960 at 5% p.a. On which day was the bill discounted?

Solution: F.V. = Rs 29200 C.V. = Rs 28960

Now B.D. = F.V. – C.V.

= 29200 − 28960

= Rs. 240

Date of drawing = 15th June

Period = 6 months

Nominal due date = 15th December

Legal due date = 18th December

B.D. = Interest on F.V. for n at 5% p.a.

$$\therefore \quad 240 = \frac{29200 \times \frac{n}{365} \times 5}{100}$$

$$\therefore \quad n = \frac{240 \times 365}{5 \times 292}$$

= 60 days

∴ Date of discounting is 60 days before 18th December

Dec.	Nov.	Oct.	Total
18	30	12	60

Date of discounting = (31 – 12) = 19th Oct.

∴ Date of discounting is 19th October.

Ex.8: Find the true discount, banker's discount and banker's gain on a bill of Rs.64,800 due 3 months hence discounted at 5% p.a.

Solution: S.D. = 64,800,

$$n = 3 \text{ months} = \frac{3}{12} = \frac{1}{4} \text{ years}$$

$r = 5\%$ p.a.

$$\text{Now} \quad \text{B.D.} = \frac{\text{S.D.} \times n \times r}{100}$$

$$= 64,800 \times \frac{1}{4} \times \frac{5}{100}$$

= Rs. 810

Let T.D. = Rs. x

B.D. = T.D. + Interest on T.D. for $\frac{1}{4}$ year at 5% p.a.

$$810 = x + \left(x \times \frac{1}{4} \times \frac{5}{100}\right)$$

$$= x + \frac{x}{80}$$

$$= \frac{81x}{80}$$

$$\therefore x = \frac{810 \times 80}{81}$$

$$= \text{Rs.}800$$

Banker's gain = banker's discount − true discount

i.e. B.G. = B.D. − T.D.

= 810 − 800

= Rs.10

∴ Banker's gain is Rs.10

Ex.9: The difference between true discount and banker's discount on a bill due 6 months hence at 4% is Rs.160. Calculate true discount, banker's discount and amount of bill.

Solution: Let T.D. = Rs. x , n = $\frac{6}{12} = \frac{1}{2}$ years

B.G. = B.D. – T.D.

= Interest on T.D for 6 months at 4% p.a.

$$\therefore 160 = x \times \frac{1}{2} \times \frac{4}{100} = \frac{x}{50}$$

$$\therefore x = 8{,}000$$

B.D. = B.G. + T.D.

= 160 + 8,000

= 8,160

∴ Banker's Discount = Rs.8,160

B.D. = interest on F.V. for 6 months at 4% p.a.

Let the face value (F.V.) be y.

$$\therefore \text{B.D.} = y \times \frac{1}{2} \times \frac{4}{100}$$

$$8{,}160 = \frac{y}{50}$$

∴ y = 4,08,000

∴ Amount of bill is Rs.4,08,000

Ex.10: A banker's discount calculated for 1 year is 13.5 times its banker's gain. Find the rate of interest.

Solution: Let the banker's gain = Rs. x

∴ B.D. = 13.5 × x

= 13.5 x

Now B.G. = B.D – T.D.

x = 13.5x – T.D.

T.D. = 12.5 x

But B.G. = Interest on T.D. for 1 year

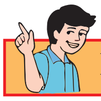
$$= \frac{\text{T.D.} \times n \times r}{100}$$

$$x = \frac{12.5x \times 1 \times r}{100}$$

$$100\,x = 12.5\,xr$$

$$r = 8$$

∴ Rate of interest is 8% p.a.



Let's Remember

- Invoice Price = List Price (catalogue price) – Trade Discount

  i.e. I.P. = L.P. – T.P.

- Net Selling Price = Invoice Price – Cash Discount

  i.e. N.S.P. = I.P. – C.D.

- Profit = Net Selling price (NSP) – Cost Price (C.P)

- Loss = Cost Price (C.P) – Net Selling Price (NSP)

- Sum due = Present Worth + True Discount

  i.e. S.D = P.W. + T.D.

- T.D. = $\frac{\text{P.W.} \times n \times r}{100}$

- B.D. = $\frac{\text{S.D.} \times n \times r}{100}$

- B.G. = B.D. − T.D.

- B.G. = $\frac{\text{T.D.} \times n \times r}{100}$

- Cash value = S.D. − B.D.

EXERCISE 1.2

1. What is the present worth of a sum of Rs.10,920 due six months hence at 8% p.a. simple interest?
2. What is sum due of Rs.8,000 due 4 months hence at 12.5% simple interest?
3. True discount on the sum due 8 months hence at 12% p.a. is Rs.560. Find the sum due and present worth of the bill.
4. The true discount on a sum is $\frac{3}{8}$ of the sum due at 12% p.a. Find the period of the bill.
5. 20 copies of a book can be purchased for a certain sum payable at the end of 6 months and 21 copies for the same sum in ready cash. Find the rate of interest.
6. Find the true discount, Banker's discount and Banker's gain on a bill of Rs.4,240 due 6 months hence at 9% p.a.
7. True discount on a bill is Rs.2,200 and bankers discount is Rs.2,310. If the bill is due 10 months, hence, find the rate of interest.
8. A bill of Rs.6,395 drawn on 19th January 2015 for 8 months was discounted on 28th February 2015 at 8% p.a. interest. What is the banker's discount? What is the cash value of the bill?
9. A bill of Rs.8,000 drawn on 5th January 1998 for 8 months was discounted for Rs.7,680 on a certain date. Find the date on which it was discounted at 10% p.a.
10. A bill drawn on 5th June for 6 months was discounted at the rate of 5% p.a. on 19th October. If the cash value of the bill is Rs 43,500, find face value of the bill.
11. A bill was drawn on 14th April for Rs.7,000 and was discounted on 6th July at 5% p.a. The Banker paid Rs.6,930 for the bill. Find period of the bill.
12. If difference between true discount and banker's discount on a sum due 4 months hence is Rs 20. Find true discount, banker's discount and amount of bill, the rate of simple interest charged being 5%p.a.
13. A bill of Rs.51,000 was drawn on 18th February 2010 for 9 months. It was encashed on 28th June 2010 at 5% p.a. Calculate the banker's gain and true discount.
14. A certain sum due 3 months hence is $\frac{21}{20}$ of the present worth, what is the rate of interest?
15. A bill of a certain sum drawn on 28th February 2007 for 8 months was encashed on 26th March 2007 for Rs.10,992 at 14% p.a. Find the face value of the bill.

MISCELLANEOUS EXERCISE - 1

I) Choose the correct alternative.

1. An agent who gives guarantee to his principal that the party will pay the sale price of goods is called

   a. Auctioneer   b. Del Credere Agent

   c. Factor   d. Broker

2. An agent who is given the possession of goods to be sold is known as

   a. Factor   b. Broker

   c. Auctioneer   d. Del Credere Agent

3. The date on which the period of the bill expires is called

   a. Legal Due Date

   b. Grace Date

   c. Nominal Due Date

   d. Date of Drawing

4. The payment date after adding 3 days of grace period is known as

   a. The legal due date

   b. The nominal due date

   c. Days of grace

   d. Date of drawing

5. The sum due is also called as

   a. Face value   b. Present value

   c. Cash value   d. True discount

6. P is the abbreviation of
   a. Face value b. Present worth
   c. Cash value d. True discount

7. Banker's gain is simple interest on
   a. Banker's discount
   b. Face Value
   c. Cash value
   d. True discount

8. The marked price is also called as
   a. Cost price b. Selling price
   c. List price d. Invoice price

9. When only one discount is given then
   a. List price = Invoice price
   b. Invoice price = Net selling price
   c. Invoice price = Cost price
   d. Cost price = Net selling price

10. The difference between face value and present worth is called
    a. Banker's discount
    b. True discount
    c. Banker's gain
    d. Cash value

II) Fill in the blanks.

1. A person who draws the bill is called ____
2. An ______ is an agent who sells the goods by auction.
3. Trade discount is allowed on the ___ price.
4. The banker's discount is also called ______.
5. The banker's discount is always ______ than the true discount.
6. The difference between the banker's discount and the true discount is called _______.
7. The date by which the buyer is legally allowed to pay the amount is known as ______.
8. A ______ is an agent who brings together the buyer and the seller.
9. If buyer is allowed both trade and cash discounts, ______ discount is first calculated on _______ price.
10. _______ = List price (catalogue Price) – Trade Discount.

III) State whether each of the following is True or False.

1. Broker is an agent who gives a guarantee to seller that the buyer will pay the sale price of goods.
2. Cash discount is allowed on list price.
3. Trade discount is allowed on catalogue price.
4. The buyer is legally allowed 6 days grace period.
5. The date on which the period of the bill expires is called the nominal due date.
6. The difference between the banker's discount and true discount is called sum due.
7. The banker's discount is always lower than the true discount.
8. The bankers discount is also called as commercial discount.
9. In general cash discount is more than trade discount.
10. A person can get both, trade discount and cash discount.

IV) Solve the following problems.

1. A salesman gets a commission of 6.5% on the total sales made by him and bonus of 1% on sales over Rs.50,000. Find his total income on a turnover of Rs.75,000.
2. A shop is sold at 30% profit, the amount of brokerage at the rate of $\frac{3}{4}$% amounts to Rs.73,125. Find cost of the shop.

3. A merchant gives 5% commission and 1.5% del credere to his agent. If the agent sells goods worth Rs.30,600 how much does he get? How much does the merchant receive?

4. After deducting commission at $7\frac{1}{2}\%$ on first Rs.50,000 and 5% on balance of sales made by him, an agent remits Rs.93750 to his principal. Find the value of goods sold by him.

5. The present worth of Rs.11,660 due 9 months hence is Rs.11,000. Find the rate of interest.

6. An article is marked at Rs.800, a trader allows a discount of 2.5% and gains 20% on the cost. Find the cost price of the article

7. A salesman is paid fixed monthly salary plus commission on the sales. If on sale of Rs.96,000 and Rs.1,08,000 in two successive months he receives in all Rs.17,600 and Rs.18,800 respectively. Find his monthly salary and rate of commission paid to him.

8. A merchant buys some mixers at 15% discount on catalogue price. The catalogue price is Rs.5,500 per piece of mixer. The freight charges amount to $2\frac{1}{2}\%$ on the catalogue price. The merchant sells each mixer at 5% discount on catalogue price. His net profit is Rs.41,250. Find number of mixers.

9. A bill is drawn for Rs.7000 on 3rd May for 3 months and is discounted on 25th May at 5.5%. Find the present worth.

10. A bill was drawn on 14th April 2005 for Rs.3,500 and was discounted on 6th July 2005 at 5% per annum. The banker paid Rs.3,465 for the bill. Find the period of the bill.

11. The difference between true discount and banker's discount on 6 months hence at 4% p.a. is Rs.80. Find the true discount, banker's discount and amount of the bill.

12. A manufacturer makes clear profit of 30% on cost after allowing 35% discount. If the cost of production rises by 20%, by what percentage should he reduce the rate of discount so as to make the same rate of profit keeping his list prices unaltered.

13. A trader offers 25% discount on the catalogue price of a radio and yet makes 20% profit. If he gains Rs.160 per radio, what must be the catalogue price of the radio?

14. A bill of Rs.4,800 was drawn on 9th March 2006 at 6 months and was discounted on 19th April 2006 for $6\frac{1}{4}\%$ p.a. How much does the banker charge and how much does the holder receive?

15. A bill of Rs.65,700 drawn on July 10 for 6 months was discounted for Rs.65,160 at 5% p.a. On what day was the bill discounted?

16. An agent sold a car and charged 3% commission on sale value. If the owner of the car received Rs.48,500, find the sale value of the car. If the agent charged 2% from the buyer, find his total remuneration.

17. An agent is paid a commission of 4% on cash sales and 6% on credit sales made by him. If on the sale of Rs.51,000 the agent claims a total commission of Rs.2,700, find the sales made by him for cash and on credit.

Activities

1) The value of the goods sold = Rs. x

Commission @ 7.5% on first Rs.10,000

= Rs. $\boxed{}$

Commission @ 5% on the balance

$$\text{Rs.}(x-10{,}000) = \frac{5}{100} \times \boxed{}$$

$$= \text{Rs. } \boxed{}$$

An Agent remits Rs.33950 to his Principal

$$\therefore \quad x - \boxed{} - \boxed{} = 33{,}950$$

$$\frac{95x}{100} = 33950 + \boxed{}$$

$$\frac{19x}{\boxed{}} = 34{,}200$$

$$x = \text{Rs. } \boxed{}$$

2) Rate of discount = 15% and other charges = 2.5% on list price.

List price of tricycle in Mumbai = Rs. 600

Net Selling price = List Price – Discount

$\boxed{}$ + Other charges

$$= 600 - \frac{\boxed{}}{100} \times 600 + \frac{2.5}{100} \times \boxed{}$$

List price of tricycle in Nashik = Rs. 750

Rate of discount = 10%

Net Selling price = List Price – Discount

$$= \boxed{} - \frac{\boxed{}}{100} \times 750$$

= Rs.675

A merchant bought tricycles from Mumbai and sold it in Nashik and made a profit of Rs.13,500

$$\therefore \quad \text{Profit per tricycles} = 675 - \boxed{}$$

$$= \text{Rs.}150$$

$$\text{No. of tricycles bought} = \frac{\text{Total Profit}}{\text{Profit per tricycles}}$$

$$= \frac{13500}{\boxed{}}$$

$$= \boxed{}$$

3) Cost Price = Rs.100

A manufacturer makes a profit of 30% on cost after allowing 35% discount.

$$\therefore \quad \text{Selling price} = \boxed{} + \text{profit}$$

$$= 100 + \frac{30}{100} \times \boxed{}$$

$$= \text{Rs.}130$$

$$\therefore \quad \text{Selling price} = \text{List price} - \text{Discount}$$

$$\therefore \quad 130 = \text{List price} - \frac{35}{100} \times \boxed{}$$

$$\therefore \quad 130 = \frac{65}{100} \times \boxed{}$$

$$\therefore \quad \text{List price} = \frac{130 \times 100}{\boxed{}}$$

$$= \text{Rs.}200$$

Now the cost of production rises by 20%

$$\therefore \quad \text{New cost price} = 100 + \frac{\boxed{}}{100} \times 100$$

$$= \text{Rs.}120$$

New list price = Rs 200

Rate of discount = x%

$$\therefore \quad \text{Selling price} = \boxed{} + \text{profit}$$

$$= 120 + \frac{\boxed{}}{100} \times 120$$

$$= \text{Rs.}156$$

$$\therefore \quad \text{List price} = \text{Selling price} + \text{Discount}$$

$$\therefore \quad \boxed{} = \boxed{} + \frac{x}{100} \times 200$$

$$\therefore \quad 2x = 200 - \boxed{}$$

$$\therefore \quad 2x = 44$$

$$\therefore \quad x = \boxed{}$$

$\therefore$ Reduction in the rate of discount

$$= \boxed{} - 22$$

$$= 13\%$$

4) Face Value (S) = Rs 4,015 $r = 8\%$ p.a.

Date of drawing bill = 19th January 2018

Period of the bill = 8 months

Nominal Due date = ☐

Legal Due date = 22nd September 2018

Date of discounting the bill = 28th February 2018

Number of days from date of discounting to legal due date

March	April	May	June	July	Aug.	Sept.	Total
31	☐	☐	30	☐	31	☐	206 days

$$\therefore \quad \text{B.D.} = \frac{s \times n \times r}{100} = 4015 - \boxed{\quad}$$

= Rs.3833.70

5. Face value (S) = Rs.7,300, $r = 12\%$ p.a.

Cash value (C.V.) = Rs.7,108

$\therefore$ B.D. = S − ☐ = Rs.192

Date of drawing the bill = 7th June 2017

Date of discounting the bill = 22nd October 2017

Number of days from date of discounting to legal due date = x

$$\therefore \quad \text{B.D.} = \frac{\text{S.D.} \times n \times r}{100}$$

$$\therefore \quad \boxed{\quad} = \boxed{\quad} \times \frac{x}{\boxed{\quad}} = \frac{12}{100}$$

$\therefore$ x = ☐ days

$\therefore$ Legal due date is 80 days after the date of discounting the bill.

Oct.	Nov.	Dec.	Jan.	Total
☐	30	☐	☐	80

$\therefore$ Legal due date = 10th Jan. ☐

Nominal due date = 7th Jan. ☐

$\therefore$ period of the bill = ☐ months

◆ ◆ ◆

2 Insurance and Annuity

Let's Learn

- Fire, Marine and Accident Insurance
- Annuity
  - Terminology of Annuity
  - Annuity Due
  - Sinking Fund

Introduction

Life is full of risk. Risk is due to uncertainty. It involves a loss or some other undesirable or negative outcome. All of us face some type of risk in every-thing we do and every decision we make. We often look for ways to avoid risk by taking steps to prevent it. However, it is impossible to completely prevent every possible risk. As an alternative, we try to minimize the impact of risk by having insurance or some other type of protection from loss. Insurance is a way of managing the risk in order to protect our life, property, vehicle, or other financial assets against possible loss or damage due to contingencies like burglary, fire, flood, earthquake, etc.

The verb "to insure" means to arrange for compensation in the event of damage or total loss of property or injury or the death of someone, in exchange of regular payments to a company or to the state. The word "insurance" means creation of some security or monetary protection against a possible damage or loss. Insurance is a legal contract between an insurance company (insurer) and a person covered by the insurance (insured). An insurance policy is a legal document of the contract or agreement between the two parties, the insured and the insurer.

Insurance is of two types: Life Insurance and General Insurance.

(1) Life insurance

A person who wishes to be insured for life agrees to pay the insurance company a certain amount of money periodically. This amount is called the premium. The period of the payment can be a month, a quarter, half-year, or a year. In return, the insurance company agrees to pay a definite amount of money in the event of death of the insured or maturity of the policy, that is, at end of the contract period. This amount is called the policy value.

(2) General Insurance

General insurance covers all types of insurance except life insurance. General insurance allows a person to insure properties like buildings, factories, and goodowns containing goods against a possible loss (total or partial) due to fire, flood, earthquake, etc.

Vehicles can be insured to cover the risk of possible damage due to accidents.

In case of loss or damage, the insurance company pays compensation in the same proportion that exists between the policy value and the property value.

All contracts of general insurance are governed by the principle of indemnity, which states that an insured may not be compensated by the insurance company in an amount exceeding the insured's economic loss. As a result, an insured person cannot make profit from an insurance policy.

Let's Learn

2.1 Fire, Marine, and Accident Insurance

(1) Fire Insurance

Fire insurance is property insurance that covers damage and losses caused by fire to property like buildings, godowns containing goods, factories, etc. It is possible to insure the entire property or only its part. The value of the property is called Property Value. The value of the insured part of property is called Policy Value. The amount paid to the insurance company to insure the property is called premium.

Primium = Rate of Premium × Policy Value

The period of a fire insurance policy is one year and the premium is expressed as percentage of the value of the insured property.

In case of damage to the property due to fire, the insurance company agrees to pay compensation in the proportion that exist between policy value and property value. The value of the damage is called "loss" and the amount that the insured can demand under the policy is called claim.

$$\therefore \quad \text{Claim} = \text{Loss} \times \frac{\text{PolicyValue}}{\text{PropertyValue}}$$

(2) Accident Insurance

Personal accident insurance is a policy that can reimburse your medical costs, provide compensation in case of disability or death caused by accidents. Accident insurance allows insuring vehicles like cars, trucks, two wheelers, etc. against to a vehicle due to accidents. This policy also covers the liability of the insured person to third parties involved in the accident. The period of such policies is one year.

(3) Marine Insurance

Marine Insurance covers goods, freight, cargo, etc. against loss or damage during transit by road, rail, sea or air. Shipments are protected from the time they leave the seller's warehouse till the time they reach the buyer's warehouse. Marine insurance offers complete protection during transit goods and compensates in the events of any loss.

The party responsible for insuring the goods is determined by the sales contract. The amount of claim is calculated by the same method that is used in the case of fire insurance.

SOLVED EXAMPLES

Ex. 1: A building worth Rs.50,00,000 is insured for $(\frac{4}{5})^{th}$ of its value at a premium of 5%. Find the amount of premium. Also, find commission of the agent if the rate of commission is 3%.

Solution:

Property value = Rs. 5000000

Policy value $= \frac{4}{5} \times 5000000$

= Rs. 4000000

Rate of premium = 5%

∴ Amount of premium $= 4000000 \times \frac{5}{100}$

∴ Premium amount = Rs. 200000

∴ Commission at 3% $= 200000 \times \frac{3}{100}$

= Rs.6000

∴ Agent's commission = Rs.6000

Ex. 2: A shopkeeper insures his shop valued Rs. 20 lakh for 80% of its value. He pays a premium of Rs.80000. Find the rate of premium. If the agent gets commission at 12%, find the agent's commission.

Solution:

Property value = Rs. 2000000

Insured value = 80% of property value

$= 2000000 \times \frac{80}{100}$

= Rs. 1600000

Now, the premium paid = Rs.80000

$\therefore$ Rate of premium $= \dfrac{100\times 80000}{1600000}$

$\therefore$ Rate of premium = 5%

Commission paid at 12% of premium

$= 80000 \times \dfrac{12}{100}$

= Rs. 9600.

$\therefore$ Agent's commission is Rs.9600.

Ex. 3: A car worth Rs.5,40,000 is insured for Rs.4,50,000. The car is damaged to the extent of Rs.2,40,000 is an accident. Find the amount of compensation that can be claimed under the policy.

Solution:

Value of the car = Rs. 540000

Insured value = Rs. 450000

Damage = Rs. 240000

$$\text{Claim} = \frac{\text{Insured value}}{\text{Property value}} \times \text{loss}$$

$$= \frac{450000}{540000} \times 240000$$

= Rs.200000

$\therefore$ A compensation of Rs. 2,00,000 can be claimed under the policy.

Ex. 4: 10000 copies of a book, priced Rs. 80 each were insured for $\left(\dfrac{3}{5}\right)^{\text{th}}$ of their value. Some copies of the book were damaged in transit, and were therefore reduced to 60% of their value. If the amount recovered against the damage was Rs.24000, find the number of damaged copies of the book.

Solution :

Total number of copies = Rs. 10000

Cost of one book = Rs.80

Insured value $= \dfrac{3}{5} \times$ Property value

Insurance claim = Rs.24000

$$\text{Now, claim} = \frac{\text{Insured value}}{\text{Property value}} \times \text{loss}$$

$\therefore$ 24000 $= \dfrac{3}{5} \times$ loss

$\therefore$ loss $= 24000 \times \dfrac{5}{3}$

= Rs.40000

This amount was equal to 40% of the damage.

$\therefore$ `Total damage $= 40000 \times \dfrac{100}{40}$

= Rs.100000

Since cost of one book was Rs. 80

The number of books damaged $= \dfrac{100000}{80}$

$\therefore$ 1250 books were damaged.

Ex. 5: A cargo valued at Rs.10,00,000 was insured for Rs.7,00,000 during a voyage. If the rate of premium is 0.4%. find (i) the amount of premium, (ii) The amount that can be claimed if the cargo worth Rs.6,00,000 is destroyed, (iii) the amount that can be claimed, if cargo worth Rs.6,00,000 is destroyed completely and the remaining cargo is so damaged that its value is reduced to 40%.

Solution:

Property value = Rs. 10,00,000

Policy value = Rs. 7,00,000

Rate of premium = 0.4%

i) Premium = 0.4% of policy value

$= 700000 \times \dfrac{0.4}{100}$

= Rs. 2800

$\therefore$ Total Premium = Rs. 2800

ii) Claim $= \text{loss} \times \dfrac{\text{PolicyValue}}{\text{PropertyValue}}$

$= 600000 \times \dfrac{700000}{1000000}$

= Rs. 420000

iii) Total value of cargo = Rs. 1000000

Value of the cargo completely destroyed = Rs.600000

$\therefore$ Value of remaining cargo = Rs.400000

Loss on value of remaining cargo = 40% of the value of remaining cargo

$$= \frac{40}{100} \times 400000$$

$$= \text{Rs.}160000$$

$$\therefore \quad \text{Total loss} = 600000 + 160000$$

$$= \text{Rs.}760000$$

$$\therefore \quad \text{Claim} = \text{loss} \times \frac{\text{PolicyValue}}{\text{PropertyValue}}$$

$$= 760000 \times \frac{700000}{1000000}$$

$$= \text{Rs. } 532000$$

Ex. 6: A property worth Rs.4,00,000 is insured with three companies X, Y and Z for amounts Rs.1,20,000, Rs.80,000, and Rs. 1,00,000 respectively. A fire caused a loss of Rs. 2,40,000, Calculate the amounts that can be claimed from the three companies.

Solution:

$$\text{Loss} = \text{Rs.}2,40,000$$

$$\text{Claim} = \text{Loss} \times \frac{\text{PolicyValue}}{\text{PropertyValue}}$$

$$\text{Claim from company X} = 2,40,000 \times \frac{120000}{400000}$$

$$= \text{Rs. } 72,000$$

$$\text{Claim from company Y} = 2,40,000 \times \frac{80000}{400000}$$

$$= \text{Rs. } 48,000$$

$$\text{Claim from company Z} = 2,40,000 \times \frac{100000}{400000}$$

$$= \text{Rs. } 60,000$$

Ex. 7: An agent places insurance for Rs. 4,00,000 on life of a person. The premium is to be paid annually at the rate of Rs.35 per thousand per annum. Find the agent's commission at 15% on the first premium.

Policy value = Rs. 4,00,000

Rate of premium = Rs.35 per thousand p.a.

$$\therefore \text{ Amount of premium} = \frac{35}{1000} \times 4,00,000$$

$$= \text{Rs.}14,000$$

Rate of commission = 15%

$$\therefore \text{ Amount of commission} = 14000 \times \frac{15}{100}$$

$$= \text{Rs.}2100$$

Ex. 8: A person takes a life policy of Rs. 2,00,000 for 15 years. The rate of premium is Rs. 55 per thousand per annum. If the bonus is paid at the average rate of Rs. 6 per thousand, what is the benefit to the insured?

Policy value = Rs. 2,00,000

Rate of premium = Rs.55 per thousand p.a.

$$\therefore \text{ Amount of premium} = \frac{55}{1000} \times 2,00,000$$

$$= \text{Rs.}11,000$$

The insured pays premium for 15 years.

$$\therefore \quad \text{Total premium paid} = 11000 \times 15$$

$$= \text{Rs.}165000$$

Rate of bonus is Rs.6 per thousand per annum on the policy value. Therefore, on the policy Rs. 2,00,000

$$\text{bonus for 1 year} = 6 \times 200$$

$$= \text{Rs.}1200$$

$$\therefore \text{ bonus for 15 year} = 1200 \times 15$$

$$= \text{Rs.}18000$$

Hence, when the policy matures,

$$\text{the insured gets} = 2,00,000 + 18000$$

$$= \text{Rs.}2,18,000$$

$$\therefore \quad \text{Benefit} = 218000 - 165000$$

$$= \text{Rs.}53000$$

EXERCISE 2.1

1. Find the premium on a property worth Rs. 25,00,000 at 3% if (i) the property is fully insured, (ii) the property is insured for 80% of its value.

2. A shop is valued at Rs. 3,60,000 for 75% of its value. If the rate of premium is 0.9%, find the premium paid by the owner of the shop. Also, find the agents commission if the agent gets commission at 15% of the premium.

3. A person insures his office valued at Rs. 5,00,000 for 80% of its value. Find the rate of premium if he pays Rs.13,000 as premium. Also. find agent's commission at 11%

4. A building is insured for 75% of its value. The annual premium at 0.70 per cent amounts to Rs.2625. If the building is damaged to the extent of 60% due to fire, how much can be claimed under the policy?

5. A stock worth Rs.7,00,000 was insured for Rs.4,50,000. Fire burnt stock worth Rs.3,00,000 completely and damaged the remaining stock to the extent of 75 % of its value. What amount can be claimed under the policy?

6. A cargo of rice was insured at 0.625 % to cover 80% of its value. The premium paid was Rs.5250. 1f the price of rice is Rs.21 per kg. find the quantity of rice (in kg) in the cargo.

7. 6000 articles costing Rs.200 per dozen were insured against fire for Rs. 2,40,000. If 20% of the articles were burnt and 7200 of the remaining articles were damaged to the extent of 80% of their value, find the amount that can be claimed under the policy.

8. The rate of premium is 2% and other expenses are 0.075%. A cargo worth Rs.3,50,100 is to be insured so that all its value and the cost of insurance will be recovered in the event of total loss.

9. A property worth Rs. 4.00.000 is insured with three companies. A. B. and C. The amounts insured with these companies are Rs.1,60,000, Rs.1,00,000 and Rs.1,40,000 respectively. Find the amount recoverable from each company in the event of a loss to the extent of Rs. 9.000.

10. A car valued at Rs.8,00,000 is insured for Rs.5,00,000. The rate of premium is 5% less 20%. How much will the owner bear including the premium if value of the car is reduced to 60 % of its original value.

11. A shop and a godown worth Rs.1,00,000 and Rs.2,00,000 respectively were insured through an agent who was paid 12% of the total premium. If the shop was insured for 80% and the godown for 60% of their respective values, find the agent's commission, given that the rate of premium was 0.80% less 20% .

12. The rate of premium on a policy of Rs. 1,00,000 is Rs.56 per thousand per annum. A rebate of Rs.0.75 per thousand is permitted if the premium is paid annually. Find the net amount of premium payable if the policy holder pays the premium annually.

13. A warehouse valued at Rs.40,000 contains goods worth Rs.240,000. The warehouse is insured against fire for Rs.16,000 and the goods to the extent of 90% of their value. Goods worth Rs.80,000 are completely destroyed, while the remaining goods are destroyed to 80% of their value due to a fire. The damage to the warehouse is to the extent of Rs.8,000. Find the total amount that can be claimed.

14. A person takes a life policy for Rs.2,00,000 for a period of 20 years. He pays premium for 10 years during which bonus was declared at an average rate of Rs.20 per year per thousand. Find the paid up value of the policy if he discontinues paying premium after 10 years.

Let's Study

2.2 Annuity

When you deposit some money in a bank, you are entitled to receive more money (in the form of interest) from the bank than you deposit, after a certain period of time. Similarly, when people borrow money for a certain period of time, they pay back more money (again, in the form of interest). These two examples show how money has a time value. A rupee today is worth more than a rupee after one year. The time value of money explains why interest is paid or earned. Interest, whether it is on a bank deposit or a loan, compensates the depositor or lender for the time value of money. When financial transactions occur at different points of time, they must be brought to a common point in time to make them comparable.

Consider the following situation. Ashok deposits Rs.1000 every year in his bank account for 5 years at a compound interest rate of 10 per cent per annum. What amount will Ashok receive at the end of five years? In other words, we wish to know the future value of the money Ashok deposited annually for five years in his bank account.

Assuming that Rs.1000 are deposited at end of every year, the future value is given by

$$1000\left[\left(1+\frac{1}{10}\right)^4+\left(1+\frac{1}{10}\right)^3+\left(1+\frac{1}{10}\right)^2+\left(1+\frac{1}{10}\right)+1\right]$$

$$= 6105$$

In the same situation, the present value of the amount that Ashok deposits in his bank account is given by

$$\frac{1000}{(1.10)^1}+\frac{1000}{(1.10)^2}+\frac{1000}{(1.10)^3}+\frac{1000}{(1.10)^4}+\frac{1000}{(1.10)^5}$$

$$= \text{Rs.}4329.48$$

We also come across a situation where a financial company offers to pay Rs.8,000 after 12 years for Rs.1000 deposited today. In such situations, we wish to know the interest rate offered by the company.

Studies of this nature can be carried out by studying Annuity.

An annuity is a sequence of payments of equal amounts with a fixed frequency. The term "annuity" originally referred to annual payments (hence the name), but it is now also used for payments with any frequency. Annuities appear in many situations: for example, interest payments on an investment can be considered as an annuity. An important application is the schedule of payments to pay off a loan. The word "annuity" refers in everyday language usually to a life annuity. A life annuity pays out an income at regular intervals until you die. Thus, the number of payments that a life annuity makes is not known. An annuity with a fixed number of payments is called an annuity certain, while an annuity whose number of payments depend on some other event (such as a life annuity) is a contingent annuity. Valuing contingent annuities requires the use of probabilities.

Definition

An annuity is a series of payments at fixed intervals, guaranteed for a fixed number of years or the lifetime of one or more individuals. Similar to a pension, the money is paid out of an investment contract under which the annuitant(s) deposit certain sums (in a lump sum or in installments) with an annuity guarantor (usually a government agency or an insurance company). The amount paid back includes principal and interest.

1.2.1 Terminology of Annuity

Four parties to an annuity

Annuitant - A person who receives an annuity is called the annuitant.

Issuer - A company (usually an insurance company) that issues an annuity.

Owner - An individual or an entity that buys an annuity from the issuer of the annuity and makes contributions to the annuity.

Beneficiary - A person who receives a death benefit from an annuity at the death of the annuitant.

Two phases of an annuity

Accumulation phase - The accumulation (or investment) phase is the time period when money is added to the annuity. An annuity can be purchased in one single lump sum (known as a single premium annuity) or by making investments periodically over time.

Distribution phase - The distribution phase is when the annuitant receiving distributions from the annuity. There are two options for receiving distributions from an annuity. The first option is to withdraw some or all of the money in the annuity in lump sums. The second option (commonly known as guaranteed income or annuitization option) provides a guaranteed income for a specific period of time or the entire lifetime of the annuitant.

Types of Annuities - There are three types of annuities.

(i) **Annuity Certain.** An annuity certain is an investment that provides a series of payments for a set period of time to a person or to the person's beneficiary. It is an investment in retirement income offered by insurance companies. The annuity may also be taken as a lump sum.

Because it has a set expiration date, a annuity certain generally pays a higher rate of return than lifetime annuity. Typical terms are 10, 15, or 20 years.

Contingent Annuity. Contingent annuity is a form of annuity contract that provides payments at the time when the named contingency occurs. For instance, upon death of one spouse, the surviving spouse will begin to receive monthly payments. In a contingent annuity policy the payment will not be made to the annuitant or the beneficiary until a certain stated event occurs.

Perpetual Annuity or Perpetuity. A perpetual annuity, also called a perpetuity promises to pay a certain amount of money to its owner forever.

Though a perpetuity may promise to pay you forever, its value isn't infinite. The bulk of the value of a perpetuity comes from the payments that you receive in the near future, rather than those you might receive 100 or even 200 years from now.

Classification of Annuities - Annuities are classified in three categories according to the commencement of income. These three categories are: Immediate Annuity, Annuity Due, and Deferred Annuity.

Immediate Annuity or Ordinary Annuity - The immediate annuity commences immediately after the end of the first income period. For instance, if the annuity is to be paid annually, then the first installment will be paid at the expiry of one year. Similarly in a half- yearly annuity, the payment will begin at the end of six months. The annuity can be paid either yearly, half-yearly, quarterly or monthly.

The purchase money (or consideration) is in a single amount. Evidence of age is always asked for at the time of entry.

Annuity Due - Under this annuity, the payment of installment starts from the time of contract. The first payment is made as soon as the contract is finalized. The premium is generally paid in single amount but can be paid in installments as is discussed in the deferred annuity. The difference between the annuity due and immediate annuity is that the payment for each period is paid in its beginning under the annuity due contract while at the end of the period in the immediate annuity contract.

Deferred Annuity - In this annuity contract the payment of annuity starts after a deferment period or at the attainment by the annuitant of a specified age. The premium may be paid as a single premium or in instalments.

The premium is paid until the date of commencement of the instalments.

We shall study only immediate annuity and annuity due.

Present value of an annuity - The present value of an annuity is the current value of future payments from an annuity, given a specified rate of return or discount rate. The annuity's future cash flows are discounted at the discount rate. Thus the higher the discount rate, the lower the present value of the annuity.

Future value of an annuity - The future value of an annuity represents the amount of money that will be accumulated by making consistent investments over a set period, assuming compound interest. Rather than planning for a guaranteed amount of income in the future by calculating how much must be invested now, this formula estimates the growth of savings given a fixed rate of investment for a given amount of time.

The present value of an annuity is the sum that must be invested now to guarantee a desired payment in the future, while the future value of an annuity is the amount to which current investments will grow over time.

Note:

(1) We consider only uniform and certain annuities.

(2) If the type of an annuity is not mentioned, we assume that the annuity is immediate annuity.

(3) If there is no mention of the type of interest, then it is assumed that the interest is compounded per annum.

If payments are made half-yearly (that is, twice per year), then r is replaced by $\frac{r}{2}$ (the compounding rate) and n is replaced by $2n$ (the number of time periods).

If payments are made quarterly (that is, four times per year), then r is replaced by $\frac{r}{4}$ (the compounding rate) and n is replaced by $4n$ (the number of time periods).

If payments are made monthly (that is, 12 times per year), then r is replaced by $\frac{r}{12}$ (the compounding rate) and n is replaced by $12n$ (the number of time periods).

Immediate Annuity - Payments are made at the end of every time period in immediate annuity.

Basic formula for an immediate annuity - The accumulated value A of an immediate annuity for n annual payments of an amount C at an interest rate r per cent per annum, compounded annually, is given by.

$$A = \frac{C}{i}\left[\left(1+i\right)^{n} - 1\right] \text{, where } i = \frac{r}{100}$$

Also, the present value P of such an immediate annuity is given by

$$P = \frac{C}{i}\ [1 - (1 + i)^{-n}]$$

The present value and the future value of an annuity have the following relations.

$$A = P(1 + i)^{n},$$

$$\frac{1}{p} - \frac{1}{A} = \frac{i}{C}.$$

SOLVED EXAMPLES

Ex. 1: Find the accumulated value after 3 years of an immediate annuity of Rs. 5000 p.a. with interest compounded at 4% p.a. [given $(1.04)^{3} = 1.12490$]

Solution:

The problem states that C = 5000, r = 4% p.a., and n = 3 years.

Then, accumulated value is given by

$$A \quad = \frac{C}{i}\left[\left(1+i\right)^{n} - 1\right] \quad \left\{i = \frac{r}{100} = 0.04\right\}$$

$$= \frac{5000}{0.04}\,[(1{+}0.04)^3 - 1]$$

$$= \frac{5000}{0.04}\,[(1.04)^3 - 1]$$

$$= \frac{5000}{0.04}\,[1.1249 - 1]$$

$$\therefore \quad A = \frac{5000}{0.04}\,[1.1249 - 1]$$

$$\therefore \quad A = \frac{5000}{0.04}\,(0.1249)$$

$$\therefore \quad A = 5000(3.12250)$$

$$\therefore \quad A = 15{,}612.50$$

The accumulated value of the annuity is Rs. 15,612.50

Ex. 2: A person plans to accumulate a sum of Rs. 5,00,000 in 5 years for higher education of his son. How much should he save every year if he gets interest compounded at 10% p.a.? [Given : $(1.10)^5 = 1.61051$]

Solution: From the problem, we have

A = Rs.5,00,000, r = 10%p.a., and n = 5 years.

$$\therefore \quad i = \frac{r}{100} = \frac{10}{100} = 0.10$$

We find C as follows,

$$A = \frac{C}{i}\,[(1+i)^n - 1]$$

$$\therefore \quad 5{,}00{,}000 = \frac{C}{0.1}\,[(1+0.1)^5 - 1]$$

$$\therefore \quad 5{,}00{,}000 = \frac{C}{0.1}\,[(1.1)^5 - 1]$$

$$\therefore \quad 5{,}00.000 = \frac{C}{0.1}\,[1.61051 - 1]$$

$$\therefore \quad 5{,}00.000 = \frac{C}{0.1}\,(0.61051)$$

$$\therefore \quad C = \frac{500000 \times 0.1}{0.61051}$$

$$\therefore \quad C = \frac{500000}{0.61051} = 81898.74$$

That is, the person should save Rs.81898.74 every year for 5 years to get Rs.5,00,000 at the end of 5 years.

Ex. 3: Mr. X saved Rs.5000 every year for some years. At the end of this period, he received an accumulated amount of Rs.23205. Find the number of years if the interest was compounded at 10% p.a. [Given : $(1.1)^4 = 1.4641$]

Solution: From the problem, we have A = Rs.23205, C = Rs.5000 and r = 10% p.a.

$$\therefore \quad i = \frac{r}{100} = \frac{10}{100} = 0.10$$

The value of n is found as follows.

$$A = \frac{C}{i}\,[(1+i)^n - 1]$$

$$\therefore \quad 23205 = \frac{5000}{0.1}\,[(1+0.1)^n - 1]$$

$$\therefore \quad \frac{23205 \times 0.1}{5000} = (1.1)^n - 1$$

$$\therefore \quad \frac{2320.5}{5000} = (1.1)^n - 1$$

$$\therefore \quad 0.4641 + 1 = (1.1)^n$$

$$\therefore \quad 1.4641 = (1.1)^n$$

$$\text{since } (1.1)^4 = 1.4641$$

$$\therefore \quad n = 4 \text{ years}$$

Ex. 4: Find the rate of interest compounded annually if an immediate annuity of Rs.20,000 per year amounts to Rs.41,000 in 2 years.

Solution: From the problem, we have A = Rs.41,000, C = Rs.20,000 and n = 2 years. The value of n is then found as follows.

$$A = \frac{C}{i}\,[(1+i)^n - 1]$$

$$\therefore \quad 41000 = \frac{20000}{i}\,[(1+i)^2 - 1]$$

$$\therefore \quad \frac{41000}{20000} = \frac{1 + 2i + i^2 - 1}{i}$$

$$\therefore \quad 2.05 = \frac{2i+i^2}{i}$$

$$\therefore \quad 2.05 = 2 + i$$

$$\therefore \quad i = 2.05 - 2$$

$$\therefore \quad i = 0.05 \text{ but } i = \frac{r}{100}$$

$$\therefore \quad r = i \times 100$$

$$\therefore \quad r = 5$$

∴ the rate of interest is 5% p.a.

Ex. 5: A person deposited Rs.15,000 every six months for 2 years. The rate of interest is 10% p.a. compounded half yearly. Find the amount accumulated at the end of 2 years.

[Given : $(1.05)^4 = 1.2155$]

Solution:

Since an amount of Rs.15,000 is deposited every six months. It is a case of immediate annuity. The problem states that C = Rs. 15000. Rate of interest is 10% p.a. Therefore, it is 5% for six months. That is, $r = 5\%$.

$$\therefore \quad i = \frac{r}{100} = \frac{5}{100} = 0.05$$

The number of half years in 2 years is 4. and therefore $n = 4$.

Now we use the formula for accumulated value A and get

$$A = \frac{C}{i}\left[(1+i)^n - 1\right]$$

$$\therefore \quad A = \frac{15000}{0.05}\left[(1+0.05)^4 - 1\right]$$

$$= 3{,}00{,}000\,[1.2155 - 1]$$

$$= 3{,}00{,}000\,[0.2155]$$

$$= 64650$$

∴ The accumulated amount after 2 years is Rs.64650.

Ex. 6: A person deposits Rs.3000 in a bank every quarter. The interest is 8% compounded every quarter. Find the accumulated amount at the end of 1 year. [*Given* : $(1.02)^4 = 1.0824$]

Solution:

Since the amount is deposited every quarter, it is an immediate annuity. From the given problem. C = 3000. The rate of interest is 8% p.a. and hence it is 2% per quarter. That is, $r = 0.02$ and hence $i = 0.02$. The number of quarters in a year is 4. That is, $n = 4$. We use the formula for accumulated amount to obtain

$$A = \frac{C}{i}\left[(1+i)^n - 1\right]$$

$$= \frac{3000}{0.02}\left[(1+0.02)^4 - 1\right]$$

$$= 1{,}50{,}000\,[(1.02)^4 - 1]$$

$$= 1{,}50{,}000\,[1.0824 - 1]$$

$$= 1{,}50{,}000\,[0.0824]$$

$$= 12360.$$

∴ Accumulated amount after 1 year is Rs.12,360

Ex. 7: Find the present value of an immediate annuity of Rs.50,000 per annum for 4 years with interest compounded at 10% p.a.

[*Given* : $(1.1)^{-4} = 0.6830$]

Solution:

From the problem, we get $C = 50{,}000$, $n = 4$ years, $r = 10\%$ p.a., so that $i = 0.1$. We use the formula for present value and get

$$P = \frac{C}{i}\left[1 - (1+i)^{-n}\right]$$

$$\therefore \quad P = \frac{50000}{0.1}\left[1 - (1+0.1)^{-4}\right]$$

$$\therefore \quad P = \frac{50000}{0.1}\left[1 - (1.10)^{-4}\right]$$

$$= \frac{50000}{0.1}\left[1 - 0.68300\right]$$

$$= 500000[0.3170]$$

$$= 1{,}58{,}493.$$

∴ Present value of the given annuity is Rs.1,58,493.

Ex. 8: The present value of an immideiate annuity paid for 4 years with interest accumulated at 10% p.a. is Rs.20,000. What is its accumulated value after 4 years? [*Given* : $(1.1)^4 = 1.4641$]

Solution:

It is given that $P = 20{,}000$. $n = 4$ years, $r = 10\%$ p.a., so that $i = 0.1$. Using the formula for the relation between A and P. we obtain

$$\begin{aligned} A &= P(1+i)^n \\ \therefore\ A &= 20{,}000(1+0.1)^4 \\ &= 20{,}000(1.1)^4 \\ &= 20{,}000(1.4641) \\ &= 29282. \end{aligned}$$

$\therefore$ Accumulated value after 4 years is Rs.29,282

Ex. 9: The present and accumulated values of an immediate annuity paid for some years at interest compounded at 10% p.a. are Rs.4,000 and Rs.8,000 respectively. Find the amount of every annuity paid.

Solution:

It is given that $P =$ Rs.4,000, $A =$ Rs.8,000, $r = 10\%$ p.a., so that $i = 0.1$.

We use the following formula for finding C.

$$\frac{1}{p} - \frac{1}{A} = \frac{i}{C}$$

$$\begin{aligned} \therefore\ \frac{1}{4000} - \frac{1}{8000} &= \frac{0.1}{C} \\ \therefore\ 0.00025 - 0.000125 &= \frac{0.1}{C} \\ \therefore\ 0.000125 &= \frac{0.1}{C} \\ \therefore\ C &= \frac{0.1}{0.000125} \\ &= \frac{100000}{125} \\ &= 800. \end{aligned}$$

$\therefore$ Every annuity paid is Rs. 800.

Annuity Due

Payments are made at the beginning of every time period in annuity due.

Basic formula for an annuity due

Let C denote the amount paid at the beginning of each of n years and let r denote the rate of interest per cent per annum.

Let $i = \dfrac{r}{100}$

The accumlated value A' is given by

$$A' = \frac{C(1+i)}{i}\left[(1+i)^n - 1\right]$$

The present value P' is given by

$$P' = \frac{C(1+i)}{i}\left[1 - (1+i)^{-n}\right]$$

A' and P' have the following relations.

$$A' = P'(1+i)^n$$

$$\frac{1}{P'} - \frac{1}{A'} = \frac{i}{C(1+i)}$$

Note:

We can use the formula of immediate annuity for anuity due only by replacing C by $C(1+i)$

SOLVED EXAMPLES

Ex. 1: Find an accumulated value of an annuity due of Rs 2,000 per annum for 4 years at 10% p.a. [*Given* $(1.1)^4 = 1.4641$]

Solution: Given $C =$ Rs 2,000, $n = 4$ years,

$r = 10\%$ p.a. so that $i = 0.1$

We use the formula for accumulated value A' of an annuity due to get

$$\begin{aligned} A' &= \frac{C(1+i)}{i}\left[(1+i)^n - 1\right] \\ \therefore\ A' &= \frac{2000(1+0.1)}{0.1}\left[(1+0.1)^4 - 1\right] \\ \therefore\ A' &= \frac{2000(1.1)}{0.1}\left[1.4641 - 1\right] \end{aligned}$$

$$= 22000\,(0.4641)$$

$$= 10210.2$$

$\therefore$ Accumulated value is Rs 10,210.2

Ex.2: Find the present value of an annuity due of Rs.5,000 to be paid per quarter at 16% p.a. compounded quarterly for 1 year. [Given: $(1.04)^{-4} = 0.8548$]

Solution:

Given C = Rs. 5000. Rate of interest is 16% p.a. and hence it is 4% per quarter. This gives $i = 0.04$. Finally, $n = 4$.

We use the formula for present value of annuity due to obtain

$$P' = \frac{C(1+i)}{i}\,[1-(1+i)^{-n}]$$

$$\therefore\quad P' = \frac{5000(1+0.04)}{0.04}\,[1-(1+0.04)^{-4}]$$

$$= \frac{5000(1.04)}{0.04}\,[1-(0.8548)]$$

$$= 1{,}25{,}000\,(1.04)(0.1452)$$

$$= 18876$$

$\therefore$ Present Value is Rs 18,876

Sinking Fund:

A sinking fund is a fund established by financial organization by setting aside revenue over a period of time compounded annually, to fund a future capital expense, or repayment of a long-term debt.

SOLVED EXAMPLES

Ex. The cost of a machine is Rs.10 lakh and its effective life is 12 years. The scrap realizes only Rs.50,000. What amount should be kept aside at the end of every year to accumulate Rs.9,50,000 at compound interest 5% p.a.? [Given $(1.05)^{12} = 1.796$]

Solution: Here A = Rs. 10 lakh, $r = 5\%$,

$$i = \frac{r}{100} = 0.05, \text{ and } n = 12.$$

$$A = 10{,}00{,}000 - 50{,}000$$

$$= 9{,}50{,}000$$

$$\text{Now } A = \frac{C}{i}\,[(1+i)^n - 1]$$

$$\therefore\quad 9{,}50{,}000 = \frac{C}{i}\,[(1+i)^n - 1]$$

$$\therefore\quad C = \frac{950000 \times 0.05}{0.796}$$

$$= 59673.40$$

$\therefore$ An amount of Rs.59673.40 should be kept aside at the end of every year for 12 years to get Rs.10 lakh at the end of 12 years.

EXERCISE 2.2

1. Find the accumulated (future) value of annuity of Rs.800 for 3 years at interest rate 8% compounded annually.
   [*Given* $(1.08)^3 = 1.2597$]
2. A person invested Rs.5,000 every year in finance company that offered him interest compounded at 10% p.a., what is the amount accumulated after 4 years?
   [*Given* $(1.1)^4 = 1.4641$]
3. Find the amount accumulated after 2 years if a sum of Rs.24,000 is invested every six months at 12% p.a. compounded half yearly. [*Given* $(1.06)^4 = 1.2625$]
4. Find accumulated value after 1 year of an annuity immediate in which Rs.10,000 are invested every quarter at 16% p.a. compounded quarterly.
   [*Given* $(1.04)^4 = 1.1699$]
5. Find the present value of an annuity immediate of Rs.36,000 p.a. for 3 years at 9% p.a. compounded annually.
   [*Given* $(1.09)^{-3} = 0.7722$]
6. Find the present value of an ordinary annuity of Rs.63,000 p.a. for 4 years at 14% p.a. compounded annually.
   [*Given* $(1.14)^{-4} = 0.5921$]

7. A lady plans to save for her daughter's marriage. She wishes to accumulate a sum of Rs.4,64,100 at the end of 4 years. What amount should she invest every year if she gets an interest of 10% p.a. compounded annually? [*Given* $(1.1)^4 = 1.4641$]

8. A person wants to create a fund of Rs.6,96,150 after 4 years at the time of his retirement. He decides to invest a fixed amount at the end of every year in a bank that offers him interest of 10% p.a. compounded annually. What amount should he invest every year? [*Given* $(1.1)^4 = 1.4641$]

9. Find the rate of interest compounded annually if an annuity immediate at Rs.20,000 per year amounts to Rs.2,60,000 in 3 years.

10. Find the number of years for which an annuity of Rs.500 is paid at the end of every year, if the accumulated amount works out to be Rs. 1,655 when interest is compounded annually at 10% p.a.

11. Find the accumulated value of annuity due of Rs.1000 p.a. for 3 years at 10% p.a. compounded annually. [*Given* $(1.1)^3 = 1.331$]

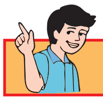
12. A person plans to put Rs.400 at the beginning of each year for 2 years in a deposit that gives interest at 2% p.a. compounded annually. Find the amount that will be accumulated at the end of 2 years.

13. Find the present value of an annuity due of Rs.600 to be paid quarterly at 32% p.a. compounded quarterly. [*Given* $(1.08)^{-4} = 0.7350$]

14. An annuity immediate is to be paid for some years at 12% p.a. The present value of the annuity is Rs.10,000 and the accumulated value is Rs.20,000. Find the amount of each annuity payment.

15. For an annuity immediate paid for 3 years with interest compounded at 10% p.a., the present value is Rs.24,000. What will be the accumulated value after 3 years?
[*Given* $(1.1)^3 = 1.331$]

16. A person sets up a sinking fund in order to have Rs.1,00,000 after 10 years. What amount should be deposited bi-annually in the account that pays him 5% p.a. compounded semi-annually? [*Given* $(1.025)^{20} = 1.675$]



Let's Remember

- Premium is paid on insured value.
- Agent's commission is paid on premium.
- $\text{Claim} = \dfrac{\text{PolicyValue}}{\text{PropertyValue}} \times \text{loss}$
- **Immediate Annuity**

Amount of accumulated (future) Value = A

$$A = \frac{C}{i}\left[(1+i)^n - 1\right]$$

$$\text{Present value } P = \frac{C}{i}\left[1-(1+i)^{-n}\right]$$

$$A = P(1+i)^n$$

$$\frac{1}{P} - \frac{1}{A} = \frac{i}{C}$$

- **Annuity Due**

$$\text{Accumulated value } A' = \frac{C(1+i)}{i}\left[(1+i)^n - 1\right]$$

$$\text{Present value } P' = \frac{C(1+i)}{i}\left[1-(1+i)^{-n}\right]$$

$$A' = P'(1+i)^n$$

$$\frac{1}{P'} - \frac{1}{A'} = \frac{i}{C(1+i)}$$

MISCELLANEOUS EXERCISE - 2

I) Choose the correct alternative.

1. “A contract that pledges payment of an agreed upon amount to the person (or his/her nominee) on the happening of an event covered against” is technically known as

   a. Death coverage b. Savings for future

   c. Life insurance d. Provident fund

2. Insurance companies collect a fixed amount from their customers at a fixed interval of time. This amount is called

   a. EMI b. Installment

   c. Contribution d. Premium

3. Following are different types of insurance.

   I. Life insurance

   II. Health insurance

   III. Liability insurance

   (a) Only I (b) Only II

   (c) Only III (d) All the three

4. By taking insurance, an individual

   a. Reduces the risk of an accident

   b. Reduces the cost of an accident

   c. Transfers the risk to someone else.

   d. Converts the possibility of large loss to certainty of a small one.

5. You get payments of Rs.8,000 at the beginning of each year for five years at 6%, what is the value of this annuity?

   a. Rs 34,720 b. Rs 39,320

   c. Rs 35,720 d. Rs. 40,000

6. In an ordinary annuity, payments or receipts occur at

a. Beginning of each period

b. End of each period

c. Mid of each period

d. Quarterly basis

7. Amount of money today which is equal to series of payments in future is called

   a. Normal value of annuity

   b. Sinking value of annuity

   c. Present value of annuity

   d. Future value of annuity

8. Rental payment for an apartment is an example of

   a. Annuity due b. Perpetuity

   c. Ordinary annuity d. Installment

9. ______ is a series of constant cashflows over a limited period of time.

   a. Perpetuity b. Annuity

   c. Present value d. Future value

10. A retirement annuity is particularly attractive to someone who has

    a. A severe illness

    b. Risk of low longevity

    c. Large family

    d. Chance of high longevity

II) Fill in the blanks

1. An installment of money paid for insurance is called __________.
2. General insurance covers all risks except __________.
3. The value of insured property is called __________.
4. The proportion of property value to insured value is called __________.
5. The person who receives annuity is called __________.
6. The payment of each single annuity is called __________.
7. The intervening time between payment of two successive installments is called as __________.
8. An annuity where payments continue forever is called __________.

9. If payments of an annuity fall due at the beginning of every period, the series is called annuity __________.

10. If payments of an annuity fall due at the end of every period, the series is called annuity __________.

III) State whether each of the following is True or False.

1. General insurance covers life, fire, and theft.
2. The amount of claim cannot exceed the amount of loss.
3. Accident insurance has a period of five years.
4 Premium is the amount paid to the insurance company every month.
5. Payment of every annuity is called an installment.
6. Annuity certain begins on a fixed date and ends when an event happens.
7. Annuity contingent begins and ends on certain fixed dates.
8. The present value of an annuity is the sum of the present value of all installments.
9. The future value of an annuity is the accumulated values of all installments.
10. Sinking fund is set aside at the beginning of a business.

IV) Solve the following problems.

1. A house valued at Rs. 8.00.000 is insured at 75% of its value. If the rate of premium is 0.80 %. Find the premium paid by the owner of the house. If agent's commission is 9% of the premium, find agent's commission.

2. A shopkeeper insures his shop and godown valued at Rs.5,00,000 and Rs.10,00.000 respectively for 80 % of their values. If the rate of premium is 8 %, find the total annual premium.

3 A factory building is insured for $\left(\frac{5}{6}\right)^{th}$ of its value at a rate of premium of 2.50%. If the agent is paid a commission of Rs.2812.50, which is 7.5% of the premium, find the value of the building.

4 A merchant takes fire insurance policy to cover 80 % of the value of his stock. Stock worth Rs.80,000 was completely destroyed in a fire. while the rest of stock was reduced to 20% of its value. If the proportional compensation under the policy was Rs.67,200, find the value of the stock.

5. A 35-year old person takes a policy for Rs.1,00,000 for a period of 20 years. The rate of premium is Rs.76 and the average rate of bonus is Rs.7 per thousand p.a. If he dies after paying 10 annual premiums, what amount will his nominee receive?

6. 15,000 articles costing Rs.200 per dozen were insured against fire for Rs.1,00,000. If 20 % of the articles were burnt completely and 2400 of other articles were damaged to the extent of 80% of their value, find the amount that can be claimed under the policy.

7. For what amount should a cargo worth Rs.25,350 be insured so that in the event of total loss, its value as well as the cost of insurance may be recovered when the rate of premium is 2.5 %.

8. A cargo of grain is insured at $\left(\frac{3}{4}\right)$ % to cover 70% of its value. Rs.1,008 is the amount of premium paid. If the grain is worth Rs. 12 per kg, how many kg of the grain did the cargo contain?

9. 4000 bedsheets worth Rs.6,40,000 were insured for $\left(\frac{3}{7}\right)^{th}$ of their value. Some of the bedsheets were damaged in the rainy season and were reduced to 40% of their value. If the amount recovered against damage was Rs.32,000. find the number of damaged bedsheets.

10. A property valued at Rs.7,00,000 is insured to the extent of Rs.5,60,000 at $\left(\frac{5}{8}\right)$% less 20%. Calculate the saving made in the premium. Find the amount of loss that the owner must bear, including premium, if the property is damaged to the extent of 40 % of its value.

11. Stocks in a shop and godown worth Rs.75,000 and Rs.1,30,000 respectively were insured through an agent who receive 15% of premium as commission. If the shop was insured for 80% and godown for 60% of the value, find the amount of agent's commission when the premium was 0.80% less 20%. If the entire stock in the shop and 20% stock in the godown is destroyed by fire, find the amount that can be claimed under the policy.

12. A person holding a life policy of Rs.1,20,000 for a term of 25 years wants to discontinue after paying premium for 8 years at the rate of Rs.58 per thousand p. a. Find the amount of paid up value he will receive on the policy. Find the amount he will receive if the surrender value granted is 35% of the premiums paid, excluding the first year's premium.

13. A godown valued at Rs.80,000 contained stock worth Rs. 4,80,000. Both were insured against fire. Godown for Rs.50,000 and stock for 80% of its value. A part of stock worth Rs.60,000 was completely destroyed and the rest was reduced to 60% of its value. The amount of damage to the godown is Rs. 40,000. Find the amount that can be claimed under the policy.

14. Find the amount of an ordinary annuity if a payment of Rs. 500 is made at the end of every quarter for 5 years at the rate of 12% per annum compounded quarterly.

15. Find the amount a company should set aside at the end of every year if it wants to buy a machine expected to cost Rs.1,00,000 at the end of 4 years and interest rate is 5% p. a. compounded annually.

16. Find the least number of years for which an annuity of Rs. 3,000 per annum must run in order that its amount exceeds Rs. 60,000 at 10% compounded annually. [$(1.1)^{11} = 2.8531$, $(1.1)^{12} = 3.1384$]

17. Find the rate of interest compounded annually if an ordinary annuity of Rs. 20,000 per year amounts to Rs. 41,000 in 2 years.

18. A person purchases a television by paying Rs.20,000 in cash and promising to pay Rs. 1000 at end of every month for the next 2 years. If money is worth 12% p. a., converted monthly. find the cash price of the television.

19. Find the present value of an annuity immediate of Rs. 20,000 per annum for 3 years at 10% p.a. compounded annually.

20. A man borrowed some money and paid back in 3 equal installments of Rs.2160 each. What amount did he borrow if the rate of interest was 20% per annum compounded annually? Also find the total interest charged.

21. A company decides to set aside a certain amount at the end of every year to create a sinking fund that should amount to Rs. 9,28,200 in 4 years at 10% p.a. Find the amount to be set aside every year.

22. Find the future value after 2 years if an amount of Rs. 12,000 is invested at the end of every half year at 12% p. a. compounded half yearly.

23. After how many years would an annuity due of Rs. 3.000 p.a. accumulated Rs.19.324.80 at 20% p. a. compounded half yearly? [Given $(1.2)^4 = 2.0736$]

24. Some machinery is expected to cost 25% more over its present cost of Rs. 6,96,000 after 20 years. The scrap value of the machinery will realize Rs.1,50,000. What amount should be set aside at the end of every year at 5% p.a. compound interest for 20 years to replace the machinery?

    [*Given* $(1.05)^{20} = 2.653$]

Activities

1) Property Value = Rs.1,00,000

Policy value = 10% of property value

$= \boxed{}$

Rate of premium = 0.4%

Amount of premium $= \dfrac{0.4}{100} \times \boxed{}$

= Rs. 280

Property worth Rs. 60,000 is destroyed

$\therefore$ Claim = loss $\times \dfrac{\text{PolicyValue}}{\text{PropertyValue}}$

$= \boxed{} \times \dfrac{7}{\boxed{}}$

= Rs. 42,000

Now, the property worth Rs.60,000 is totally destroyed and in addition the remaining property is so damaged as to reduce its value by 40%

$\therefore$ Loss $= 60{,}000 + \dfrac{40}{100} \times \boxed{}$

$= 60{,}000 + 16{,}000$

$=$ Rs. $\boxed{}$

$\therefore$ Claim $= \boxed{} \times \dfrac{7}{10}$

= Rs. 53,200

2. Policy value = Rs. 70,000

Period of policy = 15 years

Rate of premium = Rs. 56.50 per thousand p.a.

$\therefore$ Amount of premium $= \dfrac{56.50}{1000} \times \boxed{}$

= Rs.3955

$\therefore$ Total premium paid $= 3955 \times \boxed{}$

= Rs.59,325

Rate of bonus = Rs. 6 per thousand p.a.

$\therefore$ Amount of bonus $= 6 \times \boxed{}$

= Rs.420

$\therefore$ Bonus for 15 years $= 420 \times \boxed{}$

= Rs.6,300

$\therefore$ The person gets Rs. $= \boxed{} + 6300$

= Rs.76,300

$\therefore$ Benefit $= \boxed{} - 59{,}325$

= Rs.16,975.

3. For an immediate annuity,

$P =$ Rs.2000, $A =$ Rs.4000

$r = 10\%$ p.a.

$\therefore \quad i = \dfrac{r}{100} = \dfrac{\boxed{}}{100} = 0.1$

$\therefore \quad \frac{1}{P} - \frac{1}{A} = \frac{i}{C}$

$\therefore \quad \frac{1}{\boxed{\quad}} - \frac{1}{\boxed{\quad}} = \frac{0.1}{C}$

$\therefore \quad \frac{\boxed{\quad}}{4000} = \frac{0.1}{C}$

$\therefore \quad C = \text{Rs.} \boxed{\quad}$

4. For an annuity due,

C = Rs.2000, rate = 16% p.a. compounded quarterly for 1 year

∴ Rate of interest per quarter = $\frac{\boxed{\quad}}{4}$

$= 4$

$\therefore \quad r = 4\%$

$\therefore \quad i = \frac{r}{100} = \frac{\boxed{\quad}}{100} = 0.04$

n = Number of quarters

$= 4 \times 1$

$= \boxed{\quad}$

$\therefore \quad P' = \frac{C(1+i)}{i}\,[1 - (1+i)^{-n}]$

$\therefore P' = \frac{\boxed{\quad}\left(1+\boxed{\quad}\right)}{0.04}\left[1-\left(\boxed{\quad}+0.04\right)^{-\boxed{\quad}}\right]$

$= \frac{2000\left(\boxed{\quad}\right)}{\boxed{\quad}}\left[1-\left(\boxed{\quad}\right)^{-4}\right]$

$= 50{,}000\ (\boxed{\quad})\ (1 - 0.8548)$

$= 50{,}000\ (1.04)\ (\boxed{\quad})$

$= \text{Rs.}7{,}550.40$

5. The cost of machinery = Rs.10,00,000

Effective life of machinery = 12 years

Scrap value of machinery = Rs. 50,000

$r = 5\%$ p.a.

$\therefore \quad i = \frac{r}{100} = \frac{\boxed{\quad}}{\boxed{\quad}} = 0.05$

$A = 10{,}00{,}000 - 50{,}000$

$= \boxed{\quad}$

For an immediate annuity

$$A = \frac{C}{i}\,[(1+i)^n - 1]$$

$\therefore \quad \boxed{\quad} = \frac{C}{\boxed{\quad}}\left[(1+0.05)^{\boxed{\quad}} - 1\right]$

$\therefore \quad 9{,}50{,}000 = \frac{C}{0.05}\left[1.797 - \boxed{\quad}\right]$

$\therefore \quad C = \frac{950000 \times \boxed{\quad}}{0.797}$

$= \text{Rs.}59{,}598.40$

◆ ◆ ◆

3 Linear Regression

Let's Study

- Meaning and Types of Regression
- Fitting Simple Linear Regression
    - Least Square Method
    - Regression of Y on X
    - Regression of X on Y
- Properties of Regression Coefficients

Let's Recall

- Concept of Correlation
- Coefficient of Correlation
- Interpretation of Correlation

Introduction

We have already learns that correlation is used to measure the strength and direction of association between two variables. In statistics, correlation denotes association between two quantitative variables. It is assumed that this association is linear. That is, one variable increase or decreases by a fixed amount for every unit of increase or decrease in the other variable. Consider the relationship between the two variables in each of the following examples.

1. Advertising and sales of a product. (Positive correlation)
2. Height and weight of a primary school student. (Positive correlation)
3. The amount of fertilizer and the amount of crop yield. (Positive correlation)
4. Duration of exercise and weight loss. (Positive correlation)
5. Demand and price of a commodity. (Positive correlation)
6. Income and consumption. (Positive correlation)
7. Supply and price of a commodity. (Negative correlation)
8. Number of days of absence (in school) and performance in examination. (Negative correlation)
9. The more vitamins one consumes, the less likely one is to have a deficiency.(Negative correlation)

Correlation coefficient measures association between two variables but cannot determine the value of one variable when the value of the other variable is known or given. The technique used for predicting the value of one variable for a given value of the other variable is called regression. Regression is a statistical tool for investigating the relationship between variables. It is frequently used to predict the value and to identify factors that cause an outcome. Karl Pearson defined the coefficient of correlation known as *Pearson's Product Moment correlation coefficient.* Carl Friedrich Gauss developed the method known as the *Least Squares Method* for finding the linear equation that best describes the relationship between two or more variables. R.A. Fisher combined the work of Gauss and Pearson to develop the complete theory of least squares estimation in linear regression. Due to Fisher's work, linear regression is used for prediction and understanding correlations.

Note: Some statistical methods attempt to determine the value of an unknown quantity, which may be a parameter or a random variable. The method used for this purpose is called estimation if the unknown quantity is a parameter, and prediction if the unknown quantity is a variable.

Let's :Learn

3.1 Meaning and Types of Regression

Meaning of Regression

Linear regression is a method of predicting the value of one variable when the values of all other variables are known or specified. The variable being predicted is called the response or dependent variable. The variables used for predicting the response or dependent variable are called predictors or independent variables. Linear regression proposes that the relationship between two or more variables is described by a linear equation. The linear equation used for this purpose is called a linear regression model. A linear regression model consists of a linear equation with unknown coefficients. The unknown coefficients in the linear regression model are called parameters of the linear regression model. Observed values of the variables are used to estimate the unknown parameters of the model. The process of developing a linear equation to represent the relationship between two or more variables using the available sample data is known as fitting the linear regression model to observed data. Correlation analysis is used for measuring the strength or degree of the relationship between the predictors or independent variables and the response or dependent variable. The sign of correlation coefficient indicates the direction (positive or negative) of the relationship between the variables, while the absolute value(that is, magnitude) of correlation coefficient is used as a measure of the strength of the relationship. Correlation analysis, however,does not go beyond measuring the direction and strength of the relationship between predictor or independent variables and the response or dependent variable. The linear regression model goes beyond correlation analysis and develops a formula for predicting the value of the response or dependent variable when the values of the predictor or independent variables are known. Correlation analysis is therefore a part of regression analysis and is performed before performing regression analysis. The purpose of correlation analysis is to find whether there is a strong correlation between two variables. Linear regression will be useful for prediction only if there is strong correlation between the two variables.

Types of Linear Regression

The primary objective of a linear regression is to develop a linear equation to express or represent the relationship between two or more variables. Regression equation is the mathematical equation that provides prediction of values of the dependent variable based on the known or given values of the independent variables.

When the linear regression model represents the relationship between the dependent variable (Y) and only one independent variable (X), then the corresponding regression model is called a simple linear regression model. When the linear regression model represents the relationship between the dependent variable and two or more independent variables, then the corresponding regression model is called a multiple linear regression model.

Following examples illustrate situations for simple linear regression.

1. A firm may be interested in knowing the relationship between advertising (X) and sales of its product (Y), so that it can predict the amount of sales for the allocated advertising budget.
2. A botanist wants to find the relationship between the ages (X) and heights (Y) of seedling in his experiment.
3. A physician wants to find the relationship between the time since a drug is administered (X) and the concentration of the drug in the blood-stream (Y).

Following examples illustrate situations for multiple linear regression

1. The amount of sales of a product (dependent variable) is associated with

several independent variables such as price of the product, amount of expenditure on its advertisement, quality of the product, and the number of competitors.

2. Annual savings of a family (dependent variable) are associated with several independent variables such as the annual income, family size, health conditions of family members, and number of children in school or college.

3. The blood pressure of a person (dependent variable) is associated with several independent variables such as his or her age, weight, the level of blood cholesterol, and the level of blood sugar.

The linear regression model assumes that the value of the dependent variable changes in direct proportion to changes in the value of an independent variable, regardless of values of other independent variables. Linear regression is the simplest form of regression and there are more general and complicated regression models. We shall restrict our attention only to linear regression model in this chapter.

3.2 Fitting Simple Linear Regression

Consider an example where we wish to predict the amount of crop yield (in kg. per acre) as a linear function of the amount of fertilizer applied (in kg. per acre). In this example, the crop yield is to be predicted. Therefore, It is dependent variable and is denoted by *Y*. The amount of fertilizer applied is the variable used for the purpose of making the prediction. Therefore, it is the independent variable and is denoted by *X*.

Amount of fertilizer (*X*) (Kgs. In per acre)	Yield (*Y*) (in '00 kg)
30	43
40	45
50	54
60	53
70	56
80	63

Table : 3.1

Table 3.1 shows the amount of fertilizer and the crop yield for six cases. These pairs of observations are used to obtain the scatter diagram as shown in Fig. 2.1

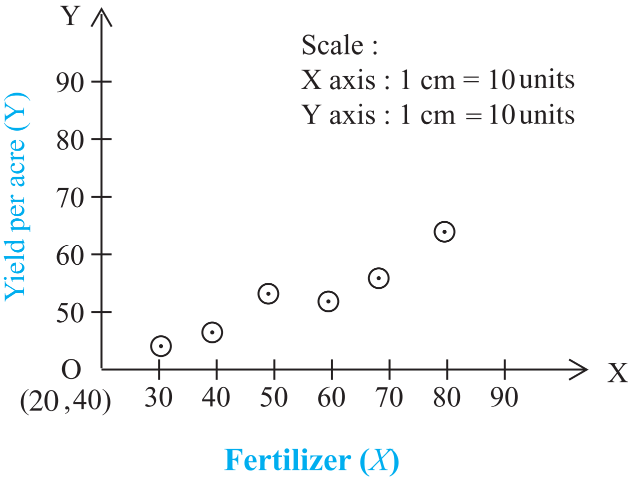


Fig. 3.1: Scatter digram of the yield of grain and amount of fertilizer used.

We want to draw a straight line that is closest to the points in the scatter diagram (Fig. 3.1). If all the points were collinear (that is, on a straight line), there would have been no problem in drawing such a line. There is a problem because all the points are not on a straight line.

Since the points in the scatter diagram do not form a straight line, we want to draw a straight line that is closest to these points. Theoretically, the number of possible line is unlimited. It is therefore necessary to specify some condition in order to ensure that we draw the straight line that is closest to all the data points in the scatter diagram. The method of least squares provide the line of best fit because it is closest to the data points in the scatter diagram according to the least squares principle.

3.2.1 Method of Least Squares

The principle used in obtaining *the line of best fit* is called the **method of least squares**. The method of least squares was developed by Adrien-Maire Lagendre and Carl Friedrich Gauss independently of each other. Let us understand the central idea behind the principle of least squares.

Suppose the data consists of n pairs of values $(x_1, y_1), (x_2, y_2), \dots, (x_n, y_n)$ and suppose that the line that fits best to the given data is written as follows. $\boxed{\hat{Y} = a + bX}$ (Here, $\hat{Y}$ is to be read as *Y cap*.)This equation is called the prediction equation. That is using the same values of constants a and b, the predicted value of Y are given by $\hat{Y}_i = a + bx_i$, where x_i is the value of the independent variable and $\hat{Y}_i$ is the corresponding predicted value of Y. Note that the observed value y_i of the independent variable Y is different from the predicted value $\hat{Y}_i$.

The observed valued (y_i) and predicted values $\left(\hat{Y}_i\right)$ of Y do not match perfectly because the observations do not fall on a straight line. The Difference between the observed values and the predicted values are called errors or residuals. Mathematically speaking the quantities

$y_1 - \hat{Y}_1$, $y_2 - \hat{Y}_2$ -------- $y_n - \hat{Y}_n$ or equivalently,

the quantities $y_1 - (a + bx_1)$, $y_2 - (a + bx_2)$, ..., $y_n - (a + bx_n)$ are deviations of observed values of Y from the corresponding predicted values and are therefore called errors or residuals. We write $e_i = y_i - \hat{Y}_1 = y_i - (a + bx_i)$, for $i = 1, 2, ..., n$.

Geometrically, the residual e_i, which is given by $y_i - (a + bx_i)$, denotes the vertical distance (which may be positive or negative) between the observed value (y_i) and the predicted value $\left(\hat{Y}_i\right)$.

The principle of the method of least squares can be stated as follows.

Among all the possible straight lines that can be drawn on the scatter diagram, the line of best fit is defined as the line that minimizes the sum of squares of residuals, that is, the sum of squares of deviations of the predicted y-values from the observed y-values. In other words, the line of the best fit is obtained by determining the constant a and b so that

$$\sum_{i=1}^{n} e_i^2 = \sum_{i=1}^{n} (y_i - \hat{y}_i)^2 = \sum_{i=1}^{n} [y_i - (a + bx_i)]^2$$

is minimum.

The straight line obtained using this principle is called *the least regression line.*

Symbolically, we write

$$S^2 = \sum_{i=1}^{n} (y_i - \hat{y}_i)^2$$

as the sum of squared errors. It can also be written as

$$S^2 = \sum_{i=1}^{n} [y_i - (a + bx_i)]^2$$

We want to determine the constants a and b in such a way that S^2 is minimum.

Note that S^2 is a continuous and differentiable function of both a and b. We differentiate S^2 with respect to a (assuming b to be constant) and with respect to b (assuming a constant) and with respect to b (assuming a to be constant). We then equate both these derivatives to zero in order to minimize S^2. As the result, we get the following two linear equations in two unknowns a and b.

$$\sum_{i=1}^{n} y_i = na + b\sum_{i=1}^{n} x_i$$

$$\sum_{i=1}^{n} x_i y_i = a\sum_{i=1}^{n} x_i + b\sum_{i=1}^{n} x_i^2$$

When we solve these two linear equations, the values of a and b that minimize S^2 are given by

$$a = \bar{y} - b\bar{x}\ ,\ b = \frac{\operatorname{cov}(X, Y)}{\sigma_x^2},$$

where

$$\operatorname{cov}(X, Y) = \frac{1}{n}\sum_{i=1}^{n} x_i y_i - \bar{x}\bar{y},$$

and

$$\sigma_x^2 = \frac{1}{n}\sum_{i=1}^{n} x_i^2 - \bar{x}^2$$

Substituting the values of a and b obtained as indicated above in the regression equation

$Y = a + bX$,

we get the equation

$Y - \bar{y} = b(X - \bar{x})$

Note: The constant b is called the regression coefficient (or the slope of the regression line) and the constant a is called the Y-intercept (that is, the Y value when $X = 0$). Recall that the equation $Y = a + bX$ defining a straight line is called the slope intercept formula of the straight line.

When observations on two variables, X and Y, are available, it is possible to fit a linear regression of Y on X as well as a linear regression of X on Y. Therefore, we consider both the models in order to understand the difference between the two and also the relationship between the two.

3.2.2 Regression of *Y* on *X*.

We now introduce notation b_{YX} for b when Y is the dependent variable and X is the independent variable.

Linear regression of Y on X assumes that the variable X is the independent variable and the variable Y is the dependent variable. In order to make this explicit, we express the linear regression model as follows.

$$Y - \bar{y} = b_{YX}(X - \bar{x}),$$

or

$$Y = b_{YX} X.$$

Here, note that b is replaced by b_{YX}.

$$\begin{aligned} b_{YX} &= \frac{\operatorname{cov}(X,Y)}{\operatorname{var}(X)} \\ &= \frac{\sum\left(x_i - \bar{x}\right)\left(y_i - \bar{y}\right)}{\sum\left(x_i - \bar{x}\right)^2} \\ &= \frac{\sum x_i y_i - n\bar{x}\bar{y}}{\sum x_i^2 - n\left(\bar{x}\right)^2} \end{aligned}$$

2.2.3 Regression of *X* on *Y*

The notation b_{XY} stands for b when X is the dependent variable and Y is the independent variable.

Linear regression of X on Y assumes that the variable Y is the independent variable and the variable X is the dependent variable. In order to make it clear that this model is different from the linear regression of Y on X, we express the linear regression model as follows.

$$X = a' + b'Y.$$

The method of least squares, when applied to this model leads to the following expressions for the constant a' and b'.

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

and

$$b' = \frac{\operatorname{cov}(X,Y)}{\operatorname{var}(Y)}$$

Substituting the values of a' and b' in the linear regression model, we get

$$\begin{aligned} X &= a' + b'Y \\ X &= \bar{x} - b'\bar{y} + b'Y \\ \text{i.e. } (X - \bar{x}) &= b'(Y - \bar{y}). \end{aligned}$$

Note: The constant b′ in the above equation is called the regression coefficient of X on Y. In order to make this explicit, it will henceforth be written as b_{XY} instead of b′. The least squares regression of X on Y will therefore be written as

$$(X - \bar{x}) = b_{YX}(Y - \bar{y}).$$

Note that the linear regression of X on Y is expressed as

$$X = a' + b_{YX} Y.$$

Here note that b is replaced by b_{XY}. This can be written as

$$Y = \frac{1}{b_{XY}}(X - a')$$

Showing that the constant $\left(\frac{1}{b_{XY}}\right)$ is the slope of the line of regression of X on Y.

Further, note that the regression coefficient b_{XY} involved in the linear regression of X on Y is given by

$$\begin{aligned} b_{XY} &= \frac{\operatorname{cov}(X,Y)}{\operatorname{var}(Y)} \\ &= \frac{\frac{1}{n}\sum\left(x_i - \bar{x}\right)\left(y_i - \bar{y}\right)}{\frac{1}{n}\sum\left(y_i - \bar{y}\right)^2} \end{aligned}$$

$$= \frac{\frac{1}{n}\sum x_i y_i - \bar{x}\bar{y}}{\frac{1}{n}\sum y_i^2 - n\bar{y}^2}$$

Also,

$$b_{XY} = \frac{\sum\left(x_i - \bar{x}\right)\sum\left(y_i - \bar{y}\right)}{\sum\left(x_i - \bar{x}\right)^2}$$

$$= \frac{\sum x_i y_i - n\bar{x}\bar{y}}{\sum x_i^{\,2} - n\left(\bar{x}\right)^2}$$

Observed that the point $\left(\bar{x}, \bar{y}\right)$ satisfies equation of both the lines of regression. Therefore, the point $\left(\bar{x}, \bar{y}\right)$ is the point of intersection of the two lines regression.

SOLVED EXAMPLES

Ex. 1 : For the data on fertilizer application and yield of grain is given in the table 3.2

i) Obtain the line of regression of yield of grain on the amount of fertilizer used.

ii) Draw the least squares regression line.

iii) Estimate the yield of grain when 90kgs. of fertilizer is applied.

Solution :

Amount of fertilizer $X = x_i$	Yield $Y = y_i$	x_i	$x_i y_i$
30	43	900	1290
40	45	1600	1800
50	54	2500	2700
60	53	3600	3180
70	56	4900	3920
80	63	6400	5040
330	314	19900	17930

Table : 3.2

i) Since $n = 6, \ \sum_{i=1}^{6} x_i = 330, \ \sum_{i=1}^{6} y_i = 314.$

$$\sum_{i=1}^{6} x_i^2 = 19900 \ \text{and} \ \sum_{i=1}^{6} x_i y_i = 17930$$

$$\bar{x} = \frac{\sum x_i}{n} = 55, \ \bar{y} = \frac{\sum y_i}{n} = 52.3$$

$$b_{XY} = \frac{\sum x_i y_i - n\bar{x}\bar{y}}{\sum x_i^2 - n\bar{x}^2}$$

$$= \frac{17930 - 6 \times 55 \times 52.3}{19900 - 6 \times 3025}$$

$$= \frac{671}{1750} = 0.38$$

and $a = \bar{y} - b_{XY} \,.\, \bar{x} = 31.4.$

Finally, the line of regression of Y on X is given by

$$\underset{\text{Predicted yield}}{Y} = \underset{\text{Fixed yield}}{31.4} + \underset{\text{change in the yield per one kg. increase in fertilizer}}{0.38}\ \underset{\text{kgs. of fertilizer used}}{X}$$

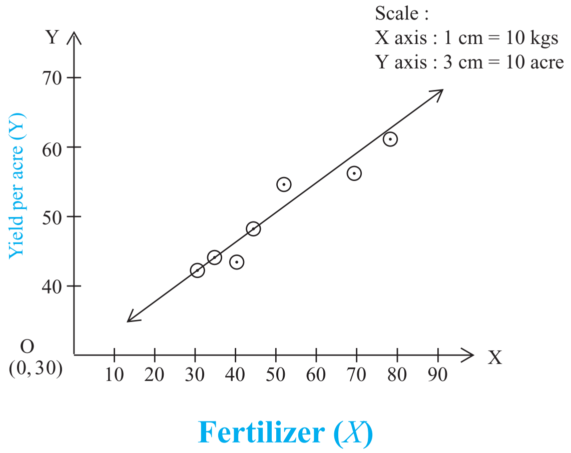


Fig. 3.2 : Scatter diagram and the line of regression of yield on amount of fertilizer

ii) To draw the least squares regression line, we pick any two convenient values of X and find the corresponding values of Y.

For $x = 35.$ $\quad y = 44.7$

$x = 45,$ $\quad y = 48.5$

Joining the two points (35,44.8) and (45,48.6), we get the line in Fig 3.2

iii) Putting $x = 90$ in the regression equation

$$\hat{Y} = 31.4 + 0.38 \times 90 = 65.6$$

Ex. 2: A departmental store gives in service training to the salesmen followed by a test. It is experienced that the performance regarding sales of any salesmen is linearly related to the scores secured by him. The following data give test scores and sales made by nine salesmen during fixed period.

Test scores (X)	16	22	28	24	29	25	16	23	24
Sales ('00Rs.) (Y)	35	42	57	40	54	51	34	47	45

i) Obtain the line of regression of Y on X.

ii) Estimate Y when $X = 17$.

Solution : To show the calculations clearly, it is better to prepare the following table

$X=x_i$	$Y=y_i$	$x_i-\bar{x}$	$y_i-\bar{y}$	$(x_i-\bar{x})^2$	$(x_i-\bar{x})$ $(y_i-\bar{y})$
16	35	−7	−10	49	70
22	42	−1	−3	1	3
28	57	5	12	25	60
24	40	1	−5	1	−5
29	54	6	9	36	54
25	51	2	6	4	12
16	34	−7	−11	49	77
23	47	0	2	00	00
24	45	1	0	1	00
207	405	00	00	166	271

i) $n = 9$ and $\sum_{i=1}^{9} x_i = 207, \; \sum_{i=1}^{9} y_i = 405.$

$$\bar{x} = \frac{\sum x_i}{n} = 23, \quad \bar{y} = \frac{\sum y_i}{n} = 45$$

Since the means of X and Y are whole numbers, it is preferable to use the formula $\frac{\sum(x_i-\bar{x})(y_i-\bar{y})}{\sum(x_i-\bar{x})^2}$ for the calculation of b_{YX}.

Line of regression of Y on X is

$$Y = a + b_{XY}X$$

where
$$b_{XY} = \frac{\text{cov}(X,Y)}{\sigma_Y^2}$$

$$= \frac{\frac{\sum(x_i-\bar{x})(y_i-\bar{y})}{n}}{\frac{\sum(x_i-\bar{x})^2}{n}}$$

$$= \frac{\sum(x_i-\bar{x})(y_i-\bar{y})}{\sum(x_i-\bar{x})^2}$$

$$= \frac{271}{166}$$

$$= 1.6325$$

and
$$a = \bar{y} - b_{XY}\,\bar{x}$$

$$= 7.4525$$

Here, the line of regression of Y on X is

$$Y = 7.4525 + 1.6325X$$

ii) Estimate of Y when $X = 17$ is

$$Y = 7.4525 + (1.6325)\,(17)$$

$$= 35.205$$

Ex 3: The management of a large furniture store would like to determine sales (in thousands of Rs.) (X) on a given day on the basis of number of people (Y) that visited the store on that day. The necessary records were kept, and a random sample of ten days was selected for the study. The summary results were as follows:

$\sum x_i = 370,\ \sum y_i = 580,\ \sum x_i^2 = 17200,$

$\sum y_i^2 = 41640,\ \sum x_i y_i \ = 11500,\ n = 10$

Obtain the line of regression of X on Y.

Solution:

Line of regression of X on Y is

$$X = a' + b_{XY} Y$$

where $$b_{XY} = \frac{\text{cov}(X,Y)}{\sigma_Y^2}$$

$$= \frac{\frac{\sum x_i y_i}{n} - \overline{x}\,\overline{y}}{\frac{\sum y_i^2}{n} - \left(\overline{y}\right)^2}$$

$$= \frac{\left(\frac{11500}{10}\right) - \left(\frac{370}{10}\right)\left(\frac{580}{10}\right)}{\left(\frac{41640}{10}\right) - \left(\frac{580}{10}\right)^2}$$

$$= \frac{1150 - 37 \times 58}{4164 - (58)^2}$$

$$= -\frac{996}{800}$$

$$= -1.245$$

and $$a' = \overline{x} - b_{XY}\ \overline{y}$$

$$= 37 - (-1.245)\,(58)$$

$$= 109.21$$

$\therefore$ Line of regression of X on Y is

$$X = 109.21 - 1.245Y$$

EXERCISE 3.1

1. The HRD manager of a company wants to find a measure which he can use to fix the monthly income of persons applying for the job in the production department. As an experimental project, he collected data of 7 persons from that department referring to years of service and their monthly incomes.

Years of service (X)	11	7	9	5	8	6	10
Monthly Income (Rs.1000's) (Y)	10	8	6	5	9	7	11

(i) Find the regression equation of income on years of service.

(ii) What initial start would you recommend for a person applying for the job after having served in similar capacity in another company for 13 years?

2. Calculate the regression equations of X on Y and Y on X from the following data:

X	10	12	13	17	18
Y	5	6	7	9	13

3. For a certain bivariate data on 5 pairs of observations given

$\sum x = 20,\ \sum y = 20,\ \sum x^2 = 90,$

$\sum y^2 = 90,\ \sum xy = 76$

Calculate (i) cov(x,y) (ii) b_{yx} and b_{xy}, (iii) r

4. From the following data estimate y when $x = 125$

X	120	115	120	125	126	123
Y	13	15	14	13	12	14

5. The following table gives the aptitude test scores and productivity indices of 10 workers selected at random.

Aptitude score (X)	60	62	65	70	72	48	53	73	65	82
Productivity Index (Y)	68	60	62	80	85	40	52	62	60	81

Obtain the two regression equations and estimate:

(i) The productivity index of a worker whose test score is 95.

(ii) The test score when productivity index is 75.

6. Compute the appropriate regression equation for the following data:

X [Independent Veriable]	2	4	5	6	8	11
Y [Dependent Veriable]	18	12	10	8	7	5

7. The following are the marks obtained by the students in Economics (X) and Mathematics (Y)

X	59	60	61	62	63
Y	78	82	82	79	81

Find the regression equation of Y on X.

8. For the following bivariate data obtain the equations of two regression lines:

X	1	2	3	4	5
Y	5	7	9	11	13

9. From the following data obtain the equation of two regression lines:

X	6	2	10	4	8
Y	9	11	5	8	7

10. For the following data, find the regression line of Y on X

X	1	2	3
Y	2	1	6

Hence find the most likely value of y when $x = 4$.

11. From the following data, find the regression equation of Y on X and estimate Y when $X = 10$.

X	1	2	3	4	5	6
Y	2	4	7	6	5	6

12. The following sample gives the number of hours of study (X) per day for an examination and marks (Y) obtained by 12 students.

X	3	3	3	4	4	5	5	5	6	6	7	8
Y	45	60	55	60	75	70	80	75	90	80	75	85

Obtain the line of regression of marks on hours of study.

3.3 Properties of Regression Coefficients

The line of regression of Y on X is given by $Y = a + b_{YX}X$ and the line of regression of X on Y is given by $X = a' + b_{XY}Y$.

Here, $b_{YX} \frac{\operatorname{cov}(X,Y)}{\operatorname{var}(X)}$, the slope of the line of regression of Y on X is called the regression coefficient of Y on X. Simillary,

$b_{XY} = \frac{\operatorname{cov}(X,Y)}{\operatorname{var}(Y)}$, the slope of the line of regression of X on Y is called the regression coefficient of X on Y. These two regression coefficients have the following property.

(a) $$b_{XY} \,.\, b_{YX} = r^2$$

where r is the correlation coefficient between X and Y,

b_{XY} is the regression coefficient of X on Y. and b_{YX} is the regression coefficient of Y on X.

Proof: Note that

$$b_{XY} \,.\, b_{YX} = \frac{\operatorname{cov}(X,Y)}{\operatorname{var}(Y)} \,.\, \frac{\operatorname{cov}(X,Y)}{\operatorname{var}(X)}$$

$$= \left[\frac{\operatorname{cov}(X,Y)}{\sigma_X . \sigma_Y}\right]^2$$

$$= r^2.$$

Can it be said that the correlation coefficient is the square root of the product of the two regression coefficients?

(b) If $b_{YX} > 1$, then $b_{XY} < 1$.

Proof: Let, if possible, $b_{XY} > 1$ and $b_{YX} > 1$.

Then, using the above result, $b_{XY} \cdot b_{YX} > 1$, implies that $r^2 > 1$, which is impossible. (Can you provide the reason?)

This shows that our assumption must be invalid. That is, both the regression coefficients cannot simultaneously exceed unity.

We already know that the two variances σ_x^2, σ_y^2, and the correlation coefficient r satisfy the relation.

$$\text{cov}(X, Y) = \text{r} \cdot \sigma_X \cdot \sigma_Y$$

$$\therefore \quad r = \frac{\text{cov}(X,Y)}{\sigma_X . \sigma_Y}$$

The regression coefficients can also be written as follows.

$$b_{YX} = \frac{\text{cov}(X.Y)}{\sigma_X^2} = \frac{r.\sigma_X.\sigma_Y}{\sigma_x^2} = r \,.\, \frac{\sigma_Y}{\sigma_X}$$

and

$$b_{XY} = \frac{\text{cov}(X.Y)}{\sigma_y^2} = \frac{r.\sigma_X.\sigma_Y}{\sigma_y^2} = r \,.\, \frac{\sigma_X}{\sigma_Y}$$

(c) $\left|\dfrac{b_{YX} + b_{XY}}{2}\right| \geq |r|$

Proof: We have already seen that

$$b_{YX} = r\,\frac{\sigma_Y}{\sigma_X} \quad \text{and } b_{XY} = r\,\frac{\sigma_X}{\sigma_Y},$$

where σ_X and σ_Y are the standard deviations of X and Y, respectively.

Therefore,

$$b_{YX} + b_{XY} = r\,\frac{\sigma_Y}{\sigma_X} + r\,\frac{\sigma_X}{\sigma_Y}$$

$$= \left[\frac{\sigma_Y}{\sigma_X} + \frac{\sigma_X}{\sigma_Y}\right]$$

$$= r\left[\frac{\sigma_Y + \sigma_X}{\sigma_X . \sigma_Y}\right] \qquad (1)$$

But $(\sigma_X - \sigma_Y)^2 > 0$ and therefore

$$\sigma^2_X - \sigma^2_Y - 2\sigma_X\sigma_Y \geq 0$$

$$\sigma^2_X + \sigma^2_Y \geq 2\sigma_X\sigma_Y$$

$$\frac{\sigma_y^2 + \sigma_x^2}{\sigma_X . \sigma_Y} \geq 2$$

$$\therefore \ \text{r}\,\frac{\sigma_y^2 + \sigma_x^2}{\sigma_X . \sigma_Y} \geq 2r. \qquad (2)$$

From (1) and (2), we have

$$b_{YX} + b_{XY} \geq 2r.$$

$$\therefore \ \frac{b_{YX} + b_{XY}}{2} \geq r.$$

this result shows that the arithmetic mean of the two regression coefficients, namely b_{YX} and b_{XY} is greater than or equal to r. This result, however, holds only when $b_{YX,}$ b_{XY} and r are positive. (Can you find the reason?)

Consider the case where $b_{YX} = -0.8$ and $b_{XY} = -0.45$. In this case, we have $r = -0.6$. (Can you find the reason?)

Note that $b_{YX} + b_{XY} = -1.25$, and $2r = -1.2$. This shows that $b_{YX} + b_{XY} \leq 2r$.

It may be interesting to note that

$$b_{YX} = \frac{\text{cov}(X.Y)}{\sigma_X^2}$$

$$b_{XY} = \frac{\text{cov}(X.Y)}{\sigma_y^2}$$

$$r = \frac{\text{cov}(X,Y)}{\sigma_X . \sigma_Y}$$

It is evident from the above three equations that all the coefficients have the same numerator and this numerator determines their sign. As the result, all these coefficients have the same sign. In other words, if $r > 0$, then $b_{YX} > 0$, and $b_{XY} > 0$. Similary. if $r < 0$. then $b_{YX} < 0$, and $b_{XY} < 0$. Finally. if $r = 0$, then $b_{YX} = b_{XY} = 0$.

(d) b_{YX} and b_{XY} are not affected by change of origin, but are affected by change of scale. This property is known as *invariance* property.

The invariance property states that b_{YX} and b_{XY} are invariant under change of origin, but are not invariant under change of scale.

Proof: Let $U = \dfrac{X-a}{h}$ and $V = \dfrac{Y-b}{k}$,

where a, b, h and k are constants with the condition that $h, k \neq 0$

We have already proved that $\sigma_X^2 = h^2\sigma_U^2$, $\sigma_Y^2 = k^2\sigma_V^2$, and cov($X$, Y) = hkcov(U, V).

Therefore,

$$b_{YX} = \frac{\text{cov}(X.Y)}{\sigma_X^2}$$

$$= \frac{hk\,\text{cov}(U,V)}{h^2\sigma_U^2}$$

$$= \frac{k}{h}\,\frac{\text{cov}(U,V)}{\sigma_U^2}$$

that is,

$$\boxed{b_{YX} = \frac{k}{h}b_{VU}}$$

Similatly,

$$\boxed{b_{XY} = \frac{k}{h}b_{UV}}$$

These two results show that regression coefficients are invariant under change of origin, but are not invariant under change of scale.

SOLVED EXAMPLES

Ex. 1 : The table below gives the heights of fathers (X) and heights of their sons (Y) respectively.

Heights of fathers (inches) (X)	64	62	66	63	67	61	69	65	67	66
Heights of sons (inches) (Y)	67	65	67	64	68	65	67	64	70	66

(i) Find the regression line of Y on X.

(ii) Find the regression line of X on Y.

(iii) Predict son's height if father's height is 68 inches.

(iv) Predict father's height if son's height is 59 inches.

Solution :

Let us use the change of origin for computations of regression coefficients.

Let $u_i = x_i - 65$ and $v_i = y_i - 67$

x_i	y_i	u_i	v_i	u_i^2	v_i^2	$u_i v_i$
64	67	−1	0	1	0	0
62	65	−3	−2	9	4	6
66	67	1	0	1	0	0
63	64	−2	−3	4	9	6
67	68	2	1	4	1	2
61	65	−4	−2	16	4	8
69	67	4	0	16	0	0
65	64	0	−3	0	9	0
67	70	2	3	4	9	6
66	66	1	−1	1	1	−1
Total		0	−7	56	37	27

Here, $n = 10$, $\sum u_i = 0$, $\sum v_i = -7$,

$\sum u_i^2 = 56$, $\sum v_i^2 = 37$, and $\sum u_i v_i = 27$

$$\bar{u} = \frac{\sum u_i}{n} = \frac{0}{10} = 0,$$

$$\sigma_u^2 = \frac{\sum u_i^2}{n} - \left(\bar{u}\right)^2$$

$$= \frac{56}{10} - 0^2$$

$$\sigma_u^2 = 5.6$$

$$\bar{v} = \frac{\sum v_i}{n}, \qquad \sigma_v^2 = \frac{\sum v_i^2}{n} - \left(\bar{v}\right)^2$$

$$= \frac{-7}{10} \qquad = \frac{37}{10} - (-0.7)^2$$

$$= -0.7 \qquad = 3.21$$

$$\text{cov}(u, v) = \frac{\sum u_i v_i}{n} - \bar{u}\ \bar{v}$$

$$= \frac{27}{10} - 0 \times -0.7$$

$$= 2.7$$

Now, you know that, regression coefficients are independent of change of origin.

$$\therefore\ b_{XY} = b_{UV} = \frac{\text{cov}(u,v)}{\sigma_V} = \frac{2.7}{3.21} = 0.84$$

$$\text{and } b_{YX} = b_{VU} = \frac{\text{cov}(u,v)}{\sigma_U} = \frac{2.7}{5.6} = 0.48$$

You are also aware that mean is affected by change of origin.

$$\therefore\ \bar{x} = \bar{u} + 65 = 0 + 65 = 65$$

$$\text{and } \bar{y} = \bar{v} + 67 = -0.7 + 67 = 66.3$$

(i) Line of regression of Y on X is

$$(Y - \bar{y}) = b_{YX}(X - \bar{x})$$

$$\therefore\ (Y - 66.3) = 0.48\ (X - 65)$$

$$\therefore\ Y = 0.48.\ X + 35.1$$

ii) Regression line of X on Y is

$$(X - \bar{x}) = b_{XY}(Y - \bar{y})$$

$$\therefore\ (X - 65) = 0.84\ (Y - 66.3)$$

$$\therefore\ X = 0.84.\ Y + 9.31$$

iii) Estimate of sons height Y for $X = 68$

$$Y = 0.48 \times 68 + 35.1$$

$$= 67.74 \text{ inches}$$

iv) Estimate of fathers height X for $Y = 59$

$$X = 0.84 \times 59 + 9.31$$

$$= 58.87 \text{ inches}$$

v) Correlation coefficient

$$r = \sqrt{b_{YX}.b_{XY}}$$

$$= \sqrt{0.84 \times 0.48}$$

$$= 0.635$$

We choose positive square root ! (why?)

Ex. 2: Compute regression coefficient from the following data on the variable weight (X) and height (Y) of 8 individuals :

$$n = 8,\ \sum(x_i - 45) = 48$$

$$\sum(x_i - 45)^2 = 4400,$$

$$\sum(y_i - 150) = 280,$$

$$\sum(y_i - 150)^2 = 167432,$$

$$\sum(x_i - 45).(y_i - 150) = 21680$$

Solution: Let $u_i = x_i - 45$ and $v_i = y_i - 150$

$$\text{So} \quad \sum u_i = 48, \quad \sum u_i^2 = 4400,$$

$$\sum v_i = 280, \ \sum v_i^2 = 167432,$$

$$\sum u_i v_i = 21680$$

$$\therefore\ \bar{u} = \frac{\sum u_i}{n} = \frac{48}{8} = 6$$

$$\bar{v} = \frac{\sum v_i}{n} = \frac{280}{8} = 35$$

$$\sigma_u^2 = \frac{\sum u_i^2}{n} - \left(\bar{u}\right)^2$$

$$= \frac{4400}{8} - (6)^2 = 514$$

$$\sigma_v^{\,2} = \frac{\sum v_i^2}{n} - \left(\bar{v}\right)^2$$

$$= \frac{167432}{8} - (35)^2 = 19704$$

$$\text{cov}\,(u, v) = \frac{\sum u_i v_i}{n} - \bar{u}\ \bar{v}$$

$$= \frac{21680}{8} - (6)\,(35)$$

$$= 2500$$

From the properties of regression coefficients, you know they are independent of change of origin.

$$\therefore\ b_{YX} = b_{VU} = \frac{\text{cov}(u,v)}{\sigma_U} = \frac{2500}{514} = 4.86$$

$$\text{and } b_{XY} = b_{UV} = \frac{\text{cov}(u,v)}{\sigma_V} = \frac{2500}{19704} = 0.12$$

(Have you noticed $b_{YX} > 1$ and $b_{XY} < 1$?)

Ex. 3: The following results were obtained from records of age (X) and systolic blood pressure (Y) of a group of 10 women.

	X	Y
Mean	53	142
Variance	130	165

$$\sum\left(x_i - \bar{x}\right)\left(y_i - \bar{y}\right) = 1170$$

Find the appropriate regression equation and use it to estimate the blood pressure of a woman with age 47 years.

Solution:

Here, we need to find line of regression of Y on X, which is given as

$$Y = a + b_{YX} X$$

$$\text{where,}\quad b_{YX} = \frac{\text{cov}(X,Y)}{\sigma_x^2}$$

$$= \frac{\dfrac{\sum\left(x_i - \bar{x}\right)\left(y_i - \bar{y}\right)}{n}}{\sigma_x^2}$$

$$= \frac{\left(\dfrac{1170}{10}\right)}{130}$$

$$= 0.9$$

$$\text{and}\quad a = \bar{y} - b_{YX}\ \bar{x}$$

$$= 142 - (0.9) \times (53)$$

$$= 94.3$$

Therefore, regression equation of Y on X is

$$Y = 94.3 + 0.9\ X$$

Now, the estimate of blood pressure of women with age 47 years is

$$Y = 94.3 + 0.9 \times 47$$

$$= 136.6$$

Ex. 4: Given the following data, obtain the linear regression & estimate of X for $Y = 10$

$\bar{x} = 7.6$, $\bar{y} = 14.8$, $\sigma_X = 3.5$, $\sigma_Y = 28$ and $r = 0.8$

Solution:

Here, we need to obtain line of regression of X on Y which can be expressed as

$$X = a' - b_{YX}\, Y$$

$$\text{where}\quad b_{XY} = \frac{\text{cov}(X,Y)}{\sigma_Y^2}$$

$$= r\,\frac{\sigma_X}{\sigma_Y}$$

$$= 0.8\,\frac{(3.5)}{(28)}$$

$$= 0.1$$

and $a' = \bar{x} - b_{YX}\ \bar{y}$

$= 7.6 - (0.1)\,(14.8)$

$= 6.12$

$\therefore$ Line of regression of X on Y is

$X = 6.12 + 0.1\ Y$

Estimate of X for $Y = 10$ is

$X = 6.12 + 0.1 \times 10$

$X = 7.12$

EXERCISE 3.2

1. For a bivariate data.

$\bar{x} = 53,\ \ \bar{y} = 28,\ b_{YX} = -1.2,\ b_{XY} = -0.3$

Find

i) Correlation coefficient between X and Y.

ii) Estimate of Y for $X = 50$

iii) Estimate of X for $Y = 25$

2. From the data of 20 pairs of observation on X and Y, following results are obtained.

$\bar{x} = 199,\ \ \bar{y} = 94,$

$$\sum\left(x_i - \bar{x}\right)^2 = 1200 \quad \sum\left(y_i - \bar{y}\right)^2 = 300$$

$$\sum\left(x_i - \bar{x}\right)\left(y_i - \bar{y}\right) = -250$$

Find

i) The line of regression of Y on X.

ii) The line of regression of X on Y.

iii) Correlation coefficient between X and Y

3. From the data of 7 pairs of observations on X and Y, following results are obtained.

$$\sum\left(x_i - 70\right) = -35 \quad \sum\left(y_i - 60\right) = -7$$

$$\sum\left(x_i - 70\right)^2 = 2989 \quad \sum\left(y_i - 60\right)^2 = 476$$

$$\sum\left(x_i - 70\right)\left(y_i - 60\right) = 1064$$

[*Given* $\sqrt{0.7884} = 0.8879$]

Obtain

i) The line of regression of Y on X.

ii) The line of regression of X on Y

iii) The correlation coefficient between X and Y.

4. You are given the following information about advertising expenditure and sales.

	Advertisment expenditure (Rs.in lakh) (X)	Sales (Rs.in lakh) (Y)
Arithmetic mean	10	90
Standard deviation	3	12

Correlation coefficient between X and Y is 0.8

(i) Obtain the two regression equations.

(ii) What is the likely sales when the advertising budget is Rs 15 lakh?

(iii) What should be the advertising budget if the company wants to attain sales target of Rs.120 lakh?

5. Bring out inconsistency if any, in the following :

(i) $b_{YX} + b_{XY} = 1.30$ and $r = 0.75$

(ii) $b_{YX} = b_{XY} = 1.50$ and $r = -0.9$

(iii) $b_{YX} = 1.9$ and $b_{XY} = -0.25$

(iv) $b_{YX} = 2.6$ and $b_{XY} = \dfrac{1}{2.6}$

6. Two samples from bivariate populations have 15 observations each. The sample means of X and Y are 25 and 18 respectively. The corresponding sum of squares of deviations from respective means are 136 and 150. The sum of product of deviations from respective means is 123. Obtain the equation of line of regression of X on Y.

7. For a certain bivariate data

	X	Y
Mean	25	20
S.D.	4	3

And $r = 0.5$. estimate y when $x = 10$ and estimate x when $y = 16$

8. Given the following information about the production and demand of a commodity obtain the two regression lines :

	Production (X)	Demand (Y)
Mean	85	90
S.D.	5	6

Coefficient of correlation between X and Y is 0.6. Also estimate the production when demand is 100.

9. Given the following data, obtain linear regression estimate of X for $Y = 10$

$\bar{x} = 7.6$, $\bar{y} = 14.8$, $\sigma_X = 3.2$, $\sigma_Y = 16$ and $r = 0.7$

10. An inquiry of 50 families to study the relationship between expenditure on accommodation (Rs. x) and expenditure on food and entertainment (Rs. y) gave the following results. :

$\sum x = 8500$, $\sum y = 9600$, $\sigma_X = 60$, $\sigma_Y = 20$, $r = 0.6$

Estimate the expenditure on food and entertainment when expenditure on accommodation is Rs 200.

11. The following data about the sales and advertisement expenditure of a firms is given below (in Rs. Crores)

	Sales	Adv. Exp.
Mean	40	6
S.D.	10	1.5

i) Estimate the likely sales for a proposed advertisement expenditure of Rs.10 crores.

ii) What should be the advertisement expenditure if the firm proposes a sales target Rs.60 crores.

12. For a certain bivariate data the following information are available.

	X	Y
A.M.	13	17
S.D.	3	2

Correlation coefficient between x and y is 0.6. estimate x when $y = 15$ and estimate y when $x = 10$.

SOLVED EXAMPLES

Ex.1: The equations of the two lines of regression are $3x + 2y - 26 = 0$ and $6x + y - 31 = 0$

(i) Find the means of X and Y.

(ii) Obtain correlation coefficient between X and Y.

(iii) Estimate Y for $X = 2$.

Solution:

(i) We know that the co-ordinates of the point of intersection of the two lines are $\bar{x}$ and $\bar{y}$, the means of X and Y.

The regression equations are

$$3x + 2y - 26 = 0$$

and $$6x + y - 31 = 0$$

Solving these equations simultaneously, we get

$$\begin{array}{l} 6x + 4y - 52 = 0 \\ 6x + y - 31 \quad = 0 \\ (-) \quad (-) \quad (+) \\ \hline 3y - 21 = 0 \\ \therefore 3y = 21 \\ \text{i.e. } y = 7 \\ \text{and } x = 4 \end{array}$$

Hence the means of X and Y are $\bar{x} = 4$ and $\bar{y} = 7$

(ii) Now, to find correlation coefficient, we have to find the regression coefficients b_{YX} and b_{XY}

For this, we have to choose one of the lines as that of line of regression of Y on X and other the line of regression of X on Y

Let $3x + 2y - 26 = 0$ be the line of regression on Y on X this gives

$$Y = -\frac{3}{2}X + 13$$

The coefficient of X in this equation is $b_{YX} = -\frac{3}{2}$

Then the other equation is that of line of regression of X on Y which can be written as

$$X = -\frac{1}{6}Y + \frac{31}{6}$$

Here, the regression coefficient $b_{XY} = -\frac{1}{6}$.

Now, you know that

$$r^2 = b_{XY} \cdot b_{YX}$$
$$= 0.25$$
$$\therefore \quad r = \pm 0.5$$

The correlation coefficient has the sign as that of b_{YX} and b_{XY}

$$\therefore \quad r = -0.5$$

Note: We choose arbitrarily the lines as that of regression of Y on X or X on Y. if the product $b_{YX} \cdot b_{XY}$ is less than unity, our choice is correct. Fortunately, there are only two choices.

Ex 2.: The regression equation of Y on X is $y = \frac{4}{3}x$ and the regression equation X on Y is

$$x = \frac{y}{3} + \frac{5}{3}$$

Find

(i) Correlation coefficient between X and Y.

(ii) σ_Y^2 if $\sigma_X^2 = 4$

Solution:

Here, the regression lines are specified.

So $b_{YX} = \frac{4}{3}$ and $b_{XY} = \frac{1}{3}$

(i) $\therefore \; r^2 = b_{YX} \cdot b_{XY}$

$$= \frac{4}{3} \cdot \frac{1}{3}$$
$$= \frac{4}{9}$$

$\therefore \; r = +\frac{2}{3}$ (why $+\frac{2}{3}$ only?)

(ii) You know that

$$b_{YX} = r \cdot \frac{\sigma_Y}{\sigma_X}$$

$$\therefore \quad \frac{4}{3} = \frac{2}{3} \cdot \frac{\sigma_Y}{2}$$

$$\therefore \quad \sigma_Y = 4$$

$$\therefore \quad \sigma_Y^2 = 16$$

EXERCISE 3.3

1. From the two regression equations find r, $\bar{x}$ and $\bar{y}$.

   $4y = 9x + 15$ and $25x = 4y + 17$

2. In a partially destroyed laboratory record of an analysis of regression data, the following data are legible:

   Variance of $X = 9$

   Regression equations:

   $8x - 10y + 66 = 0$

   and $40x - 18y = 214$.

Find on the basis of above information

(i) The mean values of X and Y.

(ii) Correlation coefficient between X and Y.

(iii) Standard deviation of Y.

3. For 50 students of a class, the regression equation of marks in statistics (X) on the marks in Accountancy (Y) is $3y - 5x + 180 = 0$. The mean marks in accountancy is 44 and the variance of marks in statistics is $\left(\frac{9}{16}\right)^{th}$ of the variance of marks in accountancy. Find the mean marks in statistics and the correlation coefficient between marks in two subjects.

4. For a bivariate data, the regression coefficient of Y on X is 0.4 and the regression coefficient of X on Y is 0.9. Find the value of variance of Y if variance of X is 9.

5. The equations of two regression lines are

$$2x + 3y - 6 = 0$$

and $$2x + 2y - 12 = 0$$

Find (i) Correlation coefficient

(ii) $\frac{\sigma_X}{\sigma_Y}$

6. For a bivariate data: $\bar{x} = 53$, $\bar{y} = 28$, $b_{YX} = -1.5$ and $b_{XY} = -0.2$. Estimate Y when $X = 50$.

7. The equations of two regression lines are $x - 4y = 5$ and $16y - x = 64$. Find means of X and Y. Also, find correlation coefficient between X and Y.

8. In a partially destroyed record, the following data are available variance of $X = 25$. Regression equation of Y on X is $5y - x = 22$ and Regression equation of X on Y is $64x - 45y = 22$ Find

(i) Mean values of X and Y

(ii) Standard deviation of Y

(iii) Coefficient of correlation between X and Y.

9. If the two regression lines for a bivariate data are $2x = y + 15$ (x on y) and $4y = 3x + 25$ (y on x), find

(i) $\bar{x}$, (ii) $\bar{y}$, (iii) b_{YX},

(iv) b_{XY}, (v) r $\left[Given\sqrt{0.375} = 0.61\right]$

10. The two regression equations are

$$5x - 6y + 90 = 0 \text{ and } 15x - 8y - 130 = 0.$$

Find $\bar{x}$, $\bar{y}$, r.

11. Two lines of regression are $10x + 3y - 62 = 0$ and $6x + 5y - 50 = 0$ Identify the regression of x on y. Hence find $\bar{x}$, $\bar{y}$ and r.

12. For certain X and Y series, which are correlated the two lines of regression are $10y = 3x + 170$ and $5x + 70 = 6y$. Find the correlation coefficient between them. Find the mean values of X and Y.

13. Regression equations of two series are

$$2x - y - 15 = 0 \text{ and } 3x - 4y + 25 = 0$$

Find $\bar{x}$, $\bar{y}$ and regression coefficients. Also find coefficients of correlation.

$\left[Given\sqrt{0.375} = 0.61\right]$

14. The two regression lines between height (X) in inches and weight (Y) in kgs of girls are,

$$4y - 15x + 500 = 0$$

and $$20x - 3y - 900 = 0$$

Find mean height and weight of the group. Also, estimate weight of a girl whose height is 70 inches.

Let's Remember

- Line of regression of Y on X is

$$Y = a + bX$$

where $b = b_{YX}$ = regression coefficient of Y on X

$$b_{YX} = \frac{\operatorname{cov}(X,Y)}{\operatorname{var}(X)}$$

$$= \frac{\dfrac{\sum\left(x_i-\bar{x}\right)\left(y_i-\bar{y}\right)}{n}}{\dfrac{\sum\left(x_i-\bar{x}\right)^2}{n}}$$

$$= \frac{\dfrac{\sum x_i y_i}{n}-\bar{x}.\bar{y}}{\dfrac{\sum x_i^2}{n}-\left(\bar{x}\right)^2}$$

$$= \frac{\sum x_i y_i - n\bar{x}.\bar{y}}{\sum x_i^2 - n\left(\bar{x}\right)^2}$$

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

- Line of regression of X on Y is

$$X = a' + b'y$$

where $b' = b_{XY}$ = regression coefficient of Y on X

$$b_{XY} = \frac{\operatorname{cov}(X,Y)}{\operatorname{var}(Y)}$$

$$= \frac{\dfrac{\sum\left(x_i-\bar{x}\right)\left(y_i-\bar{y}\right)}{n}}{\dfrac{\sum\left(y_i-\bar{y}\right)^2}{n}}$$

$$= \frac{\dfrac{\sum x_i y_i}{n}-\bar{x}.\bar{y}}{\dfrac{\sum y_i^2}{n}-\left(\bar{y}\right)^2}$$

$$= \frac{\sum x_i y_i - n\bar{x}.\bar{y}}{\sum y_i^2 - n\left(\bar{y}\right)^2}$$

and $a^{l} = \bar{x} - b'\,\bar{y}$

- Line of regression of Y on X is also given as

  $(Y - \bar{y}) = b(X - \bar{x})$

- Line of regression of X on Y is also given as

  $(X - \bar{x}) = b'(Y - \bar{y})$

- $r^2 = b_{YX} \,.\, b_{XY} = b \,.\, b'$
- If $b_{YX} > 1$ then $b_{XY} < 1$
- $\left|\dfrac{b_{yx}+b_{xy}}{2}\right| \geq |r|$
- Regression coefficients are independent of change of origin but not of scale.
- Lines of regression have a point of intersection$(\bar{x}, \bar{y})$

MISCELLANEOUS EXERCISE - 3

I) Choose the correct alternative.

1. Regression analysis is the theory of

   a) Estimation b) Prediction

   c) Both a and b d) Calculation

2. We can estimate the value of one variable with the help of other known variable only if they are

   a) Correlated

   b) Positively correlated

   c) Negatively correlated

   d) Uncorrelated

3. There are __________ types of regression equations.

   a) 4 b) 2 c) 3 d) 1

4. In the regression equation of Y on X
   a) X is independent and Y is dependent.
   b) Y is independent and X is dependent.
   c) Both X and Y are independent.
   d) Both X and Y are dependent.

5. In the regression equation of X on Y
   a) X is independent and Y is dependent.
   b) Y is independent and X is dependent.
   c) Both X and Y are independent.
   d) Both X and Y are dependent.

6. b_{XY} is ___________
   a) Regression coefficient of Y on X
   b) Regression coefficient of X on Y
   c) Correlation coefficient between X and Y
   d) Covariance between X and Y

7. b_{YX} is ___________
   a) Regression coefficient of Y on X
   b) Regression coefficient of X on Y
   c) Correlation coefficient between X and Y
   d) Covariance between X and Y

8. ‘r’ is ___________
   a) Regression coefficient of Y on X
   b) Regression coefficient of X on Y
   c) Correlation coefficient between X and Y
   d) Covariance between X and Y

9. $b_{XY} . b_{YX}$ ___________
   a) $v(x)$ (b) σ_x (c) r^2 (d) $(\sigma_y)^2$

10. If $b_{yx} > 1$ then b_{xy} is ___________
    a) >1 (b) < 1 (c) > 0 (d) < 0

11. $| b_{xy} + b_{yx}| \geq$ ___________
    a) $|r|$ (b) $2|r|$ (c) r (d) $2r$

12. b_{xy} and b_{yx} are ___________
    a) Independent of change of origin and scale
    b) Independent of change of origin but not of scale
    c) Independent of change of scale but not of origin
    d. Affected by change of origin and scale

13. If $u = \dfrac{x-a}{c}$ and $v = \dfrac{y-b}{d}$ then $b_{yx} =$_____
    a) $\dfrac{d}{c}b_{vu}$ b) $\dfrac{c}{d}b_{vu}$
    c) $\dfrac{a}{b}b_{vu}$ d) $\dfrac{b}{a}b_{vu}$

14. If $u = \dfrac{x-a}{c}$ and $v = \dfrac{y-b}{d}$ then $b_{xy} =$_____
    a) $\dfrac{d}{c}b_{uv}$ b) $\dfrac{c}{d}b_{uv}$
    c) $\dfrac{a}{b}b_{uv}$ d) $\dfrac{b}{a}b_{uv}$

15. Corr $(x,x) =$ ___________
    a) 0 (b) 1 (c) -1 (d) can't be found

16. Corr $(x,y) =$ ___________
    a) corr(x,x) (b) corr(y,y)
    c) corr(y,x) (d) cov(y,x)

17. Corr $\left(\dfrac{x-a}{c}, \dfrac{y-b}{d}\right) = -\,corr(x,y)$ if,
    a) c and d are opposite in sign
    b) c and d are same in sign
    c) a and b are opposite in sign
    d) a and b are same in sign

18. Regression equation of X on Y is ____
    a) $y - \bar{y} = b_{yx}(x - \bar{x})$
    b) $x - \bar{x} = b_{xy}(y - \bar{y})$
    c) $y - \bar{y} = b_{xy}(x - \bar{x})$
    d) $x - \bar{x} = b_{yx}(y - \bar{y})$

19. Regression equation of Y on X is ______

a) $y - \bar{y} = b_{yx}(x - \bar{x})$

b) $x - \bar{x} = b_{xy}(y - \bar{y})$

c) $y - \bar{y} = b_{xy}(x - \bar{x})$

d) $x - \bar{x} = b_{yx}(y - \bar{y})$

20. b_{yx} = __________

a) $r\dfrac{\sigma_x}{\sigma_y}$ b) $r\dfrac{\sigma_y}{\sigma_x}$

c) $\dfrac{1}{r}\dfrac{\sigma_y}{\sigma_x}$ d) $\dfrac{1}{r}\dfrac{\sigma_x}{\sigma_y}$

21. b_{xy} = __________

a) $r\dfrac{\sigma_x}{\sigma_y}$ b) $r\dfrac{\sigma_y}{\sigma_x}$

c) $\dfrac{1}{r}\dfrac{\sigma_y}{\sigma_x}$ d) $\dfrac{1}{r}\dfrac{\sigma_x}{\sigma_y}$

22. Cov (x,y) = __________

a) $\sum\left(x-\bar{x}\right)\left(y-\bar{y}\right)$

b) $\dfrac{\sum\left(x-\bar{x}\right)\left(y-\bar{y}\right)}{n}$

c) $\dfrac{\sum xy}{n} - \bar{x}\bar{y}$

d) b and c both

23. If $b_{xy} < 0$ and $b_{yx} < 0$ then 'r' is __________

a) > 0 (b) < 0 (c) >1 (d) not found

24. If equations of regression lines are $3x + 2y - 26 = 0$ and $6x + y - 31 = 0$ then means of x and y are __________

a) (7,4) b) (4,7) c) (2,9) d) (−4,7)

II) Fill in the blanks :

1. If $b_{xy} < 0$ and $b_{yx} < 0$ then 'r' is __________

2. Regression equation of Y on X is__________

3. Regression equation of X on Y is __________

4. There are __________ types of regression equations.

5. Corr $(x, -x)$ = __________

6. If $u = \dfrac{x-a}{c}$ and $v = \dfrac{y-b}{d}$ then b_{xy} =______

7. If $u = \dfrac{x-a}{c}$ and $v = \dfrac{y-b}{d}$ then b_{yx} =______

8. $|\, b_{xy} + b_{yx}| \geq$ __________

9. If $b_{yx} > 1$ then b_{xy} is __________

10. $b_{xy}.b_{yx}$ = __________

III) State whether each of the following is True or False.

1. Corr (x,x) = 1

2. Regression equation of X on Y is $y - \bar{y} = b_{yx}(x - \bar{x})$

3. Regression equation of Y on X is $y - \bar{y} = b_{yx}(x - \bar{x})$

4. Corr (x,y) = Corr (y,x)

5. b_{xy} and b_{yx} are independent of change of origin and scale.

6. 'r' is regression coefficient of Y on X

7. b_{yx} is correlation coefficient between X and Y

8. If $u = x - a$ and $v = y - b$ then $b_{xy} = b_{uv}$

9. If $u = x - a$ and $v = y - b$ then $r_{xy} = r_{uv}$

10. In the regression equation of Y on X, b_{yx} represents slope of the line.

IV) Solve the following problems.

1. The data obtained on X, the length of time in weeks that a promotional project has been in progress at a small business, and Y, the percentage increase in weekly sales over the period just prior to the beginning of the campaign.

X	1	2	3	4	1	3	1	2	3	4	2	4
Y	10	10	18	20	11	15	12	15	17	19	13	16

Find the equation of regression line to predict the percentage increase in sales if the campaign has been in progress for 1.5 weeks.

2. The regression equation of y on x is given by $3x + 2y - 26 = 0$ Find b_{yx}.

3. If for a bivariate data $\bar{x} = 10$, $\bar{y} = 12$, v(x) = 9, $\sigma_y = 4$ and $r = 0.6$. Estimate y when $x = 5$.

4. The equation of the line of regression of y on x is $y = \frac{2}{9}x$ and x on y is $x = \frac{y}{2} + \frac{7}{6}$. Find

   (i) r (ii) σ_y^2 if $\sigma_x^2 = 4$.

5. Identify the regression equations of x on y and y on x from the following equations,

   $2x + 3y = 6$ and $5x + 7y - 12 = 0$

6. (i) If for a bivariate data $b_{yx} = -1.2$ and $b_{xy} = -0.3$ then find r.

   (ii) From the two regression equations $y = 4x - 5$ and $3x = 2y + 5$, find $\bar{x}$ and $\bar{y}$.

7. The equations of the two lines of regression are $3x + 2y - 26 = 0$ and $6x + y - 31 = 0$ Find

   (i) Means of X and Y

   (ii) Correlation coefficient between X and Y

   (iii) Estimate of Y for X = 2

   (iv) var (X) if var (Y) = 36

8. Find the line of regression of X on Y for the following data:

$$n = 8,\ \sum\left(x_i - \bar{x}\right)^2 = 36,\ \sum\left(y_i - \bar{y}\right)^2 = 44$$

$$\sum\left(x_i - \bar{x}\right)\left(y_i - \bar{y}\right) = 24$$

9. Find the equation of line of regression of Y on X for the following data:

$$n = 8,\ \sum\left(x_i - \bar{x}\right)\left(y_i - \bar{y}\right) = 120,$$

$$\bar{x} = 20,\ \bar{y} = 36,\ \sigma_x = 2,\ \sigma_y = 3.$$

10. The following results were obtained from records of age (X) and systolic blood pressure (Y) of a group of 10 men.

	X	*Y*
Mean	50	140
Variance	150	165

and $\sum\left(x_i - \bar{x}\right)\left(y_i - \bar{y}\right) = 1120$

Find the prediction of blood pressure of a man of age 40 years.

11. The equations of two regression lines are $10x - 4y = 80$ and $10y - 9x = -40$

    Find:

    (i) $\bar{x}$ and $\bar{y}$

    (ii) b_{YX} and b_{XY}

    (iii) If var (Y) = 36, obtain var (X)

    (iv) r

12. If $b_{YX} = -0.6$ and $b_{XY} = -0.216$ then find correlation coefficient between X and Y. Comment on it.

Activities

1) Consider a group of 70 students of your class to take their heights in cm (x) and weights kg (y). Hence find both the regression equations.

2) The age in years of 7 young couples is given below:

Husband (x)	21	25	26	24	22	30	20
Wife (y)	19	20	24	20	22	24	18

i) Find the equation of regression line of age of husband on age of wife.

ii) Draw the regression line of y on x

iii) Predict the age of wife whose husband's age is 27 years.

3) The equations of two regression lines are

$10x - 4y = 80$ .... (1)

$10y - 9x = -40$ .... (2)

$\therefore$ $(\bar{x}, \bar{y})$ is the point of intersection of both the regression lines.

$\therefore$ Solve equations (i) and (ii), we get

$\bar{x} = \boxed{}$ and $\bar{y} = \boxed{}$

Now, consider $10x - 4y = 80$

$\therefore\ a =, \quad b =$

$\therefore$ slope$(m_1) = -\dfrac{a}{b} = \dfrac{\boxed{}}{\boxed{}}$

Consider, $10y - 9x = -40$

$\therefore\ a = \quad , b =$

$\therefore$ slope$(m_2) = -\dfrac{a}{b} = \dfrac{\boxed{}}{\boxed{}}$

$\therefore |m_1| > |m_2|$

$\therefore\ b_{yx} = \boxed{}$ and $b_{xy} = \dfrac{1}{\boxed{}}$

$\therefore$ $10x - 4y = 80$ is the regression equations of $\boxed{}$ on $\boxed{}$ and

$\therefore$ $10y - 9x = -40$ is the regression equations of $\boxed{}$ on $\boxed{}$.

Now, $r = \pm\sqrt{b_{yx}.b_{xy}}$

$\therefore\ r = \pm\sqrt{\dfrac{\boxed{}}{\boxed{}} \times \dfrac{4}{10}}$

$r = \boxed{}$

If $V(y) = 36$ then $\sigma_y = \sqrt{\boxed{}} = \boxed{}$

$\therefore\ b_{xy} = r\dfrac{\sigma_y}{\sigma_x}$

$\therefore\ \dfrac{\boxed{}}{\boxed{}} = \boxed{} \times \dfrac{\boxed{}}{\sigma_x}$

$\therefore\ \sigma_x = (\quad)$

$\therefore\ V(x) = \sigma_x^2 = \boxed{}$

4) Given $n = 8$, $\sum\left(x_i - \bar{x}\right)^2 = 36$,

$\sum\left(y_i - \bar{y}\right)^2 = 40$,

$\sum\left(x_i - \bar{x}\right)\left(y_i - \bar{y}\right) = 24$

$$\therefore\ b_{yx} = \frac{\sum\left(x_i - \bar{x}\right)\left(y_i - \bar{y}\right)}{\sum\left(x_i - \bar{x}\right)^2}$$

$$= \frac{\boxed{}}{\boxed{}} = \frac{\boxed{}}{\boxed{}}$$

$$\therefore\ b_{xy} = \frac{\sum\left(x_i - \bar{x}\right)\left(y_i - \bar{y}\right)}{\sum\left(y_i - \bar{y}\right)^2}$$

$$= \frac{\boxed{}}{\boxed{}} = \frac{\boxed{}}{\boxed{}}$$

$\therefore$ Regression equation of Y on X :

$y - \bar{y} = b_{yx}(x - \bar{x})$

$y - \boxed{} = \dfrac{\boxed{}}{\boxed{}}(x - \boxed{})$

$\therefore$ Regression equation of X on Y :

$x - \bar{x} = b_{xy}(y - \bar{y})$

$x - \boxed{} = \dfrac{\boxed{}}{\boxed{}}(y - \boxed{})$

5) Consider, given

$$n = 8 \; \sum\left(x_i - \bar{x}\right)\left(y_i - \bar{y}\right) = 120,$$

$$\bar{y} = 36, \; \sigma_x = 2, \; \sigma y = 3$$

$$\therefore \; \operatorname{cov}(x, y) = \frac{\sum\left(x - \bar{x}\right)\left(y - \bar{y}\right)}{n}$$

$$= \frac{\boxed{}}{\boxed{}} = \boxed{}$$

$$\therefore \; b_{yx} = \frac{\boxed{}}{\sigma_x^4} = \frac{150}{\boxed{}}$$

$$b_{xy} = \frac{\operatorname{cov}(x, y)}{\boxed{}}$$

$$= \frac{\boxed{}}{9} = \frac{\boxed{}}{\boxed{}}$$

∴ Regression equation of Y on X :

$$y - \boxed{} = b_{yx}(x - \boxed{})$$

$$y - \boxed{} = \boxed{}(x - \boxed{})$$

∴ Regression equation of Y on X :

$$x - \bar{x} = \boxed{}(y - \bar{y})$$

$$x - \boxed{} = \boxed{}(y - \boxed{})$$

◆ ◆ ◆

4 Time Series

Let's Study

- Uses of time series analysis.
- Components of a time series.
    - Secular Trend
    - Seasonal Variation
    - Cyclical Variation
    - Irregular Variation
- Mathematical Models
    - Additive Model
    - Multiplicative Model
- Measurement of Secular Trend
    - Graphical Method
    - Method of Moving Averages
    - Method of Least Squares

Introduction

A manufacturing company wants to predict demand for its product for next year to make a production plan. An investor wants to know fluctuations in share prices so that he can decide if he should purchase or sell certain shares. These and many other situations involve a variable that changes with time. A variable observed over a period of time is called a time series. Analysis of time series is useful in understanding the patterns of changes in the variable over time. Let us now define a time series.

Definition

Time Series is a sequence of observations made on a variable at regular time intervals over a specified period of time.

Data collected arbitrarily or irregularly does not form a time series. Time series analysis involves the use of statistical methods to analyze time series data in order to extract meaningful statistics and understand important characteristics of the observed data.

Time Series Analysis helps us understand the underlying forces leading to a particular pattern in the time series and helps us in monitoring and forecasting data with help of appropriate statistical models.

Analysis of time series data requires maintaining records of values of the variable over time.

Some examples from day-to-day life may give a better idea of time series.

1. Monthly, quarterly, or yearly production of an industrial product.
2. Yearly GDP (Gross Domestic Product) of a country.
3. Monthly sales in a departmental store.
4. Weekly prices of vegetables.
5. Daily closing price of a share at a stock exchange.
6. Hourly temperature of a city recorded by the Meteorological Department.

4.1 Uses of Time Series Analysis

The main objective of time series analysis is to understand, interpret and assess chronological changes in values of a variable in the past, so that reliable predictions can be made about its future values. For example, the government may be interested in predicting population growth in near future for planning its welfare schemes, the agricultural ministry may be interested in predicting annual crop yield before declaring the MSP (minimum support price) of agricultural produce or an industrialist may be interested in predicting the weekly demand for his product for making the production schedule. Following are considered to be some of the important uses of time series ananlysis.

1. It is useful for studying the past behaviour of a variable.

   In a time series, the past observations on a variable are arranged in an orderly manner over a period of time. By simple observation of such a series, one can understand the nature of changes that have taken place in values of the variable during the course of time. Further, by applying appropriate technique of analysis to the series, one can study the general tendency of the variable in addition to seasonal changes, cyclical changes, and irregular or accidental changes in values of the variable.

2. It is useful for forecasting future behaviour of a variable.

   Analysis of a time series reveals the nature of changes in the value of a variable during the course of times. This can be useful in forecasting the future values of the variable. Thus, with the help of observations on an appropriate time series, future plans can be made relating to certain matters like purchase, production, sales, etc. This is how a planned economy makes plans for the future development on the basis of time series analysis of the relevant data.

3. It is useful in evaluating the performance.

   Evaluation of the actual performances in comparison with predetermined targets is necessary to judge efficiency of the work. For example, the achievements of Five-Year Plans are evaluated by determining the annual rate of growth in the gross national product. Similarly, the national policy of controlling inflation and price rises is evaluated with the help of different price indices. All these are made possible by analysis of time series of the relevant variables.

4. It is useful in making a comparative study.

   A comparative study of data relating to two or more periods, regions, or industries reveals a lot of valuable information that can guide management in taking a proper course of action. A time series itself provides a scientific basis for making comparisons between two or more related sets of data. Note that data are arranged chronologically in such a series, and the effects of its various components are gradually isolated, analyzed, and interpreted.

4.2 Components of Time Series

A graphical representation of time series data shows continuous changes in its values over time, giving an impression of fluctuating nature of data. A close look of the graph, however, reveals that the fluctuations are not totally arbitrary, and a part of these fluctuations has a steady behavior and can be related to time. This part is the systematic part of the time series and the remaining part is non systematic or irregular. The systematic part is further divided in the following broad categories: (i) secular trend (T), (ii) seasonal variation (S), and (iii) cyclical variation (C). The non systematic part is also called (iv) irregular variation (I). Every time series has some or all of these components. Of course, only the systematic components of a time series are useful in forecasting its future values.

We now discuss the four components of a time series in detail.

4.2.1. Secular Trend (T)

The secular trend is the long term pattern of a time series. The secular trend can be positive or negative depending on whether the time series exhibits an increasing long term pattern or a decreasing long term pattern. The secular trend shows a smooth and regular long term movement of the time series. The secular trend does not include short term fluctuations, but only consists of a steady movement over a long period of time. It is the movement that the series would take if there are no seasonal, cyclical or irregular variations. It is the effect of factors that are more or less constant for a long time or that change very gradually and slowly over time.

If a time series does not show an increasing or decreasing pattern, then the series is stationary around the mean.

SOLVE EXAMPLES

1. The following table shows annual sales (in lakh Rs.) of a departmental store for years 2011 to 2018.

Year	2011	2012	2013	2014
Sales	26.2	28.9	33.7	32.1
Year	2015	2016	2017	2018
Sales	39.8	38.7	45.4	42.6

The following graph shows the above time series.

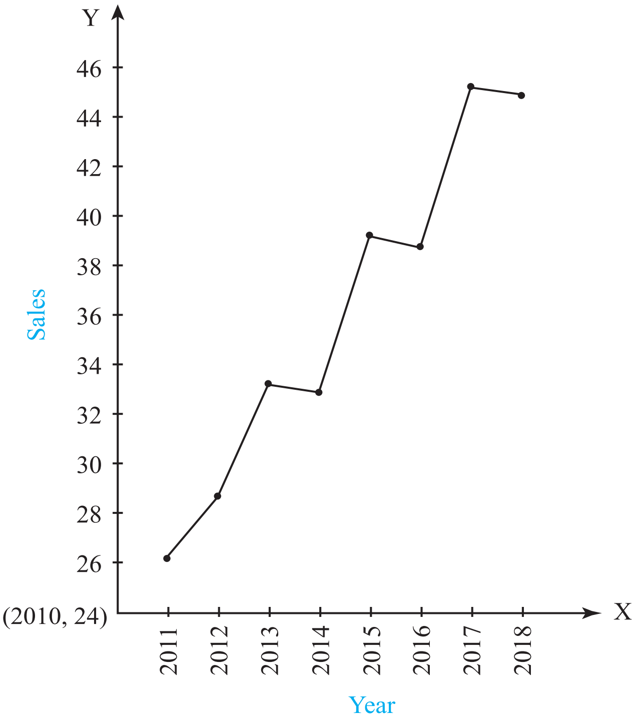


Fig. 4.1

There are ups and downs in the graph, but the time series shows an **upward trend** in long run.

2. The following table shows production (in '000 tonnes) of a commodity during years 2001-2008.

Year	2001	2002	2003	2004
Production	50.0	36.5	43.0	44.5
Year	2005	2006	2007	2008
Production	38.9	38.1	32.6	33.7

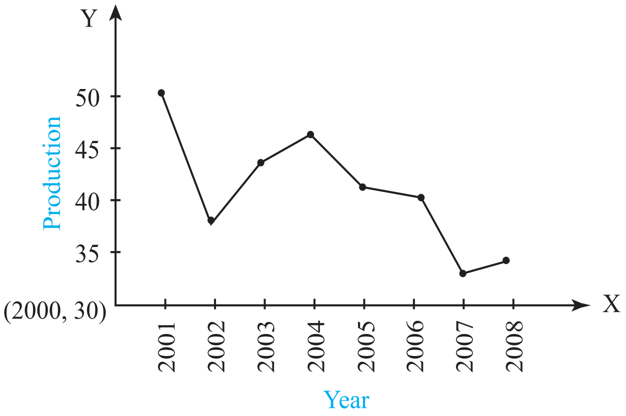


Fig. 4.2

The above graph shows a **downward trend.**

4.2.2 Seasonal Variation (S)

Many time series related to financial, economic, and business activities consist of monthly or quarterly data. It is observed very often that these time series exhibit seasonal variation in the sense that similar patterns are repeated from year to year. Seasonal variation is the component of a time series that involves patterns of change within a year that repeat from year to year.

Several commodities show seasonal fluctuations in their demand. Warm clothes and woolen products have a market during the winter season. Fans, coolers, cold drinks and ice creams are in great demand during summer. Umbrellas and raincoats are in great demand during the rainy season. Different festivals are associated with different commodities and every festival season is associated with an increase in demand for related commodities. For example, clothes and firecrackers are in great demand during Diwali. Most of the seasonal variations in demand reflect changes in climatic conditions or customs and habits of people.

All the above examples have one year as the period of seasonal variation. However, the period of seasonal variation can be a month, a week, a day, or even an hour, depending on the nature of available data. For example, cash withdrawals in a bank show seasonal variation among the days of a month, the number of books borrowed by readers from a library show seasonal variation according to days of a week, passenger traffic at a railway station has seasonal variation during hours of a day, and the temperature recorded in a city exhibits seasonal variation over hours of a day, in addition to seasonal variation with changing seasons in a year.

Seasonal variation is measured with help of seasonal indices, which are useful for short term forecasting. Such short term forecasts are useful for a departmental store in planning its inventory according to months of a year. A bank manager can use such short term forecasts in managing cash flow on different days of a week or a month.

3. The following table shows quarterly sales (in lakh Rs.) of woolen garments in four consecutive years.

Year	I				II			
Quarter	1	2	3	4	1	2	3	4
Sales	11	8	16	28	19	17	32	38
Year	III				IV			
Quarter	1	2	3	4	1	2	3	4
Sales	33	23	39	52	41	37	44	58

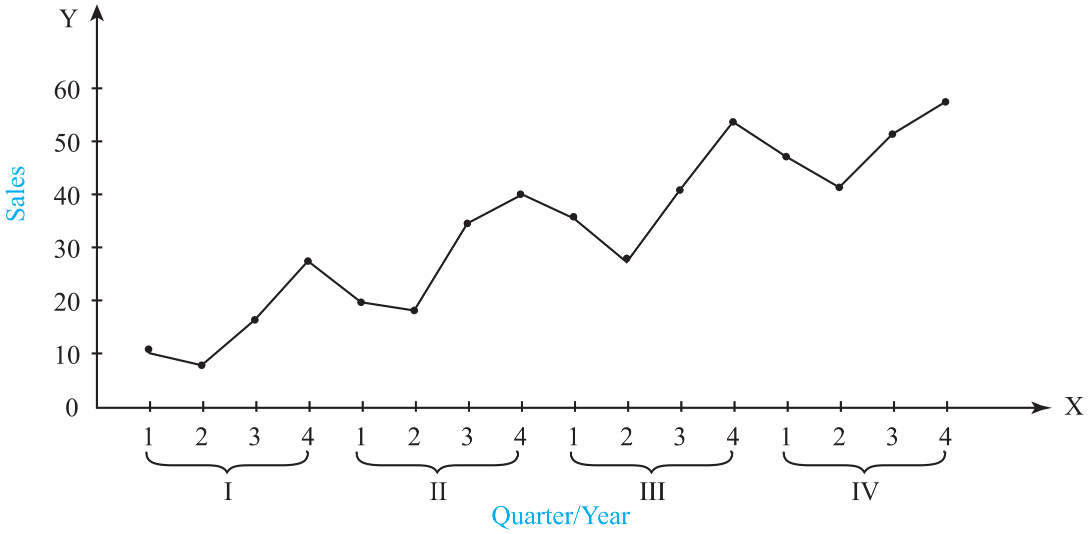


Fig. 4.3

Figure 4.3 shows a pattern that is repeated year after year. The values are lowest (in the year) in second quarter and highest (in the year) in fourth quarter of every year. Although the overall graph of the time series in Figure 4.3 shows an increasing trend, the seasonal variation within every year is very clearly visible in the graph.

4.2.3 Cyclical Variation (C)

Cyclical variation is a long term oscillatory movement in values of a time series. Cyclical variation occurs over a long period, usually several years, if seasonal variation occurs within a year. One complete round of oscillation is called a cycle. Cyclical variations need not be periodic in the sense that the length of a cycle or the magnitude of variation within a cycle can change from one cycle to another.

Cyclical variations are observed in almost all time series related to economic or business activities, where a cycle is known as a business cycle or trade cycle. Recurring ups and downs in a business are the main causes of cyclical variation.

A typical business cycle consists of the following four phases: (i) prosperity, (ii) recession, (iii) depression, (iv) recovery. Figure 4.4 depicts these four phases of a business cycle, where every phase changes to the next phase gradually in the order mentioned above.

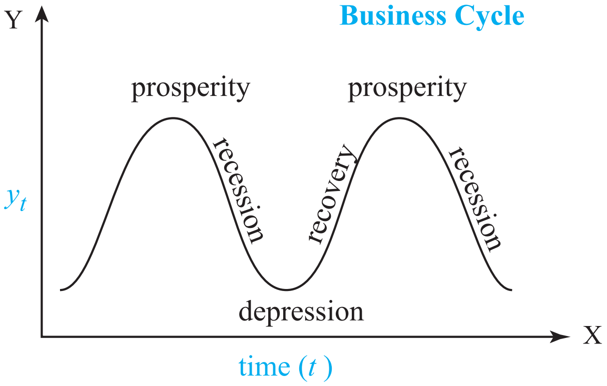


Fig. 4.4

Cyclical variations can consist of a period of 5 years, 10 years, or even longer duration. The period often changes from one cycle to another. Cyclical variation may be attributed to internal organizational factors such as purchase and inventory policies or external factors such as financial market conditions and government policies.

4.2.4 Irregular Variation (I)

Irregular variations are unexpected variations in time series caused by unforeseen events that can include natural disasters like floods or famines, political events line strikes or agitations, or international events like wars or others conflicts. As the name suggests, irregular variations do not follow any patterns and are, therefore, totally unpredictable. For this reason, irregular variation are also known as unexplained or unaccounted variations.

4.3 Mathematical Models of Time Series

Let X_t denote the value of the variable at time t. The time series is denoted by the collection of values, $\{X_t, t = 0, 1, \ldots, T\}$ where T is the total duration of observation. There are two standard mathematical models for time series based on the four components mentioned earlier, namely, secular trend (T), seasonal variation (S), cyclical variation (C), and irregular variation (I).

4.3.1 Additive Model

The additive model assumes that the value X_t at time t is the sum of the four components at time t. Thus,

$$X_t = T_t + S_t + C_t + I_t$$

The additive model assumes that the four components of the time series are independent of one another. It is also important to remember that all the four components in the additive model must be measured in the same unit of measurement. The magnitude of the seasonal variation does not depend on the value of the time series in the additive model. In other words, the magnitude of the seasonal variation does no change as the series goes up or down.

The assumption of independence of the components is often not realistic. In such situations, the multiplicative model can be used.

4.3.2 Multiplicative Model

The multiplicative model that the value X_t at the time t is obtained by multiplication of the four components at time t. That is,

$$X_t = T_t \times S_t \times C_t \times I_t$$

The multiplicative model does not assume independence of the four components of the series and is, therefore, more realistic. Values of the trend are expressed in units of measurements and other components are expressed as percentage or relative values, and hence are free from units of measurements.

It is recommended to choose the multiplicative model when the magnitude of the seasonal variation in the data depends on the magnitude of the data. In other words, the magnitude of the seasonal variation increases as the data values increase, and decreases as the data values decrease.

4.4 Measurement of Secular Trend

4.4.1 Method of Freehand Curve (Graphical Method)

In this method, a graph is drawn for the given time series by plotting X_t (on Y-axis) against t (on X-axis). Then a free hand smooth curve is plotted on the same graph to indicate the general trend.

This method is simple and does not require any mathematical calculation. But, in this method, different researchers may draw different trend lines for the same set of data. Forecasting using this method is therefore risky if the person drawing free hand curve is not efficient and experienced. On the other hand, this method is quite flexible and can be used for all types of trends, linear as well as non – linear, and involves minimum amount of work.

SOLVED EXAMPLES

1. Fit a trend line to the following data using the graphical method.

Year	Number of crimes ('000)	Year	Number of crimes ('000)
1981	40	1987	43
1982	42	1988	46
1983	43	1989	47
1984	42	1990	45
1985	44	1991	46
1986	44		

Solution:

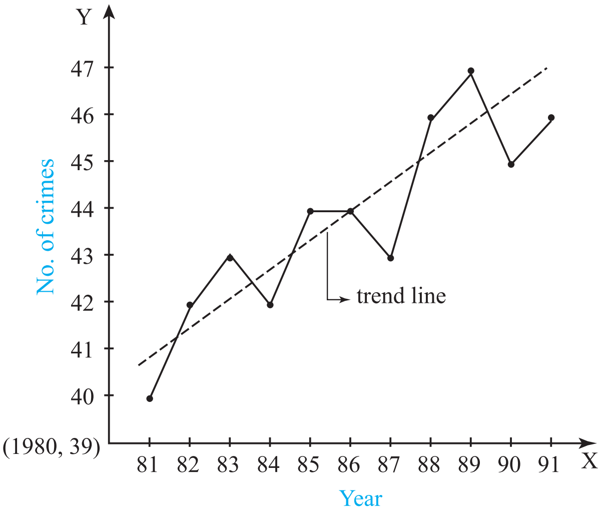


Fig. 4.5

2. The publisher of a magazine wants to determine the rate of increase in the number of subscribers. The following table shows the subscription information for eight consecutive years.

Year	1976	1977	1978	1979
No. of subscribers (in millions)	12	11	19	17
Year	1980	1981	1982	1983
No. of subscribers (in millions)	19	18	20	23

Fit a trend line by the graphical method.

Solution:

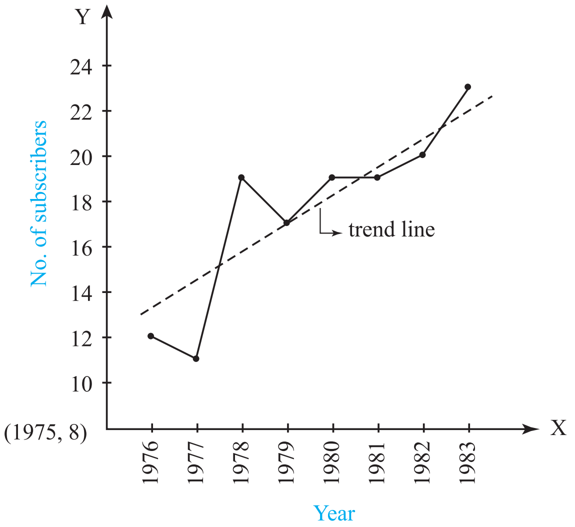


Fig. 4.6

4.4.2 Method of moving Averages

The moving average of period k of a time series forms a time series of arithmetic means of k successive observations from the original time series. The method begins with the first k observations and finds the arithmetic mean of these k observations. The next step leaves the first observation and includes observation number $k + 1$ and finds the arithmetic mean of these k observations. This process continues till the average of the last k observations is found. In other words, the method of moving averages finds the following.

First moving average $= \dfrac{X_1 + X_2 + \ldots + X_k}{k}$

Second moving average $= \dfrac{X_2 + X_3 + \ldots + X_{k+1}}{k}$

Third moving average $= \dfrac{X_3 + X_4 + \ldots + X_{k+2}}{k}$

and so on.

Each of these averages is written against the time point that is the middle term in the sum. As a result, when k is an odd integer, moving average values correspond to observed values of the given time series. On the other hand, when k is an even integer, the moving averages fall mid-way between two observed values of the given time series. In this case, a subsequent two-unit moving average is calculated to make the resulting moving average values correspond to observed values of the given time series.

A moving average with an appropriate period smooths out cyclical variations from the given time series and provides a good estimate of the trend. Cyclical fluctuations with a uniform period and a uniform amplitude can be completely eliminated by taking the period of moving averages that is equal to or a multiple of the period of the cycles as long as the trend is linear.

The method of moving averages is flexible in the sense that even if a few observations are added to the given series, the moving averages calculated earlier are not affected and remain unchanged. However, the method of moving averages does not provide a mathematical equation for the time series and hence cannot be used for the purpose of forecasting. Another drawback of the method of moving averages is that some of the trend values at each end of the given series cannot be estimated by this method.

SOLVED EXAMPLES

3. The following table shows gross capital formation (in crore Rs) for years 1966 to 1975.

Year	1966	1967	1968	1969	1970
Gross Capital Formation	19.3	20.9	17.8	16.1	17.6
Year	1971	1972	1973	1974	1975
Gross Capital Formation	17.8	18.3	17.3	21.4	19.3

(i) Obtain trend values using 5-yearly moving averages.

(ii) Plot the original time series and trend values obtained in (i) on the same graph.

Solution :

Table 4.3

Year	X_t	5-yearly moving total	5-yearly moving averages (trend value)
1966	19.3	-	-
1967	20.9	-	-
1968	17.8	91.7	18.34
1969	16.1	90.2	18.04
1970	17.6	87.6	17.52
1971	17.8	87.1	17.42
1972	18.3	92.4	18.48
1973	17.3	94.1	18.82
1974	21.4	-	-
1975	19.3	-	-

Note that 5-yearly average is not available for the first 2 years and last 2 years.

The following graph shows the original time series and the trend values obtained in Table 4.3.

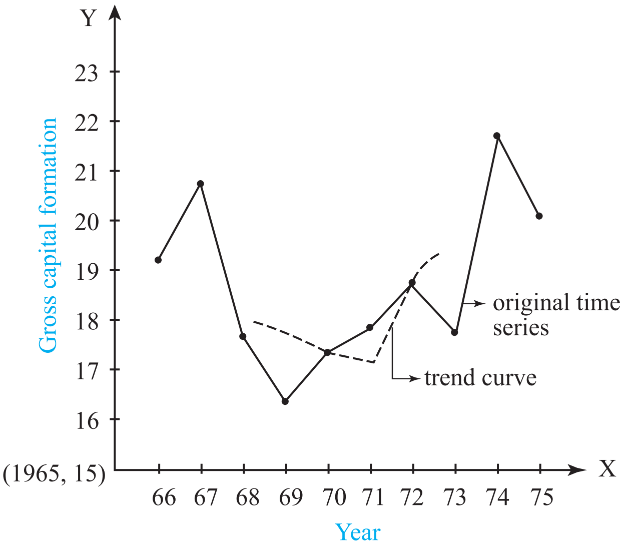


Fig. 4.7

4. Obtain 4-yearly centered moving averages for the following time series.

Year	1987	1988	1989	1990	1991
Annual sales (in lakh Rs.)	3.6	4.3	4.3	3.4	4.4
Year	1992	1993	1994	1995	
Annual sales (in lakh Rs.)	5.2	3.8	4.9	5.4	

Solution:

Year	X_i	4-yearly moving total	4-yearly moving average	2-unit moving total	4-yearly centred moving average (trend value)
1987	3.6				
1988	4.3				
		15.6	3.9		
1989	4.3			8	4
		16.4	4.1		
1990	3.4			8.425	4.2125
		17.3	4.325		
1991	4.4			8.525	4.2625
		16.8	4.2		
1992	5.2			8.775	4.3875
		18.3	4.575		
1993	3.8			9.4	4.7
		19.3	4.825		
1994	4.9				
1995	5.4				

Note: Entries in the third and fourth columns are *between* tabulated time periods, while entries in fifth and sixth columns are *in front of tabulated* time periods.

4.4.3 Method of Least Squares

This is the most objective and perhaps the best method of determining trend in a given time series. The method begins with selection of an appropriate form of trend equation and then proceeds with estimation of the unknown constants in this equation. It is a common practice to choose a polynomial of a suitable degree and then to determine its unknown (but constant) coefficients by the method of least squares. The choice of the degree of polynomial is often based on the graphical representation of the given data.

In a linear trend, the equation is given by

$$X_t = a + bt$$

The method of least squares involves solving the following set of linear equations, commonly known as normal equations.

$$\sum X_t = na + b\sum t$$

$$\sum tX_t = a\sum t + b\sum t^2$$

Where n is the number of the time periods for which data is available, whereas,

$\sum X_t$, $\sum tX_t$, $\sum t$ and $\sum t^2$ are obtained

from the data. The least squares estimates of a and b are obtained by solving the two equations in the two unknowns, namely a and b. The required equation of the trend line is the obtained by substituting these estimates in equation $X_t = a + bt$

SOLVED EXAMPLES

5. Fit a trend line by the method of least squares to the time series in Example 1 (Section 4.4.1). Also, obtain the trend value for the number of crimes in the year 1993.

Solution:

Let the trend line be represented by the equation $y_t = a + bt$ Calculations can be made simpler by transforming from t to u using the formula

$$u = \frac{t - \text{middle } t \text{ value}}{h} \quad \text{......... if } n \text{ is odd}$$

$$= \frac{t - \text{mean of two middle } t \text{ values}}{h} \quad \text{......... if } n \text{ is even}$$

Where h is the difference between successive t values.

In the given problem, $n = 11$ (odd), middle t value is 1986, and $h = 1$.

$\therefore$ we use the transformation

$$u = \frac{t - 1986}{1} = t - 1986.$$

The equation of the trend line then becomes

$$y_t = a' + b'u$$

The two normal equations are then given by

$$\sum y_t = na' + b' \sum u \, ,$$

$$\sum uy_t = a' \sum u + b' \sum u^2$$

We obtain $\sum u$, $\sum u^2$, $\sum y_t$ and $\sum uy_t$ from the following table.

Year t	y_t	u	u^2	uy_t	Trend Value
1981	40	−5	25	−200	41.0457
1982	42	−4	16	−168	41.6002
1983	43	−3	9	−129	42.1547
1984	42	−2	4	−84	42.7092
1985	44	−1	1	−44	43.2637
1986	44	0	0	0	43.8182
1987	43	1	1	43	44.3727
1988	46	2	4	92	44.9272
1989	47	3	9	141	45.4817
1990	45	4	16	180	46.0362
1991	46	5	25	230	46.5907
Total	482	0	110	61	

The normal equations are

$$482 = 11a' + b'(0)$$

$$61 = a'(0)\ 110\ b'$$

$$\therefore \quad a' = \frac{482}{11} = 43.8182$$

$$b' = \frac{61}{110} = 0.5545$$

The equation of the trend line is then given by

$$y_t = 43.8182 + (0.5545)\, u,$$

where $u = t - 1986$

Trend values computed using this result are given in the last column of the table.

6. Fit a trend line by the method of least squares to the time series in Example 2 (Section 4.4.1). Also obtain the trend value for the number of subscribers in the year 1984.

Solution :

Let the equation of the trend line, if the n is even

$$y_t = a + bt$$

In the given problem, $n = 8$ (even), two middle t – values are 1979 and 1980, and $h = 1$.

We use the transformation

$$u = \frac{t - 1979.5}{\frac{1}{2}}$$

$$= 2t - 3959.$$

The equation of the trend line then becomes

$$y_t = a' + b'u$$

The two normal equations are

$$\sum y_t = na' + b' \sum u \, , \qquad (1)$$

$$\sum uy_t = a' \sum u + b' \sum u^2 \qquad (2)$$

We obtain $\sum u$, $\sum u^2$, $\sum y_t$ and $\sum uy_t$

from the following table.

Table 4.2

Year t	y_t	u	u^2	uy_t	Trend Value
1976	12	−7	49	−84	12.3336
1977	11	−5	25	−55	13.774
1978	19	−3	9	−57	15.2144
1979	17	−1	1	−17	16.6548
1980	19	1	1	19	18.0952
1981	18	3	9	54	19.5356
1982	20	5	25	100	20.976
1983	23	7	49	161	22.4164
Total	139	0	168	121	

From (1) and (2),

$$139 = 8a' + b'(0) \qquad (3)$$

$$121 = a'(0) + b'(168) \qquad (4)$$

From (3), $a' = 17.375$.

From (4), $b' = 0.7202$.

The equation of the trend line is then given by

$$y_t = 17.375 + (0.7202)u,$$

where $u = 2t - 3959$.

Trend values computed using this result are given in the last column of Table 4.2.

Estimation of the trend value for the year 1984 is obtained as shown below.

For $t = 1984$, $u = 2(1984) - 3959 = 9$.

And hence

$$\hat{y}_t = 17.375 + (0.7202)(9)$$

$$= 23.8568 \ (\textit{in millions})$$

EXERCISE 4.1

1. The following data gives the production of bleaching powder (in '000 tonnes) for the years 1962 to 1972.

Year	1962	1963	1964	1965	1966	
Production	0	0	1	1	4	
Year	1967	1968	1969	1970	1971	1972
Production	2	4	9	7	10	8

Fit a trend line by graphical method to the above data.

2. Use the method of least squares to fit a trend line to the data in Problem 1 above. Also, obtain the trend value for the year 1975.

3. Obtain the trend line for the above data using 5 yearly moving averages.

4. The following table shows the index of industrial production for the period from 1976 to 1985, using the year 1976 as the base year.

Year	1976	1977	1978	1979	1980
Index	0	2	3	3	2
Year	1981	1982	1983	1984	1985
Index	4	5	6	7	10

Fit a trend line to the above data by graphical method.

5. Fit a trend line to the data in Problem 4 above by the method of least squares. Also, obtain the trend value for the index of industrial production for the year 1987.

6. Obtain the trend values for the data in problem 4 using 4-yearly centered moving averages.

7. The following table gives the production of steel (in millions of tonnes) for years 1976 to 1986.

Year	1976	1977	1978	1979	1980	1981
Production	0	4	4	2	6	8
Year	1982	1983	1984	1985	1986	
Production	5	9	4	10	10	

Fit a trend line to the above data by the graphical method.

8. Fit a trend line to the data in Problem 7 by the method of least squares. Also, obtain the trend value for the year 1990.

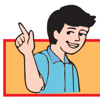
9. Obtain the trend values for the above data using 3-yearly moving averages.

10. The following table shows the production of gasoline in U.S.A. for the years 1962 to 1976.

Year	Production (million Barrels)	Year	Production (million Barrels)
1962	10	1970	6
1963	0	1971	7
1964	1	1972	8
1965	1	1973	9
1966	2	1974	8
1967	3	1975	9
1968	4	1976	10
1969	5		

(i) Obtain trend values for the above data using 5-yearly moving averages.

(ii) Plot the original time series and trend values obtained above on the same graph.



Let's Remember

- A time series is a set of observations made at regular intervals of time and therefore arranged chronologically (that is, according to time).
- A time series has the following four components:
  (i) Secular trend
  (ii) Seasonal variation
  (iii) Cyclical variation
  (iv) Irregular variation
- Secular trend of a time series is its smooth and regular long-term movement.
- Seasonal variation involves patterns of change within a year that are repeated from year to year. Seasonal variations often reflect changes in underlying climatic conditions, or customs and habits of people.
- Cyclical variation is a long-term oscillatory movement that occurs over a long period of time, mostly more than two years or more.
- A standard business cycle consists of the following four phases:
  (i) Prosperity (ii) Decline
  (iii) Depression (iv) Recovery
- Irregular variations are caused by either random factors or unforeseen events like floods, famines, earthquakes, strikes, wars, etc.
- Additive Model:

$$X_t = T_t + S_t + C_t + I_t$$

where X_t is the value of the variable X at time t and T_t, S_t, C_t and I_t are secular trend, seasonal variation, cyclical variation, and irregular variation at time t, respectively.

- Multiplicative Model:

$$X_t = T_t \times S_t \times C_t \times I_t$$

where the notation is same as in the additive model.

- Secular trend can be measured using
  (i) Graphical method
  (ii) Method of moving averages
  (iii) Method of least squares
- The moving averages of period k of a time series form a new series of arithmetic means, each of k successive observations of the given time series.
- Method of least squares involves solving the following two normal equations.

$$\sum X_t = na + b\sum t$$

$$\sum tX_t = a\sum t + b\sum t^2$$

MISCELLANEOUS EXERCISE - 4

I) Choose the correct alternative.

1. Which of the following can't be a component of a time series?

(a) Seasonality (b) Cyclical
(c) Trend (d) Mean

2. The first step in time series analysis is to
   (a) Perform regression calculations
   (b) Calculate a moving average
   (c) Plot the data on a graph
   (d) Identify seasonal variation

3. Time-series analysis is based on the assumption that
   (a) Random error terms are normally distributed.
   (b) The variable to be forecast and other independent variables are correlated.
   (c) Past patterns in the variable to be forecast will continue unchanged into the future.
   (d) The data do not exhibit a trend.

4. Moving averages are useful in identifying
   (a) Seasonal component
   (b) Irregular component
   (c) Trend component
   (d) Cyclical component

5. We can use regression line for past data to forecast future data. We then use the line which
   (a) Minimizes the sum of squared deviations of past data from the line
   (b) Minimizes the sum of deviations of past data from the line.
   (c) Maximizes the sum of squared deviations of past data from the line
   (d) Maximizes the sum of deviations of past data from the line.

6. Which of the following is a major problem for forecasting, especially when using the method of least squares?
   (a) The past cannot be known
   (b) The future is not entirely certain
   (c) The future exactly follows the patterns of the past
   (d) The future may not follow the patterns of the past

7. An overall upward or downward pattern in an annual time series would be contained in which component of the times series
   (a) Trend (b) Cyclical
   (c) Irregular (d) Seasonal

8. The following trend line equation was developed for annual sales from 1984 to 1990 with 1984 as base or zero year.
   $Y1 = 500 + 60X$ (in 1000 Rs). The estimated sales for 1984 (in 1000 Rs) is:
   (a) Rs 500 (b) Rs 560
   (c) Rs 1,040 (d) Rs 1,100

9. What is a disadvantage of the graphical method of determining a trend line?
   (a) Provides quick approximations
   (b) Is subject to human error
   (c) Provides accurate forecasts
   (d) Is too difficult to calculate

10. Which component of time series refers to erratic time series movements that follow no recognizable or regular pattern.
    (a) Trend (b) Seasonal
    (c) Cyclical (d) Irregular

II) Fill in the blanks

1. ___ components of time series is indicated by a smooth line.
2. ___ component of time series is indicated by periodic variation year after year.
3. ___ component of time series is indicated by a long wave spanning two or more years.
4. ___ component of time series is indicated by up and down movements without any pattern.
5. Addictive models of time series ___ independence of its components.
6. Multiplicative models of time series ___ independence of its components.
7. The simplest method of measuring trend of time series is ____
8. The method of measuring trend of time series using only averages is ___

9. The complicated but efficient method of measuring trend of time series is _____.

10. The graph of time series clearly shows _____ of it is monotone.

III) State whether each of the following is True or False.

1. The secular trend component of time series represents irregular variations.
2. Seasonal variation can be observed over several years.
3. Cyclical variation can occur several times in a year.
4. Irregular variation is not a random component of time series.
5. Additive model of time series does not require the assumption of independence of its components.
6. Multiplicative model of time series does not require the assumption of independence of its components.
7. Graphical method of finding trend is very complicated and involves several calculations.
8. Moving average method of finding trend is very complicated and involves several calculations.
9. Least squares method of finding trend is very simple and does not involve any calculations.
10. All the three methods of measuring trend will always give the same results.

IV) Solve the following problems.

1. The following table shows the production of pig-iron and ferro-alloys ('000 metric tonnes)

Year	1974	1975	1976	1977	1978
Production	0	4	9	9	8
Year	1979	1980	1981	1982	
Production	5	4	8	10	

Fit a trend line to the above data by graphical method.

2. Fit a trend line to the data in Problem IV (1) by the method of least squares.
3. Obtain trend values for data in Problem IV (1) using 5-yearly moving averages.
4. Following table shows the amount of sugar production (in lac tonnes) for the years 1971 to 1982.

Year	Production	Year	Production
1971	1	1977	3
1972	0	1978	6
1973	1	1979	5
1974	2	1980	1
1975	3	1981	4
1976	2	1982	10

Fit a trend line to the above data by graphical method.

5. Fit a trend line to data in Problem 4 by the method of least squares.
6. Obtain trend values for data in Problem 4 using 4-yearly centered moving averages.
7. The percentage of girls' enrollment in total enrollment for years 1960-2005 is shown in the following table.

Year	1960	1965	1970	1975	1980
Production	0	3	3	4	4
Year	1985	1990	1995	2000	2005
Production	5	6	8	8	10

Fit a trend line to the above data by graphical method.

8. Fit a trend line to the data in Problem 7 by the method of least squares.
9. Obtain trend values for the data in Problem 7 using 4-yearly moving averages.
10. Following data shows the number of boxes of cereal sold in years 1977 to 1984.

Year	1977	1978	1979	1980
No. of boxes in ten thousands	1	0	3	8
Year	1981	1982	1983	1984
No. of boxes in ten thousands	10	4	5	8

Fit a trend line to the above data by graphical method.

11. Fit a trend line to data in Problem 10 by the method of least squares.

12. Obtain trend values for data in Problem 10 using 3-yearly moving averages.

13. Following table shows the number of traffic fatalities (in a state) resulting from drunken driving for years 1975 to 1983.

Year	1975	1976	1977	1978	1979
No. of deaths	0	6	3	8	2
Year	1980	1981	1982	1983	
No. of deaths	9	4	5	10	

Fit a trend line to the above data by graphical method.

14. Fit a trend line to data in Problem 13 by the method of least squares.

15. Obtain trend values for data in Problem 13 using 4-yearly moving averages.

16. Following table shows the all India infant mortality rates (per '000) for years 1980 to 2000.

Year	1980	1985	1990	1995
IMR	10	7	5	4
Year	2000	2005	2010	
IMR	3	1	0	

Fit a trend line to the above data by graphical method.

17. Fit a trend line to data in Problem 16 by the method of least squares.

18. Obtain trend values for data in Problem 16 using 3-yearly moving averages.

19. Following tables shows the wheat yield ('000 tonnes) in India for years 1959 to 1968.

Year	Yield	Year	Yield
1959	0	1964	0
1960	1	1965	4
1961	2	1966	1
1962	3	1967	2
1963	1	1968	10

Fit a trend line to the above data by the method of least squares.

20. Obtain trend values for data in Problem 19 using 3-yearly moving averages.

Activities

Note: You may change the origin and scale in the following problems according to your convenience.

1. Daily SENSEX index values at opening are given for fifty days in the following table. Plot a graph from the data. Find the trend graphically, using moving averages, and by the method of least squares.

Date	Index	Date	Index
1-Jan-19	36161.8	2-Jan-19	36198.13
3-Jan-19	35934.5	4-Jan-19	35590.79
7-Jan-19	35971.18	8-Jan-19	35964.62
9-Jan-19	36181.37	10-Jan-19	36258
11-Jan-19	36191.87	14-Jan-19	36113.27
15-Jan-19	35950.08	16-Jan-19	36370.74
17-Jan-19	36413.6	18-Jan-19	36417.58
21-Jan-19	36467.12	22-Jan-19	36649.92
23-Jan-19	36494.12	24-Jan-19	36146.55
25-Jan-19	36245.77	28-Jan-19	36099.62
29-Jan-19	35716.72	30-Jan-19	35819.67
31-Jan-19	35805.51	1-Feb-19	36311.74
4-Feb-19	36456.22	5-Feb-19	36573.04
6-Feb-19	36714.54	7-Feb-19	37026.56
8-Feb-19	36873.59	11-Feb-19	36585.5
12-Feb-19	36405.72	13-Feb-19	36279.63
14-Feb-19	36065.08	15-Feb-19	35985.68
18-Feb-19	35831.18	19-Feb-19	35543.24
20-Feb-19	35564.93	21-Feb-19	35837
22-Feb-19	35906.01	25-Feb-19	35983.8
26-Feb-19	35975.75	27-Feb-19	36138.83
28-Feb-19	36025.72	1-Mar-19	36018.49
5-Mar-19	36141.07	6-Mar-19	6544.86
7-Mar-19	36744.02	8-Mar-19	36753.59
8-Mar-19	36753.59	12-Mar-19	37249.65

2. Onion prices (per quintal) in a market are given for fifteen days. Plot a graph of given data. Find the trend graphically, using moving averages, and by the method of least squares.

Date	Price
24/01/2019	400
31/01/2019	650
3/2/19	650
7/2/19	700
24/02/2019	550
28/02/2019	550
5/3/19	500
7/3/19	600
14/03/2019	600
28/03/2019	600
7/4/19	800
11/4/19	801
14/04/2019	800
25/04/2019	800
2/5/19	800

3. Following table gives the number of persons injured in road accidents for 11 years. Plot a graph from the data. Find the trend graphically, using moving averages, and by the method of least squares.

Year	2006	2007	2008	2009	2010	2011
No. of injured	29955	27464	32224	32144	25400	28366
Year	2012	2013	2014	2015	2016	
No. of injured	27185	26314	24488	23825	22072	

4. Following table gives the number of road accidents due to over-speeding in Maharashtra for 9 years. Plot a graph from the data. Find the trend graphically, using moving averages, and by the method of least squares.

Year	2008	2009	2010	2011	2012
No. of accidents	38680	18090	21238	28489	27054
Year	2013	2014	2015	2016	
No. of accidents	26931	22925	24622	22071	

◆ ◆ ◆

5

Index Numbers

Let's Study

- Definition of Index Numbers
- Types of Index Numbers
- Terminology and Notation
- Construction of Index Numbers
    - Simple Aggregate Method
    - Weighted Aggregate Method
- Cost of Living Index Number
    - Aggregative Expenditure Method
    - Family Budget Method
- Uses of Cost of Living Index Number

Introduction

The value of money does not remain the same for all the time. It cannot be observed directly, but can be understood by observing the general level of prices. A rise in the price level indicates a fall in the value of money and a fall in the price level indicates a rise in the value of money. Changes in the value of money are reflected in changes in general level of prices over a period of time. Changes in the value of money are found to be inversely related to changes in price levels. So, changes in the value of money can be understood by observing changes in the general level of prices over a specified time period. Changes in the general level of prices are measured using a statistical tool known as **index numbers**. Index numbers provide one of the most popular statistical tools used in economics.

Index numbers cannot be measured directly, but are constructed with help of some mathematical formula. Index numbers are not expressed in terms of any units of measurement because they are ratios. Index numbers are usually expressed as percentages.

Maslow describes an index number as a numerical value characterizing the change in a complex economic phenomenon over a period of time. According to Spiegel, an index number is a statistical measure designed to show changes in a variable or a group of related variables with respect to time, geographical location or some other characteristic. Gregory and Ward describe it as a measure designed to show an average change, over time, in the price, quantity or value of a group of items. Croxton and Cowden say that an index number is a device that measures differences in the magnitude of a group of related variables. B. L. Bowley describes an index number as a series that reflects in its trend and fluctuations the movements of some quantity to which it is related. Blair puts an index number as a special kind of average.

Let's Learn

5.1 Definition of Index Numbers.

Index Numbers are defined in different ways by different experts. Some of the most popular definitions of Index Numbers are given below.

1. An Index Number is a statistical measure of changes in a variable or a group of variables with respect to time, geographical location, or some other characteristic such as production, income, etc.
2. An Index Number is used for measuring changes in some quantity that can not be measured directly.
3. An Index Number is a single ratio, usually expressed as percentage, that measures aggregate (or average) change in several

variables between two different times, places, or situations.

After reading the above definitions, we can conclude that an Index Number is an *'economic indicator'* of business activities.

Examples of index numbers.

NIFTY:

The NIFTY 50 index is National Stock Exchange of India's benchmark broad based stock market index for the Indian equity market. It represents the weighted average of 50 Indian company stocks in 13 sectors and is one of the two main stock indices used in India, the other being the BSE Sensex.

SENSEX:

The BSE SENSEX (also known as the S&P Bombay Stock Exchange Sensitive Index or simply the SENSEX) is a free-float market-weighted stock market index of 30 well-established and financially sound companies listed on Bombay Stock Ex change

5.2 Types of Index Numbers

Following are three major types ot index numbers.

1. Price Index Number

Price index numbers measure changes in the level of prices in the economy. It compares the price of the current year, with that of the base year to indicate the relative variation. It is a very good measure of inflation in the economy.

2. Quantity Index Number

As the name suggests, quantity index numbers measure changes in the quantities of goods between the two specified years. This can be the number of goods produced, sold, consumed, etc. It is a good indication of the output of an economy.

3. Value Index Number

A value index number is the ratio of the aggregate value of a given commodity (or a group of commodities) in the current year and its value in the base year. A value index number combines prices and quantities by taking the product of price and quantity as the value. The value index number thus measures the percentage change in the value of a commodity or a group of commodities during the current year in comparison to its value during the base year.

5.3 (a) : Terminology.

Base Period: The base period of an index number is the period against which comparisons are made. For example, the Central Statistical Organisation (CSO) is constructing the Consumer Price Index by taking 2010 as the base year. It means that the prices in 2015 are compared with 2010 prices by taking them as 100. The base period is indicated by subscript Zero.

Current Period : The present period is called the current period of an index number. An index number measures the changes between the base period and the current period. The current period is indicated by subscript 1.

Note:

The period used in index numbers can be a day, a month, or a year. We shall use a year as the period in our study.

5.3 (b) : Notation.

p_0 : Price of a commodity in the base year.

q_0 : Quantity (produced, purchased, or consumed) of a commodity in the base year.

p_1 : Price of a commodity in the current year.

q_1 : Quantity (produced, purchased, or consumed) of a commodity in the current year.

w : Weight assigned to a commodity according to its relative importance in the group.

I : Simple index number. It is also called the price relative. It is given by

$$I = \frac{p_1}{p_0} \times 100 \qquad (1)$$

P_{01}: Price index for the current year with respect to the base year.

Q_{01}: Quantity index for the current year with respect to the base year.

V_{01} : Value index for the current year with respect to the base year.

5.4 Construction of Index Numbers

Index number are constructed by the following two methods

1. Simple Aggregate Method.
2. Weighted Aggregate Method.

Let us now learn how index numbers are constructed by these two methods.

5.4.1 Method 1: Simple Aggregate Method

This is the simplest method of constructing index numbers. This method assumes that every commodity is equally important.

(a) Simple Aggregate Method to find Price Index Number

The procedure of calculating Price Index Number by the Simple Aggregate Method is as follows.

Step I : Prices of all commodities are added for the base year. This total is denoted by $\sum p_0$

Step II : Prices of all commodities are added for the current year. This total is denoted by $\sum p_1$.

Step III : The total obtained in Step II is divided by the total obtained in Step I. The ratio is then multiplied by 100.

Thus, the required price index number is given by

$$P_{01} = \frac{\sum p_1}{\sum p_0} \times 100.$$

(b) Simple Aggregate Method to find Quantity Index Number

Quantity Index Number can be calculated by the same procedure as above, only replacing prices by quantities.

Step I : Quantities of all commodities are added for the base year. This total is denoted by $\sum q_0$

Step II : Quantities of all commodities are added for the current year. This total is denoted by $\sum q_1$

Step III : The total obtained in Step II is divided by the total obtained in Step I. The ratio is then multiplied by 100.

Thus, the required quantity index number is given by

$$Q_{01} = \frac{\sum q_1}{\sum q_0} \times 100.$$

(c) Simple Aggregate Method to find Value Index Number

Value of a commodity is defined as the product of its price and quantity. Value Index Number is then calculated using the same procedure as above, where price or quantity is replaced by value.

Step I : Values (that is, products of prices and quantities) of all commodities are added for the base year. This total is denoted by $\sum p_0 q_0$

Step II : Values (that is, products of prices and quantities) of all commodities are added for the current year. This total is denoted by $\sum p_1 q_1$

Step III : The total obtained in Step II is divided by the total obtained in Step I. The ratio is then multiplied by 100.

Thus, the required value index number is given by

$$V_{01} = \frac{\sum p_1 q_1}{\sum p_0 q_0} \times 100.$$

SOLVED EXAMPLES

1. Calculate the price index number for the following data using the Simple Aggregate Method. Take 2000 as the base year.

Commodities	A	B	C	D	E
Price (in Rs.) for 2000	30	35	45	55	25
Price (in Rs.) for 2003	30	50	70	75	40

Solution :

Let us first tabulate the data in the following tabular form.

Table 5.1

Commodities	Price in 2000 (Base year) p_0	Price in 2003 (Current year) p_1
A	30	40
B	35	50
C	45	70
D	55	75
E	25	40
Total	$\sum p_0 = 190$	$\sum p_1 = 275$

Price Index Number is then given by

$$P_{01} = \frac{\sum p_1}{\sum p_0} \times 100$$

$$= \frac{275}{190} \times 100$$

$$= 144.74$$

Interpretation:

If the price of a commodity was Rs.100 in the year 2000, then the price of the same commodity is approximately Rs.145 in the year 2003. Hence, the overall increase in the price level is 45% in three years.

2. Calculate the Quantity Index Number for the following data using Simple Aggregate Method. take year 2000 as the base year.

Commodity	I	II	III	IV	V	VI
Quantity in 2000	30	55	65	70	40	90
Quantity in 2004	40	60	70	90	55	95

Solution:

We first tabulate the data in the following tabular form.

Table 5.2

Commodities	Quantity in 2000 (Base year) q_0	Quantity in 2004 (Current year) q_1
I	30	40
II	55	60
III	65	70
IV	70	90
V	40	55
VI	90	95
Total	$\sum q_0 = 350$	$\sum q_1 = 410$

Quantity Index Number is then given by

$$Q_{01} = \frac{\sum q_1}{\sum q_0} \times 100$$

$$= \frac{410}{350} \times 100$$

$$= 117.14.$$

This means that the output in terms of quantity rose by approximately 17% in year 2004 from year 2000.

3. Calculate the Value Index Number for the following data using the Simple Aggregate Method.

Commodities	Base Year		Current Year	
	Price Rs. p_0	Quantity (units) q_0	Price Rs. p_1	Quantity (units) q_1
p	10	6	60	7
Q	20	4	70	6
R	30	7	80	8
S	40	8	90	9
T	50	3	100	5

Solution : First, prepare the following table.

Table 5.3

Commodity	Base Year		Current Year			
	p_0	q_0	p_1	q_1	$p_0 q_0$	$p_1 q_1$
p	10	6	60	7	60	420
Q	20	4	70	6	80	420
R	30	7	80	8	210	640
S	40	8	90	9	320	810
T	50	3	100	5	150	500
Total					820	2790

Note : that $\sum p_0 q_0 = 820$, $\sum p_1 q_1 = 2790$, and, therefore, Value Index Number is given by

$$V_{01} = \frac{\sum p_1 q_1}{\sum p_0 q_0} \times 100$$

$$= \frac{2790}{820} \times 100$$

$$= 340.24$$

4. Find x in the following table if the Aggregate Price Index Number for year 1998 with respect to Base Year 1995 is 120.

Commodity	I	II	III	IV
Price in 1995	6	15	x	4
Price in 1998	8	18	28	6

Solution:

Table 5.4

Commodity	I	II	III	IV	Total
Price in 1995	6	15	x	4	$25 + x$
Price in 1998	8	18	28	6	60

From the above table, we have

$\sum p_0 = 25 + x$, $\sum p_1 = 60$, and

$P_{01} = 120$.

The value of x is then found from the formula

$$P_{01} = \frac{\sum p_1}{\sum p_0} \times 100$$

So that we obtain

$$120 = \frac{60}{25+x} \times 100$$

$$\therefore \quad 12 = \frac{60}{25+x} \times 10$$

$$\therefore 12(25+x) = 600$$

$$\therefore 300 + 12x = 600$$

$$\therefore \quad 12x = 600 - 300$$

$$\therefore \quad 12x = 300$$

$$\therefore \quad x = 25.$$

Hence, $x = 25$.

5. The Price Index Number for year 2004, with respect to year 2000 as base year, is known to be 130. Find the missing numbers in the following table if

$$\sum p_0 = 320$$

Commodity	A	B	C	D	E	F
Price (in Rs.) in 2000	40	50	30	x	60	100
Price (in Rs.) in 2005	50	70	30	85	y	115

Solution:

We first tabulate the given data as shown in the following table.

Table 5.5

Commodities	Price in 2000 (Base year) p_0	Price in 2005 (Current year) p_1
A	40	50
B	50	70
C	30	30
D	x	85
E	60	y
F	100	115

From the above table, we have

$\sum p_0 = 280 + x, \ \sum p_1 = 350 + y,$

But it is given that $\sum p_0 = 320$, so that

$$280 + x = 320$$

$$\therefore \quad x = 40$$

Further, using the formula

$$P_{01} = \frac{\sum p_1}{\sum p_0} \times 100$$

We have

$$130 = \frac{350 + y}{320} \times 100$$

$$\therefore \quad 416 = 350 + y$$

$$\therefore \quad y = 66.$$

EXERCISE 5.1

Find the Price Index Number using Simple Aggregate Method in each of the following examples.

1. Use 1995 as base year in the following problem.

Commodity	P	Q	R	S	T
Price (in Rs.) in 1995	15	20	24	22	28
Price (in Rs.) in 2000	27	38	32	40	45

2. Use 1995 as base year in the following problem.

Commodity	A	B	C	D	E
Price (in Rs.) in 1995	42	30	54	70	120
Price (in Rs.) in 2005	60	55	74	110	140

3.

Commodity	Unit	Base Year Price (in Rs.)	Current Year Price (in Rs.)
Wheat	kg	28	36
Rice	kg	40	56
Milk	litre	35	45
Clothing	meter	82	104
Fuel	litre	58	72

4. Use 2000 as base year in the following problem.

Commodity	Price (in Rs.) for year 2000	Price (in Rs.)for year 2006
Watch	900	1475
Shoes	1760	2300
Sunglasses	600	1040
Mobile	4500	8500

5. Use 1990 as base year in the following problem.

Commodity	Unit	Price (in Rs.) for 1990	Price (in Rs.) for 1997
Butter	kg	27	33
Cheese	kg	30	36
Milk	litre	25	29
Bread	loaf	10	14
Eggs	doz	24	36
Ghee	tin	250	320

6. Assume 2000 to be base year in the following problem.

Fruit	Unit	Price (in Rs.) in 2000	Price (in Rs.) in 2007
Mango	doz	250	300
Banana	doz	12	24
Apple	kg	80	110
Peach	kg	75	90
Orange	doz	36	65
Sweet Lime	doz	30	45

7. Use 2005 as base year in the following problem.

Vegetable	Unit	Price (in Rs.) in 2005	Price (in Rs.) in 2012
Ladies Finger	kg	32	38
Capsicum	kg	30	36
Brinjal	kg	40	60
Tomato	kg	40	62
Potato	kg	16	28

Find the Quantity Index Number using Simple Aggregate Method in each of the following examples.

8.

Commodity	I	II	III	IV	V
Base Year Quantities	140	120	100	200	225
Current Year Quantities	100	80	70	150	185

9.

Commodity	A	B	C	D	E
Base Year Quantities	360	280	340	160	260
Current Year Quantities	440	320	470	210	300

Find the Value Index Number using Simple Aggregate Method in each of the following examples.

10.

Commodity	Base Year		Current Year	
	Price	Quantity	Price	Quantity
A	30	22	40	18
B	40	16	60	12
C	10	38	15	24
D	50	12	60	16
E	20	28	25	36

11.

Commodity	Base Year		Current Year	
	Price	Quantity	Price	Quantity
A	50	22	70	14
B	70	16	90	22
C	60	18	105	14
D	120	12	140	15
E	100	22	155	28

12. Find x if the Price Index Number by Simple Aggregate Method is 125.

Commodity	P	Q	R	S	T
Base Year Price (in Rs.)	8	12	16	22	18
Current Year Price (in Rs.)	12	18	x	28	22

13. Find y if the Price Index Number by Simple Aggregate Method is 120, taking 1995 as base year

Commodity	A	B	C	D
Price (in Rs.) for 1995	95	y	80	35
Price (in Rs.) for 2003	116	74	92	42

5.4.2 Method 2: Weighted Aggregate Method

This method assigns suitable weights to different commodities before aggregating their prices, quantities, or values. These weights indicate relative importance of various commodities in the group. If w denotes the weight attached to a commodity, then the Price Index Number is given by

$$P_{01} = \frac{\sum p_1 w}{\sum p_0 w} \times 100$$

Weights are usually defined in terms of quantities in the weighted aggregate method. Index numbers constructed by the weighted aggregate method are known by names of the developers of these index numbers. Following are most popular price index numbers constructed by the weighted aggregate method.

(a) Laspeyre's Price Index Number

$$P_{01}(L) = \frac{\sum p_1 q_0}{\sum p_0 q_0} \times 100$$

Note: This construction uses base year quantities as weights.

(b) Paasche's Price Index Number

$$P_{01}(P) = \frac{\sum p_1 q_1}{\sum p_0 q_1} \times 100$$

Note: This construction uses current year quantities as weights.

(c) Dorbish-Bowley's Price Index Number

$$P_{01}(D-\mathrm{B}) = \frac{\frac{\sum p_1 q_0}{\sum p_0 q_0} + \frac{\sum p_1 q_1}{\sum p_0 q_1}}{2} \times 100$$

(d) Fisher's Ideal Price Index Number

$$P_{01}(F) = \sqrt{\frac{\sum p_1 q_0}{\sum p_0 q_0} \times \frac{\sum p_1 q_1}{\sum p_0 q_1}} \times 100$$

Question:

Can you find any relation among Laspeyre's, Paasche's, Dorbish-Bowley's and Fisher's Price Index Number?

(e) Marshall-Edgeworth's Price Index Number

$$P_{01}(M-E) = \frac{\sum p_1 \left(q_0 + q_1\right)}{\sum p_0 \left(q_0 + q_1\right)} \times 100$$

$$= \frac{\sum p_1 q_0 + \sum p_1 q_1}{\sum p_0 q_0 + \sum p_0 q_1} \times 100$$

(f) Walsh's Price Index Number

$$P_{01}(W) = \frac{\sum p_1 \sqrt{q_0 q_1}}{\sum p_0 \sqrt{q_0 q_1}} \times 100$$

SOLVED EXAMPLES

1. Calculate (a) Laspeyre's, (b) Paasche's, (c) Dorbish-Bowley's and Marshall-Edgeworth's Price Index Number for the following data.

Commodity	Base Year		Current Year	
	Price	Quantity	Price	Quantity
p	12	20	18	24
Q	14	12	21	16
R	8	10	12	18
S	16	15	20	25

Solution:

Let us first prepare the following table.

Table 5.6

Commodity	Base Year		Current Year		$p_0 q_0$	$p_1 q_0$	$p_0 q_1$	$p_1 q_1$
	p_0	q_0	p_1	q_1				
P	12	20	18	24	240	360	288	432
Q	14	12	21	16	168	252	224	336
R	8	10	12	18	80	120	144	216
S	16	15	20	25	240	300	400	500
Total					728	1032	1056	1484

From the above table, we have

$$\sum p_0 q_0 = 728, \quad \sum p_1 q_0 = 1032$$

$$\sum p_0 q_1 = 1056 \quad \sum p_1 q_1 = 1484$$

(a) Laspeyre's Price Index Number is then

$$P_{01}(L) = \frac{\sum p_1 q_0}{\sum p_0 q_0} \times 100$$

$$= \frac{1032}{728} \times 100$$

$$\therefore P_{01}(L) = 141.76$$

(b) Paasche's Price Index Number is given by

$$P_{01}(P) = \frac{\sum p_1q_1}{\sum p_0q_1} \times 100$$

$$= \frac{1484}{1056} \times 100$$

$$\therefore\ P_{01}(P) = 140.53$$

(c) Dorbish-Bowley's Price Index Number is given by

$$P_{01}(D-B) = \frac{P_{01}(L)+P_{01}(P)}{2}$$

$$= \frac{141.76+140.53}{2}$$

$$= 141.15$$

(d) Marchall-Edgeworth's Price Index Number is given by

$$P_{01}(M-E) = \frac{\left(\sum p_1q_0+\sum p_1q_1\right)}{\left(\sum p_0q_0+\sum p_0q_1\right)} \times 100$$

$$= \frac{(1032+1484)}{(728+1056)} \times 100$$

$$= \frac{2516}{1784} \times 100 = 141.03$$

$$\therefore\ P_{01}(M-E) = 141.03$$

2. Calculate Walsh's Price Index Number for the following data.

Commodity	Base Year		Current Year	
	Price	Quantity	Price	Quantity
A	20	9	30	4
B	10	5	50	5
C	40	8	10	2
D	30	4	20	1

Solution: Let us prepare the following table.

Table 5.7

Comm-odity	Base Year		Current Year		$\sqrt{q_0q_1}$	$p_0\sqrt{q_0q_1}$	$P_1\sqrt{q_0q_1}$
	p_0	q_0	p_1	q_1			
A	20	9	30	4	6	120	180
B	10	5	50	5	5	50	250
C	40	8	10	2	4	160	40
D	30	4	20	1	2	60	40
Total						390	510

From the above table, we get

$$\sum p_0\sqrt{q_0q_1} = 390$$

$$\sum p_1\sqrt{q_0q_1} = 510$$

Walsh's Price Index Number is given by

$$P_{01}(W) = \frac{\sum p_1\sqrt{q_0q_1}}{\sum p_0\sqrt{q_0q_1}} \times 100$$

$$= \frac{510}{390} \times 100$$

$$= \frac{5100}{39}$$

$$\therefore\ P_{01}(W) = 130.77$$

3. If $P_{01}(L) = 225, P_{01}(P) = 144$, then calculate $P_{01}(F)$ and $P_{01}(D-B)$

Solution :

Given $P_{01}(L)=225, P_{01}(P)=144$, we obtain

$$P_{01}(F) = \sqrt{P_{01}(L)\times P_{01}(P)}$$

$$= \sqrt{225\times 144}$$

$$= 15 \times 12$$

$$\therefore\ P_{01}(F) = 180$$

Next,

$$P_{01}(D-B) = \frac{P_{01}(L)+P_{01}(P)}{2}$$

$$= \frac{225+144}{2}$$

$$\therefore\ P_{01}(D-B) = 184.50$$

Example 4:

Find the missing price in the following table if Laspeyre's and Paasche's Price Index Numbers are the same.

Commodity	Base Year		Current Year	
	Price	Quantity	Price	Quantity
A	1	10	2	5
B	1	5	-	2

Solution:

Let us denote the missing value by x, and reconstruct the table as follows.

Table 5.8

Comm-odity	Base Year		Current Year		p_0q_0	p_1q_0	p_1q_1	p_0q_1
	p_0	q_0	p_1	q_1				
A	1	10	2	5	10	20	10	5
B	1	5	x	2	5	$5x$	$2x$	2
Total					15	$20+5x$	$10+2x$	7

The above table gives

$$\sum p_0q_0 = 15, \qquad \sum p_1q_0 = 20+5x$$

$$\sum p_0q_1 = 7 \qquad \sum p_1q_1 = 10+2x$$

It is given that

$$P_{01}(L) = P_{01}(P)$$

$$\frac{\sum p_1q_0}{\sum p_0q_0} \times 100 = \frac{\sum p_1q_1}{\sum p_0q_1} \times 100$$

$$\therefore \frac{5x+20}{15} = \frac{2x+10}{7}$$

$$\therefore \frac{5(x+4)}{15} = \frac{2x+10}{7}$$

$$\therefore 7(x+4) = 3(2x+10)$$

$$\therefore 7x+28 = 6x+30$$

$$\therefore x = 2.$$

The missing price is 2.

5. If $\sum p_0q_0 = 120$, $\sum p_0q_1 = 200$

$\sum p_1q_1 = 300$, and $P_{01}(L) = 150$, find

$P_{01}(M-E)$.

Solution: Note that

$$P_{01}(L) = \frac{\sum p_1q_0}{\sum p_0q_0} \times 100$$

$$\therefore 150 = \frac{\sum p_1q_0}{120} \times 100$$

$$\therefore \sum p_0q_1 = 15 \times 12$$

$$\therefore \sum p_0q_1 = 180.$$

Now,

$$P_{01}(M-E) = \frac{\left(\sum p_1q_0 + \sum p_1q_1\right)}{\left(\sum p_0q_0 + \sum p_0q_1\right)} \times 100$$

$$= \frac{180+300}{120+200} \times 100$$

$$= \frac{480}{320} \times 100$$

$$\therefore P_{01}(M-E) = 150.$$

6. If $\sum p_0q_0 = 180, \sum p_0q_1 = 200$

$\sum p_1q_1 = 280$, and $P_{01}(M-E) = 150$,

find $P_{01}(P)$.

Solution:

Let us denote $\sum p_0q_1$ by x. Then, using the fact that

$$P_{01}(M-E) = \frac{\left(\sum p_1q_0 + \sum p_1q_1\right)}{\left(\sum p_0q_0 + \sum p_0q_1\right)} \times 100$$

$$\therefore 150 = \frac{200+280}{180+x} \times 100$$

$$\therefore 15(180+x) = 4800$$

$$\begin{aligned}
\therefore\ 180 + x &= \frac{4800}{15} \\
\therefore\ 180 + x &= 320 \\
\therefore\ x &= 140 \\
\therefore \sum p_0q_1 &= 140.
\end{aligned}$$

Now,

$$\begin{aligned}
P_{01}(P) &= \frac{\sum p_1q_1}{\sum p_0q_1} \times 100 \\
&= \frac{280}{140} \times 100 \\
\therefore\ P_{01}(P) &= 200
\end{aligned}$$

EXERCISE 5.2

Calculate Laspeyre’s, Paasche’s, Dorbish-Bowley’s and Marshall-Edgeworth’s Price Index Numbers in Problems 1 and 2

1.

Commodity	Base Year		Current Year	
	Price	Quantity	Price	Quantity
A	8	20	11	15
B	7	10	12	10
C	3	30	5	25
D	2	50	4	35

2.

Commodity	Base Year		Current Year	
	Price	Quantity	Price	Quantity
I	10	9	20	8
II	20	5	30	4
III	30	7	50	5
IV	40	8	60	6

Calculate Walsh’s Price Index Number in Problem 3 and 4.

3.

Commodity	Base Year		Current Year	
	Price	Quantity	Price	Quantity
L	4	16	3	19
M	6	16	8	14
N	8	28	7	32

4.

Commodity	Base Year		Current Year	
	Price	Quantity	Price	Quantity
I	10	12	20	9
II	20	4	25	8
III	30	13	40	27
IV	60	29	75	36

5. If $P_{01}(L) = 90$, and $P_{01}(P) = 40$, find $P_{01}(D-B)$ and $P_{01}(F)$

6. If $\sum p_0q_0 = 140, \sum p_0q_1 = 200$, $\sum p_0q_1 = 350$, $\sum p_1q_1 = 460$, find Laspeyre’s, Paasche’s, Dorbish-Bowley’s and Marshall-Edgeworth’s Price Index Numbers.

7. Given that Laspeyre’s and Dorbish-Bowley’s Price Index Numbers are 160.32 and 164.18 respectively. Find Paasche’s Price Index Number.

8. Given that $\sum p_0q_0 = 220, \sum p_0q_1 = 380$, $\sum p_1q_1 = 350$ and Marshall-Edgeworth’s Price Index Number is 150, find Laspeyre’s Price Index Number.

9. Find x in the following table if Laspeyre’s and Paasche’s Price Index Numbers are equal.

Commodity	Base Year		Current Year	
	Price	Quantity	Price	Quantity
A	2	10	2	5
B	2	5	x	2

10. If Laspeyre's Price Index Number is four times Paasche's Price Index Number, then find the relation between Dorbish-Bowley's and Fisher's Price Index Numbers.

11. If Dorbish-Bowley's and Fisher's Price Index Numbers are 5 and 4, respectively, then find

    Laspeyre's and Paasche's Price Index Numbers.

5.5 Cost of Living Index Number

Cost of Living Index Number, also known as Consumer Price Index Number, is an index number of the cost of buying goods and services in day-to-day life for a specific consumer class. Different classes of consumers show different patterns of consumption of goods and services. As a result, a general index number cannot reflect changes in cost of living for a specific consumer class. For example, cost of living index numbers for rural population are different from cost of living index numbers for urban population. The goods and services consumed by members of different consumer classes can be different and therefore cost of living index numbers calculated for different consumer classes can be based on costs of different sets of goods and services.

Steps in Construction of Cost of Living Index Numbers

Construction of cost of living index numbers involves the following steps:

1. Choice of Base Year:

The first step in preparing cost of living index numbers is choice of base year. Base years defined as that year with reference to which price changes in other years are compared and expressed as percentages. The base year should be a normal year. It should be free from abnormal conditions like wars, famines, floods, political instability, etc.

Base year can be chosen in two ways:

(a) Using fixed base method, where the base year remains fixed; and

(b) Using chain base method, where the base year goes on changing. For example, 1979 will be the base year for 1980, it will be 1978 for 1979, and so on.

2. Choice of Commodities:

The second step in construction of cost of living index numbers is choosing the commodities. Since all commodities cannot be included, only representative commodities should be chosen according to the purpose of the index number.

In choosing commodities, the following points must be kept in mind:

(a) The commodities must represent the tastes, habits and customs of the people.

(b) Commodities should be recognizable.

(c) Commodities should have the same quality over different periods and places.

(d) The economic and social importance of different commodities should be taken in consideration.

(e) The commodities should be sufficiently large in number,

(f) All varieties of a commodity should be included that are in common use and are stable in nature.

3. Collection of Prices:

After choosing the commodities, the next step is collection of their prices. The following points are important while collecting prices of commodities chosen for constructing cost of living index numbers.

(a) From where prices are to be collected.

(b) Whether to collect wholesale prices or retail prices.

(c) Whether to include taxes in prices.

Following points are to be noted while collecting prices:

(a) Prices must be collected from places where a particular commodity is traded in large quantities

(b) If published information on prices is available, it must be used,

(c) Care should be taken while collecting price quotations from individuals or institutions that they provide correct information.

(d) Choice of wholesale or retail prices depends on the purpose of preparing index numbers. Wholesale prices are used in the construction of general price index, while retail prices are used in the construction of cost of living index.

(e) Prices must be averaged if collected from several sources.

4. Choice of Average:

Since the index numbers are a specialized average, it is important to choose a suitable average. Geometric mean is theoretically the best, but arithmetic mean is used in practice because it is easier to calculate.

5. Choice of Weights:

Generally, all the commodities included in the construction of index numbers are not equally important. Therefore, proper weights must be assigned to the commodities according to their relative importance. For example, cost of living index for teachers will assign higher weightage to prices of books than cost of living index for workers. Weights should be chosen rationally and not arbitrarily.

6. Purpose of Index Numbers:

The most important consideration in the construction of index numbers is their objective. All other steps are to be viewed in light of the purpose for which a particular index number is being prepared. Since every index number is prepared with a specific purpose, no single index number can be 'all purpose' index number. It is important to have a clear idea about the purpose of the index number before it is constructed.

Methods of constructing Cost of Living Index Numbers

5.5.1 Aggregative Expenditure Method (Weighted Aggregate Method)

This method uses quantities consumed in base year as weights, so that Cost of Living Index Number is defined as follows.

$$\text{CLI} = \frac{\text{Total expendutre in current year}}{\text{Total expendutre in base year}}$$

$$= \frac{\sum p_1q_0}{\sum p_0q_0} \times 100$$

The above formula is similar to that of a weighted Index Number. Do you recognize that Index Number?

5.5.2 Family Budget Method (Weighted Relative Method)

Cost of Living Index Number is defined as follows.

$$\text{CLI} = \frac{\sum IW}{\sum W}$$

where

$$I = \frac{p_1}{p_0} \times 100$$

= price relative for current year

and

$$W = p_0q_0$$

= base year weightage.

Do you find the above two methods of calculating the Cost of Living Index Numbers to be same?

SOLVED EXAMPLES

1. Construct the Cost of Living Index Number for the following data.

Group	Base Year		Current Year
	Price	Quantity	Price
Food & Clothing	40	3	70
Fuel & Lighting	30	5	60
House Rent	50	2	50
Miscellaneous	60	3	90

Solution:

We shall begin by preparing the following table.

Group	Base Year		Current Year	p_1q_0	p_0q_0
	p_0	q_0	p_1		
Food & Clothing	40	3	70	210	120
Fuel & Lighting	30	5	60	300	150
House Rent	50	2	50	100	100
Miscellaneous	60	3	90	270	180
Total				880	550

We shall use Aggregative Expenditure Method since p_0, q_0 and p_1 are given.

$$\text{CLI} = \frac{\sum p_1q_0}{\sum p_0q_0} \times 100$$

$$= \frac{880}{550} \times 100$$

$$\therefore \text{CLI} = 160.$$

Interpretation. A person earning Rs 100 in the base year, should earn Rs 160 in the current year to maintain the same standard of living.

2. The following table gives the base year weightage (W) and current year price relative (I) for five commodities. Calculate the Cost of Living Index Number.

Group	Food	Clothing	Fuel & Lighting	House Rent	Miscellaneous
I	120	100	140	160	150
W	3	6	5	2	4

Solution:

We use Family Budget Method since I and W are given. For this, we prepare the following table.

Table 5.10

Group	I	W	IW
Food	120	3	360
Clothing	100	6	600
Fuel & Lighting	140	5	700
House Rent	160	2	320
Miscellaneous	150	4	600
Total	-	20	2580

The above table shows that $\sum W$ =20 and $\sum IW$ = 2580, and

$$\therefore \text{CLI} = \frac{\sum IW}{\sum W}$$

$$= \frac{2580}{20}$$

$$\therefore \text{CLI} = 129.$$

3. Find x in the following table if the Cost of Living Index Number is 121.

Group	Food	Clothing	Fuel & Lighting	House Rent	Miscellaneous
I	100	125	174	x	90
W	13	12	10	8	7

Solution:

First, we prepare the following table.

Group	I	W	IW
Food	100	13	1300
Clothing	125	12	1500
Fuel & Lighting	174	10	1740
House Rent	x	8	$8x$
Miscellaneous	90	7	630
Total	-	50	$5170+8x$

It can be found from the above table that

$\sum W = 50$ and $\sum IW = 5170 + 8x$

$$\therefore \quad \text{CLI} = \frac{\sum IW}{\sum W}$$

$$\therefore \quad 121 = \frac{5170+8x}{50}$$

$$\therefore \quad 6050 = 5170 + 8x$$

$$\therefore \quad x = 110.$$

4. Cost of Living Index Numbers for the years 2000 and 2005 are 120 and 200, respectively. If a person has monthly earnings of Rs 10800 in year 2000, what should be his monthly earnings in year 2005 in order to maintain same standard of living?

Solution:

For the year 2000, it is given that CLI = 120, and Income = Rs 10800. These two give us real income as follows.

$$\text{Real Income} = \frac{\text{Income}}{\text{CLI}} \times 100$$

$$= \frac{10800}{120} \times 100$$

$$\therefore \text{Real Income} = 9000.$$

This shows that the real income is Rs 9000.

The CLI for year 2005 is 220.

$$\text{Real Income} = \frac{\text{Income}}{\text{CLI}} \times 100$$

$$\therefore \quad 9000 = \frac{\text{Income}}{220} \times 100$$

$$\therefore \quad \text{Income} = 19800.$$

This shows that the monthly income of the person should be Rs 19800 in year of 2005 in order to maintain the same standard of living as in year 2000

5. Calculate the Cost of Living Index Number for the year 1999 by Family Budget Method from the following data. Also, find the expenditure of a person in year 1999 if his expenditure in year 1995 was 800.

Group	Price in year 1995	Price in year 1999	W
Food	8	24	6
Clothing	18	36	12
Fuel & Lighting	20	40	8
House Rent	15	30	4
Miscellaneous	10	22	10

Solution:

Let us first prepare the following table.

Table 5.12

Group	Price in year 1995	Price in year 1999	$I = \frac{p_1}{p_0} \times 100$	W	IW
Food	8	24	300	6	1800
Clothing	18	36	200	12	2400
Fuel & Lighting	20	40	200	8	1600
House Rent	15	30	200	4	800
Miscellaneous	10	22	220	10	2200
Total	-	-	-	40	8800

By Family Budget Method,

$$\therefore \quad \text{CLI} = \frac{\sum IW}{\sum W}$$

$$= \frac{8800}{40}$$

$$\therefore \quad \text{CLI} = 220$$

Now, the expenditure in 1995 was Rs 800. In other words, the expenditure is Rs 800 when CLI is 100. The question is to find expenditure when CLI is 220 in 1999.

$$\therefore \text{Expenditure in 1999} = \frac{220}{100} \times 800$$

$$= 1760$$

Thus, the expenditure in 1999 is Rs 1760.

EXERCISE 5.3

Calculate the cost of living index in problems 1 to 3.

1.

Group	Base Year		Current Year
	Price	Quantity	Price
Food	120	15	170
Clothing	150	20	190
Fuel & Lighting	130	30	220
House Rent	160	10	180
Miscellaneous	200	12	200

2.

Group	Base Year		Current Year
	Price	Quantity	Price
Food	40	15	45
Clothing	30	10	35
Fuel & Lighting	20	17	25
House Rent	60	22	70
Miscellaneous	70	25	80

3.

Group	Base Year		Current Year
	Price	Quantity	Price
Food	132	10	170
Clothing	154	12	160
Fuel & Lighting	164	20	180
House Rent	175	18	195
Miscellaneous	128	5	120

Base year weights (W) and current year price relatives (I) are given in Problems 4 to 8. Calculate the cost of living index in each case

4.

Group	Food	Clothing	Fuel & Lighting	House Rent	Miscellaneous
I	70	90	100	60	80
W	5	3	2	4	6

5.

Group	Food	Clothing	Fuel & Lighting	House Rent	Miscellaneous
I	400	300	150	120	100
W	3	3	4	5	2

6.

Group	Food	Clothing	Fuel & Lighting	House Rent	Miscellaneous
I	200	150	120	180	160
W	30	20	10	40	50

7. Find x if the cost of living index is 150.

Group	Food	Clothing	Fuel & Lighting	House Rent	Miscellaneous
I	180	120	300	100	160
W	4	5	6	x	3

8. Find y if the cost of living index is 200

Group	Food	Clothing	Fuel & Lighting	House Rent	Miscellaneous
I	180	120	160	300	200
W	4	5	3	y	2

9. The Cost of Living Index Number for years 1995 and 1999 are 140 and 200 respectively. A person earns Rs. 11,200 per month in the year 1995. What should be his monthly earnings in the year 1999 in order to maintain his standard of living as in the year 1995 ?

5.6 Uses of Cost of Living Index Number

1. Cost of Living Index Number is used to regulate the dearness allowance or the grant of bonus to employees in order to enable them bear the increased cost of living.
2. Cost of Living Index Number is used for settling dispute related to salaries and wages.
3. Cost of Living Index Number is used in calculating purchasing power of money.

   Purchasing power of money

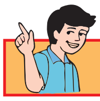
   $$= \frac{1}{\text{Cost of Living Index Number}}$$

4. Cost of Living Index Number is used in determining real wages.

   Real Wages

   $$= \frac{\text{Money wages}}{\text{Cost of Living Index Number}} \times 100$$

5. Cost of Living Index Numbers are widely used in negotiations of wages in wage contracts.



Let's Remember

- There are three types of index numbers.
  - (i) Price Index Number
  - (ii) Quantity Index Number
  - (iii) Value Index Number
- There are two methods of constructing index numbers.
  - (i) Simple Aggregate Method
  - (ii) Weighted Aggregate Method
- Price Index Number using Simple aggregate method is calculated by the following formula.

$$P_{01} = \frac{\sum p_1}{\sum p_0} \times 100$$

Where

P_{01} : Price index Number for the current year with respect to base year

P_1 : Price of the commodity in current year

P_0 : Price of the commodity in base year

- Price Index Number using Weighted Aggregate Method is calculated by the following formula.

$$P_{01} = \frac{\sum p_1 w}{\sum p_0 w} \times 100$$

Where

w : Weight assigned to a commodity

- **Laspeyre's Price Index Number**

$$P_{01}(L) = \frac{\sum p_1 q_0}{\sum p_0 q_0} \times 100$$

- **Paasche's Price Index Number**

$$P_{01}(P) = \frac{\sum p_1 q_1}{\sum p_0 q_1} \times 100$$

- **Dorbish-Bowley's Price Index Number**

$$P_{01}(D\text{–}B) = \frac{\frac{\sum p_1 q_0}{\sum p_0 q_0} + \frac{\sum p_1 q_1}{\sum p_0 q_1}}{2} \times 100$$

- **Fisher's Price Index Number**

$$P_{01}(F) = \sqrt{\frac{\sum p_1 q_0}{\sum p_0 q_0} \times \frac{\sum p_1 q_1}{\sum p_0 q_1}} \times 100$$

- **Marshall-Edgeworth's Price Index Number**

$$P_{01}(M-E) = \frac{\sum p_1 (q_0 + q_1)}{\sum p_0 (q_0 + q_1)} \times 100$$

- **Walsh's Price Index Number**

$$P_{01}(W) = \frac{\sum p_1 \sqrt{q_0 q_1}}{\sum p_0 \sqrt{q_0 q_1}} \times 100$$

- **Cost of Living Index Number using Aggregate Expenditure Method**

$$\text{CLI} = \frac{\sum p_1 q_0}{\sum p_0 q_0} \times 100$$

- **Cost of Living Index Number using Weighted Relative Method**

$$\text{CLI} = \frac{\sum IW}{\sum W}$$

where $I = \frac{p_1}{p_0} \times 100$

and $w = p_0 q_0$

MISCELLANEOUS EXERCISE - 5

I) Choose the correct alternative.

1. Price Index Number by Simple Aggregate Method is given by

(a) $\sum \frac{p_1}{p_0} \times 100$

(b) $\sum \frac{p_0}{p_1} \times 100$

(c) $\frac{\sum p_1}{\sum p_0} \times 100$

(d) $\frac{\sum p_0}{\sum p_1} \times 100$

2. Quantity Index Number by Simple Aggregate Method is given by

(a) $\sum \frac{q_1}{q_0} \times 100$

(b) $\sum \frac{q_0}{q_1} \times 100$

(c) $\frac{\sum q_1}{\sum q_0} \times 100$

(d) $\frac{\sum q_0}{\sum q_1} \times 100$

3. Value Index Number by Simple Aggregate Method is given by

(a) $\sum \frac{p_1 q_0}{p_0 q_1} \times 100$

(b) $\sum \frac{p_0 q_1}{p_0 q_0} \times 100$

(c) $\frac{\sum p_1 q_1}{\sum p_1 q_0} \times 100$

(d) $\frac{\sum p_1 q_1}{\sum p_0 q_0} \times 100$

4. Price Index Number by Weighted Aggregate Method is given by

(a) $\sum \frac{p_1 w}{p_0 w} \times 100$

(b) $\sum \frac{p_0 w}{p_1 w} \times 100$

(c) $\frac{\sum p_1 w}{\sum p_0 w} \times 100$

(d) $\frac{\sum p_0 w}{\sum p_1 w} \times 100$

5. Quantity Index Number by Weighted Aggregate Method is given by

(a) $\sum \frac{q_1 w}{q_0 w} \times 100$

(b) $\sum \frac{q_0 w}{q_1 w} \times 100$

(c) $\frac{\sum q_1 w}{\sum q_0 w} \times 100$

(d) $\frac{\sum q_0}{\sum q_1} \times 100$

6. Value Index Number by Weighted Aggregate Method is given by

(a) $\sum \frac{p_1q_0w}{p_0q_0w} \times 100$

(b) $\sum \frac{p_0q_1w}{p_0q_0w} \times 100$

(c) $\frac{\sum p_1q_1w}{\sum p_0q_1w} \times 100$

(d) $\frac{\sum p_1q_1w}{\sum p_0q_0w} \times 100$

7. Laspeyre's Price Index Number is given by

(a) $\frac{\sum p_0q_0}{\sum p_1q_0} \times 100$

(b) $\frac{\sum p_0q_1}{\sum p_1q_1} \times 100$

(c) $\frac{\sum p_1q_0}{\sum p_0q_0} \times 100$

(d) $\frac{\sum p_1q_1}{\sum p_0q_1} \times 100$

8. Paasche's Price Index Number is given by

(a) $\frac{\sum p_0q_0}{\sum p_1q_0} \times 100$

(b) $\frac{\sum p_0q_1}{\sum p_1q_1} \times 100$

(c) $\frac{\sum p_1q_0}{\sum p_0q_0} \times 100$

(d) $\frac{\sum p_1q_1}{\sum p_0q_1} \times 100$

9. Dorbish-Bowley's Price Index Number is given by

(a) $\frac{\frac{\sum p_1q_0}{\sum p_0q_1} + \frac{\sum p_0q_1}{\sum p_1q_0}}{2} \times 100$

(b) $\frac{\frac{\sum p_1q_1}{\sum p_0q_0} + \frac{\sum p_0q_0}{\sum p_1q_1}}{2} \times 100$

(c) $\frac{\frac{\sum p_1q_0}{\sum p_0q_0} + \frac{\sum p_1q_1}{\sum p_0q_1}}{2} \times 100$

(d) $\frac{\frac{\sum p_0q_0}{\sum p_1q_0} + \frac{\sum p_0q_1}{\sum p_1q_1}}{2} \times 100$

10. Fisher's Price Number is given by

(a) $\sqrt{\frac{\sum p_1q_0}{\sum p_0q_0} \times \frac{\sum p_1q_1}{\sum p_0q_1}} \times 100$

(b) $\sqrt{\frac{\sum p_0q_0}{\sum p_1q_0} \times \frac{\sum p_0q_1}{\sum p_1q_1}} \times 100$

(c) $\sqrt{\frac{\sum p_0q_1}{\sum p_0q_0} \times \frac{\sum p_1q_1}{\sum p_1q_0}} \times 100$

(d) $\sqrt{\frac{\sum p_1q_0}{\sum p_1q_1} \times \frac{\sum p_0q_0}{\sum p_0q_1}} \times 100$

11. Marshall-Edgeworth's Price Index Number is given by

(a) $\frac{\sum p_1(q_0+q_1)}{\sum p_0(q_0+q_1)} \times 100$

(b) $\frac{\sum p_0(q_0+q_1)}{\sum p_1(q_0+q_1)} \times 100$

(c) $\frac{\sum q_1(p_0+p_1)}{\sum q_0(p_0+p_1)} \times 100$

(d) $\frac{\sum q_1(p_0+p_1)}{\sum q_0(p_0+p_1)} \times 100$

12. Walsh's Price Index Number is given by

(a) $\frac{\sum p_1\sqrt{q_0q_1}}{\sum p_0\sqrt{q_0q_1}} \times 100$

(b) $\dfrac{\sum p_0\sqrt{q_0q_1}}{\sum p_1\sqrt{q_0q_1}} \times 100$

(c) $\dfrac{\sum q_1\sqrt{p_0p_1}}{\sum q_0\sqrt{p_0p_1}} \times 100$

(d) $\dfrac{\sum q_0\sqrt{p_0p_1}}{\sum q_1\sqrt{p_0p_1}} \times 100$

13. The Cost of Living Index Number using Aggregate Expenditure Method is given by

(a) $\dfrac{\sum p_1q_0}{\sum p_0q_0} \times 100$

(b) $\sum \dfrac{p_1q_1}{p_0q_1} \times 100$

(c) $\dfrac{\sum p_1q_1}{\sum p_0q_1} \times 100$

(d) $\sum \dfrac{p_1q_0}{p_0q_0} \times 100$

14. The Cost of Living Index Number using Weighted Relative Method is given by

(a) $\dfrac{\sum IW}{\sum W}$

(b) $\sum \dfrac{W}{IW}$

(c) $\dfrac{\sum W}{\sum IW}$

(d) $\sum \dfrac{IW}{W}$

II) Fill in the blanks.

1. Price Index Number by Simple Aggregate Method is given by ________________.
2. Quantity Index Number by Simple Aggregate Method is given by __________.
3. Value Index Number by Simple Aggregate Method is given by _________________.
4. Price Index Number by Weighted Aggregate Method is given by ________________.
5. Quantity Index Number by Weighted Aggregate Method is given by ________________.
6. Value Index Number by Weighted Aggregate Method is given by ___________________.
7. Laspeyre's Price Index Number is given by ______________.
8. Paasche's Price Index Number is given by ________________.
9. Dorbish-Bowley's Price Index Number is given by _______________.
10. Fisher's Price Index Number is given by _______________.
11. Marshall-Edgeworth's Price Index Number is given by __________________.
12. Walsh's Price Index Number is given by __________________.

III) State whether each of the following is True or False.

1. $\dfrac{\sum p_1}{\sum p_0} \times 100$ is the Price Index Number by Simple Aggregate Method. .
2. $\dfrac{\sum q_0}{\sum q_1} \times 100$ is the Quantity Index Number by Simple Aggregate Method.
3. $\sum \dfrac{p_0q_0}{p_1q_1} \times 100$ is Value Index Number by Simple Aggregate Method.
4. $\sum \dfrac{p_1q_0}{p_0q_0} \times 100$ is Paasche's Price Index Number.
5. $\dfrac{\sum p_1q_1}{\sum p_0q_1} \times 100$ is Laspeyre's Price Index Number.

6. $\frac{\sum p_1q_0}{\sum p_0q_0} \times \frac{\sum p_1q_0}{\sum p_0q_0} \times 100$ is Dorbish-Bowley's Price Index Number.

7. $\frac{1}{2}\left[\sqrt{\frac{\sum p_1q_0}{\sum p_0q_0}} + \frac{\sqrt{p_1q_1}}{\sqrt{p_0q_1}}\right] \times 100$ is Fisher's Price Index Number.

8. $\frac{\sum p_0(q_0+q_1)}{\sum p_1(q_0+q_1)} \times 100$ is Marshall-Edgeworth's Price Index Number.

9. $\frac{\sum p_0\sqrt{q_0q_1}}{\sum p_1\sqrt{q_0q_1}} \times 100$ is Walsh's Price Index Number.

10. $\sqrt{\frac{p_1q_0}{\sum p_0q_0}} \times \sqrt{\frac{\sum p_1q_1}{\sum p_0q_1}} \times 100$ is Fisher's Price Index Number.

IV) Solve the following problems.

1. Find the Price Index Number using Simple Aggregate Method. Consider 1980 as base year.

Commodity	Price in 1980 (in Rs.)	Price in 1985 (in Rs.)
I	22	46
II	38	36
III	20	28
IV	18	44
V	12	16

2. Find the Quantity Index Number using Simple Aggregate Method.

Commodity	Based year quantity	Current year quantity
A	100	130
B	170	200
C	210	250
D	90	110
E	50	150

3. Find the Value Index Number using Simple Aggregate Method.

Commodity	Base Year		Current Year	
	Price	Quantity	Price	Quantity
I	20	42	22	45
II	35	60	40	58
III	50	22	55	24
IV	60	56	70	62
V	25	40	30	41

4. Find x if the Price Index Number using Simple Aggregate Method is 200

Commodity	P	Q	R	S	T
Base Year Price	20	12	22	23	13
Current Year Price	30	x	38	51	19

5. Calculate Laspeyre's and Paasche's Price Index Number for the following data.

Commodity	Base Year		Current Year	
	Price p_0	Quantity q_0	Price p_1	Quantity q_1
A	20	18	30	15
B	25	8	28	5
C	32	5	40	7
D	12	10	18	10

6. Calculate Dorbish-Bowley's Price Index Number for the following data.

Commodity	Base Year		Current Year	
	Price p_0	Quantity q_0	Price p_1	Quantity q_1
I	8	30	11	28
II	9	25	12	22
III	10	15	13	11

7. Calculate Marshall-Edgeworth's Price Index Number for the following data.

Commodity	Base Year		Current Year	
	Price p_0	Quantity q_0	Price p_1	Quantity q_1
X	12	35	15	25
Y	29	50	30	70

8. Calculate Walsh's Price Index Number for the following data.

Commodity	Base Year		Current Year	
	Price p_0	Quantity q_0	Price p_1	Quantity q_1
I	8	30	12	25
II	10	42	20	16

9. Calculate Laspeyre';s and Paasche's Price Index Number for the following data.

Commodity	Base Year		Current Year	
	Price p_0	Quantity q_0	Price p_1	Quantity q_1
I	8	30	12	25
II	10	42	20	16

10. Find x if Laseyre's Price Index Number is same as Paasche's Price Index Number for the following data.

Commodity	Base Year		Current Year	
	Price p_0	Quantity q_0	Price p_1	Quantity q_1
A	3	x	2	5
B	4	6	3	5

11. If find x is Walsh's Price Index Number is 150 for the following data.

Commodity	Base Year		Current Year	
	Price p_0	Quantity q_0	Price p_1	Quantity q_1
A	5	3	10	3
B	x	4	16	9
C	15	5	23	5
D	10	2	26	8

12. Find x if Paasche's Price Index Number is 140 for the following data.

Commodity	Base Year		Current Year	
	Price p_0	Quantity q_0	Price p_1	Quantity q_1
A	20	8	40	7
B	50	10	60	10
C	40	15	60	x
D	12	15	15	15

13. Given that Laspeyre's and Paasche's Price Index Numbers are 25 and 16 respectively. Find Dorbish-Bowley's and Fisher's Price Index Number.

14. If Laspeyre's and Dorbish's Price Index Numbers are 150.2 and 152.8 respectively, find Paasche's Price Index Number.

15. If $\sum p_0q_0 = 120$, $\sum p_0q_1 = 160$, $\sum p_1q_1 = 140$, and $\sum p_0q_1 = 200$ find Laspeyre's, Paasche's, Dorbish-Bowley's, and Marshall-Edgeworth's Price Index Numbers.

16. Given that $\sum p_0q_0 = 130$, $\sum p_1q_1 = 140$, $\sum p_0q_1 = 160$, and $\sum p_0q_1 = 200$, find Laspeyre's, Paasche's, Dorbish-Bowley's, and Marshall-Edgeworth's Price Index Numbers.

17. Given that $\sum p_1q_1 = 300$, $\sum p_0q_1 = 320$, $\sum p_0q_0 = 120$, and Marshall-Edgeworth's Price Index Number is 120, find Paasche's Price Index Number.

18. Calculate the cost of living number for the following data.

Group	Base Year		Current Year
	Price p_0	Quantity q_0	Price p_1
Food	150	13	160
Clothing	170	18	150
Fuel & Lighting	175	10	190
House Rent	200	12	210
Miscellaneous	210	15	260

19. Find the cost living index number by the weighted aggregate method.

Group	Food	Clothing	Fuel & Lighting	House Rent	Miscellaneous
I	78	80	110	60	90
W	5	3	4	2	6

20. Find the cost of living index number by Family Budget Method for the following data. Also, find the expenditure of a person in the year 2008 if his expenditure in the year 2005 was Rs. 10,000.

Group	Base Year (2005) Price	Current Year (2005) Price	Weight
Food	12	60	25
Clothing	10	45	20
Fuel & Lighting	20	35	15
House Rent	25	20	30
Miscellaneous	16	48	10

21. Find x if the cost of living index number is 193 for the following data.

Group	Food	Clothing	Fuel & Lighting	House Rent	Miscellaneous
I	221	198	171	183	161
W	35	14	x	8	20

22. The cost of living index number for year 2000 and 2003 are 150 and 210 respectively. A person earns Rs. 13,500 per month in the year 2000. What should be his monthly earning in the year 2003 in order to maintain the same standard of living ?

Activities

Try each of the following activities for better understanding of index numbers.

1. Find weekly prices of any five vegetables for at least six months. Taking the first week of observation as the base period, find price index numbers for the remaining five months for every vegetable.
2. Note the SENSEX for six months. Talking the first mouth as the base period, find price index numbers for the remaining five months.
3. Note inflation rate for six months. Taking the first month as the base period. Find price index numbers for the remaining five months.
4. Note petrol prices for six months. Taking the first months as the base period, find price index numbers for the remaining five months.
5. Note gold prices for six months. Taking the first month as the base period, find price index numbers for the remaining five months.

◆ ◆ ◆

6 Linear Programming

Let's Study

- Meaning of L.P.P.
- Mathematical formula of L.P.P.

  - Solution of L.P.P. by graphical Method

Let's Recall

- Linear inequations

A linear equation in two variables namely $ax+by+c=0$, where $a,b,c \in R$ and $(a,b) \neq (0,0)$, represents a straight line. A straight line makes three disjoint parts of the plane : the points lying on the straight line and two half planes on either side, which are represented by $ax+by+c<0$ or $ax+by+c>0$

The set of points $\{ (x, y) | ax + by + c < 0\}$ and $\{ (x,y) \; | ax+by+c>o \}$ are two open half planes The two sets have the common boundary $\{(x,y) \mid ax + by + c = 0\}$.

In the earlier classes, we have studied graphical solution of linear equations and linear inequations in two variables. In this chapter, we shall study these graphical solutions to find the maximum/minimum value of a linear expression.

Let's Learn

6.1 Linear Programming Problem (L.P.P.)

Linear programming is used in industries and government sectors where attempts are made to increase the profitability or efficiency and to reduce wastage. These problems are related to efficient use of limited resources like raw materials, man-power, availability of machine time, cost of material and so on.

Linear Programming is a mathematical technique designed to help for planning and decision making. Linear Programming problems are also known as optimization problems. Mathematical programming involves optimization of a certain function, called objective function, subject to given conditions or restrictions known as constraints.

Meaning of L.P.P.

Linear implies all the mathematical functions containing variables of index one. A L.P.P. may be defined as the problem of maximizing or minimizing a linear function subject to linear constraints.

These constraints may be equations or inequations.

Now, we formally define the terms related to L.P.P. as follows.

1) Decision Variables: The variables involved in L.P.P. are called decision variables
2) Objective function: A linear function of decision variables which is to be optimized, i.e. either maximized or minimized, is called objective function.
3) Constraints: Conditions under which the objective function is to be optimized, are called constraints. These are in the form of equations or inequations.
4) Non-negativity constraints: In some situations, the values of the variables under considerations may be positive or zero due to the imposed conditions. These constraints are referred as non-negativity constraints.

6.2 Mathematical Formulation of L.P.P.

Step 1: Identify the decision variables as (x,y) or (x_1, x_2)

Step 2: Identify the objective function and write it as mathematical expression in terms of decision variables.

Step 3: Identify the different constraints and express them as mathematical equations or inequations.

The general mathematical form of L.P.P.

The L.P.P. can be put in the following form. Maximize $z = c_1x_1 + c_2x_2$ .... (1).

subject to the constraints.

$$\left.\begin{array}{l} a_{11}x_1 + a_{12}x_2 \le b_1 \\ a_{21}x_1 + a_{22}x_2 \le b_2 \\ \ldots\ldots\ldots\ldots\ldots\ldots \\ \ldots\ldots\ldots\ldots\ldots\ldots \\ a_{m1}x_1 + a_{m2}x_2 \le b_m \end{array}\right\} \qquad (2)$$

and each $x_i \ge 0$ for $i = 1, 2$ (3)

1) The linear function in (1) is called the objective function.
2) Conditions in (3) are called non-negativity constraints.

Note:

i) We shall study L.P.P. with only two variables.

ii) We shall restrict ourselves to L.P.P. involving non-negativity constraints.

SOLVED EXAMPLES

Ex. 1:

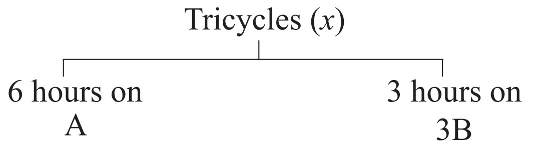
A manufacturer produces bicycles and tricycles, each of which must be processed through two machines, A and B. Machine A has maximum of 120 hours available and machine B has a maximum of 180 hours available. Manufacturing a bicycle requires 4 hours on machine A and 10 hours on machine B. Manufacturing a tricycle requires 6 hours on machine A and 3 hours on machine B. If profits are Rs. 65 for a bicycle and Rs. 45 for a tricycle, formulate L.P.P. to maximize profit.

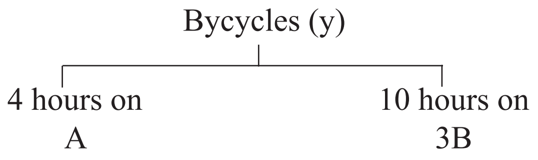
Solution:

Let Z be the profit, which can be made by manufacturing and selling x tricycles and y bicycles. $x \ge 0, y \ge 0$.

Total profit $= z = 45x + 65y$

Maximize $Z = 45x + 65y$





Mchine	Tricyles (x)	Bicycles (y)	Availability
A	6	4	120
B	3	10	180

From the above table, remaining conditions are

$$\left.\begin{array}{l} 6x + 4y \le 120, \\ 3x + 10y \le 180, \end{array}\right\}$$

$\therefore$ The required formulated L.P.P. is as follows.

Maximize

$z = 45x + 65y$ (objective function)

Subject to $\left.\begin{array}{l} 6x + 4y \le 120, \\ 3x + 10y \le 180, \end{array}\right\}$ (Constraints)

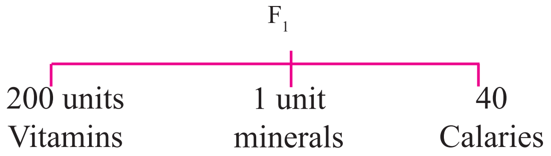
$x, y \ge 0 \rightarrow$ (non negativy Constraints)

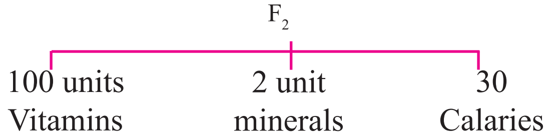
Ex. 2:

Diet for a sick person must contain at least 4000 units of vitamins, 50 units of minerals and 1500 calories. Two foods F1 and F2 cost Rs. 50 and Rs. 75 per unit respectively. Each unit of food F1 contains 200 units of Vitamins, 1 unit of minerals and 40 calories, whereas each unit of food F2 contain 100 units of vitamins, 2 units of minerals and 30 calories. Formulate the above problem as L.P.P. to satisfy sick person's requirements at minimum cost.

Solution: Let x units of food F1 and y units of food F2 be fed to sick persons to meet his requirements at minimum cost

$\therefore x \geq 0 , y \geq 0$





Food/ Product	F_1 (x) Per Unit	F_2 (y) Per Unit	Minimum requirement
Vitamin	200	100	4000
Minerals	1	2	50
Calories	40	30	1500
Cost/Unit Rs.	50	75	

sick person's problem is to determine x and y so as to minimize the total cost.

Total cost $= z = 50x + 75y$

Minimize $z = 50x + 75y$

The remaining conditions are

$200x + 100y \geq 4000$

$x + 2y \geq 50$

$40x + 30y \geq 1500$

where x, y denote units of food F_1 and F_2 respectively.

$\therefore \quad x, y \geq 0$

$\therefore$ The L.P.P. is as follows.

Minimize $z = 50x + 75y$ subject to the constraints

$200x + 100y \geq 4000,$

$x + 2y \geq 50,$

$40x + 30y \geq 1500,$

$x \geq 0, y \geq 0.$

Ex. 3:

Rakesh wants to invest at most Rs. 45000/- in savings certificates and fixed deposits. He wants to invest at least Rs. 5000/- in savings certificates and at least Rs. 15000/- in fixed deposits. The rate of interest on savings certificates is 4% p. a. and that on fixed deposits is 7% p.a. Formulate the above problem as L.P.P. to determine maximum yearly income.

Solution:

Let Rakesh invest Rs. x in savings certificate and Rs. y in fixed deposits

$\therefore x \geq 0, y \geq 0$

Since he has at most Rs. 4500/- to invest, from the given conditions, $x + y \leq 45{,}000$.

$x \geq 5000$ and $y \geq 15000$

The rate of interest on savings certificate is 4% p.a. and that on fixed deposits is 7% p.a.

$\therefore$ Total annual income $= z = 0.04x + 0.07y$

$\therefore$ The L.P.P. is

Maximize $z = 0.04x + 0.07y$

subject to

$x \geq 5000, y \geq 15000,$

$x + y \leq 45000,$

$x \geq 0, y \geq 0.$

EXERCISE 6.1

1) A manufacturing firm produces two types of gadgets A and B, which are first processed in the foundry and then sent to machine shop for finishing. The number of man hours of labour required in each shop for production of A and B and the number of man hours available for the firm are as follows.

Gadgets	Foundry	Machine Shops
A	10	5
B	6	4
Time available (hours)	60	35

Profit on the sale of A is Rs. 30 and B is Rs. 20 per unit Formulate the LPP to have maximum profit.

2) In a cattle breeding firm, it is prescribed that the food ration for one animal must contain 14, 22 and 1 unit of nutrients A, B and C respectively Two different kinds of fodder are available. Each unit weight of these two contains the following amounts of these three nutrients.

Fodder Nutrients	Fodder 1	Fodder 2
Nutrients A	2	1
Nutrients B	2	3
Nutrients C	1	1

The cost of fodder 1 is Rs. 3 per unit and that of fodder 2 is Rs. 2 per unit. Formulate the LPP to minimize the cost.

3) A Company manufactures two types of chemicals A and B. Each chemical requires two types of raw material P and Q. The table below shows number of units of P and Q required to manufacture one unit of A and one unit of B.

Chemical Raw Material	A	B	Availability
P	3	2	120
Q	2	5	160

The company gets profits of Rs. 350/- and Rs. 400/- by selling one unit of A and one unit of B respectively. Formulate the problem as LPP to maximize the profit.

4) A printing company prints two types of magazines A and B. The company earns Rs. 10 and Rs. 15 on magazine A and B per copy. These are processed on three machine I, II, III. Magazine A requires 2 hours on machine I, 5 hours on machine II, and 2 hours on machine III. Magazine B requires 3 hours on machine I, 2 hours on machine II and 6 hours on machine III, Machines I, II, III are available for 36, 50, 60 hours per week respective. Formulate the Linear programming problem to maximize the profit.

5) A manufacturer produces bulbs and tubes. Each of these must be processed through two machines. M_1 and M_2.

A package of bulbs require 1 hour of work on machine M_1 and 3 hours of work on M_2. A package of tubes require 2 hours on machine M_1 and 4 hours machine M_2.

He earns a profit of Rs. 13.5 per package of bulbs and Rs. 55 per package of tubes. Formulate the LPP to maximize the profit.

6) A Company manufactures two types of fertilizers F1 and F_2. Each type of fertilizer requires two raw materials A and B. The number of units of A and B required to manufacture one unit of fertilizer F_1 and F_2

and availability of the raw materials A and B per day are given in the table below

Fertilizer Raw Materials	F_1	F_2	Availability
A	2	3	40
B	1	4	70

By selling one unit of F_1 and one unit of F_2, company get a profit of Rs. 500 and Rs. 750 respectively. Formulate the problem as LPP to maximize the profit.

7) A doctor has prescribed two different kinds of feeds A and B to form a weekly diet for sick person. The minimum requirement of fats, carbohydrates and proteins are 18, 28, 14 units respectively. One unit of food A has 4 units of fat, 14 units of carbohydrates and 8 units of protein. One unit of food B has 6 units of fat, 12 units of carbohydrates and 8 units of protein. The price of food A is Rs. 4.5 per unit and that of food B is Rs. 3.5 per unit. Form the LPP so that the sick person's diet meets the requirements at a minimum cost.

8) If John drives a car at a speed of 60 kms/hour he has to spend Rs. 5 per km on petrol. If he drives at a faster speed of 90 km/hour, the cost of petrol increases to Rs. 8 per km. He has Rs. 600 to spend on petrol and wishes to travel the maximum distance within an hour. Formulate the above problem as LPP.

9) The company makes concrete bricks made up of cement and sand. The weight of a concrete brick has to be at least 5 kg. cement costs Rs. 20 per kg. and sand cost Rs. 6 per kg. strength consideration dictate that a concrete brick should contain minimum 4kg of cement and not more that 2 kg of sand. Formulate the LPP for the cost to be minimum.

6.2.1 Convex set and feasible region.

Definition: A set of points in a plane is said to be a convex set if the line segment joining any two points of the set entirely lies within the same set.

The following sets are convex sets

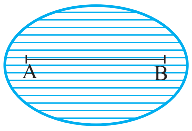


Fig. 6.1

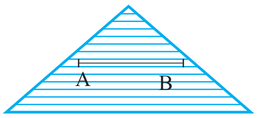


Fig. 6.2

The following sets are not convex sets

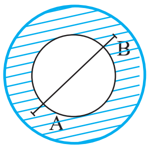


Fig. 6.3

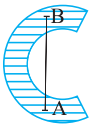


Fig. 6.4

Note :

i) The convex sets may be bounded. Following are bounded convex sets.

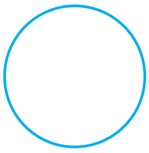


Fig. 6.5

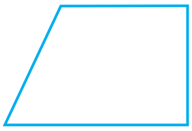


Fig. 6.6

ii) Convex sets may be unbounded Following are unbounded conves sets

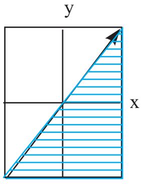


Fig. 6.7

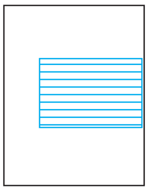


Fig. 6.8

Solution of LLP:

There are two methods to find the solution of L.P.P. 1) Graphical method 2) Simplex method.

Note : We shall restrict ourselves to graphical method.

Some definitions:

1) **Solution :** A set of values of variable which satisfies all the constraints of the LPP, is called the solultion of the LPP.

2) **Feasible Solution:** Solution which satisfy all constraints is called feasible solution.

3) **Optimum feasible solution:** A feasible solution which optimizes i.e. either maximizes or minimizes the objective function of LPP is called optimum feasible solution.

4) **Feasible Region:** The common region determined by all the constraints and non-negativity restrictions of the linear programming problem is called the feasible region.

Note: The boundaries of the region may or may not be included in the feasible region.

Theorems (without proof)

Theorem 1 : The set of all feasible solutions of LPP is a convex set.

Convex polygon theorem:

Theorem 2 : The objective function of LPP attains its optimum value (either maximum or minimum) at, at least one of the vertices of convex polygon.

Note: If a LPP has optimum solutions at more than one point then the entire line joining those two points will give optimum solutions. Hence the problem will have infinite solutions.

Solution of LPP by Corner point method (convex polygon theorem) Algorithm:

Step:

i) Convert all inequation of the constraints into equations.

ii) Draw the lines in *xy* plane, by using *x* intercept and y intercept of the line from its equation.

iii) Locate common region indicated by the constraints. This common region is called feasible region.

iv) Find the vertices of the feasible region.

v) Find the value of the objective function *z* at all vertices of the feasible region.

vi) If the objective function is of maximization (or minimization) type, then the co-ordinates, of the vertex (Vertices) for which *z* is maximum (or minimum) gives (give) the optimum solution/solutions.

SOLVED EXAMPLES

Ex.1. Maximize $z = 9x + 13y$ Subject to

$2x + 3y \leq 18$, $2x + y \leq 10$, $x \geq 0$, $y \geq 0$

Solution: To draw $2x + 3y \leq 18$, and $2x + y \leq 10$

Draw lines $2x + 3y = 18$, and $2x + y = 10$.

Equation of line	Intercept	Constraint type	Feasible Region
$2x + 3y = 18$	$x : 9$ $y : 6$	$\leq$	Originside
$2x + y = 10$	$x : 5$ $y : 10$	$\leq$	Originside

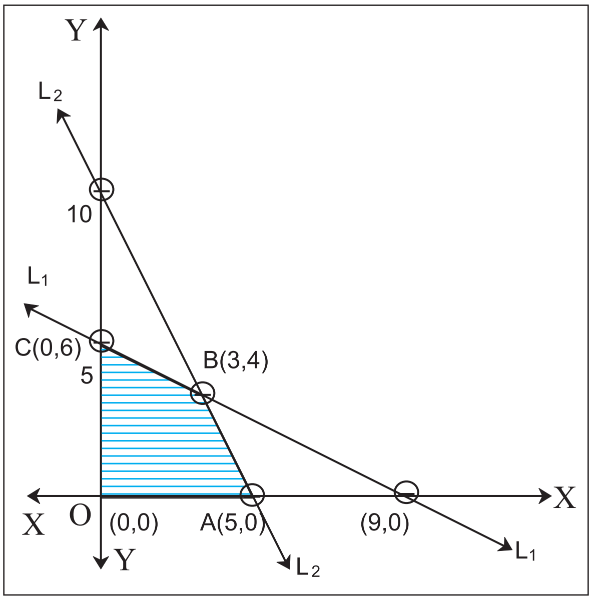


Fig. 6.9

The common shaded region OABCO is the feasible region with vertices 0(0,0), A(5,0), B(3,4), C (0,6).

Vertex	**Lines through vertex**	**Value of objective**
A (5, 0)	$2x + y = 10$ $y = 0$	45
B (3, 4)	$2x + 3y = 18$ $2x + y = 10$	79 Maximum
C (0, 6)	$2x + 3y = 18$ $x = 0$	78
O (0, 0)	$x = 0$ $y = 0$	0

From the table, maximum value of $z = 79$ occurs at B (3, 4) i.e. when $x = 3$, $y = 4$.

Ex. 2. Solve graphically the following LPP Minimize $z = 5x + 2y$ subject to

$5x + y \geq 10$, $x + y \geq 6$, $x \geq 0$, $y \geq 0$.

Solution: To draw $5x + y \geq 10$, and $x + y \geq 6$

Draw lines $5x + y = 10$, $x + y = 6$.

Equation of line	Intercept	Constraint type	Feasible Region
$5x + y = 10$	$x : 2$ $y : 10$	$\geq$	Non-originside
$x + y = 6$	$x : 6$ $y : 6$	$\geq$	Non-originside

The common shaded region is feasible region with vertices A(6,0), B (1, 5), C (0,10)

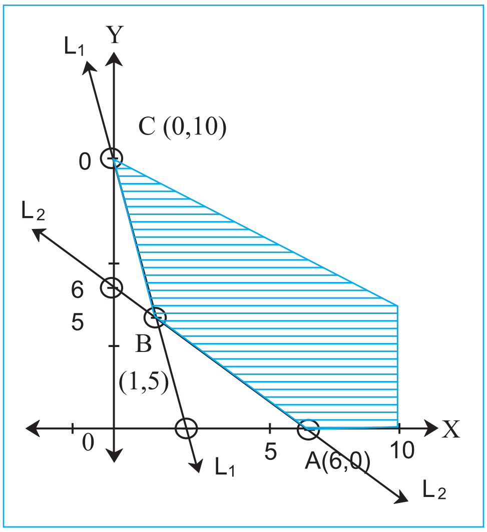


Fig. 6.10

(x, y)	Value of $z = 5x + 2y$ at (x, y)
A (6, 0)	30
B (1, 5)	15
C (0, 10)	20

From the table, we observe the following.

The minimum value of $z = 15$ occurs at B (1,5) ie. when $x = 1, y = 5$.

Ex. 3. Maximize $z = 3x + 4y$ Subject to

$x - y \geq -1,\ 2x - y \leq 2, x \geq 0, y \geq 0.$

Solution: To draw $x - y \geq -1, 2x - y \leq 2,$

Draw lines $x - y = -1, 2x - y = 2.$

Vertex	Lines through vertex	Value of objective
(3, 4)	$x - y = -1$ $2x - y = 2$	25
(1, 0)	$2x - y = 2$ $y = 0$	3

From graph, we can see that the common shaded area is feasible region.

which is unbounded (not a closed polygon)

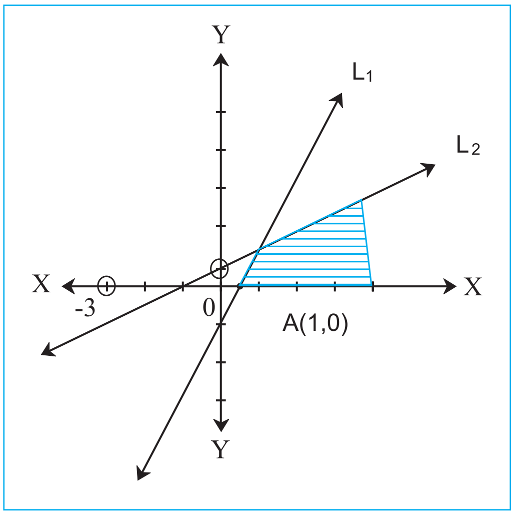


Fig. 6.11

∴ There is no finite maximum value of z since the feasible region is unbounded.

EXERCISE 6.2

Solve the following LPP by graphical method

1. Maximize $z = 11x + 8y$ Subject to
   $x \leq 4, y \leq 6\ x + y\ \leq 6, x \geq 0, y \geq 0.$
2. Maximize $z = 4x + 6y$ Subject to
   $3x + 2y \leq 12\ x + y \geq 4, x, y \geq 0.$
3. Maximize $z = 7x + 11y$ Subject to
   $3x + 5y \leq 26,\ \ 5x + 3y \leq 30, x \geq 0, y \geq 0.$
4. Maximize $z = 10x + 25y$
   Subject to $0 \leq x \leq 30 \leq y \leq 3, x + y \leq 5.$
5. Maximize $z = 3x + 5y$
   Subject to $x + 4y \leq 24,\ 3x + y \leq 21$
   $x + y \leq 9, x \geq 0, y \geq 0.$
6. Minimize $z = 7x + y$ Subject to
   $5x + y \geq 5, x + y \geq 3, x \geq 0, y \geq 0$
7. Minimize $z = 8x + 10y$ Subject to
   $2x + y \geq 7,\ 2x + 3y \geq 15, y \geq 2, x \geq 0, y \geq 0$
8. minimize $z = 6x + 2y$ Subject to $x + 2y \geq 3,$
   $x + 4y \geq 4\ 3x + y \geq 3\ x \geq 0, y \geq 0,$

Working Rule to formulate the LPP.

Step1: Identify the decision variables and assign the symbols x, y or x_1, x_2 to them. Introduce non-negativity constraints.

Step2: Identify the set of constraints and express them as linear inequations in terms of the decision variables.

Step3: Identify the objective function to be optimized (ie. maximized on minimized) and express it as a linear function of decisions variables.

* Let R be the feasible region (convex polygon) for a LPP and let z = ax + by be the objective function then the optimum value (maximum or minimum) of z occurs at, at least one of the corner points (vertex) of the feasible region.

Corner point method for solving the LPP graphically.

Step1: Find the feasible region of the LPP.

Step2: Determine the vertices of the feasible region either by inspection or by solving the two equations of the lines intersecting at that points.

Step 3: Find the value of the objective function z, at all vertices of feasible region.

Step 4: Determine the feasible solution which optimizes the value of the objective function.

Working rule to formulate and solve the LPP Graphically.

Identify the decision variables and assign the symbols x, y or x_1, x_2 to them.

Identify the objective function (maximized or minimized) and express it as a linear function of decision variables.

Convert inequations (constraints) into equations, find out intercept points on them.

Draw the graph.

Identify the feasible region (convex polygon) of the L.P.P. and shade it.

Find all corner points of the feasible region.

Find the value of z at all the corner points.

State the optimum value of z (maximum or minimum).

MISCELLANEOUS EXERCISE - 6

I Choose the correct alternative.

1. The value of objective function is maximize under linear constraints.
   a) at the centre of feasible region
   b) at (0, 0)
   c) at any vertex of feasible region.
   d) The vertex which is at maximum distance from (0, 0).

2. Which of the following is correct?
   a) Every LPP has on optional solution
   b) Every LPP has unique optional solution.
   c) If LPP has two optional solution the it has infinitely many solutions.
   d) The set of all feasible solutions of LPP may not be a convex set.

3. Objective function of LPP is
   a) a constraint
   b) a function to be maximized or minimized
   c) a relation between the decision variables
   d) a feasible region.

4. The maximum value of $z = 5x + 3y$.
   subject to the constraints $3x + 5y = 15$;
   $5x + 2y \leq 10, x, y \geq 0$ is.
   a) 235 b) 235/9
   c) 235/19 d) 235/3

5. The maximum value of $z = 10x + 6y$, subjected to the constraints $3x + y \leq 12$,
   $2x + 5y \leq 34,\ x \geq 0, y \geq 0$ is.
   a) 56 b) 65
   c) 55 d) 66

6. The point at which the maximum value of $z = x + y$ subject to the constraints $x + 2y \leq 70,\ 2x + y \leq 15,\ x \geq 0,\ y \geq 0$ is
   a) (36, 25) b) (20, 35)
   c) (35, 20) d) (40, 15)

7. Of all the points of the feasible region the optimal value of z is obtained at a point
   a) inside the feasible region.
   b) at the boundary of the feasible region.
   c) at vertex of feasible region.
   d) on x - axis.

8. Feasible region; the set of points which satify.
   a) The objective function.
   b) All of the given function.
   c) Some of the given constraints
   d) Only non-negative constrains

9. Solution of LPP to minimize $z = 2x + 3y$ st. $x \geq 0, y \geq 0, 1 \leq x + 2y \leq 10$ is
   a) $x = 0, y = 1/2$ b) $x = 1/2, y = 0$
   c) $x = 1, y = -2$ d) $x = y = 1/2$.

10. The corner points of the feasible region given by the inequations $x + y \leq 4$, $2x + y \leq 7, x \geq 0, y \geq 0$, are
   a) (0, 0), (4, 0), (3, 1), (0, 4).
   b) (0, 0), (7/2, 0), (3, 1), (0, 4).
   c) (0, 0), (7/2, 0), (3, 1), (5, 7).
   d) (6, 0), (4, 0), (3, 1), (0, 7).

11. The corner points of the feasible region are (0, 0), (2, 0), (12/7, 3/7) and (0,1) then the point of maximum $z = 6.5x + y = 13$
   a) (0, 0) b) (2, 0)
   c) (11/7, 3/7) d) (0, 1)

12. If the corner points of the feasible region are (0, 0), (3, 0), (2, 1) and (0, 7/3) the maximum value of $z = 4x + 5y$ is .
   a) 12 b) 13
   c) 35/2 d) 0

13. If the corner points of the feasible region are (0, 10), (2, 2), and (4, 0) then the point of minimum $z = 3x + 2y$ is.
   a) (2, 2) b) (0, 10)
   c) (4, 0) d) (2, 4)

14. The half plane represented by $3x + 2y \leq 0$ constraints the point.
   a) (1, 5/2) b) (2,1)
   c) (0, 0) d) (5, 1)

15. The half plane represented by $4x + 3y \geq 14$ contains the point
   a) (0, 0) b) (2, 2)
   c) (3, 4) d) (1, 1)

II Fill in the blanks.

1) Graphical solution set of the in equations $x \geq 0, y \geq 0$ is in ....quadrant

2) The region represented by the in equations $x \geq 0,\ y \geq 0$ lines in .... quadrants

3) The optimal value of the objective function is attained at the .....points of feasible region.

4) The region represented by the inequality $y \leq 0$ lies in ....quadrants

5) The constraint that a factory has to employ more women (y) than men (x) is given by........

6) A garage employs eight men to work in its showroom and repair shop. The constants that there must be not least 3 men in showroom and repair shop. The constrains that there must be at least 3 men in showroom and at least 2 men in repair shop are.........and ....respectively

7) A train carries at least twice as many first class passengers (y) as second class passengers (x) The constraint is given by......

8) A dish washing machine holds up to 40 pieces of large crockery (x) This constraint is given by...........

III State whether each of the following is True or False.

1) The region represented by the inequalities $x \geq 0$, $y \geq 0$ lies in first quadrant.

2) The region represented by the in qualities $x \leq 0$, $y \leq 0$ lies in first quadrant.

3) The optimum value of the objective function of LPP occurs at the center of the feasible region.

4) Graphical solution set of $x \leq 0, y \geq 0$ in xy system lies in second quadrant.

5) Saina wants to invest at most Rs. 24000 in bonds and fixed deposits. Mathematically this constraints is written as $x + y \leq 24000$ where x is investment in bond and y is in fixed deposits.

6) The point (1, 2) is not a vertex of the feasible region bounded by $2x + 3y \leq 6$, $5x + 3y \leq 15$, $x \geq 0$, $y \geq 0$.

7) The feasible solution of LPP belongs to only quadrant I The Feasible region of graph $x + y \leq 1$ and $2x + 2y \geq 6$ exists.

IV) Solve the following problems.

1) Maximize $z = 5x_1 + 6x_2$ Subject to

$2x_1 + 3x_2 \leq 18,\ 2x_1 + x_2 \leq 12,\ x \geq 0,\ y \geq 0$

2) Minimize $z = 4x + 2y$ Subject to

$3x + y \geq 27,\ x + y \geq 21,\ x \geq 0,\ y \geq 0$

3) Maximize $z = 6x + 10y$ Subject to

$3x + 5y \leq 10,\ 5x + 3y \leq 15,\ x \geq 0,\ y \geq 0$

4) Minimize $z = 2x + 3y$ Subject to

$x - y \leq 1,\ x + y \geq 3,\ x \geq 0,\ y \geq 0$

5) Maximize $z = 4x_1 + 3x_2$ Subject to

$3x_1 + x_2 \leq 15,\ 3x_1 + 4x_2 \leq 24,\ x \geq 0,\ y \geq 0$

6) Maximize $z = 60x + 50y$ Subject to

$x + 2y \leq 40,\ 3x + 2y \leq 60,\ x \geq 0,\ y \geq 0$

7) Minimize $z = 4x + 2y$ Subject to

$3x + y \geq 27,\ x + y \geq 21,\ x + 2y \geq 30$

$x \geq 0,\ y \geq 0$

8) A carpenter makes chairs and tables profits are Rs. 140 per chair and Rs. 210 per table Both products are processed on three machine, Assembling, Finishing and Polishing the time required for each product in hours and availability of each machine is given by following table.

Product / Machines	Chair (x)	Table (y)	Available time (hours)
Assembling	3	3	36
Finishing	5	2	50
Polishing	2	6	60

Formulate and solve the following Linear programming problems using graphical method.

9) A company manufactures bicyles and tricycles, each of which must be processed through two machines A and B Maximum availability of machine A and B is respectively 120 and 180 hours. Manufacturing a bicycle requires 6 hours on machine A and 3 hours on machine B Manufacturing a tricycle requires 4 hours on machine A and 10 hours on machine B. If profits are Rs. 180 for a bicycle and Rs. 220 on a tricycle, determine the number of bicycles and tricycles that should be manufacturing in order to maximize the profit.

10) A factory produced two types of chemicals A and B The following table gives the units of ingredients P & Q (per kg) of Chemicals A and B as well as minimum requirements of P and Q and also cost per kg. of chemicals A and B.

Chemicals Units / Ingredients per kg.	A (x)	B (y)	Minimum requirements in
P	1	2	80
Q	3	1	75
Cost (in Rs.)	4	6	

Find the number of units of chemicals A and B should be produced so as to minimize cost.

11) A Company produces mixers and processors Profit on selling one mixer and one food processor is Rs. 2000 and Rs. 3000 respectively. Both the products are processed through three machines A, B, C The time required in hours by each product and total time available in hours per week on each machine are as follows:

Product / Machine	Mixer per unit	Food processor per unit	Available time
A	3	3	36
B	5	2	50
C	2	6	60

How many mixers and food processors should be produced to maximize the profit?

12) A Chemical company produces a chemical containing three basic elements A, B, C so that it has at least 16 liters of A, 24 liters of B and 18 liters of C. This chemical is made by mixing two compounds I and II. Each unit of compound I has 4 liters of A, 12 liters of B, 2 liters of C Each unit of compound II has 2 liters of A, 2 liters of B and 6 liters of C. The cost per unit of compound I is Rs. 800 and that of compound II is Rs. 640 Formulate the problem as LPP. and solve it to minimize the cost.

13) A person makes two types of gift items A and B requiring the services of a cutter and a finisher. Gift item A requires 4 hours of cutter's time and 2 hours of finisher's time. B requires 2 hours of cutters time, 4 hours of finishers time. The cutter and finisher have 208 hours and 152 hours available times respectively every month. The profit of one gift item of type A is Rs. 75 and on gift item B is Rs. 125. Assuming that the person can sell all the items produced, determine how many gift items of each type should be make every month to obtain the best returns?

14) A firm manufactures two products A and B on which profit earned per unit is Rs. 3 and Rs. 4 respectively. The product A requirs one minute of processing time on M_1 and 2 minutes on M_2. B requires one minutes on M_1 and one minute on M_2. Machine M_1 is available for use for 450 minutes while M_2 is available for 600 minutes during any working day. Find the number of units of product A and B to be manufactured to get the maximum profit.

15) A firm manufacturing two types of electrical items A and B, can make a profit of Rs. 20 per unit of A and Rs. 30 per unit of B. Both A and B make use of two essential components, a motor and a transformer. Each unit of A requires 3 motors and 2 transformers and each unit of B requires 2 motors and 4 transformers. The total supply of components per month is restricted to 210 motors and 300 transformers. How many units of A and B should be manufacture per month to maximize profit? How much is the maximum profit?

Activities

1) Find the graphical solution for the following system of linear inequations.

$8x + 5y \leq 40, \;\; 4x + 5y \leq 40, \;\; x \geq 0, y \geq 0$

Solution to draw $8x + 5y \leq 40$

Draw line L_i $8x + 5y = 40$

x	y	(x, y)	Sign	Region
	0	(, 0)	$\leq$	On origin side of line
0		(0,)		

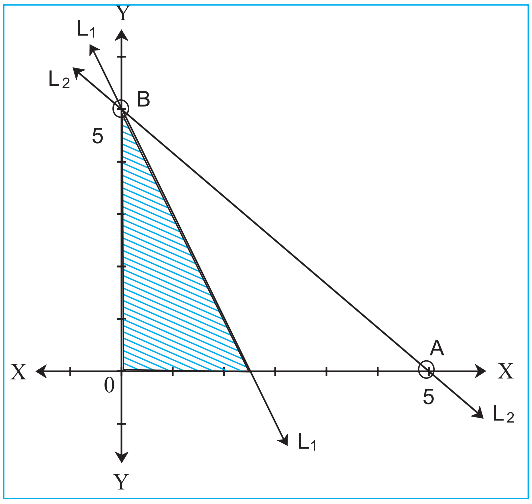


Fig. 6.12

To draw $4x + 5y \le 40$

Draw line $L_2 : 4x + 5y = 40$

x	y	(x, y)	Sign	Region
	0	(, 0)	$\le$	On origin side of line
0		(0,)		

The common shaded region OABO is graphical solution, with vertices 0(,), A (,) , B (,)

2) Find the graphical solution for the following system of linear inequations $3x + 5y \ge 15$, $2x + 5y \le 15$, $2x + 3y \le 18$, $x \ge 0$, $y \ge 0$

Solution : To Draw $3x + 5y \ge 15$

Draw line $3x + 5y = 15$.

x	y	(x, y)	Sign	Region
	0	(, 0)	$\le$	On origin side of line
0		(0,)		

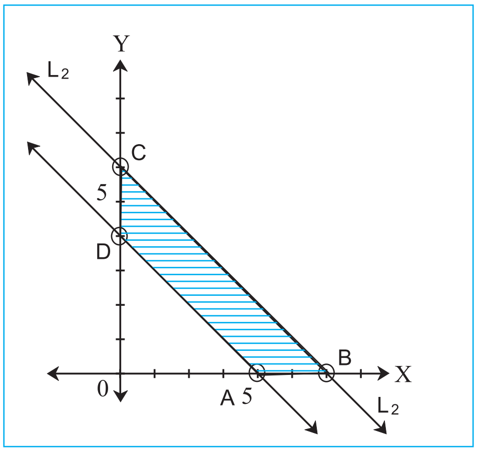


Fig. 6.13

The common shaded region OABO is graphical solution, with vertices 0 (,), A (,), B (,)

3) Shraddha wants to invest at most 25,000/- in savings certificates and fixed deposits. She wants to invest at least Rs. 10,000/- in savings certificate and at least Rs. 15,000/- in fixed deposits. The rate of interest on saving certificate is 5% per annum and that on fixed deposits is 7% per annum. Formulate the above problem as LPP to determine maximum yearly income.

Solution: Let x_1 amount (in Rs.) invest in saving cerficate

x_2: amount (in Rs.) invest in fixed deposits.

$x_1 \ge 0, x_2 \ge 0$

From given conditions $x_1 + x_2 \boxed{} 25,000$

She wants to invest at least Rs. 10000/- in saving certificate

$\therefore x_1 \boxed{} 10,000$

Shradha want to invest at least Rs. 15,000/- in fixed deposits.

$x_2 \boxed{} 15,000$

Total interest = z = ...

Maximize z = ... Subject to.

....

....

4) The graphical solution of LPP is shown by following figure. Find the maximum value of $z = 3x + 2y$ subject to the conditions given in graphical solution.

Solution: From Fig. 6.14. The common Shaded region OABCO is feasible region with vertices 0(,), A(,), B(4, 3) C(,)

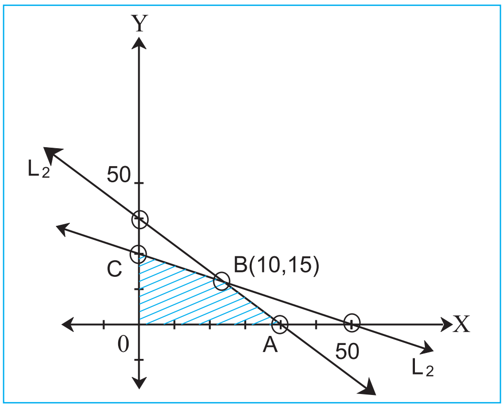


Fig. 6.14

Sr. No	(x, y)	Value of $z = 3x + 2y$ at (x, y)
1)	0(0, 0)	$z = 6$
2)	A(5, 0)	$z =$
3)	B(,)	$z =$
4)	C(0, 3)	$z = 10$

From above table, maximum value of

$z =$ ....occurs at point .... that is when $x =$ ......, $y =$ ....

5) Formulate and solve the following LPP. A company manufactures bicycles and tricycles, each of which must be processed through two machines A and B. Machine A has maximum of 120 hours available and machine B has a maximum of 180 hours available. Manufacturing a Bicycle require 6 hours on machine A and B hours on machine B. Manufacturing a tricycle requires 4 hours on machine A and to hours on machine B. If profits are Rs. 180/- for a bicycle and Rs. 220/- for a tricycle, determine the number of bicycles and tricycles that should be manufactured in order to maximize the profit.

Sol. Let x no of bicycles and y no. of tricycles be manufactured $x \geq 0$, $y \geq 0$ 1

Total profit = $z =$

Maximize $z =$

The remaining conditions are............

$\therefore$ LPP is maximize $z =$ ....

subject to $x \geq 0, y \geq 0, \ldots$

To draw $6x + 4y \leq 120$

Draw line $L_1 : 6x + 4y = 120$

x	y	(x, y)	Sign	Region
	0	(, 0)	$\leq$	origin side of line L_1
0		(0,)		

To draw $3x + 10\,y \leq 180$

Draw line $L_2 : 3x + 10y = 180$

x	y	(x, y)	Sign	Region
	0	(, 0)	$\leq$	--
0		(0,)		

The common shaded region ☐ is feasible region with vertices 0(0, 0), A(,), B(10, 15), C(0, 18).

Sr. No	(x, y)	Value of $z = 3x + 2y$ at (x, y)
1)	0(0, 0)	$z = 0$
2)	A(, 0)	$z =$
3)	B(,)	$z =$
4)	C(0, 18)	$z =$

Maximum value of $z =$ ☐ occurs at point .... that is when $x =$, $y =$

Thus company gets maximum profit $z =$ Rs. ☐ when $x =$ ☐ no of bicycles and $y =$ ☐ no of tricycles are manufactured.

◆ ◆ ◆

7 Assignment Problem and Sequencing

Let's Study

- Definition of Assignment Problem
- Assignment model
- Hungarian method of solving Assignment Problem
- Special cases of Assignment Problem
- Sequencing Problem
- Types of Sequencing Problem
- Finding an optimal sequence

Let's Recall

Linear Programming Problem

Let's :Learn

Introduction to Assignment Problem:

We often come across situations in which we have to assign n jobs to n workers. All n workers are capable of doing all jobs, but with a varying cost. Hence our task is to find the best possible assignment that gives maximum efficiency and minimum cost e.g. assigning activities to students, subjects to teachers, different routes of pizza delivery boys, salesmen to different regions, jobs to machines, products to factories, research problems to teams, vehicles and drivers to different routes etc. A problem of this nature is called an assignment problem.

7.1 Definition of Assignment Problem:

Assignment problem is a special type of problem which deals with allocation of various resources to various activities on one to one basis. It is done in such a way that the total cost or time involved in the process is minimum or the total profit is maximum.

Conditions:

i) Number of jobs is equal to number of machines or workers.

ii) Each worker or machine is assigned to only one job.

iii) Each worker or machine is independently capable of handling any job.

iv) Objective of the assignment is clearly specified (minimizing cost or maximizing profit)

Assignment Model:

Given n workers and n jobs with the cost of every worker for every job, the problem is to assign each worker to one and only one job so as to optimize the total cost.

Let C_{ij} be the cost of assigning i^{th} worker to j^{th} job, x_{ij} be the assignment of i^{th} worker to j^{th} job

and $x_{ij} = 1$, if i^{th} worker is assigned to j^{th} job

$= 0$, otherwise

Following table represents the cost of assigning n workers to n jobs.

Worker	Jobs					
	1	2	3	...	...	n
1	C_{11}	C_{12}	C_{13}	...	...	C_{1n}
2	C_{21}	C_{22}	C_{23}	...	...	C_{2n}
3	C_{31}	C_{32}	C_{33}	...	...	C_{3n}
⋮	⋮	⋮	⋮	⋮	⋮	⋮
n	C_{nl}	C_{n2}	...	...	...	C_{nn}

The objective is to make assignments that minimize the total cost.

Thus, an assignment problem can be represented by n x n matrix which covers all the n! possible ways of making assignments.

Assignment Problem is a special case of Linear Programming Problem.

Assignment problem can be expressed symbolically as follows:

Minimize $Z = \sum_{i=1}^{n}\sum_{i=1}^{n} C_{ij}x_{ij}$

Subject to constrains

$\sum_{j=1}^{n} x_{ij} = 1;\ i = 1, 2, 3, n$

(exactly one job is assigned to ithworker)

$\sum_{i=1}^{n} x_{ij} = 1;\ i = 1, 2, 3, n$

(exactly one worker is assigned to jth job

where x_{ij} takes a value 0 or 1.

Let's Discuss ...

Let us consider the following problem.

Due to neglect, your home is in serious need of repair. You approach 3 contractors for remodeling and repairing. Suppose you get quotations as shown below.

Contractors	Home Repairs (Cost)
	Price (in Rs)
Amar	27,980
Akabar	31,640
Anthony	29,330

Naturally, we may think of Amar as he is giving the best overall price. However, if we think of individual items and get the prices per repair item from our contractors, this would be more beneficial in two ways:

1. We may save time since all contractors will be working on different repair items simultaneously.
2. We may get a better price by hiring contractors based on their lowest item cost.

Take a look at the following table containing prices of individual items:

Contractors	**Home Repairs (Cost)**			**Total**
	Price (in Rs)			
	Flooring	Painting	Aluminum Sliding Window	
Amar	14440	8500	5040	Rs. 27,980
Akbar	13840	13300	4500	Rs. 31,640
Anthony	14080	11200	4050	Rs. 29,330

We wish to hire one contractor for one job to minimize both the time and cost. For example, we may choose the following:

Contractors	**Home Repairs (Cost)**			**Total**
	Price (in Rs)			
	Flooring	Painting	Aluminum Sliding Window	
Amar	14440	8500	5040	Rs. 26,390
Akbar	13840	13300	4500	
Anthony	14080	11200	4050	

By this strategy, we can minimize the total cost and also the total time required to complete the job. Though this looks simple, the problem is difficult to solve for larger number of contractors and many more repairs.

An assignment problem can be represented by n × n matrix which constitutes n! possible ways of making assignments. Finding an optimal solution by writing all the n! possible arrangements is time consuming. Hence there is a need of an efficient computational technique for solving such problems.

There is an interesting and easy method to solve this type of problems called **Hungarian Method.**

The Hungarian Method is an optimization algorithm that solves an Assignment Problem.

7.2 Hungarian Method:

Hungarian method is based on the following properties:

1) If a constant (positive or negative) is added to every element of any row or column in the given cost matrix, an assignment that minimizes the total cost in the original

matrix also minimizes the total cost in the revised matrix.

2) In an assignment problem, a solution having zero total cost of assignment is an optimal solution.

The Hungarian algorithm can be explained with the help of the following example.

Consider an example where 4 jobs need to be performed by 4 workers, one job per worker. The matrix below shows the cost of assigning a certain worker to a certain job. The objective is to minimize the total cost of assignment.

Workers	Jobs			
	J_1	J_2	J_3	J_4
W_1	62	63	50	72
W_2	57	35	49	60
W_3	21	49	15	56
W_4	18	19	78	23

Let us solve this problem by Hungarian method.

Step 1: Subtract the smallest element of each row from every element of that row.

Workers	Jobs			
	J_1	J_2	J_3	J_4
W_1	12	13	0	22
W_2	22	0	14	25
W_3	6	34	0	41
W_4	0	1	60	5

Step 2: Subtract the smallest element of each column from every element of that column.

Workers	Jobs			
	J_1	J_2	J_3	J_4
W_1	12	13	0	17
W_2	22	0	14	20
W_3	6	34	0	36
W_4	0	1	60	0

Step 3: Assign through zeros.

Workers	Jobs				
			√		
W_1	12	13	$\boxed{0}$	17	√
W_2	22	$\boxed{0}$	14	20	
W_3	6	34	$\cancel{0}$	36	√
W_4	$\boxed{0}$	1	60	$\cancel{0}$	

Observe that third row does not contain an assignment.

Step 4 :

1. Mark (√) the row (R_3).
2. Mark (√) the columns (C_3) having zeros in the marked rows.
3. Mark (√) the row (R_1) which contains assignment in marked column.
4. Draw straight lines through **marked columns** and **unmarked rows**.

Workers	Jobs				
			√		
W1	12	13	$\boxed{0}$	17	√
W2	22	$\boxed{0}$	14	20	
W3	6	34	$\cancel{0}$	36	√
W4	$\boxed{0}$	1	60	$\cancel{0}$	

All zeros can be covered using 3 lines.

Therefore, number of lines required = 3 and order of matrix = 4

Hence, the number of lines required ≠ order of matrix.

Therefore we continue with the next step to create additional zeros.

Step 4:

(i) Find the smallest uncovered element (6)

(ii) Subtract this number from all uncovered elements and add it to all elements which lie at the intersection of two lines and other elements on the lines remain unchanged.

Workers	Jobs			
	J_1	J_2	J_3	J_4
W_1	6	7	0	11
W_2	22	0	20	20
W_3	0	28	0	30
W_4	0	1	66	0

Step 5: Assigning through zeros we get.

Workers	Jobs			
	J_1	J_2	J_3	J_4
W_1	6	7	$\boxed{0}$	11
W_2	22	$\boxed{0}$	20	20
W_3	$\boxed{0}$	28	$\cancel{0}$	30
W_4	$\cancel{0}$	1	66	$\boxed{0}$

Now, each row and each column contains an assignment.

Hence, optimal solution is obtained and the optimal assignment is as follows.

Worker 1 should perform job 3, worker 2 job 2, worker 3 job 1 worker 4 job 4 i.e. $W_1 \to J_3$, $W_2 \to J_2$, $W_3 \to J_1$, $W_4 \to J_4$

Total Minimum Cost = 50+35+21+23 = Rs. 129

Steps of the Hungarian Method :

Following steps describe the Hungarian Method.

Step1. Subtract the minimum cost in each row of the cost matrix from all the elements in the respective row.

Step2. Subtract the minimum cost in each column of the cost matrix from all the elements in the respective column

Step3. Starting with the first row, examine the rows one by one until a row containing exactly single zero is found. Make an assignment by marking (☐) that zero. Then cross (×) all other zeros in the column in which the assignment was made. This eliminates the possibility of making further assignments in that column.

Step 4. After examining all the rows, repeat the same procedure for columns. i.e. examine the columns one by one until a column containing exactly one zero is found. Make an assignment by marking (☐) that zero. Then cross (×) all other zeros in the row in which the assignment was made.

Step 5. Continue these successive operations on rows and columns until all the zeros have been either assigned or crossed out and there is exactly one assignment in each row and in each column. In such case optimal solution is obtained.

Step 6. There may be some rows (or columns) without assignments i.e. the total number of marked zeros is less than the order of the cost matrix. In such case, proceed to step 7.

Step 7. Draw the least possible number of horizontal and vertical lines to cover all zeros. This can be done as follows:

i) Mark (✓) the rows in which no assignment has been made.

ii) Mark (✓) the column having zeros in the marked rows.

iii) Mark (✓) rows which contain assignments in marked columns.

iv) Repeat 2 and 3 until the chain of marking is completed.

v) Draw straight lines through marked columns.

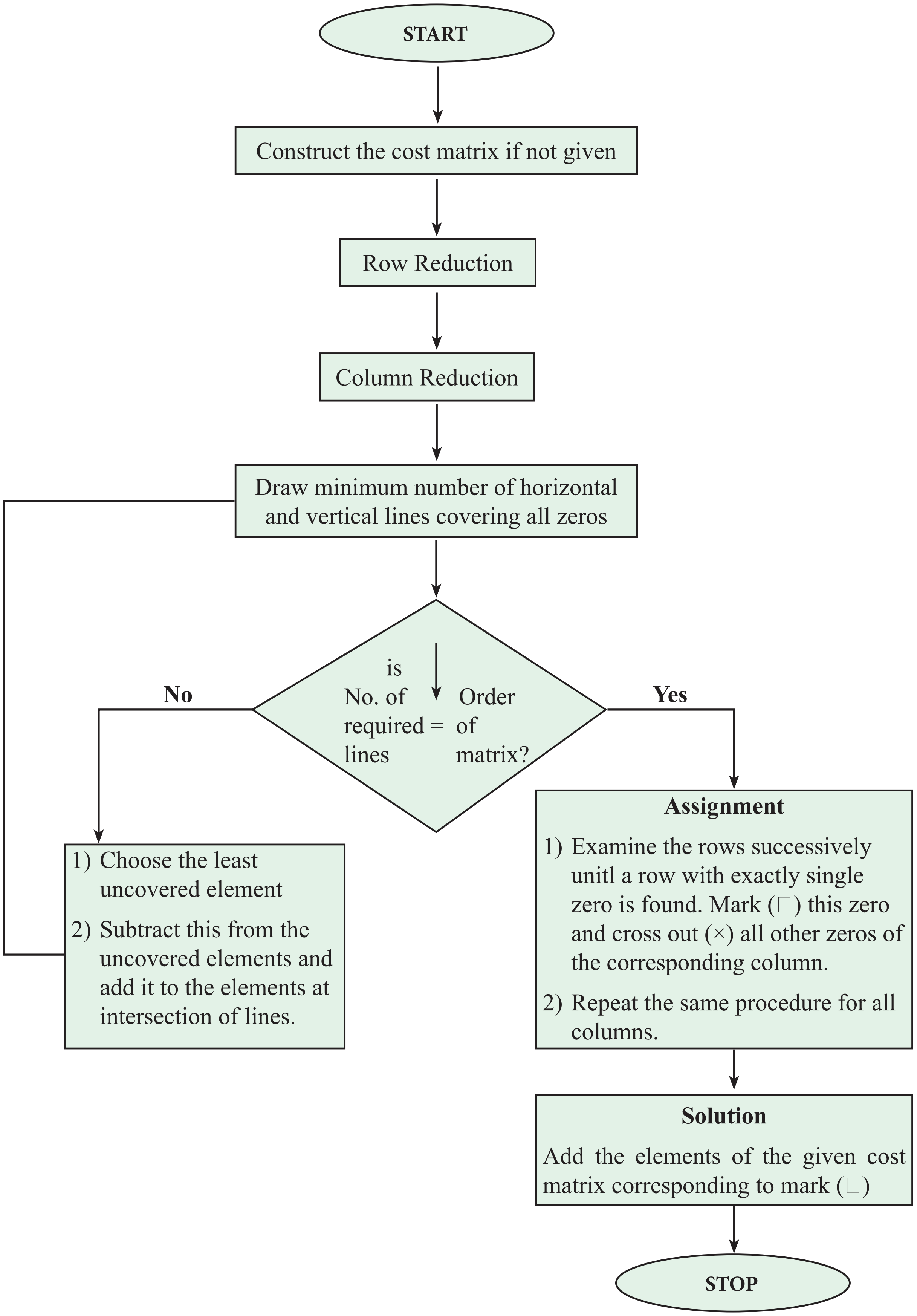
vi) Draw straight lines through unmarked rows.

By this way we draw the minimum number of horizontal and vertical lines required to cover all zeros. If the number of lines is less than the order of matrix, then there is no solution. And if the minimum number of lines is equal to the order of matrix, then there is a solution and it is optimal.

Step 8. If minimum number of lines < order of matrix, then

a) Select the smallest element not covered by any of the lines of the table.

Flow Chart of Hungarian Metod:


START
Construct the cost matrix if not given
Row Reduction
Column Reduction
Draw minimum number of horizontal and vertical lines covering all zeros
is No. of required lines = Order of matrix?
No
Yes
1) Choose the least uncovered element
2) Subtract this from the uncovered elements and add it to the elements at intersection of lines.
Assignment
1) Examine the rows successively unitl a row with exactly single zero is found. Mark (□) this zero and cross out (×) all other zeros of the corresponding column.
2) Repeat the same procedure for all columns.
Solution
Add the elements of the given cost matrix corresponding to mark (□)
STOP

b) Subtract this value from all the uncovered elements in the matrix and add it to all those elements which lie at the intersection of horizontal and vertical lines.

Step 9. Repeat steps 4, 5 and 6 until we get the number of lines covering all zeros equal to the order of matrix. In this case, optimal solution can be obtained.

Step 10. We now have exactly one marked (☐) zero in each row and each column of the cost matrix. The assignment schedule corresponding to these zeros is the optimal assignment.

Let's Note :-

The Hungarian Method was developed and published in 1955 by Harold Kuhn, who gave the name 'Hungarian Method' as the algorithm was largely based on the earlier works of two Hungarian mathematicians : Dènes Kőnig and Jenő Egervàry.

SOLVED EXAMPLES

Ex.1.

A departmental store has four workers to pack their goods. The times (in minutes) required for each worker to complete the packings per item sold is given below. How should the manager of the store assign the jobs to the workers, so as to minimize the total time of packing.

Workers	Packing of			
	Books	Toys	Crockery	Cutlery
A	0	11	10	8
B	13	2	12	2
C	3	4	6	1
D	4	15	4	9

Solution:

Let us solve this problem by Hungarian method.

Step1: Subtract the smallest element of each row from every element of that row.

Workers	Packing of			
	Books	Toys	Crockery	Cutlery
A	0	8	7	5
B	11	0	10	⊗
C	2	3	5	0
D	⊗	11	0	5

Step 2: Since all column minimums are zero, no need to subtract anything from columns.

Step 3: Assigning through zeros we get,

Workers	Packing of			
	Books	Toys	Crockery	Cutlery
A	$\boxed{0}$	8	7	5
B	11	$\boxed{0}$	10	⊗
C	2	3	5	$\boxed{0}$
D	⊗	11	$\boxed{0}$	5

∴ Optimal assignment schedule is :

A →Books, B→ Toys, C→ Cutlery, D→Crockery. Total Minimum Time = 3 + 2 + 4 + 1 = 10 minutes.

Ex.2.

Solve the following assignment problem for minimization.

Operator	Machine				
	I	**II**	**III**	**IV**	**V**
1	18	24	19	20	23
2	19	21	20	18	22
3	22	23	20	21	23
4	20	18	21	19	19
5	18	22	23	22	21

Solution:

Let us solve this problem by Hungarian method.

Step1: Subtract the smallest element of each row from every element of that row.

	I	**II**	**III**	**IV**	**V**
1	0	6	1	2	4
2	1	3	2	0	3
3	2	3	0	1	3
4	2	0	3	1	1
5	0	4	5	4	3

Step2: Subtract the smallest element of each column from every element of that column.

	I	II	III	IV	V
1	0	6	1	2	4
2	1	3	2	0	3
3	2	3	0	1	2
4	2	0	3	1	0
5	0	4	5	4	2

Step3: Draw minimum number of lines (horizontal and vertical) that are required to cover all zeros in the matrix.

	I	II	III	IV	V	
1	$\boxed{0}$	6	1	2	4	√
2	1	3	2	$\boxed{0}$	3	
3	2	3	$\boxed{0}$	1	2	
4	2	$\boxed{0}$	3	1	⊗	
5	⊗	4	5	4	2	√

Here, minimum number of lines (4) < order of matrix (5). Therefore we continue with the next step to create additional zeros.

Step 4:

(i) Find the smallest uncovered element (1)

(ii) Subtract this number from all uncovered elements and add it to all elements which lie at the intersection of two lines.

	I	II	III	IV	V	
1	⊗	5	⊗	1	3	√
2	2	3	2	$\boxed{0}$	3	
3	3	2	$\boxed{0}$	1	2	√
4	3	$\boxed{0}$	3	1	⊗	
5	$\boxed{0}$	3	4	3	1	√

Here, minimum number of lines (4) < order of matrix (5). Therefore we continue with the next step to create additional zeros.

Step5:

(i) Find the smallest uncovered element (1)

(ii) Subtract this number from all uncovered elements and add it to all elements which lie at the intersection of two lines.

(iii) Then assign through zeros.

	I	II	III	IV	V
1	$\boxed{0}$	4	⊗	⊗	2
2	3	3	3	$\boxed{0}$	3
3	3	1	$\boxed{0}$	⊗	1
4	4	$\boxed{0}$	4	⊗	⊗
5	⊗	2	4	2	$\boxed{0}$

Optimal Solution: 1→I, 2→IV, 3→III,

4→II, 5→V

Minimum Value = 18 + 18 + 20 +18 + 21 = 95

7.3 Special Cases of Assignment Problem:

The assignment problem is generally defined as a problem of minimization. In practice, some situations are like Assignment Problem but with some variations. The following four variations are more common and can be solved using the Hungarian method.

I. Unbalanced assignment problem:

An unbalanced assignment problem is one in which the number of resources is not equal to the number of activities i.e. the cost matrix of an assignment problem is not a square matrix (no. of rows ≠ no. of columns).

An unbalanced assignment problems can be balanced by adding dummy resources/tasks (row/column) with zero costs.

II. Maximization Problem:

Sometimes the assignment problem may deal with maximization of the objective function. To solve such a problem, we need to convert it to minimization so that we can solve it using Hungarian Method. This conversion to minimization problem can be done in either of the following ways:

(i) by subtracting all the elements from the largest element of the matrix

(ii) by multiplying all the elements of the matrix by '-1'

Then this equivalent minimization problem can be solved using Hungarian method.

III. Restricted assignment problem:

An assignment problem involving restrictions on allocation due to personal, technical, legal or other reasons is called a restricted assignment problem. A restricted assignment problem does not allow some worker(s) to be assigned to some job(s). It can be solved by assigning a very high cost (or infinite cost) to the restricted cells where assignment cannot be made.

IV. Alternative optimal solutions:

An alternate (multiple) solution exists for an assignment problem when the final assignment matrix contains more than the required number of zeros. In this case, assignments can be made through zeros arbitrarily, keeping in mind that each row and each column can contain only one assignment.

SOLVED EXAMPLES

Ex.1. [Unbalanceed assignment problem]

A departmental head has four subordinates, and three tasks to be performed. The subordinates differ in efficiency. Estimated time for that task would take to perform each given in the matrix below How should the tasks be allotted so as to minimize the total man hours?

Job	Man			
	M_1	M_2	M_3	M_4
A	7	2	6	3
B	3	7	5	4
C	5	4	3	7

Solution:

Step1: Observe that the number of rows is not equal to number of columns in the above matrix. Therefore it is an unbalanced assignment problem. It can be balanced by introducing a dummy job D with zero cost.

Job	Man			
	M_1	M_2	M_3	M_4
A	7	2	6	3
B	3	7	5	4
C	5	4	3	7
D	0	0	0	0

Step2: Subtract the smallest element of each row from every elements of that row.

Job	Man			
	M_1	M_2	M_3	M_4
A	5	0	4	1
B	0	4	5	1
C	2	1	0	4
D	0	0	0	0

Step3: Since all the column minimums are zeros, no need to subtract anything from columns.

Step4: Assigning through zeros we get,

Job	Man			
	M_1	M_2	M_3	M_4
A	5	$\boxed{0}$	4	1
B	$\boxed{0}$	4	2	1
C	2	1	$\boxed{0}$	4
D	⊗	⊗	⊗	$\boxed{0}$

Optimal Solution:

Job	Man	Man hours
A	M_2	2
B	M_1	3
C	M_3	3
	Total	8

Ex.2. [Maximization Case and Alternative Optimal Solutions]

A marketing manager has list of salesmen and towns. Considering the capabilities of the salesmen and the nature of towns, the marketing manager estimates amounts of sales per month (in thousand rupees) for each salesman in each town. Suppose these amounts are as follows:

Salesman	Town				
	T_1	T_2	T_3	T_4	T_5
S_1	37	43	45	33	45
S_2	45	29	33	26	41
S_3	46	32	38	35	42
S_4	27	43	46	41	41
S_5	34	38	45	40	44

Find the assignment of salesmen to towns that will result in maximum sale.

Solution: The above maximization problem can be converted into the equivalent minimization problem by subtracting all the matrix elements from the largest element which is 46. Then the resulting matrix is.

Salesman	**Town**				
	T_1	T_2	T_3	T_4	T_5
S_1	9	3	1	13	1
S_2	1	17	13	20	5
S_3	0	14	8	11	4
S_4	19	3	0	5	5
S_5	12	8	1	6	2

Now, we can solve this problem by Hungarian method.

Step1: Subtract the smallest element of each row from every element of that row.

Salesman	**Town**				
	T_1	T_2	T_3	T_4	T_5
S_1	8	2	0	12	0
S_2	0	16	12	19	4
S_3	0	14	8	11	4
S_4	19	3	0	5	5
S_5	11	7	0	5	1

Step2: Subtract the smallest element of each column from every element of that column.

	T_1	T_2	T_3	T_4	T_5
S_1	8	0	0	7	0
S_2	0	14	12	14	4
S_3	0	12	8	6	4
S_4	19	1	0	0	5
S_5	11	5	0	0	1

Step3: Draw minimum number of lines (horizontal and vertical) that are required to cover all zeros in the matrix.

	T_1	T_2	T_3	T_4	T_5
S_1	8	0	0	7	0
S_2	0	14	12	14	4
S_3	0	12	8	6	4
S_4	19	1	0	0	5
S_5	11	5	0	0	1

Therefore, number of lines required (4) < order of matrix (5)

Therefore we continue with next step to create additional zeros.

Step 4:

(i) Find the smallest uncovered elements (4).

(ii) Subtract this number from all uncovered elements and add it to all elements which lie at the intersection of two lines.

	T_1	T_2	T_3	T_4	T_5
S_1	12	0	0	7	0
S_2	0	10	8	10	0
S_3	0	8	4	2	0
S_4	23	1	0	0	5
S_5	15	5	0	0	1

Step5: We return to step 3 i.e. again we determine the minimum number of lines required to cover all zeros in the matrix.

	T_1	T_2	T_3	T_4	T_5
S_1	12	0	0	7	0
S_2	0	10	8	10	0
S_3	0	8	4	2	0
S_4	23	1	0	0	5
S_5	15	5	0	0	1

Number of lines required (5) = Order of matrix

Therefore optimal assignment can be made.

Optimal assignment: Optimal assignment can be made through zeros.

Note that after assigning $S_1 \rightarrow T_1$, each row and column more than one zeros. Therefore alternate optimal solutions exist. Assigning through zeros in different ways, we get two different assignments:

(i)

	T_1	T_2	T_3	T_4	T_5
S_1	12	$\boxed{0}$	⊗	7	⊗
S_2	⊗	10	8	10	$\boxed{0}$
S_3	$\boxed{0}$	8	4	2	⊗
S_4	23	1	⊗	$\boxed{0}$	5
S_5	15	5	$\boxed{0}$	⊗	1

$S_1 \rightarrow T_2$, $S_2 \rightarrow T_5$, $S_3 \rightarrow T_1$, $S_4 \rightarrow T_4$, $S_5 \rightarrow T_3$

Maximum Sale = 43 + 41 + 46 + 41 + 45

= 216 thousand rupees.

(ii)

	T_1	T_2	T_3	T_4	T_5
S_1	12	$\boxed{0}$	⊗	7	⊗
S_2	$\boxed{0}$	10	8	10	⊗
S_3	⊗	8	4	2	$\boxed{0}$
S_4	23	1	$\boxed{0}$	⊗	5
S_5	15	5	⊗	$\boxed{0}$	1

$S_1 \rightarrow T_2$, $S_2 \rightarrow T_1$, $S_3 \rightarrow T_5$, $S_4 \rightarrow T_3$, $S_5 \rightarrow T_4$

Maximum Sale = 43 + 45 + 42 + 46 + 40

= 216 thousand rupees.

- **Observe that the amount of Maximum Sale is same in both the cases.**

Ex.3. [Restricted assignment problem]

Three new machines M_1, M_2, M_3 are to be installed in a machine shop. There are four vacant places A, B, C, D. Due to limited space, machine M_2 can not be placed at B.

The cost matrix (in hundred rupees) is as follows :

Machines	Places			
	A	B	C	D
M_1	13	10	12	11
M_2	15	-	13	20
M_3	5	7	10	6

Determine the optimum assignment schedule.

Solution:

Step1:

(a) Observe that the number of rows is not equal to number of columns in the above matrix. Therefore it is an unbalanced assigned problem. It can be balanced by introducing a dummy job D with zero cost.

(b) Also, it is a restricted assignment problem. So we assign a very high cost '∞' to the prohibited cell.

Machines	Places			
	A	B	C	D
M_1	13	10	12	11
M_2	15	∞	13	20
M_3	5	7	10	6
M_4	0	0	0	0

Step2: Subtract the smallest element of each row from every element of that row.

Machines	Places			
	A	B	C	D
M_1	3	0	2	1
M_2	2	∞	0	7
M_3	0	2	5	1
M_4	0	0	0	0

Step3: Since all the column minimums are zeros, no need to subtract anything from columns.

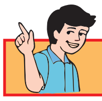
Step4: Assigning through zeros we get,

Machines	Places			
	A	B	C	D
M_1	3	$\boxed{0}$	2	1
M_2	2	∞	$\boxed{0}$	7
M_3	$\boxed{0}$	2	5	1
M_4	⊗	⊗	⊗	$\boxed{0}$

Optimal Solution:

Machnine	Place	Man hours
M_1	A	10
M_2	B	13
M_3	C	5
	Total	28

Therefore, Total Minimum Cost = 28 hundred rupees.



Let's Remember

- Assignment Problem is a special case of LPP in which every worker or machine is assigned only one job.
- Objective of the assignment is clearly specified (minimizing cost or maximizing profit).
- Hungarian Method is used to solve a minimization assignment problem.
- Special Cases of Assignment Problem:

1) **Unbalanced assignment problem:**

(No. of rows $\neq$ No of columns)

An unbalanced assignment problem can be balanced by adding dummy row/column with zero costs.

2) **Maximization Problem:**

Such problem is converted to minimization by subtracting all the elements from the largest element of the matrix. Then this can be solved by Hungarian method.

3) **Restricted assignment problem:**

It can be solved by assigning a very high cost (infinite cost) to the restricted cell.

4) **Alternative optimal solutions:**

If the final assignment matrix contains more than the required number of zeros, assign through zeros arbitrarily.

EXERCISE 7.1

1. A job production unit has four jobs A, B, C, D which can be manufactured on each of the four machines P, Q, R and S. The processing cost of each job for each machine is given in the following table:

Job	Machines (Processing Cost in Rs.)			
	I	II	III	IV
P	31	25	33	29
Q	25	24	23	21
R	19	21	23	24
S	38	36	34	40

Find the optimal assignment to minimize the total processing cost.

2. Five wagons are available at stations 1, 2, 3, 4 and 5. These are required at 5 stations I, II, III, and IV and V. The mileage between various stations are given in the table below. How should the wagons be transported so as to minimize the mileage covered?

	I	II	III	IV	V
1	10	5	9	18	11
2	13	9	6	12	14
3	3	2	4	4	5
4	18	9	12	17	15
5	11	6	14	19	10

3. Five different machines can do any of the five required jobs, with different profits resulting from each assignment as shown below:

Job	Machines (Profit in Rs.)				
	A	B	C	D	E
1	30	37	40	28	40
2	40	24	27	21	36
3	40	32	33	30	35
4	25	38	40	36	36
5	29	62	41	34	39

Find the optimal assignment schedule.

4. Four new machines M_1,M_2,M_3 and M_4 are to be installed in a machine shop. There are five vacant places A, B. C, D and E available. Because of limited space, machine M_2 cannot be placed at C and M_3 cannot be placed at A. The cost matrix is given below.

Machines	Places				
	A	B	C	D	E
M1	4	6	10	5	6
M2	7	4	-	5	4
M3	-	6	9	6	2
M4	9	3	7	2	3

Find the optimal assiment schedule.

5. A. company has a team of four salesmen and there are four districts where the company wants to start its business. After taking into account the capabilities of salesmen and the nature of districts, the company estimates that the profit per day in rupees for each salesman in each district is as below :

Salesman	District			
	1	2	3	4
A	16	10	12	11
B	12	13	15	15
C	15	15	11	14
D	13	14	14	15

Find. the assignment of salesman to various districts which will yield maximum profit.

6. In the modification of a plant layout of a factory four new machines M_1, M_2, M_3 and M_4 are to be installed in a machine shop. There are five vacant places A, B, C, D and E available. Because of limited space, machine M2 can not be placed at C and M3 can not be placed at A the cost of locating a machine at a place (in hundred rupees) is as follows.

Machines	Location				
	A	B	C	D	E
M1	9	11	15	10	11
M2	12	9	-	10	9
M3	-	11	14	11	7
M4	14	8	12	7	8

Find the optimal assignment schedule.

SEQUENCING PROBLEM

Introduction to Sequencing Problem:

Suppose we have two machines - A : Cutting and B: Sewing machine Suppose there are two items I and II to be processed on these machines in the order A-B.

The machines can handle only one job at a time and the time taken in hours by the machines to complete the jobs is given by following table.

Machine	Item	
	I	II
A	6	3
B	3	6

Then there are two ways of completing this task

(i) Processing in the order I-II

(ii) Processing in the order II-I

Case i. Let us start with item I at 0 hours. Then we get

Item	Machine			
	A		B	
	In	Out	In	Out
1	0	6	6	9

Processing of item I starts at 0 hrs and is completed at 6 hrs, Note that during this time, though machine B is idle, it can not process job II, since cutting is required completed before sewing.

Once the processing of item I is completed on machine A at 6 hours, it is shifted to machine B for sewing immediately.

Machine B being idle, item I is immediately taken for processing on machine B without any wastage of time.

Therefore, 'time in' for item I on machine B is 6 and time out is 6 + 3 = 9; where 3 hrs is the time required for the processing of item I on machine B.

While item I is being processed on machine B. Machine A is free and hence, it can take item II for processing.

Thus, item II enters machine A at 6 hrs and since it needs 3 hrs for cutting (refer to table 1) it gets out at 9 hrs from machine A.

At 9 hrs machine B is available, and hence can take, item II for processing at that time.

Item II requires 6 hrs on machine B, and will be out from machine B at 9 + 6 = 15 hrs. as shown below :

Item	Machine			
	A		B	
	In	Out	In	Out
I	0	6	6	9
II	6	9	9	15

Thus, the processing of items I and II in the order I - II takes 15 hours.

Case ii.

Let us now see what happens if we change the order of processing the two items i.e. processing item II first and item I second, (II-I)

Repeating the same process as in **case (i)** we get

Item	Machine			
	A		B	
	In	Out	In	Out
II	0	3	3	9
I	3	9	9	12

Therefore, processing of items in the order II-I takes 12 hours.

Observe that -

Order of processing the items	Time required to complete the task
I-II	15 hrs
II-I	12 hrs.

By a mere change in the order of processing of the two jobs, we could save 3 hrs.

Thus, it is very important to decide the order in which the jobs should be lined up for processing so as to complete the entire schedule in minimum time.

Such type of problem where, one has to determine the order or sequence in which the jobs are to be processed through machines so as to minimize the total processing time is called a **'sequencing problem.'**

Conditions:

1) No machine can process more than one job at a time.
2) Each job, once started, must be processed till its completion.
3) The processing times are independent of the order of processing the jobs.
4) Each machine is of different type.
5) The time required to transfer a job from one machine to another is negligible.

Terminology:

1) Total Elapsed Time:

It is the time required to complete all the jobs i.e. the entire task.

Thus, total elapsed time is the time between the beginning of the first job on the first machine till the completion of the last job on the last machine.

2) Idle Time:

Idle time is the time when a machine is available but not being used, i.e. the machine is available but is waiting for a job to be processed.

General Sequencing Problem:

Let there be `n' jobs, to be performed one at a time, on each of 'm' different machines, where the order of processing on machines and the processing time of jobs on machines is known to us. Then our aim is to find the optional sequence of processing jobs that minimizes the total processing time or cost.

Hence our job is to find that sequence out of $(n!)^m$ sequences, which minimizes the total elapsed time.

NOTATION

A_i, B_i : Processing time required by i^{th} job on machine A and machine B (i =1, 2, 3...n)

T : Total elapsed time

X_A, X_B: Idle times on machines A, B from end of $(i-1)^{th}$ job to the start of i^{th} job

Type of sequencing problems:

I. Sequencing n jobs on two machine.

II. Sequencing n jobs on three machine.

7.4.1 Sequencing n jobs on Two Machine:

Let there be 'n' job each of which is to be processed through two machines say A and B in the order AB. Let the processing time A_1, A_2, A_3, ...A_n, B_1, B_2, B_3...B_n be given.

Algorithm to find Optimal Sequence:

1) Find out Min $\{A_i, B_i\}$

2) (a) If the minimum processing time is A_r, then process r^{th} job first.

(b) If the minimum processing time is B_s, then process s^{th} job in the last.

3) Case of tie: Tie can be broken arbitrarily.

4) Cross off the jobs already placed in the sequence and repeat steps 1 to 3 till all the jobs are placed in the sequence.

5) Once the sequence is decided, prepare the work table and find total elapsed time.

SOLVED EXAMPLES

Ex.1. We have five jobs each of which has to go through the Machine M_1 and M_2 in the order M_1, M_2. Processing time (in hours) are given as:

Job	I	II	III	IV	V
Machine A	3	3	7	5	2
Machine B	6	4	2	1	5

Determine a sequence of these job that will minimize the total elapsed time T, idle time for machine M_1 and idle time for machine M_2.

Solution:

Observe that Min $\{A_i, B_i\}$ = 1, which corresponds to job IV on machine B.

Therefore, job IV is placed last in the sequence.

				IV

Then the problem reduces to :

Job	I	II	III	V
Machine A	3	3	7	2
Machine B	6	4	2	5

Now,

Min$\{A_i, B_i\}$ = 2, which corresponds to job V on machine A & job III on machine B is placed. Therefore, job V is placed first and job III is placed next to last.

V			III	IV

Then the problem reduces to:

Job	I	II
Machine A	3	3
Machine B	6	4

Now, Min$\{A_i, B_i\}$ = 3, which corresponds to job I and job II on machine A.

Therefore, job I and II can be placed after job V in any order. i.e.

V	I	II	III	IV

or

V	II	I	III	IV

Therefore, the optimal sequence is: V-I-II-III-IV or V-II-I-III-IV

(i) Total elapsed time for sequence V-I-II-III-IV

Job	Machine A		Machine B	
	In	Out	In	Out
V	0	2	2	7
I	2	5	7	13
II	5	8	13	17
III	8	15	17	19
IV	15	20	20	21

(ii) Total elapsed time for sequence V-II-I-III-IV

Job	Machine M_1		Machine M_2	
	In	Out	In	Out
V	0	2	2	7
II	5	8	11	17
I	2	5	7	11
III	8	15	17	19
IV	15	20	20	21

(Observe that, through the optimal sequence are different, total elapsed time is same i.e. 21 hrs)

$\therefore$ Total elapsed time = 21 hrs.

Idle time for machine A = T - (sum of processing times of all jobs on M_1)

= 21- 20

= 1hrs.

Idle time for machine B = T - (sum of processing times of all jobs on M_2)

= 21 - (6 + 4 + 2 + 1 + 5)

= 21 - 18

= 3 hrs.

Ex. 2. A book has one printing machine, one binding machine and manuscripts of 7 different books. The times required for performing printing and binding operations for different books are shown below:

Book	1	2	3	4	5	6	7
Printing time (hours	20	60	50	30	110	25	55
Binding time (hours)	25	40	43	24	80	35	40

Decide the optimum sequence of processing of books in order to minimize the total time required to bring out all the books.

Solution:

Let A be the printing machine and

B be the binding machine

Observe that $\text{Min}\{A_i, B_i\} = 20$, which corresponds to 1st Book on machine A

Therefore, book 1 is processed first on machine M_1.

1						

Then the problem reduces to:

Book	2	3	4	5	6	7
A	60	50	30	110	25	55
B	40	45	24	80	35	40

Now, $\text{Min}\{A_i, B_i\} = 24$, which corresponds to 4th book on machine B.

1						4

Therefore, book 4 is processed in the last.

Then the problem reduces to:

Book	2	3	5	6	7
A	60	50	110	25	55
B	40	45	80	35	40

Now, $\text{Min}\{A_i, B_i\} = 25$, which corresponds to book 6 on machine A.

1	6					4

Book	2	3	5	7
A	60	50	110	55
B	40	45	80	40

$\text{Min}\{A_i, B_i\} = 40$, which corresponds to book 2 and 7 are processed last but before 4th book in any order i.e.

1	6	3	5	2	7	4

or

1	6	3	5	7	2	4

Then the problem reduces to

Book	3	5
A	50	110
B	45	80

Now, Min $\{A_i, B_i\}$ = 45, which corresponds to book 3 on machine B. Therefor, book 3 is processed after book 6 and at the remaining.

Optional Sequence is :

1	6	3	5	2	7	4

or

1	6	3	5	7	2	4

Total elapsed time.

Job	Machine A		Machine B	
	In	Out	In	Out
1	0	20	20	45
6	20	45	45	80
3	45	95	95	140
5	95	205	205	285
2	205	265	285	325
7	265	320	325	365
4	320	350	365	389

$\therefore$ Total elapsed time = 389 hrs.

Idle time for machine A = 389 - 350 = 39 hrs.

Idle time for machine B = 20 + 15 + 65 = 100 hrs. or = 389 - 289 = 100 hrs.

7.4.2 Sequencing 'n' Jobs on Three Machines:

Let there be 'n' jobs each of which is to be processed through three machines say A, B and C in the order ABC. To solve this problem -

(i) first reduce it to the 'n job 2 machine' problem and determine the optimal sequence.

(ii) once the sequence is determined, go back to the original 3 machines and prepare the work table for 3 machines.

Conditions for reducing a 3 machine problem to a 2 machine problem:

To convert a 3 machine problem into a 2 machine problem, at least one of the following conditions must hold true.

1) The minimum processing time for machine A is greater than or equal to the maximum processing time for machine B.

   i.e. Min $A_i \geq$ Max B_i, i = 1, 2, 3, ... n

OR

2) The minimum processing time for machine C is greater than or equal to the maximum processing time for mac-hine B.

   i.e. Min $C_i, \geq$ Max B_i, i = 1, 2, 3, ... n

PROCEDURE

Step1. If either one of the above conditions holds, go to step 2. If not, the method fails.

Step 2. Introduce two fictitious machines say G and H such that

$G_i = A_i + B_i$

$H_i = B_i + C_i$, i = 1, 2, 3, ... n

Where G_i and H_i are the processing times of i^{th} job on machines G and H respectively. Now, solve the problem as n jobs 2 machines (G, H) problem as before.

SOLVED EXAMPLES

Ex.1. Determine the optimal sequence of job that minimizes the total elapsed time for the date given below (processing time on machines is given in hours). Also find total elapsed time T and the idle time for three machines.

Job	I	II	III	IV	V	VI	VII
Machine A	3	8	7	4	9	8	7
Machine B	4	3	2	5	1	4	3
Machine C	6	7	5	11	5	6	12

Solution: Here, minA=3, min C=5, and max B=5

Since min C ≥ max B is satisfied, the problem can be converted into a two machine problem.

Let G and H be two fictions machines such that G = A + B and H = B + C

Then the problem can be written as

Job	I	II	III	IV	V	VI	VII
Machine G	7	11	9	9	10	12	10
Machine H	10	10	7	16	6	10	15

Using the optimal sequence algorithm, the following optimal sequence can be obtained.

I	IV	VII	VI	II	III	V

Total elapsed time.

Job	Machine A		Machine B		Machine C	
	In	Out	In	Out	In	Out
I	0	3	3	7	7	13
IV	3	7	7	12	13	24
VII	7	14	14	17	24	36
VI	14	22	22	26	36	42
II	22	30	30	33	42	49
III	30	37	37	39	49	54
V	37	46	46	47	54	59

∴ Total elapsed time = 59 hrs.

Idle time for machine A = 59 - 46 = 13 hrs.

Idle time for machine B = 59 - 22 = 37 hrs.

Idle time for machine C = 59 - 52 = 7 hrs.

Ex. 2. Find the sequence that minimizes the total times required in performing following jobs on three machines in the order ABC. Processing times (in hrs.) are given in the following table.

Job	1	2	3	4	5
Machine A	8	10	6	7	11
Machine B	5	6	2	3	4
Machine C	4	9	8	6	5

Solution: Here, min A = 6, min C = 4 and max B = 6

Since min A ≥ max B is satisfied, the problem can be converted into a two machine problem.

Let G and h be two fictitious machines such that G = A + B and H = B + C

Then the problem can be written as.

Job	1	2	3	4	5
Machine G	13	16	8	10	15
Machine H	9	15	10	9	9

Using the optimal sequence algorithm, the following optimal sequence can be obtained.

3	2	5	4	1

Total elapsed time

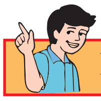
Job	Machine A		Machine B		Machine C	
	In	Out	In	Out	In	Out
3	0	6	6	8	8	16
2	6	16	16	22	22	31
5	16	27	27	31	31	36
4	27	34	34	37	37	43
1	34	42	42	47	47	51

∴ Total elapsed time = 51 hrs.

Idle time for machine A = 51 - 42 = 9 hrs.

Idle time for machine B = 6 + 8 + 5 + 3 + 5 + (51-47) = 31 hrs.

Idle time for machine C=8 + 6 +1 + 4=19hrs.



Let's Remember

- **Sequencing problem :** In sequencing problems, one has to determine the order or sequence in which the jobs are to be processed through machines so as to minimize the total processing time.
- **Total Elapsed Time:** It is the time required to complete all the jobs i.e. the entire task.
- **Idle Time:** Idle time is the time when a machine is available, but is not being used.

- **Types of sequencing problems :**
- **Sequencing n jobs on Two machines :**
- **Sequencing n jobs on Three machines:**

To convert a 3 machine problem into a 2 machine problem, at least one of the following conditions must hold true.

1) Min $A_i \geq$ Max B_i OR 2) Min $C_i \geq$ Max B_i

$i = 1, 2, 3, \ldots n$

If either one of the above conditions hold, introduce two fictitious machines say

- G and H such that : $G_i = A_i + B_i$

  $H_i = B_i + C_i$, $i = 1, 2, 3, \ldots n$

  If not, the problem cannot be solved.

EXERCISE 7.2

1. A machine operator has to perform two operations, turning and threading on 6 different jobs. The time required to perform these operations (in minutes) for each job is known. Determine the order in which the jobs should be processed in order to minimize the total time required to complete all the jobs. Also find the total processing time and idle times for turning and threading operations.

Job	1	2	3	4	5	6
Time for turning	3	12	5	2	9	11
Time for threading	8	10	9	6	3	1

2. A company has three jobs on hand. Each of these must be processed through two departments, in the order AB where

   Department A : Press shop and

   Department B : Finishing

   The table below gives the number of days required by each job in each department

Job	I	II	III
Department A	8	6	5
Department B	8	3	4

   Find the sequence in which the three jobs should be processed so as to take minimum time to finish all the three jobs. Also find idle time for both the departments.

3. An insurance company receives three types of policy application bundles daily from its head office for data entry and tiling. The time (in minutes) required for each type for these two operations is given in the following table:

Policy	1	2	3
Data Entry	90	120	180
Filing	140	110	100

   Find the sequence that minimizes the total time required to complete the entire task. Also find the total elapsed time and idle times for each operation.

4. There are five jobs, each of which must go through two machines in the order XY. Processing times (in hours) are given below. Determine the sequence for the jobs that will minimize the total elapsed time. Also find the total elapsed time and idle time for each machine.

Job	A	B	C	D	E
Machine X	10	2	18	6	20
Machine Y	4	12	14	16	8

5. Find the sequence that minimizes the total elapsed time to complete the following jobs in the order AB. Find the total elapsed time and idle times for both the machines.

Job	I	II	III	IV	V	VI	VII
Machine A	7	16	19	10	14	15	5
Machine B	12	14	14	10	16	5	7

6. Find the optimal sequence that minimizes total time required to complete the following jobs in the order ABC. The processing times are given in hrs.

(i)

Job	I	II	III	IV	V	VI	VII
Machine A	6	7	5	11	6	7	12
Machine B	4	3	2	5	1	5	3
Machine C	3	8	7	4	9	8	7

(ii)

Job	1	2	3	4	5
Machine A	5	7	6	9	5
Machine B	2	1	4	5	3
Machine C	3	7	5	6	7

7. A publisher produces 5 books on Mathematics. The books have to go through composing, printing and binding done by 3 machines P, Q, R. The time schedule for the entire task in proper unit is as follows.

Book	A	B	C	D	E
Machine P	4	9	8	6	5
Machine Q	5	6	2	3	4
Machine R	8	10	6	7	11

Determine the optimum time required to finish the entire task.

MISCELLANEOUS EXERCISE - 7

I) Choose the correct alternative.

1. In sequencing, an optimal path is one that minimizes ....
   (a) Elapsed time (b) Idle time
   (c) Both (a) and (b) (d) Ready time

2. If job A to D have processing times as 5, 6, 8, 4 on first machine and 4, 7, 9, 10 on second machine then the optimal sequence is :
   (a) CDAB (b) DBCA
   (c) BCDA (d) ABCD

3. The objective of sequencing problem is
   (a) to find the order in which jobs are to be made
   (b) to find the time required for the completing all the job on hand
   (c) to find the sequence in which jobs on hand are to be processed to minimize the total time required for processing the jobs
   (d) to maximize the cost

4. If there are n jobs and m machines, then there will be........... sequences of doing the jobs.
   (a) mn (b) m(n!)
   (c) n^m (d) $(n!)^m$

5. The Assignment Problem is solved by
   (a) Simplex method,
   (b) Hungarian method
   (c) Vector method,
   (d) Graphical method,

6. In solving 2 machine and n jobs sequencing problem, the following assumption is wrong
   (a) No passing is allowed
   (b) Processing times are known
   (c) Handling time is negligible
   (d) The time of passing depends on the order of machining

7. To use the Hungarian method, a profit maximization assignment problem requires
   (a) Converting all profits to opportunity losses
   (b) A dummy person or job
   (c) Matrix expansion
   (d) Finding the maximum number of lines to cover all the zeros in the reduced matrix

8. Using Hungarian method the optimal assignment obtained for the following assignment problem to minimize the total cost is :

Agent	Job			
	A	B	C	D
1	10	12	15	25
2	14	11	19	32
3	18	21	23	29
4	15	20	26	28

   (a) 1 — C, 2 — B , 3 — D, 4 — A
   (b) 1 — B, 2 — C, 3 — A, 4 — D
   (c) 1 — A, 2 — B, 3 — C, 4 — D
   (d) 1 — D, 2 — A, 3 — B, 4 — C

9. The assignment problem is said to be unbalance if
   - (a) Number of rows is greater than number of columns
   - (b) Number of rows is lesser than number of columns
   - (c) Number of rows is equal to number of columns
   - (d) Both (a) and (b)

10. The assignment problem is said to be balanced if
    - (a) Number of rows is greater than number of columns
    - (b) Number of rows is lesser than number of columns
    - (c) Number of rows is equal to number of columns
    - (d) If the entry of row is zero

11. The assignment problem is said to be balanced if it is a
    - (a) Square matrix (b) Rectangular matrix
    - (c) Unit matrix (d) Triangular matrix

12. In an assignment problem if number of rows is greater than number of columns then
    - (a) Dummy column is added
    - (b) Dummy row is added
    - (c) Row with cost 1 is added
    - (d) Column with cost 1 is added

13. In a 3 machine and 5 jobs problem, the least of processing times on machine A, B and C are 5, 1. and 3 hours and the highest processing times are 9, 5, and 7 respectively, then it can be converted to a 2 machine problem if order of the machines is:
    - (a) B-A-C, (b) A-B-C
    - (c) C - B - A (d) Any order

14. The objective of an assignment problem is to assign
    - (a) Number of jobs to equal number of persons at maximum cost
    - (b) Number of jobs to equal number of persons at minimum cost
    - (c) Only the maximize cost
    - (d) Only to minimize cost

II) Fill in the blanks.

1. An assignment problem is said to be unbalanced when ....
2. When the number of rows is equal to the number of columns then the problem is said to be ....assignment problem.
3. For solving an assignment problem the matrix should be a ....matrix.
4. If the given matrix is not a .... matrix, the assignment problem is called an unbalanced problem.
5. A dummy row(s) or column(s) with the cost elements as .... the matrix of an unbalanced assignment problem as a square matrix.
6. The time interval between starting the first job and completing the last. job including the idle time (if any) in a particular order by the given set of machines is called ....
7. The time for which a machine j does not have a job to process to the start of job i is called ....
8. Maximization assignment problem is transformed to minimization problem by subtracting each entry in the table from the.................. value in the table.
9. When an assignment problem has more than one solution, then it is............... optimal solution.
10. The time required for printing of four books A, B, C and D is 5, 8, 10 and 7 hours. While its data entry requires 7, 4, 3 and 6 hrs respectively. The sequence that minimizes total elapsed time is.........................

III) State whether each of the following is True or False.

1. One machine - one job is not an assumption in solving sequencing problems.

2. If there are two least processing times for machine A and machine B, priority is given for the processing time which has lowest time of the adjacent machine.
3. To convert the assignment problem into a maximization problem, the smallest element in the matrix is deducted from all other elements.
4. The Hungarian method operates on the principle of matrix reduction, whereby the cost table is reduced to a set of opportunity costs.
5. In a sequencing problem, the processing times are dependent of order of processing the jobs on machines.
6. Optimal assignments are made in the Hungarian method to cells in the reduced matrix that contain a zero.
7. Using the Hungarian method, the optimal solution to an assignment problem is found when the minimum number of lines required to cover the zero cells in the reduced matrix equals the no of persons.
8. In an assignment problem, if number of column is greater than number of rows, then a dummy column is added,.
9. The purpose of dummy row or column in an assignment problem is to obtain balance between total number of activities and total number of resources.
10. One of the assumptions made while sequencing n jobs on 2 machines is : two jobs must be loaded at a time on any machine.

PART - I

IV) Solve the following problems.

1. A plant manager has four subordinates, and four tasks to be performed. The subordinates differ in efficiency and the tasks differ in their intrinsic difficulty. This estimate of the times each man would take to perform each task is given in the effectiveness matrix below.

	I	II	III	IV
A	7	25	26	10
B	12	27	3	25
C	37	18	17	14
D	18	25	23	9

How should the tasks be allocated, one to a man, as to minimize the total man hours?

2. A dairy plant has five milk tankers, I, II, III. IV & V. These milk tankers are to be used on five delivery routes A, B, C, D & E. The distances (in kms) between the dairy plant and the delivery routes are given in the following distance matrix.

	I	II	II	IV	V
A	150	120	175	180	200
B	125	110	120	150	165
C	130	100	145	160	175
D	40	40	70	70	100
E	45	25	60	70	95

How should the milk tankers be assigned to the chilling center so as to minimize the distance travelled?

3. Solve the following assignment problem to maximize sales:

Salesmen	Territories				
	I	II	III	IV	V
A	11	16	18	15	15
B	7	19	11	13	17
C	9	6	14	14	7
D	13	12	17	11	13

4. The estimated sales (tons) per month in four different cities by five different managers are given below:

Manager	Cities			
	P	Q	R	S
I	34	36	33	35
II	33	35	31	33
III	37	39	35	35
IV	36	36	34	34
V	35	36	35	33

Find out the assignment of managers to cities in order to maximize sales.

5. Consider the problem of assigning five operators to five machines. The assignment costs are given in following table.

Operator	Machine				
	1	2	3	4	5
A	6	6	-	3	7
B	8	5	3	4	5
C	10	4	6	-	4
D	8	3	7	8	3
E	7	6	8	10	2

Operator A cannot be assigned to machine 3 and operator C cannot be assigned to machine 4. Find the optimal assignment schedule.

6. A chartered accountant's firm has accepted five new cases. The estimated number of days required by each of their five employees for each case are given below, where - means that the particular employee can not be assigned the particular case. Determine the optimal assignment of cases of the employees so that the total number of days required to complete these five cases will be minimum. Also find the minimum number of days.

Employee	Cases				
	I	II	III	IV	V
E_1	6	4	5	7	8
E_s	7	-	8	6	9
E_3	8	6	7	9	10
E_4	5	7	-	4	6
E_5	9	5	3	10	-

PART - II

1. A readymade garments manufacturer has to process 7 items through two stages of production, namely cutting and sewing. The time taken in hours for each of these items in different stages are given below:

Items	1	2	3	4	5	6	7
Time for Cutting	5	7	3	4	6	7	12
Time for Sewing	2	6	7	5	9	5	8

Find the sequence in which these items are to be processed through these stages so as to minimize the total processing time. Also find the idle time of each machine.

2. Five jobs must pass through a lathe and a surface grinder, in that order. The processing times in hours are shown below. Determine the optimal sequence of the jobs. Also find the idle time of each machine.

Job	I	II	III	IV	V
Lathe	4	1	5	2	5
Surface grinder	3	2	4	3	6

3. Find the sequence that minimizes the total elapsed time to complete the following jobs. Each job is processed in order AB.

	Jobs (Processing times in minutes						
	I	II	III	IV	V	VI	VII
Machine A	12	6	5	11	5	7	6
Machine B	7	8	9	4	7	8	3

Determine the sequence for the jobs so as to minimize the processing time. Find the total elapsed time and the idle times for both the machines.

4. A toy manufacturing company produces five types of toys. Each toy has to go through three machines A, B, C in the order ABC. The time required in hours for each process is given in the following table.

Type	1	2	3	4	5
Machine A	16	20	12	14	22
Machine B	10	12	4	6	8
Machine C	8	18	16	12	10

Solve the problem for minimizing the total elapsed time.

5. A foreman wants to process 4 different jobs on three machines: a shaping machine, a drilling machine and a tapping machine, the sequence of operations being shaping-drilling-tapping. Decide the optimal sequence for the four jobs to minimize the total elapsed time. Also find the total elapsed time and the idle time for every machine.

Job	Shaping (minutes)	Drilling (minutes)	Trapping (Minutes)
1	13	3	18
2	18	8	4
3	8	6	13
4	23	6	8

Activities... Assignment Problem

1. Given below the costs of assigning 3 workers to 3 jobs. Find all possible assignments by trial and error method.

	Jobs		
Workers	X	Y	Z
A	11	16	21
B	20	13	17
C	13	15	12

Among these assignments, find the optimal assignment that minimizes the total cost.

2. Show that the optimal solution of an assignment problem is unchanged if we add or subtract the same constant to the entries of any row or column of the cost matrix.

3. Construct a 3 × 3 cost matrix by taking the costs as the first 9 natural numbers and arranging them row wise in ascending order. Find all possible assignments that will minimize the total sum.

4. Given below the costs (in hundred rupees) of assigning 3 operators to 3 different machines. Find the assignment that will minimize the total cost. Also find the minimum cost.

Operators	Machines		
	I	II	III
A	$3i + 4j$	$2j^2 + 5i$	$5j - 3i$
B	$i^3 + 8j$	$7i + j^2$	$4i + j$
C	$2i^2 - 1$	$3i + 5j$	$i^3 - 4j$

Where, i stands for number of row and j stands for number of column.

5. A firm marketing a product has four salesmen S_1, S_2, S_3 and S_4. There are three customers C_1, C_2 and C_3. The probability of making a sale to a customer depends upon the salesman customer support. The Table below represents the probability with which each of the salesmen can sell to each of the customers.

Customers	Salesmen			
	S1	S2	S3	S4
C1	0.7	0.4	0.5	0.8
C2	0.5	0.8	0.6	0.7
C3	0.3	0.9	0.6	0.2

If only one salesman is to be assigned to one customer, what combination of salesmen and customers shall be optimal? Profit obtained by selling one unit to C_1 is Rs. 500, to C_2 is Rs.450 and to C_3 is Rs. 540. What is the total expected profit

Activities

1. Give two different examples of sequencing problems from your daily life.

2. Let there be five jobs I, I1, III, IV & V to be processed on two machines A and B in the order AB. Take the first 5 composite numbers as the processing times on machine A for jobs I. II. III, IV, V respectively and the first five odd numbers as the processing times on machine B for jobs V, IV, III, II, I respectively. Find the sequence that minimizes the total elapsed time. Also find the total elapsed time and idle times on both the machines.

3. Determine the optimal sequence of jobs that minimizes the total elapsed time. Processing times are given in hours. Also find total elapsed time and idle times for the machines.

Job	I	II	III	IV	V	VI
Machine A	$2\frac{1}{2}$	$\frac{5}{2}$	$\frac{1}{2}$	$2\frac{3}{4}$	$3\frac{1}{4}$	$\frac{9}{4}$
Machine B	$3\frac{1}{2}$	$4\frac{1}{2}$	$\frac{7}{4}$	$2\frac{1}{4}$	$\frac{5}{4}$	$1\frac{1}{2}$

4. Consider 4 jobs to be processed on 3 machines A, B and C on the order ABC. Assign processing times to jobs and find the optimal sequence that minimizes the total processing time. Also find the elapsed time and idle times for all the three machines.

5. (a) Determine the optimal sequence of jobs that minimizes the total elapsed time based on the following information. Processing time on machines is given in hours, and passing is not allowed. Find total elapsed time and idle times for the machines.

Job	A	B	C	D	E	F	G
Machine X	13	18	17	14	19	18	17
Machine Y	14	13	12	15	11	14	13
Machine Z	16	17	14	21	14	16	22

(b) What happens if we change the processing times on machine Z corresponding to jobs C and E and take them as 15 instead of 14? .

◆ ◆ ◆

8 Probability Distributions

Let's Study

- Random variables
- Types of random variables
- Probability distribution of a random variable
  - Discrete random variable
    - Probability mass function
    - Cumulative distribution function
    - Expected value and variance
  - Continuous random variable
    - Probability density function
    - Cumulative distribution function
  - Binomial distribution
  - Bernoulli trial
- Mean and variance of Binomial distribution
- Poisson distribution

Let's Recall

- A random experiment and all possible outcomes of an experiment.
- The sample space of a random experiment.

Let's Learn

8.1 Random variables:

We have already studied random experiments and sample spaces corresponding to random experiments. As an example, consider the experiment of tossing two fair coins. The sample space corresponding to this experiment contains four points, namely {HH, HT, TH, TT}.

We have already learnt to construct the sample space of any random experiment. However, the interest is not always in a random experiment and its sample space. We are often not interested in the outcomes of a random experiment, but only in some number obtained from the outcome. For example, in case of the experiment of tossing two fair coins, our interest may be only in the number of heads when two coins are tossed. In general, it is possible to associate a unique real number with very possible outcome of a random experiment. The number obtained from an outcome of a random experiment can take different values for different outcomes. This is why such a number is a variable. The value of this variable depends on the outcome of the random experiment, and is therefore called a random variable. A random variable is usually denoted by capital letters like $X, Y, Z, \ldots\ldots$.

Consider the following examples to understand the concept of random variables.

- When we throw two dice, there are 36 possible outcomes, but if we are interested in the sum of the numbers on the two dice, then there are only 11 different possible values, from 2 to 12.
- If we toss a coin 10 times, then there are $2^{10} = 1024$ possible outcomes, but if we are interested in the number of heads among the 10 tosses of the coin, then there are only 11 different possible values, from 0 to 10.
- In the experiment of randomly selecting four items from a lot of 20 items that contains 6 defective items, the interest is in the number of defective items among the selected four items. In this case, there are only 5 different possible outcomes, from 0 to 4.

In all the above examples, there is a rule to assign a unique value to every possible outcomes of the random experiment. Since this number can change from one outcome to another, it is a variable. Also, since this number is obtained from outcomes of a random experiment, it is called a random variable.

A random variable is formally defined as follows:

Definition : A random variable is a real-valued function defined on the sample space of a random experiment. In other words, the domain of a random variable is the sample space of a random experiment, while its co-domain is the real line.

Thus $X : S \rightarrow R$ is a random variable.

We often use the abbreviation *r.v.* for random variable.

Consider an experiment where three seeds are sown in order to find how many of them germinate. Every seed will either germinate or will not germinate. Let us use the letter Y when a seed germinates. The sample space of this experiment can then be written as $S = \{YYY, YYN, YNY, NYY, YNN, NYN, NNY, NNN\}$ and $n(S) = 8$.

None of these outcomes is a number. We shall try to represent every outcome by a number. Consider the number of times the letter Y appears is a possible outcome and denote it by X. Then, we have $X(YYY) = 3$, $X(YYN) = X(YNY) = X(NYY) = 2$, $X(YNN) = X(NYN) = X(NNY) = 1$, $X(NNN) = 0$.

The variable X has four possible values, namely 0, 1, 2 and 3. The set of possible values of X is called the range of X. Thus, in this example, the range of X is the set $\{0, 1, 2, 3\}$.

A random variable is denoted by a capital letter, like X and Y. A particular value taken by the random variable is denoted by the small letter x. Note that x is a real number and the set of all possible outcomes corresponding to a particular value x of X is denoted by the event $[X = x]$. For example, in the experiment of three seeds, the random variable X has four possible values, namely 0, 1, 2, 3. The four events are then defined as follows.

$[X = 0] = \{NNN\}$,

$[X = 1] = \{YNN, NYN, NNY\}$,

$[X = 2] = \{YYN, YNY, NYY\}$,

$[X = 3] = \{YYY\}$.

Note that the sample space in this experiment is finite and so is the random variable defined on it.

A sample space need not be finite. Consider, for example, the experiment of tossing a coin until a head is obtained. The sample space for this experiment is $S = \{H, TH, TTH, TTTH, ...\}$. Note that S contains an unending sequence of tosses required to get a head. Here, S is countably infinite. The random variable.

$X : S \rightarrow$ R, denoting the number of tosses required to get a head, has the range $\{1, 2, 3, ...\}$ which is also countably infinite.

8.2 Types of Random Variables:

There are two types of random variables, namely discrete and continuous.

8.2.1 Discrete Random Variable:

Definition : A random variable is a discrete random variable if its possible values form a countable set, which may be finite or infinite.

The values of a discrete random variable are usually denoted by non-negative integers, that is, 0, 1, 2, ... Examples of discrete random variables include the number of children in a family, the number of patients in a hospital ward, the number of cars sold by a dealer, and so on.

Note: The values of a discrete random variable are obtained by counting.

8.2.2 Continuous Random Variable

Definition: A random variable is a continuous random variable if its possible values form an interval of real numbers.

A continuous random variable has uncountably infinite possible values and these values form an interval of real numbers. Examples of continuous random variables include heights of trees in a forest, weights of students in a class, daily temperature of a city, speed of a vehicle, and so on.

The value of a continuous random variable is obtained by measurement. This value can be measured to any degree of accuracy, depending on the unit of measurement. This measurement can be represented by a point in an interval of real numbers.

The purpose of defining a random variable is to study its properties. The most important property of a random variable is its probability distribution. Many other properties of a random variable are obtained with help of its probability distribution. We shall now learn the probability distribution of a random variable. We shall first learn the probability distribution of a discrete random variable, and then learn the probability distribution of a continuous random variable.

8.3 Probability Distribution of a Discrete Random Variable

Let us consider the experiment of throwing two dice and noting the numbers on the uppermost faces of the two dice. The sample space of this experiment is $S = \{(1,1), (1,2), ..., (6,6)\}$ and $n(S) = 36$.

Let X denote the sum of the two numbers in a single throw. Then the set of possible values of X is $\{2, 3, ..., 12\}$. Further,

$$
\begin{aligned}
[X = 2] &= \{(1,1)\}, \\
[X = 3] &= \{(1,2), (2,1)\}, \\
& \;\;\vdots \\
[X = 12] &= \{6,6)\}
\end{aligned}
$$

Next, all of the 36 possible outcomes are equally likely if the two dice are fair. That is, each of the six faces has the same probability of being uppermost when a die is thrown.

As the result, each of these 36 possible outcomes has the probability $\frac{1}{36}$.

This leads to the following results.

$$P[X = 2] = P\{(1,1)\} = \frac{1}{36}$$

$$P[X = 3] = P\{(1,2), (2,1)\} = \frac{2}{36}$$

$$P[X = 4] = P\{(1,3), (2,2),(3,1)\} = \frac{3}{36},$$

and so on.

The following table shows the probabilities of all possible values of X.

x	2	3	4	5	6	7
$P(x)$	$\frac{1}{36}$	$\frac{2}{36}$	$\frac{3}{36}$	$\frac{4}{36}$	$\frac{5}{36}$	$\frac{6}{36}$
x	8	9	10	11	12	
$P(x)$	$\frac{5}{36}$	$\frac{4}{36}$	$\frac{3}{36}$	$\frac{2}{36}$	$\frac{1}{36}$	

Such a description of the possible values of a random variable X along with corresponding probabilities is called the probability distribution of the random variable X.

In general, the probability distribution of a discrete random variable X is defined as follows.

Definition : The probability distribution of a discrete random variable X is defined by the following system. Let the possible values of X be denoted by $x_1, x_2, x_3,$ ...., and the corresponding probabilities be denoted by $p_1, p_2, p_3,$ .... . Then, the set of ordered pairs $\{(x_1, p_1), (x_2, p_2), (x_3, p_3),$ ....$\}$ is called the **probability distribution of the random variable X**.

For example, consider the coin-tossing experiment where the random variable X is defined as the number of tosses required to get a head. Let the probability of getting head be 't' and that of not getting head be $1 - t$. The possible values of X are given by the set of natural numbers, $\{1, 2, 3, ...\}$ and $P[X = i] = (1 - t)^{i-1}t$, for $i = 1,2,3,....$. This result can be verified by noting that if head is obtained for the first time on the ith toss, then the first $i - 1$ tosses have resulted in tail. In other words, $[X = i]$ represents the event of having $i - 1$ tails followed by the first head on the ith toss.

$$p_i = P[X = x_i] \text{ for } i = 1, 2, 3, \ldots\ldots .$$

Note: A discrete random variable can have finite or infinite possible values, but they are countable.

The probability distribution of a discrete random variable is sometimes presented in a tabular form as follows.

x_i	x_1	x_2	x_3	...
$P[X = x_i]$	p_1	p_2	p_3	...

Note : If x_i is a possible value of X and $p_i = P[X = x_i]$, then there is an event E_i in the sample space S such that $p_i = P[E_i]$. Since x_i is a possible value of X, $p_i = P[X = x_i] > 0$. Also, all possible values of X cover all sample points in the sample space S, and hence the sum of their probabilities is 1. That is $p_i > 0$ for all i and $\Sigma p_i = 1$.

8.3.1 Probability Mass Function (p. m.f.)

The probability p_i of X taking the value x_i *is* sometimes a function of x_i for all possible values of X. In such cases, it is sufficient to specify all possible values of X and the function that gives probablilities of these values. Such a function is called the probability mass function (p. m. f.) of the discrete random variable X.

For example, consider the coin-tossing experiment where the random variable X is defined as the number of tosses required to get a head. Let probability of getting a head be 't' and that of not getting a head be $1 - t$. The possible values of X are given by the set of natural numbers $\{1,2,3,\ldots\ldots\}$ and $P[X = i] = (1 - t)^{i-1}t$, for $i = 1,2,3,\ldots\ldots$ is a function of i.

The probability mass function (p. m. f.) of a discrete random variable is defined as follows.

Definition. Let the possible values of a discrete random variable X be denoted by $x_1, x_2, \ldots\ldots$, with the corresponding probabilities $p_i = P[X = x_i]$, $i = 1, 2, \ldots$ If there is a function f such that $f(x_i) = p_i = P[X = x_i]$ for all possible values of X, then f is called the probability mass function (p. m. f.) of X.

For example, consider the experiment of tossing a coin 4 times and defining the random variable X as the number of heads in 4 tosses. The possible values of X are 0, 1, 2, 3, 4, and the probability distribution of X is given by the following table.

x	0	1	2	3	4
$P[X = x]$	$\frac{1}{16}$	$\frac{1}{4}$	$\frac{3}{8}$	$\frac{1}{4}$	$\frac{1}{16}$

Note that $P[X = x] = \binom{4}{x}\left(\frac{1}{2}\right)^4$, for $x = 0$, 1, 2, 3, 4, where $\binom{4}{x}$ is the number of ways of getting x heads in 4 tosses.

8.3.2 Cumulative Distribution Function (c. d. f.)

The probability distribution of a discrete random variable can be specified with help of the p. m. f. It is sometimes more convenient to use the cumulative distribution function (c. d. f.) of the random variable.

The cumulative distribution function (c. d. f.) of a discrete random variable is defined as follows.

Definition : The cumulative distribution function (c. d. f.) of a discrete random variable X is denoted by F and is defined as follows.

$$\begin{aligned} F(x) &= P[X \le x] \\ &= \sum_{x_i \le x} P[X = x_i] \\ &= \sum_{x_i \le x} f(x_i) \end{aligned}$$

where f is the probability mass function (p. m. f.) of the discrete random variable X.

For example, consider the experiment of tossing 4 coins and counting the number of heads.

We can form the following table for the probability distribution of X.

x	0	1	2	3	4
$f(x) = P[X = x]$	$\frac{1}{16}$	$\frac{1}{4}$	$\frac{3}{8}$	$\frac{1}{4}$	$\frac{1}{16}$
$f(x) = P[X \le x]$	$\frac{1}{16}$	$\frac{5}{16}$	$\frac{11}{16}$	$\frac{15}{16}$	1

As another example, consider the experiment of tossing a coin till a head is obtained. The following table shows the p. m. f. and the c. d. f. of the random variable X, defined as the number of tosses required for the first head.

x	1	2	3	4	5	......
$f(x)$	$\frac{1}{2}$	$\frac{1}{4}$	$\frac{1}{8}$	$\frac{1}{16}$	$\frac{1}{32}$	......
$f(x)$	$\frac{1}{2}$	$\frac{3}{4}$	$\frac{7}{8}$	$\frac{15}{16}$	$\frac{31}{32}$	......

It is possible to define several random variables on the same sample space. If two or more random variables are defined on the same sample space, their probability distributions need not be the same.

For example, consider the simple experiment of tossing a coin twice. The sample space of this experiment is $S = \{HH, HT, TH, TT\}$.

Let X denote the number of heads obtained in two tosses, Then X is a discrete random variable and its values for the possible outcomes of the experiment are obtained as follows.

$$X(HH) = 2, X(HT) = X(TH) = 1, X(TT) = 0.$$

Let Y denote the number of heads minus the number of tails in two tosses. Then Y is also a discrete random variable and its values for the possible outcomes of the experiment are obtained as follow.

$$Y(HH) = 2, \ Y(HT) = Y(TH) = 0,$$

$$Y(TT) = -2.$$

Let $Z = \dfrac{\text{number of heads}}{\text{number of tails} + 1}$. Then Z is also a discrete random variable and its values for the possible outcomes of the experiment are obtained as follows.

$$Z(HH) = 2, Z(HT) = Z(TH) = \frac{1}{2}, Z(TT) = 0.$$

These examples show that it is possible to define many discrete random variables on the same sample space. Possible values of a discrete random variable can be positive or negative, integer or fraction, and so on, as long as they are countable.

SOLVED EXAMPLES

Ex. 1. Two persons A and B play a game of tossing a coin thrice. If a toss results in a head. A gets Rs. 2 from B. If a toss results in tail, B gets Rs. 1.5 from A. Let X denote the amount gained or lost by A. Show that X is a discrete random variable and it can be defined as a function on the sample space of the experiment.

Solution: X is a number whose value depends on the outcome of a random experiment. Therefore, X is a random variable. Since the sample space of the experiment has only 8 possible outcomes, X is a discrete random variable. Now, the sample space of the experiment is S = $\{HHH, HHT, HTH, THH, HTT, THT, TTH, TTT\}$. The values of X corresponding to theses outcomes of the experiment are as follows.

$$\begin{aligned} X(HHH) &= 2 \times 3 \\ &= 6 \\ X(HHT) &= X(HTH) \\ &= X(THH) \\ &= 2 \times 2 - 1.50 \times 1 \\ &= 2.50 \\ X(HTT) &= X(THT) \\ &= X(TTH) \\ &= 2 \times 1 - 1.50 \times 2 \\ &= -1.00 \end{aligned}$$

$$X(TTT) = -1.50 \times 3$$
$$= -4.50$$

Here, a negative amount shows a loss to player A. This example shows that X takes a unique value for every element of the sample space and therefore X is a function on the sample space. Further, possible values of X are –4.50, –1, 2.50, 6.

Ex. 2. A bag contains 1 red and 2 green balls. One ball is drawn from the bag at random, its colour is noted, and the ball is put back in the bag. One more ball is drawn from the bag at random and its colour is also noted. Let X denote the number of red balls drawn. Derive the probability distribution of X.

Solution: Let the balls in the bag be denoted by r, g_1, g_2. The sample space of the experiment is then given by $S = \{rr, rg_1, rg_2, g_1r, g_1g_1, g_1g_2, g_2r, g_2g_1, g_2g_2\}$.

Since X is the number of red balls, we have

$$X(\{rr\}) = 2$$

$$\begin{aligned} X(\{rg_1\}) &= X(\{rg_2\}) \\ &= X(\{g_1r\}) \\ &= X(\{g_2r\}) \\ &= 1 \end{aligned}$$

$$\begin{aligned} X(\{g_1g_1\}) &= X(\{g_1g_2\}) \\ &= X(\{g_2g_1\}) \\ &= X(\{g_2g_2\}) \\ &= 0 \end{aligned}$$

Thus X is a discrete random variable with possible values, 0, 1 and 2. The probability distribution of X is obtained as follows.

x	0	1	2
$P[X = x]$	$\frac{4}{9}$	$\frac{4}{9}$	$\frac{1}{9}$

Ex. 3. Two cards are randomly drawn, with replacement, from a well shuffled deck of 52 playing cards. Find the probability distribution of the number of aces drawn.

Solution: Let X denote the number of aces among the two cards drawn with replacement. Clearly, 0, 1 and 2 are the possible values of X. Since the draws are with replacement, the outcomes of the two draws are independent of each other. Also, since there are 4 aces in the deck of 52 cards, $P[\text{an ace}] = \frac{4}{52} = \frac{1}{13}$, and $P[\text{a non-ace}] = \frac{12}{13}$.

Then

$$\begin{aligned} P[X = 0] &= P[\text{non-ace and non-ace}] \\ &= \frac{12}{13} \times \frac{12}{13} \\ &= \frac{144}{169} \end{aligned}$$

$$\begin{aligned} P[X = 1] &= P[\text{ace and non-ace}] \\ &+ P[\text{non-ace and ace}] \\ &= \frac{1}{13} \times \frac{12}{13} + \frac{12}{13} \times \frac{1}{13} \\ &= \frac{24}{169} \end{aligned}$$

$$\begin{aligned} \text{and } P[X = 2] &= P[\text{ace and ace}] \\ &= \frac{1}{13} \times \frac{1}{13} \\ &= \frac{1}{169} \end{aligned}$$

The requaried probability distribution is then as follows.

x	0	1	2
$P[X = x]$	$\frac{144}{169}$	$\frac{24}{169}$	$\frac{1}{169}$

Ex. 4. A fair die is thrown. Let X denote the number of factors of the number on the upper face. Find the probability distribution of X.

Solution: The sample space of the experiment is S = {1, 2, 3, 4, 5, 6}. The values of X for the possible outcomes of the experiment are as follows. $X(1) = 1$, $X(2) = 2$, $X(3) = 2$, $X(4) = 3$, $X(5) = 2$, $X(6) = 4$. Therefore,

$$p_1 = P[X = 1] = P[\{1\}] = \frac{1}{6}$$

$$p_2 = P[X = 2] = P[\{2,3,5\}] = \frac{3}{6}$$

$$p_3 = P[X = 3] = P[\{4\}] = \frac{1}{6}$$

$$p_4 = P[X = 4] = P[\{6\}] = \frac{1}{6}$$

The probability distribution of X is as shown in the following table.

x	1	2	3	4
$P[X = x]$	$\frac{1}{6}$	$\frac{3}{6}$	$\frac{1}{6}$	$\frac{1}{6}$

Ex. 5. Find the probability distribution of the number of doubles in three throws of a pair of dice.

Solution : Let X denote the number of doubles. The possible doubles in a single throw of a pair of dice are given by (1, 1), (2, 2), (3, 3), (4, 4), (5, 5), (6, 6).

Since the dice are thrown thrice, 0, 1, 2, and 3 are possible values of X. Probability of getting a doublet in a single throw of a pair of dice is $p = \frac{1}{6}$ and $q = 1 - \frac{1}{6} = \frac{5}{6}$.

$$
\begin{aligned}
P[X = 0] &= P\,[\text{no doublet}] \\
&= qqq = \frac{5}{6}\cdot\frac{5}{6}\cdot\frac{5}{6} = \frac{125}{216} \\
P[X = 1] &= P\,[\text{one doublets}] \\
&\ pqq + qpq + qqp \\
&= 3pq^2 = \frac{75}{216} \\
P[X = 2] &= P\,[\text{two doublets}] \\
&= ppq + pqp + qpp \\
&= 3p^2q = \frac{15}{216} \\
P[X = 3] &= P\,[\text{three doublets}] \\
&= ppp = \frac{1}{216}
\end{aligned}
$$

Ex. 6. The probability distribution of X is as follows.

x	0	1	2	3	4
$P[X = x]$	0.1	k	$2k$	$2k$	k

Find (i) k, (ii) $P[X < 2]$, (iii) $P[X \geq 3]$, (iv) $P\,[1 \leq X < 4]$, (v) $F(2)$.

Solution: The table gives a probability distribution and therefore.

$P[X = 0] + P[X= 1] + P[X = 2] + P[X = 3] + P[X = 4] = 1.$

That is, $0.1 + k + 2k + 2k + k = 1$

That is, $6k = 0.9$ Therefore $k = 0.15$.

(i) $k = 0.15$

(ii) $P[X < 2] = P[X = 0] + P[X = 1] = 0.1 + k = 0.1 + 0.15 = 0.25$

(iii) $P[X \geq 3] = P[\ X = 3] + P[X = 4] = 2k + k = 3(0.15) = 0.45$

(iv) $P[1 \leq X < 4] = P[X = 1] + P[X = 2] + P[x = 3] = k + 2k + 2k = 5k = 5(0.15) = 0.75$.

v) $F(2) = P[X \leq 2] = P[X = 0]\ P[x = 1] + P[X = 2] = 0.1 + k + 2k = 0.1 + 3k = 0.1 + 3(0.15) = 0.1 + 0.45 = 0.55$.

8.3.3 Expected value and variance of a random variable.

In many problems, the interest is in some feature of a random variable computed from its probability distribution. Some such numbers are mean, variance and standard deviation. We shall discuss mean and variance in this section. Mean is a measure of location in the sense that it is the average value of the random variable.

Definition : Let X be a random variable whose possible values $x_1, x_2, x_3, \ldots, x_n$ occur with probabilities $p_1, p_2, p_3, \ldots\ldots, p_n$ respectively. The expected value or arithmetic mean of X, denoted by $E(X)$ or μ is defined by

$$\mu = E(X) = (x_1p_1 + x_2p_2 + \ldots.. + x_np_n) = \sum_{i=1}^{n} x_ip_i$$

The mean or expected value of a random variable X is the sum of products of possible values of X and their respective probabilities.

Definition : Let X be a random variable whose possible values $x_1, x_2, \ldots, x_n$ occur with probabilities $p_1, p_2, p_3, \ldots, p_n$ respectively. The variance of X, denoted by Var (X) or σ^2 is defined as

$$\sigma^2 = Var(X) = \sum_{i=1}^{n} (x_i - \mu)^2 \, p_i .$$

The non-negative square root of Var(X) is called the standard deviation of the random variable X. That is, $\sigma = \sqrt{\text{Var}(X)}$.

Another formula to find the variance of a random variable. We can also use the simplified form of Var $(X) = \sum_{i=1}^{n} x^2{}_i p_i - \left(\sum_{i=1}^{n} x_i p_i\right)^2$

or $\text{Var}(X) = E(X^2) - [E(X)]^2$.

SOLVED EXAMPLES

Ex. 1. Three coins are tossed simultaneously. X is the number of heads. Find expected value and variance of X.

Solution. Let $S = \{HHH, HHT, HTH, THH, HTT, THT, TTH, TTT\}$ and $X = \{0, 1, 2, 3\}$.

x_i	p_i	$x_i\ p_i$	$x^2{}_i p_i$
0	$\frac{1}{8}$	0	0
1	$\frac{3}{8}$	$\frac{3}{8}$	$\frac{3}{8}$
2	$\frac{3}{8}$	$\frac{6}{8}$	$\frac{12}{8}$
3	$\frac{1}{8}$	$\frac{3}{8}$	$\frac{9}{8}$

Then $E(X) = \sum_{i=1}^{n} x_i p_i = \frac{12}{8} = 1.5$.

$$\text{Var}(X) = \sum_{i=1}^{n} x^2{}_i p_i - \left(\sum_{i=1}^{n} x_i p_i\right)^2 = \frac{24}{8} - (1.5)^2$$

$= 3 - 2.25 = 0.75$.

Ex. 2. Let a pair of dice be thrown and the random variable X be sum of numbers on the two dice. Find the mean and variance of X.

Solution : The sample space of the experiment consists of 36 elementary events in the form of ordered pairs (x_i, y_i), where $x_i = 1, 2, 3, 4, 5, 6$ and $y_i = 1, 2, 3, 4, 5, 6$. The random variable X has the possible values 2, 3, 4, 5, 6, 7, 8, 9, 10, 11, and 12.

x_i	P_i	$x_i p_i$	$x^2{}_i p_i$
2	1/36	2/36	4/36
3	2/36	6/36	18/36
4	3/36	12/36	48/36
5	4/36	20/36	100/36
6	5/36	30/36	180/36
7	6/36	42/36	294/36
8	5/36	40/36	320/36
9	4/36	36/36	324/36
10	3/36	30/36	300/36
11	2/36	22/36	242/36
12	1/36	12/36	144/36
		$\sum_{i=1}^{n} x_i p_i = \frac{252}{36} = 7$	$\sum_{i=1}^{n} x^2{}_i p_i = \frac{1974}{36} = 54.83$

$$E(X) = \sum_{i=1}^{n} x_i p_i = 7$$

$$\text{Var}(X) = \sum_{i=1}^{n} x^2{}_i p_i - \left(\sum_{i=1}^{n} x_i p_i\right)^2 = 5.83$$

Ex. 3. Find the mean and variance of the number randomly selected from 1 to 15.

Solution. The sample space of the experiment is $S = \{1, 2, 3, \ldots, 15\}$.

Let X denote the selected number. Then X is a random variable with possible values 1, 2, 3, ..., 15. Each of these numbers is equiprobable.

Therefore, $P(1) = P(2) = P(3) = \ldots = P(15) = \frac{1}{15}$.

$$\mu = E(X) = \sum_{i=1}^{n} x_i p_i = 1 \times \frac{1}{15} + 2 \times \frac{1}{15} + \ldots$$

$$+ 15 \times \frac{1}{15} = (1 + 2 + ... + 15) \times \frac{1}{15}$$

$$= \left(\frac{15\times16}{2}\right)\times \frac{1}{15} = 8.$$

$$\text{Var}(X) = \sum_{i=1}^{n} x^2_i\, p_i - \left(\sum_{i=1}^{n} x_i p_i\right)^2$$

$$= 1^2 \times \frac{1}{15} + 2^2 \times \frac{1}{15} + ... + 15^2 \times \frac{1}{15} - (8)^2$$

$$= (1^2 + 2^2 + ... + 15^2) \times \frac{1}{15} - (8)^2$$

$$= \left(\frac{15\times16\times31}{6}\right)\times \frac{1}{15} - (8)^2$$

$$= 82.67 - 64 = 18.67.$$

Ex. 4. Two cards are drawn without replacement from a well shuffled pack of 52 cards. Find the mean and variance of the number of kings drawn.

Solution: Let X denote the number of kings in a draw of two cards. X is a random variable with possible values 0, 1 or 2. Then

$$P\text{ (both cards are not king)} = \frac{\binom{48}{2}}{\binom{52}{2}} = \frac{48\times47}{52\times51} = \frac{188}{221}.$$

P (one card is king and other card is not king)

$$= \frac{\binom{4}{1}\times\binom{48}{1}}{\binom{52}{2}} = \frac{4\times48\times2}{52\times51} = \frac{32}{221}$$

$$P\text{ (both cards are king)} = \frac{\binom{4}{2}}{\binom{52}{2}} = \frac{4\times3}{52\times51} = \frac{1}{221}.$$

Therefore,

$$\mu = E(X) = \sum_{i=1}^{n} x_i\, p_i$$

$$= 0 \times \frac{188}{221} + 1 \times \frac{32}{221} + 2 \times \frac{1}{221} = \frac{34}{221},$$

$$\text{Var (X)} = \sum_{i=1}^{n} x^2_i\, p_i - \left(\sum_{i=1}^{n} x_i p_i\right)^2$$

$$= (0^2 \times \frac{188}{221} + 1^2 \times \frac{32}{221} + 2^2 \times \frac{1}{221}) - \left(\frac{34}{221}\right)^2$$

$$= \frac{36}{221} - \frac{1156}{48841} = \frac{6800}{48841} = 0.1392$$

EXERCISE 8.1

1. Let X represent the difference between number of heads and number of tails obtained when a coin is tossed 6 times. What are the possible values of X?

2. An urn contains 5 red and 2 black balls. Two balls are drawn at random. X denotes number of black balls drawn. What are the possible values of X?

3. Determine whether each of the following is a probability distribution. Give reasons for your answer.

(i)

x	0	1	2
$P(x)$	0.4	0.4	0.2

(ii)

x	0	1	2	3	4
$P(x)$	0.1	0.5	0.2	−0.1	0.3

(iii)

x	0	1	2
$P(x)$	0.1	0.6	0.3

(iv)

z	3	2	1	0	−1
$P(z)$	0.3	0.2	0.4	0.05	0.05

(v)

y	-1	0	1
$P(y)$	0.6	0.1	0.2

(vi)

x	0	1	2
$P(x)$	0.3	0.4	0.2

4. Find the probability distribution of (i) number of heads in two tosses of a coin, (ii) number of tails in three tosses of a coin, (iii) number of heads in four tosses of a coin.

5. Find the probability distribution of the number of successes in two tosses of a die if success is defind as getting a number greater than 4.

6. A sample of 4 bulbs is drawn at random with replacement from a lot of 30 bulbs which includes 6 defectives bulbs. Find the probability distribution of the number of defective bulbs.

7. A coin is blased so that the head is 3 times as likely to occur as tail. Find the probability distribution of number of tails in two tosses.

8. A random varibale X has the following probability distribution:

x	1	2	3	4	5	6	7
$P(x)$	k	$2k$	$2k$	$3k$	k^2	$2k^2$	$7k^2+k$

Determine (i) k, (ii) $P(X < 3)$,

(iii) $P(0 < X < 3)$, (iv) $P(X > 4)$.

9. Find expected value and variance of X using the following p. m. f.

x	-2	-1	0	1	2
$P(x)$	0.2	0.3	0.1	0.15	0.25

10. Find expected value and variance of X, the number on the uppermost face of a fair die.

11. Find the mean of number of heads in three tosses of a fair coin.

12. Two dice are thrown simultaneously. If X denotes the number of sixes, find the expectation of X.

13. Two numbers are selected at random (without replacement) from the first six positive integers. Let X denote the larger of the two numbers. Find $E(X)$.

14. Let X denote the sum of the numbers obtained when two fair dice are rolled. Find the variance of X.

15. A class has 15 students whose ages are 14, 17, 15, 14, 21, 17, 19, 20, 16, 18, 20, 17, 16, 19 and 20 years. If X denotes the age of a randomly selected student, find the probability distribution of X. Find the mean and variance of X.

16. 70% of the members favour and 30% oppose a proposal in a meeting. The random variable X takes the value 0 if a member opposes the proposal and the value 1 if a member is in favour. Find $E(X)$ and Var(X).

8.4 Probability Distribution of a Continuous Random Variable

The possible values of a continuous random variable form an interval of real numbers. The probability distribution of a continuous random variable is represented by a continuous function called the probability density function (p. d. f.). A continuous random variable also has a cumulative distribution function (c. d. f.).

Suppose the possible values of the continuous random variable X form the interval (a, b), where a and b are real numbers such that $a < b$. The interval (a, b) is called the support of the continuous random variable X.

We shall now define the probability density function (p. d. f.) and cumulative distribution fucntions (c. d. f.) of a continuous random variable.

8.4.1 Probability Density Function (p. d. f.)

Let X be a continuous random variable with the interval (a, b) as its support. The probability density function (p. d. f.) of X is an integrable function f that satisfies the following conditions.

1. $f(x) \geq 0$ for all $x \in (a, b)$.
2. $\int_a^b f(x)\,dx = 1$
3. For any real numbers c and d such that

$$a \leq c < d \leq b,$$

$$P[X \in (c, d)] = \int_c^d f(x)\,dx$$

It is easy to notice that the p. d. f. of a continuous random variable is different from the p. m. f. of a discrete random variable. Both the p. m. f. and p. d. f. are positive at possible values of the random variable. However, the p. d. f. is positive over an entire interval that is, over an uncountably infinite set of points.

8.4.2 : Cumulative Distribution Functions (c.d.f)

The cumulative distribution function (p. d. f.) of a continuous random variable is defined as follows.

Definition. The cumulative distribution function (c. d. f.) of a continuous random variable X is denoted by F and defined by

$$F(x) = 0 \text{ for all } x \leq a,$$
$$= \int_a^x f(x)\,dx \text{ for all } x \geq a.$$

Note: The c. d. f. of a continuous random variable is a non-decreasing continuous function. The c. d. f. of a discrete random variable is a step function, while the c. d. f. of a continuous random variable is a continuous function.

SOLVED EXAMPLES

Ex. 1. Let X be a continuous random variable with probability function (p. d. f.) $f(x) = 3x^2$ for $0 < x < 1$. Can we claim that $f(x) = P[X = x]$?

Solution. Note, for example, that $f(0.9) = 3(0.9)^2 = 2.43$, which is not a probability. This example shows that the p. d. f. of a continuous random variable does not represent the probability of a possible value of the random variable. In case of a continuous random variable, the probability of an interval is obtained by integrating the p. d. f. over the specified interval. In this case, the probability is given by the area under the curve of the p. d. f. over the interval.

Let us now verify whether f is a valid probability density function (p. d. f). This is done through the following steps.

1. $f(x) = 3x^2 > 0$ for all $x \in (0,1)$.
2. $\int_0^1 f(x)\,dx = 1$

Therefore, the function $f(x) = 3x^2$ for $0 < x < 1$ is a proper probability density function. Also, for real numbers c and d such that $0 \leq c < d \leq 1$, note that $P[c < X < d] = \int_c^d f(x)\,dx = \int_c^d 3x^2 dx = \left[x^3\right]_c^d = d^3 - c^3 > 0$.

What is the probability that X falls between 1/2 and 1?

That is, what is $P(1/2 < X < 1]$?

Take $c = 1/2$ and $d = 1$ in the above integral to obtain $P[1/2 < X < 1] = 1 - (1/2)^3 = 1 - 1/8 = 7/8$.

What is $P(X = 1/2)$?

It is easy to see that the probability is 0. This is so because $\int_c^d f(x)\,dx = \int_{1/2}^{1/2} 3x^2 dx = (1/2)^3 - (1/2)^3 = 0$.

As a matter of fact, the probability that a continuous random variable X takes any specific value x is 0. That is, when X is a continuous random variable, $P[X = x] = 0$ for all real x.

Ex. 2. Let X be a continuous random variable with probability density function $f(x) = x^3/4$ for an interval $0 < x <$ c. What is the value of the constant c that makes $f(x)$ a valid probability density function?

Solution. Not that the internal of the p. d. f. over the support of the random variable must be 1. That is $\int_0^c f(x)\,dx = 1$. That is, $\int_0^c \frac{x^3}{4} dx = 1$. But, $\int_0^c \frac{x^3}{4} dx = \frac{c^4}{16}$. Since this integral must be equal to I, we have $\dfrac{c^4}{16} = 1$, or equivalently $c^4 = 16$, so that $c = 2$ since c must be positive.

Ex. 3. Let's return to the example in which X has the following probability density function.

$$f(x) = 3x^2 \text{ for } 0 < x < 1.$$

Find the cumulative distribution function $F(x)$.

Solution : $F(x) = \int_0^x 3x^2 dx = \left[x^3 - (0)^3\right]_0^x =$ x^3 for $x \in (0,1)$.

Therefore,

$$f(x) = \begin{cases} 0 & \text{for } x \le 0, \\ x^3 & \text{for } 0 < x < 1, \\ 1 & \text{for } x \ge 1. \end{cases}$$

Ex. 4. Let's return to the example in which X has the following probability density function.

$$f(x) = \frac{x^3}{4} \text{ for } 0 < x < 2.$$

Find the cumulative distribuiton function of X.

Solution : $F(x) = \int_0^x f(x)\, dx = \int_0^x \frac{x^3}{4} dx =$

$$\frac{1}{4}\left[\frac{x^4}{4}\right]_0^x = \frac{1}{16}\left[x^4 - 0\right]_0^x = \frac{x^4}{16} \text{ for } 0 < x < 2.$$

Therefore,

$$F(x) = \begin{cases} 0 & \text{for } x < 0, \\ \dfrac{x^4}{16} & \text{for } 0 \le x < 2, \\ 1 & \text{for } x \ge 2. \end{cases}$$

Ex. 5. Verify whether each of the following functions is p. d. f. of a continuous r. v. X.

(i) $f(x) = \begin{cases} e^{-x} & \text{for } 0 < x < \infty \\ 0 & \text{otherwise.} \end{cases}$

(ii) $f(x) = \begin{cases} \dfrac{x}{2} & \text{for } -2 < x < 2, \\ 0 & \text{otherwise.} \end{cases}$

Solution :

(i) $e^{-x} \ge 0$ for all real values of x, because $e > 0$. Therefore, $e^{-x} > 0$ for $0 < x < \infty$.

Also, $\int_0^\infty f(x)\, dx = \int_0^\infty e^{-x}\, dx = \left[-e^{-x}\right]_0^\infty =$ $\left[\frac{1}{e^\infty} - e^0\right] -(0 - 1) = 1.$

Both the conditions of p. d. f. are satisfied and hence $f(x)$ is p. d. f. of a continuous r. v.

(ii) $f(x)$ is negative for $-2 < x < 0$, and therefore $f(x)$ is not a valid p. d. f.

Ex. 6. Find k if the following function is the p. d. f. of a r. v. X.

$$\begin{cases} kx^2(1-x) & \text{for } 0 < x < 1 \\ 0 & \text{otherwise} \end{cases}$$

Solution. Since $f(x)$ is the p.d.f. of r. v. X,

$$\int_0^1 kx^2(1-x)\, dx = 1$$

$$\int_0^1 k(x^2 - x^3)\, dx = 1$$

$$k\left[\frac{x^3}{3} - \frac{x^4}{4}\right]_0^1 = 1$$

$$k\left\{\left[\frac{1}{3} - \frac{1}{4}\right] - 0\right\} = 1$$

$$\therefore k \times 12 = 1$$

$$\therefore k = 12$$

Ex. 7. In each of the following cases, find (a) $P(X < 1)$ and (b) $P(|X| < 1)$.

(i) $f(x) = \begin{cases} \dfrac{x^2}{18} & \text{for } -3 < x < 3, \\ 0 & \text{otherwise} \end{cases}$

(ii) $f(x) = \begin{cases} \dfrac{x+2}{18} & \text{for } -2 < x < 4, \\ 0 & \text{otherwise} \end{cases}$

Solution:

(i) (a) $P(X<1) = \int_{-3}^{1} \frac{x^2}{18}\,dx = \left[\frac{\frac{1}{18}(x^3)}{3}\right]_{-3}^{1}$

$= \frac{1}{54}\,[1-(-3)^3] = \frac{1}{54}\,(1+27) = \frac{28}{54}$

$= \frac{14}{27}$

(b) $P(|X|<1) = P(-1<X<1) = \int_{-1}^{1}\frac{x^2}{18}\,dx$

$= \left[\frac{\frac{1}{18}(x^3)}{3}\right]_{-3}^{1} = \frac{1}{54}[1 - (-1)^3]$

$= \frac{1}{54}(1+1) = \frac{2}{54} = \frac{1}{27}$

ii) (a) $P(X<1) = \int_{-2}^{1}\frac{x+2}{18}\,dx = \frac{1}{18}\left[\frac{x^2}{2}+2x\right]_{-2}^{1}$

$= \frac{1}{18}\left\{\left(\frac{1}{2}+2\right)-\left(\frac{(-2)^2}{2}+2(-2)\right)\right\}$

$= \frac{1}{18}\left\{\left(\frac{5}{2}+2\right)\right\} = \frac{1}{18}\times\frac{9}{2} = \frac{1}{4}$

(b) $P(|X|<1) = P(-1< X<1)$

$= \int_{-1}^{1}\frac{x+2}{18}\,dx = \frac{1}{18}\left[\frac{x^2}{2}+2x\right]_{-1}^{1}$

$= \frac{1}{18}\left\{\left(\frac{1}{2}+2\right)-\left(\frac{1}{2}-2\right)\right\}$

$= \frac{1}{18}\left\{\frac{5}{2}+\frac{3}{2}\right\} = \frac{1}{18}\times 4 = \frac{2}{9}$

Ex. 8. Find the c. d. f. $F(x)$ associated with the p. d. f. $f(x)$ or r. v. X where

$$f(x) = \begin{cases} 3(1-2x^2) & \text{for } 0<x<1,\\ 0 & \text{otherwise}\end{cases}$$

Solution: Since $f(x)$ is a p. d. f., the c. d. f. is given by

$F(x) = \int_0^x 3(1-2x^2)\,dx = \left[3\left(x-\frac{2x^3}{3}\right)\right]$

$= \left[3x-2x^3\right]_0^x = 3x - 2x^3$ for $0<x<1$.

Therefore,

$$F(x) = \begin{cases} 0 & \text{for } x\le 0,\\ 3x-2x^3 & \text{for } 0<x<1,\\ 1 & \text{for } x\ge 1.\end{cases}$$

EXERCISE 8.2

1. Check whether each of the following is a p. d. f.

   i) $f(x) = \begin{cases} x & \text{for } 0\le x\le 1,\\ 2-x & \text{for } 1<x\le 2.\end{cases}$

   ii) $f(x) = 2 \quad$ for $0<x<1$.

2. The following is the p. d. f. of a r. v. X.

   $$f(x) = \begin{cases} \frac{x}{8} & \text{for } 0<x<4,\\ 0 & \text{otherwise}\end{cases}$$

   Find (i) $P(x<1.5)$, (ii) $P(1<x<2)$, (iii) $P(x>2)$.

3. It is felt that error in measurement of reaction temperature (in celesus) in an experiment is a continuous r. v. with p. d. f.

   $$f(x) = \begin{cases} \frac{x^3}{64} & \text{for } 0\le x\le 4,\\ 0 & \text{otherwise}\end{cases}$$

   (i) Verify whether $f(x)$ is a p. d. f.

   (ii) Find $P(0<x\le 1)$.

   (iii) Find probability that X is between 1 and 3.

4. Find k if the following function represents the p. d. f. of a r. v. X.

   (i) $f(x) = \begin{cases} kx & \text{for } 0<x<2,\\ 0 & \text{otherwise}\end{cases}$

   Also find $P\left[\frac{1}{4}<X<\frac{1}{2}\right]$

(ii) $f(x) = \begin{cases} kx(1-x) & \text{for } 0 < x < 1, \\ 0 & \text{otherwise} \end{cases}$

Also find (a) $P\left[\frac{1}{4} < X < \frac{1}{2}\right]$,

(b) $P\left[X < \frac{1}{2}\right]$

5. Let X be the amount of time for which a book is taken out of library by a randomly selected student and suppose that X has p. d. f.

$$f(x) = \begin{cases} 0.5x & \text{for } 0 \le x \le 2, \\ 0 & \text{otherwise} \end{cases}$$

Calculate (i) $P(X \le 1)$, (ii) $P(0.5 \le X \le 1.5)$, (iii) $P(X \ge 1.5)$.

6. Suppose X is the waiting time (in minutes) for a bus and its p. d. f. is given by

$$f(x) = \begin{cases} \frac{1}{5} & \text{for } 0 \le x \le 5, \\ 0 & \text{otherwise} \end{cases}$$

Find the probability that (i) waiting time is between 1 and 3 minutes, (ii) waiting time is more than 4 minutes.

7. Suppose error involved in making a certain measurement is a continuous r. v. X with p. d. f.

$$f(x) = \begin{cases} k(4-x^2) & \text{for } -2 \le x \le 2, \\ 0 & \text{otherwise} \end{cases}$$

Compute (i) $P(X > 0)$, (ii) $P(-1 < x < 1)$, (iii) $P(X < -0.5 \text{ or } X > 0.5)$.

8. Following is the p. d. f. of a continuous r. v. X.

$$f(x) = \begin{cases} \frac{x}{8} & \text{for } 0 < x < 4, \\ 0 & \text{otherwise} \end{cases}$$

(i) Find expression for the c. d. f. of X. (ii) Find $F(x)$ at $x = 0.5$, 1.7, and 5.

9. The p. d. f. of a continuous r. v. X is

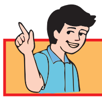
$$f(x) = \begin{cases} \frac{3x^2}{8} & \text{for } 0 < x < 2, \\ 0 & \text{otherwise} \end{cases}$$

Determine the c. d. f. of X and hence find (i) $P(X < 1)$, (ii) $P(X < -2)$, (iii) $P(X > 0)$, (iv) $P(1 < X < 2)$.

10. If a r. v. X has p. d. f.

$$f(x) = \begin{cases} \frac{c}{x} & \text{for } 1 < x < 3, \\ 0 & \text{otherwise} \end{cases}, \ c > 0..$$

Find c, $E(X)$, and Var (X). Also find $F(x)$.



Let's Remember

A random variable (r. v.) is a real-valued function defined on the sample space of a random experiment.

The domain of a random variable is the sample space of a random experiment, while its co-domain is the real line.

Thus : $X : S \rightarrow R$ is a random variable.

There are two types of random variables, namely discrete and continuous.

Discrete random variable :

Let the possible values of a discrete random variable X be denoted by $x_1, x_2, \ldots$, with the corresponding probabilities $p_i = P[X = x_i]$, $i = 1, 2, \ldots$. If there is a function f such that $p_i = P[X = x]$, $f(x_i)$ for all possible values of X, then f is called the probability mass function (p. m. f.) of X.

Note : If x_i is a possible value of X and $p_i = P[X = x_i]$, then there is an event E_i in the sample space S such that $p_i = P[E_i]$. Since x_i is a possible value of X, $p_i = P[X = x_i] > 0$. Also, all possible values of X cover all sample points in the sample space S, and hence the sum of their probabilities is 1. That is, $p_i > 0$ for all i and $\Sigma p_i = 1$.

c. d. f. ($F(x)$)

The cumulative distribution function (c. d. f.) of a discrete random variable X is denoted by F and is defined as follows.

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

$$= \sum_{x_i \leq x} P[X = x_i]$$

$$= \sum_{x_i \leq x} f(x_i).$$

Expected Value or Mean of a Discrete r. v.

Let X be a random variable whose possible values $x_1, x_2, x_3, ..., x_n$ occur with respective probabilities $p_1, p_2, p_3, ..., p_n$. The expected value or arithmetic mean of X, denoted by $E(X)$ or μ, is defined by

$$\mu = E(X) = \sum_{i=1}^{n} (x_i p_i + x_1 p_1 + x_2 p_2 + ... + x_n p_n.)$$

In other words, the mean or expectation of a random variable X is the sum of products of all possible values of X and their respective probabilities.

Variance of a Discrete r. v.

Let X be a random variable whose possible values $x_1, x_2, ..., x_n$ occur with respective probabilities p_1, p_2, p_3, p_n. The variance of X, denoted by Var(X) or σ_x^2 is defined as.

$$\sigma_x^2 = \sum_{i=1}^{n} (x_i - \mu)^2 p_i$$

The non-negative square-root of the variance $\sigma_x = \sqrt{\text{Var}(X)}$ is called the standard deviation of the random variable X.

Another formula to find the variance of a random variable. We can also use the simplified form $\text{Var}(X) = \sum_{i=1}^{n} x_i^2 p_i - \left(\sum_{i=1}^{n} x_i p_i\right)^2$, or

$$\text{Var}(X) = E(X^2) - [E(X)]^2,$$

where $E(X^2) = \sum_{i=1}^{n} x_i^2 p_i$.

Probability Density Function (p. d. f.)

Let X be a continuous random variable defined on the interval S = (a, b). A non-negative integrable function f is called the probability density function (p. d. f.) of X if it satisfies the following conditions. (i) $f(x) > 0$ for all $x \in$ S. (ii) The area under the curve $f(x)$ over S is 1. That is, $\int_S f(x)\, dx = 1$. (iii) The probability that X takes a value in some interval A is given by the integral of $f(x)$ over A. That is $P[X \in A] = \int_A f(x)\, dx$.

The **cumulative distribution function (c. d. f.)** of a continuous random variable X is defined as follows.

$$F(x) = \int_a^x f(x)\, dx \text{for a} < x < \text{b}.$$

8.5 Binomial Distribution

Let's Recall

Many experiments are dichotomous in nature. An experiment is dichotomous if it has only two possible outcomes. For example, a tossed coin shows 'head' or 'tail', the result of a student is 'pass' or 'fail', a manufactured item is 'defective' or 'non-defective', the response to a question is 'yes' or 'no', an egg has 'hatched' or 'not hatched', the decision is 'yes' or 'no' etc. In such cases, it is customary to call one of the outcomes 'success' and the other 'failure'. For example, in tossing a coin, if the occurrence of head is considered success, then occurrence of tail is failure.

8.5.1 Bernoulli trial

An experiment that can result in one of two possible outcomes is called a dichotomous experiment. One of the two outcomes is called success and the other outcome is called failure.

Definition of Bernoulli Trial : A dichotomous experiment is called a Bernoulli trial.

Every time we toss a coin or perform a dichotomous experiment, we call it a trial. If a coin is tossed 4 times, the number of trials is 4, each having exactly two possible outcomes, namely success and failure. The outcome of any trial is independent of the outcome of other trials. In all such trials, the probability of success (and hence of failure) remains the same.

Definition of Sequence of Bernoulli Trials: A sequence of dichotomous experiments is called a sequence of Bernoulli trials if it satisfies the following conditions.

- The trials are independent.
- The probability of success remains the same in all trials.

The probability of success in a Bernoulli trial is denoted by p and the probability of failure is denoted by $q = 1 - p$. For example, if we throw a die and define success as getting an even number and failure as getting an odd number, we have a Bernoulli trial. Successive throws of the die are independent trials and form a sequence of Bernoulli trials. If the die is fair, then $p = 1/2$ and $q = 1 - p = 1 - 1/2 = 1/2$.

Ex. 1. Six balls are drawn successively from an urn containing 7 red and 9 black balls. Decide whether the trials of drawing balls are Bernoulli trials if, after each draw, the ball drawn is (i) replaced (ii) not replaced in the urn.

Solution.

(i) When the drawing is done with replacement, the probability of success (say, red ball) is $p = 7/16$ which remains the same for all six trials (draws). Hence, drawing balls with replacement are Bernoulli trials.

(ii) When the drawing is done without replacement, the probability of success (i.e. red ball) in first trial is 7/16. In second trial, it is 6/15 if first ball drawn is red and is 7/15 if first ball drawn in black, and so on. Clearly probability of success is not same for all trials, and hence the trials are not Bernoulli trials.

8.5.2 Binomial distribution

Consider a sequence of Bernoulli trials. where we are interested in the number successes, X, in a given number of trials, n. The total number of possible outcomes in 2^n. We want to find the number of outcomes that result in X successes in order to find the probability of getting X successes in n Bernoulli trials. We use the binomial theorem and derive a formula for counting the number of favorable outcomes for all possible values of X.

Consider the situation where we are interested in one success in 6 trials. Clearly, six possible cases are $\{SFFFFF, FSFFFF, FFSFFF, FFFSFF, FFFFSF, FFFFFS\}$. Counting the number of outcomes with two successes is possible by listing favorable outcomes. However, it will be time-consuming to list all of these outcomes. As a result, calculating probabilities of getting 0, 1, 2, ..., 6 successes in 6 trials will be time consuming. Lengthy calculations that require listing of all favorable outcomes for calculating probabilities of number of successes in n Bernoulli trials can be avoided by deriving a formula.

For illustrating this formula, let us take the experiment made up of three Bernoulli trials with probablities p and $q = 1 - p$ of success and failure, respectively, in each trial. The sample space of the experiment is the set.

$$S = \{SSS, SSF, SFS, SFF, FSF, FFS, FFF\}.$$

The number of successes is a random variable X and can take values 0, 1, 2, and 3. The probability distribution of the number of successes is as follows.

$$\begin{aligned} P(X=0) &= P\text{ (no success)} \\ &= P(\{FFF\}) = P(F)\,P(F)\,P(F) \\ &= q\,.\,q\,.\,q = q^3, \\ &\quad \text{since trials are independent.} \end{aligned}$$

$$\begin{aligned} P(X=1) &= P\text{(one success)} \\ &= P(\{SFF, FSF, FFS\}) \\ &= P(\{SFF\}) + P(\{FSF\}) + P(\{FFS\}) \\ &= P(S)P(F)P(F) + P(F)P(S)P(F) + P(F)P(F)P(S) \\ &= p.q.q + q.p.q + q.q.p = 3pq^2 \end{aligned}$$

$$
\begin{aligned}
P(X=2) &= P(\text{ two successes}) \\
&= P(\{SSF, SFS, FSS\}) \\
&= P(\{SSF\}) + P(\{SFS\}) + P(\{FSS\}) \\
&= P(S)P(S)P(F) + P(S)P(F)P(S) + P(F)P(S)P(S) \\
&= p.p.q + p.q.p + q.p.p = 3p^2q
\end{aligned}
$$

$$
\begin{aligned}
\text{and} P(X=3) &= P(\text{three successes}) \\
&= P(\{SSS\}) \\
&= P(S).P(S).P(S) = p^3.
\end{aligned}
$$

Thus, the probability distribution of X is as shown in the following table.

x	0	1	2	3
$P(x)$	q^3	$3q^2p$	$3qp^2$	p^3

Note that the binomial expansion of $(q+p)^3$ is $q^3+3p^2q+3qp^2+p^3$ and the probabilities of 0, 1, 2 and 3 successes are respectively the 1st, 2nd, 3rd and 4th term in the expansion of $(q+p)^3$. Also, since $q+p=1$, it follows that the sum of these probabilities, as expected, is 1. Thus, we may conclude that in an experiment of n Bernoulli trials, the probabilities of 0, 1, 2, ..., n successes can be obtained as 1st, 2nd, ..., $(n+1)$th terms in the expansion of $(q+p)^n$. To prove this assertion (result), let us find the probability of x successes in an experiment of n Bernoulli trials.

In an experiment with n trials, if there are x successes (S), there will be $(n-x)$ failure (F). Now x successes (S) and $(n-x)$ failure (F) can be obtained in $\dfrac{n!}{x!(n-x)!}$ ways. In each of these ways, the probabilitiy of x succcesses and $(n-x)$ failure, that is, P(x successes). P($n-x$ failure) is given by

$(P(S).P(S) \ldots P(S)\ x \text{ times})\ (P(F).P(F) \ldots P(F)\ (n-x) \text{ times})$

$= \ (\text{p.p.p...p } x \text{ times})\ (\text{q.q.q...q } (\text{n} - x) \text{ times})$

$= \ p^x q^{n-x}.$

Thus, the probability of x successes in n Bernoulli trials is

P(x successes out of n trials)

$$\frac{n!}{x!(n-x)!}\, p^x q^{n-x} = \binom{n}{x} p^x q^{n-x}$$

Note that

$P(x \text{ successes}) = \binom{n}{x} p^x q^{n-x}$ is the $(x+1)$th term in the binomial expansion of $(q+p)^n$.

Thus, the probability distribution of number of successes in an experiment consisting of n Bernoulli trials can be obtained by the binomial expansion of $(q+p)^n$. This distribution of X successes in n Bernoulli trials can be written as follows.

X	0	1	2	...
P(X)	$\binom{n}{0} p^0 q^n$	$\binom{n}{1} p^1 q^{n-1}$	$\binom{n}{2} p^2 q^{n-2}$	...
X	x	...	n	
P(X)	$\binom{n}{x} p^x q^{n-x}$	...	$\binom{n}{n} p^n q^0$	

The above probability distribution is known as the **binomial distribution** with parameters n and p. We can find the complete probability distribution of the random variable using the given values of n and p. The fact that the r. v. X follows the binomial distribution with parameters n and p is written in short as $X \sim \text{B}(n, p)$ and read as "X follows the binomail distribution with parameters n and p."

The probability of x successes P($X = x$) is also denoted by $P(x)$ and is given by $P(x) = {}^nC_x\, q^{n-x} p^x,\ x = 0, 1, \ldots n;\ (q = 1-p)$.

This $P(x)$ is the probability mass function of the binomial distribution.

A binomial distribution with n Bernoulli trials and p as probability of success in each trial is denoted by B (n,p) or written as $X \sim \text{B}(n,p)$.

Let us solve some examples.

Ex. 2. If a fair coin is tossed 10 times, find the probability of obtaining.

(i) exactly six heads

(ii) at least six heads

(iii) at most six heads

Solution: The repeated tosses of a coin are Bernoulli trails. Let X denote the number of heads in an experiment of 10 trials.

Clearly, $X \sim B(n,p)$ with $n = 10$ and $p = \frac{1}{2}$ and $q = 1 - p = \frac{1}{2}$.

$$P(X = x) = \binom{n}{x} p^x q^{n-x}$$

$$P(X = x) = \binom{10}{x}\left(\frac{1}{2}\right)^x \left(\frac{1}{2}\right)^{10-x}$$

(i) Exactly six successes means $X = 6$.

$$\begin{aligned} P(X = 6) &= \binom{10}{6}\left(\frac{1}{2}\right)^6 \left(\frac{1}{2}\right)^{10-6} \\ &= \frac{10!}{6!(10-6)!}\left(\frac{1}{2}\right)^6 \times \left(\frac{1}{2}\right)^4 \\ &= \frac{10\times9\times8\times7}{4\times3\times2\times1}\left(\frac{1}{2}\right)^{10} \\ &= \frac{105}{512} \end{aligned}$$

(ii) At least six successes means $X \geq 6$.

$$\begin{aligned} P(X \geq 6) &= P(X = 6) + P(X = 7) + P(X = 8) + P(X = 9) + P(X = 10) \\ &= \binom{10}{6}\left(\frac{1}{2}\right)^6 \times \left(\frac{1}{2}\right)^4 + \binom{10}{7}\left(\frac{1}{2}\right)^7 \times \left(\frac{1}{2}\right)^3 + \binom{10}{8}\left(\frac{1}{2}\right)^8 \times \left(\frac{1}{2}\right)^2 + \\ &\quad \binom{10}{9}\left(\frac{1}{2}\right)^9 \times \left(\frac{1}{2}\right)^1 + \binom{10}{10}\left(\frac{1}{2}\right)^{10} \times \left(\frac{1}{2}\right)^0 \\ &= \frac{10\times9\times8\times7}{4\times3\times2\times1}\left(\frac{1}{2}\right)^{10} + \frac{10\times9\times8}{3\times2\times1}\left(\frac{1}{2}\right)^{10} \\ &\quad + \frac{10\times9}{2\times1}\left(\frac{1}{2}\right)^{10} + \frac{10}{1}\left(\frac{1}{2}\right)^{10} + \left(\frac{1}{2}\right)^{10} \\ &= (210 + 120 + 45 + 10 + 1) \times \frac{1}{1024} \\ &= \frac{386}{1024} = \frac{193}{512} \end{aligned}$$

(iii) At most six successes means $X \leq 6$.

$$\begin{aligned} P(X \leq 6) &= 1 - P(X > 6) \\ &= 1 - [P(X = 7) + P(X = 8) + P(X = 9) \\ &\quad + P(X = 10)] \\ &= 1 -[\binom{10}{7}\left(\frac{1}{2}\right)^7 \times \left(\frac{1}{2}\right)^3 + \binom{10}{8}\left(\frac{1}{2}\right)^8 \\ &\quad \times \left(\frac{1}{2}\right)^2 + \binom{10}{9}\left(\frac{1}{2}\right)^9 \times \left(\frac{1}{2}\right)^1 + \\ &\quad \binom{10}{10}\left(\frac{1}{2}\right)^{10} \times \left(\frac{1}{2}\right)^0 +] \\ &= 1 - \left[\frac{10\times9\times8}{3\times2\times1}\left(\frac{1}{2}\right)^{10} + \frac{10\times9}{2\times1}\left(\frac{1}{2}\right)^{10} \right. \\ &\quad \left. + \frac{10}{1}\left(\frac{1}{2}\right)^{10} + \left(\frac{1}{2}\right)^{10}\right] \\ &= 1 - \left[(120 + 45 + 10 + 1) \times \frac{1}{1024}\right] \\ &= 1 - \frac{176}{1024} = 1 - \frac{11}{64} = \frac{53}{64} \end{aligned}$$

Ex. 3. The eggs are drawn successively with replacement from a lot containing 10% defective eggs. Find the probability that there is at least one defective egg.

Solution: Let X denote the number of defective eggs in the 10 eggs drawn. Since the drawing is done with replacement, the trials are Bernoulli trials.

Probability of success $= \frac{1}{10}$

$$\therefore \quad P = \frac{1}{10}, q = 1 - p = 1 - \frac{1}{10} = \frac{9}{10}, \text{ and}$$

$n = 10$.

$$X \sim B\left(10, \frac{1}{10}\right).$$

$$P(X = x) = \binom{10}{x}\left(\frac{1}{10}\right)^{x}\left(\frac{9}{10}\right)^{10-x}$$

Here, $X \geq 1$.

$$\begin{aligned} P(X \geq 1) &= 1 - P(X < 1) = 1 - P(X = 0) \\ &= 1 - \binom{10}{0}\left(\frac{1}{10}\right)^{0}\left(\frac{9}{10}\right)^{10} \\ &= 1 - 1 \times \left(\frac{9}{10}\right)^{10} \\ &= 1 - \left(\frac{9}{10}\right)^{10} \end{aligned}$$

8.5.3 Mean and Variance of Binomial Distribution

Let $X \sim B(n,p)$. The mean or expected value of X is denoted by μ. It is also called expected value of X and is denoted by $E(X)$ and given by $\mu = E(X) = np$. The variance is denoted by $\mathrm{Var}(X)$ and is given by $\sigma^2 = \mathrm{Var}(X) = npq$.

Ex. 4. If $X \sim B\,(10, 0.4)$, then find $E(X)$ and $\mathrm{Var}(X)$.

Solution: Here $n = 10$, $p = 0.4$, $q = 1 - p = 1 - 0, 4 = 0.6$.

$E(X) = np = 10(0.4) = 4$

$\mathrm{Var}(X) = npq = 10(0.4)\,(0.6) = 2.4$

Ex. 5. Let the p.m. f. of the r. v. X be $P(X = x) = \binom{4}{x}\left(\frac{5}{9}\right)^{x}\left(\frac{4}{9}\right)^{4-x}$, for $x = 0, 1, 2, 3, 4$. Find $E(X)$ and $\mathrm{Var}(X)$.

Solution: X follows the binomial distribution with $n = 4$, $p = 5/9$ and $q = 4/9$.

$$\mathrm{Var}(X) = npq = 4\left(\frac{5}{9}\right)\left(\frac{4}{9}\right) = 80/81.$$

Ex. 6. If $E(X) = 6$ and $\mathrm{Var}(X) = 4.2$, find n and p.

Solution: $E(x) = 6$. Therefore, $np = 6$. $\mathrm{Var}(X) = 4.2$.

Therefore, $npq = 4.2$.

$$\frac{npq}{np} = \frac{4.2}{6}$$

$$\therefore \quad q = 0.7, p = 1 - q = 1 - 0.7 = 0.3$$

$$np = 6 \quad \therefore\ n \times 0.3 = 6 \quad \therefore\ n = \frac{6}{0.3} = 20$$

EXERCISE 8.3

1. A die is thrown 4 times. If 'getting an odd number' is a success, find the probability of

   (i) 2 successes (ii) at least 3 successes (iii) at most 2 successes.

2. A pair of dice is thrown 3 times. If getting a doublet is considered a success, find the probability of two successes.

3. There are 10% defective items in a large bulk of items. What is the probability that a sample of 4 items will include not more than one defective item?

4. Five cards are drawn successively with replacement from a well-shuffled deck of 52 cards. Find the probability that

   (i) all the five cards are spades.

   (ii) only 3 cards are spades.

   (iii) none is a spade.

5. The probability that a bulb produced by a factory will fuse after 200 days of use is 0.2. Let X denote the number of bulbs (out of 5) that fuse after 200 days of use. Find the probability of

   (i) $X = 0$, (ii) $X \leq 1$, (iii) $X > 1$, (iv) $X \geq 1$.

6. 10 balls are marked with digits 0 to 9. If four balls are selected with replacement. What is the probability that none is marked 0?

7. In a multiple choice test with three possible answers for each of the five questions, what is the probability of a candidate getting four or more correct answers by random choice?

8. Find the probability of throwing at most 2 sixes in 6 throws of a single die.

9. Given that $X \sim B(n,p)$,

   (i) if $n = 10$ and $p = 0.4$, find $E(X)$ and $\text{Var}(X)$.

   (ii) if $p = 0.6$ and $E(X) = 6$, find n and $\text{Var}(X)$.

   (iii) if $n = 25$, $E(X) = 10$, find p and Var (X).

   (iv) if $n = 10$, $E(X) = 8$, find $\text{Var}(X)$.

8.6 Poisson Distribution

Poisson distribution is a discrete probability distribution that gives the probability of number of occurrences of an event in a fixed interval of time, if these events occur with a known average rate and are independent of the time since the last occurrence. For instance, suppose someone receives 4 emails per day on an average. There will be, however, variation in the number of emails, sometimes more, sometimes fewer, once in a while no email at all. The Poisson distribution was first introduced by Simeon Denis Poisson (1781-1840) and published in 1837. The work focused on certain random variables N that count the number of discrete occurrences of an event that take place during a time-interval of given length.

Definition: A discrete random variable X is said to have the Poisson distribution with parameter $m > 0$, if its p.m. is given by

$$P(X = x) = \frac{e^{-m}m^x}{x!} = 0, 1, 2, ...$$

Note.

(i) we use the notation $X \sim P(m)$ to show that X follows the Poisson distribution with parameter m.

(ii) Observe that $P(x) > 0$ for all non-negative integers x and $\sum_{x=0}^{\infty} P(x) = 1$.

(iii) For the Poisson distribution Mean = $E(X) = m$ and Variance = $\text{Var}(X) = m$.

(iv) When n is very large and p is very small in the binomial distribution, then X follows the Poisson distribution with parameter $m = np$.

SOLVED EXAMPLES

Ex. 1. If $X \sim P(m)$ with $m = 5$ and $e^{-5} = 0.0067$, then find

(i) $P(X = 5)$, (ii) $P(X \geq 2)$.

Solution: $P(X = x) = \dfrac{e^{-m}m^x}{x!}$, for $x = 0, 1, 2, \ldots$.

(i) Here $m = 5$ and $x = 5$

$$\begin{aligned} P(X = 5) &= \frac{e^{-5}5^5}{5!} \\ &= \frac{0.0067 \times 3125}{5 \times 4 \times 3 \times 2 \times 1} \\ &= 0.1745 \end{aligned}$$

(ii)
$$\begin{aligned} P(X \geq 2) &= 1 - P(X < 2) \\ &= 1 - [P(X = 0) + P(X = 1)] \\ &= 1 - \left[\frac{e^{-5}5^0}{0!} + \frac{e^{-5}5^1}{1!}\right] \\ &= 1 - \left[\frac{0.0067 \times 1}{1} + \frac{0.0067 \times 5}{1}\right] \\ &= 1 - [0.0067 + 0.0335] \\ &= 1 - 0.0402 \\ &= 0.9598 \end{aligned}$$

Ex. 2. If $X \sim P(m)$ with $P(X = 1) = P(X = 2)$, then find the mean and $P(X = 2)$ given that $e^{-2} = 0.1353$.

Solution: Since $P(X = 1) = P(X = 2)$,

$$\therefore \quad \frac{e^{-m}m^1}{1!} = \frac{e^{-m}m^2}{2!}$$

$$\therefore \quad m = 2$$

Then

$$P(X=2) = \frac{e^{-2}2^2}{2!} = \frac{0.1353 \times 4}{2} = 0.2706.$$

Ex. 3. In a town, 10 accidents takes place in the span of 50 days. Assuming that the number of accidents follows Poisson distribution, find the probability that there will be 3 or more accidents on a day.

(given that $e^{-0.2} = 0.8187$)

Solution: Here $m = \frac{10}{50} = 0.2$, and hence $X \sim P(m)$ with $m = 0.2$. The p. m. f. of X is

$$P(X=x) = \frac{e^{-m}m^x}{x!}, x = 0, 1, 2, ...$$

$$P(X \geq 3) = 1 - P(X < 3)$$

$$= 1 - [P(X=0) + P(X=1) + P(X=2)]$$

$$= \quad 1 - \left[\frac{e^{-0.2}(0.2)^0}{0!} + \frac{e^{-0.2}(0.2)^1}{1!} + \frac{e^{-0.2}(0.2)^2}{2!}\right]$$

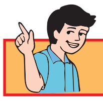
$$= 1 - \left[\frac{0.8187 \times 1}{1} + \frac{0.8187 \times 2}{1} + \frac{0.8187 \times 0.004}{2}\right]$$

$$= 1 - [0.8187 + 0.16374 + 0.016374]$$

$$= 1 - 0.9988 = 0.0012$$

EXERCISE 8.4

1. If X has Poisson distribution with m = 1, then find $P(X \leq 1)$ given $e^{-1} = 0.3678$.
2. If $X \sim P(1.2)$, then find $P(X = 3)$ given $e^{-0.5} = 0.6065$.
3. If X has Poisson distribution with parameter m and $P(X = 2) = P(X = 3)$, then find $P(X \geq 2)$. Use $e^{-3} = 0.0497$.
4. The number of complaints which a bank manager receives per day follows a Poisson distribution with parameter m = 4. find the probability that the manager receives (i) only two complaints on a given day, (ii) at most two complaints on a given day. Use $e^{-4} = 0.0183$.
5. A car firm has 2 cars, which are hired out day by day. The number of cards hired on a day follows Poisson distribution with mean 1.5. Find the probability that (i) no car is used on a given day, (ii) some demand is refused on a given day, given $e^{-1.5} = 0.2231$.
6. Defects on plywood sheet occur at random with the average of one defect per 50 sq. ft. Find the probability that such a sheet has (i) no defect, (ii) at least one defect. Use $e^{-1} = 0.3678$.
7. It is known that, in a certain area of a large city, the average number of rats per bungalow is five. Assuming that the number of rats follows Poisson distribution, find the probability that a randomly selected bungalow has (i) exactly 5 rats (ii) more than 5 rats (iii) between 5 and 7 rats, inclusive. Given $e^{-5} = 0067$.



Let's Remember

Trials of a random experiment are called Bernoulli trials, if they satisfy the following conditions.

- Each trail has two possible outcomes, success and failure.
- The probability of success remains the same in all trials.

Thus, the probability of getting x successes in n Bernoulli trials is

$$P(x \text{ successes in } n \text{ trials}) = \binom{n}{x} p^x \times q^{n-x}$$

$$= \frac{n!}{x! \times (n-x)!} p^x \times q^{n-x}.$$

Clearly, P (x successess) $= \binom{n}{x} p^x \times q^{n-x}$ is the $(x + 1)^{\text{th}}$ term in the binomial expansion of $(q + p)^n$.

Let $X \sim B(n,p)$. Then the mean or expected value of r. v. X is denoted by μ. It is also denoted by $E(X)$ and is given by $\mu = E(X) = np$. The variance is denoted by Var(X) and is given by Var(X) = npq.

A discrete random variable X is said to follow the Poisson distribution with parameter $m > 0$ if its p. m. f. is given by

$$P(X = x) = \frac{e^{-m}m^{x}}{x!}, x = 0, 1, 2, ...$$

Note.

1. We use the notation $X \sim P(m)$. That is, X follows Poisson distribution with parameter m.
2. Observe that $P(x) > 0$ for all non-negative integers x and $\sum_{x=0}^{\infty} \frac{e^{-m}m^{x}}{x!} = 1$
3. For the Poisson distribution, Mean $E(X) = m$ and Variance = Var(X) = m.
4. If n is very large and p is very small then X follows Poisson distribution with $m = np$.

MISCELLANEOUS EXERCISE - 8

I) Choose the correct alternative.

1. F(x) is c.d.f. of discrete r.v. X whose p.m.f. is given by $P(x) = k\binom{4}{x}$, for x = 0, 1, 2, 3, 4 & $P(x)$ = 0 otherwise then F(5) = ...

   (a) $\frac{1}{16}$ (b) $\frac{1}{8}$ (c) $\frac{1}{4}$ (d) 1

2. F(x) is c.d.f. of discrete r.v. X whose distribution is

X_i	−2	−1	0	1	2
P_i	0.2	0.3	0.15	0.25	0.1

   then F(−3) = ...

   (a) 0 (b) 1 (c) 0.2 (d) 0.15

3. X : is number obtained on upper most free when a fair die ..... thrown then E(x) = ...

   (a) 3.0 (b) 3.5 (c) 4.0 (d) 4.5

4. If p.m.f. of r. v. X is given below.

x	0	1	2
$P(x)$	q^2	$2pq$	p^2

   then Var(x) = ...

   (a) p^2 (b) q^2 (c) pq (d) $2pq$

5. The expected value of the sum of two numbers obtained when two fair dice are rolled is ....

   (a) 5 (b) 6 (c) 7 (d) 8

6. Given p.d.f. of a continuous r.v. X as

   $f(x) = \frac{x^2}{3}$ for $-1 < x < 2$

   = 0 otherwise then

   $F(1)$ =

   (a) $\frac{1}{9}$ (b) $\frac{2}{9}$ (c) $\frac{3}{9}$ (d) $\frac{4}{9}$

7. X is r.v. with p.d.f. $f(x) = \frac{k}{\sqrt{x}}$, $0 < x < 4$

   = 0 therwise

   then E (x) = ...

   (a) $\frac{1}{3}$ (b) $\frac{4}{3}$ (c) $\frac{2}{3}$ (d) 1

8. If X ~ B (20, $\frac{1}{10}$) then E (x) = ...

   (a) 2 (b) 5 (c) 4 (d) 3

9. If E (x) = m and Var (x) = m then X follows ....

   (a) Binomial distribution
   (b) Possion distribution
   (c) Normal distribution
   (d) none of the above

10. If E $(x) >$ Var (x) then X follows ....
    - (a) Binomial distribution
    - (b) Possion distribution
    - (c) Normal distribution
    - (d) none of the above

II) Fill in the blanks.

1. The values of discrete r.v. are generally obtained by ....

2. The values of continuous r.v. are generally obtained by ....

3. If X is discrete random variable takes the values $x_1, x_2, x_3, \dots x_n$ then $\sum_{i=1}^{n} P(x_i) =$ ....

4. If F(x) is distribution function of discrete r.v.x with p.m.f. P$(x) = \frac{x-1}{3}$ for $x = 1, 2, 3,$ & P$(x) = 0$ otherwise then F(4) = ...

5. If F(x) is distribution function of discrete r.v. X with p.m.f. P$(x) = k\binom{4}{x}$ for $x = 0, 1,$ 2, 3, 4, and P$(x) = 0$ otherwise then F(-1) = ...

6. E(x) is considered to be .... of the probability distribution of x.

7. If x is continuous r.v. and F$(x_i) = $ P$(X \le x_i)$ $= \int_{-\infty}^{x_i} f(x)\,dx$ then F(x) is called ....

8. In Binomial distribution probability of success .... from trial to trial.

9. In Binomail distribution if n is very large and probability success of p is very small such that $np = m$ (constant) then .... distribution is appplied.

III) State whether each of the following is True or False.

1. If P$(X = x) = k\binom{4}{x}$ for $x = 0,1,2,3,4,$ then F(5) $= \frac{1}{4}$ when F(x) is c.d.f.

2.

X	−2	−1	0	1	2
$P(X = x)$	0.2	0.3	0.15	0.25	0.1

If F(x) is c.d.f. of discrete r.v. X then F$(-3) = 0$.

3. X is the number obtained on upper most face when a die is thrown then E$(x) = 3.5$.

4. If p.m.f. of discrete r.v. X is

X	0	1	2
$P(X = x)$	q^2	$2pq$	p^2

then E$(x) = 2p$.

5. The p.m.f. of a r.v. X is

P$(x) = \frac{2x}{n(n+1)}$, $x = 1, 2, \dots\dots\dots$ n

$= 0$ otherwise,

Then

E$(x) = \frac{2n+1}{3}$

6. If $f(x) = k\,x\,(1-x)$ for $0 < x < 1$

$= 0$ otherwise then

k = 12

7. If X ~ B(n,p) and $n = 6$ & P$(x = 4)$ $=$ P$(x = 2)$ then p $= \frac{1}{2}$.

8. If r.v. X assumes values 1,2,3, ... n with equal probabilities then E$(x) = \frac{(n+1)}{2}$

9. If r.v. X assumes the values 1,2,3, ..., 9 with equal probabilities, E$(x) = 5$.

IV) Solve the following problems.

PART - I

1. Identify the random variable as discrete or continuous in each of the following. Identify its range if it is discrete.

   (i) An economist is interested in knowing the number of unemployed graduates in the town with a population of 1 lakh.

   (ii) Amount of syrup prescribed by a physician.

   (iii) A person on high protein diet is interested in the weight gained in a week.

   (iv) Twelve of 20 white rats available for an experiment are male. A scientist randomly selects 5 rats and counts the number of female rats among them.

   (v) A highway safety group is interested in the speed (km/hrs) of a car at a check point.

2. The probability distribution of a discrete r. v. X is as follows.

x	1	2	3	4	5	6
$P(X=x)$	k	$2k$	$3k$	$4k$	$5k$	$6k$

   (i) Determine the value of k.

   (ii) Find $P(X \leq 4)$, $P(2 < X < 4)$, $P(X \geq 3)$.

3. Following is the probability distribution of a r. v. X.

x	-3	-2	-1	0	1	2	3
$P(X=x)$	0.05	0.1	0.15	0.20	0.25	0.15	0.1

   Find the probability that

   (i) X is positive.

   (ii) X is non-negative

   (iii) X is odd.

   (iv) X is even.

4. The p. m. f of a r. v. X is given by

$$P(X=x) = \begin{cases} \binom{5}{x}\dfrac{1}{2^5}, & x = 0,1,2,3,4,5. \\ 0 & \text{otherwise} \end{cases}$$

   Show that $P(X \leq 2) = P(X \geq 3)$.

5. In the following probability distribution of a r. v. X.

x	1	2	3	4	5
$P(x)$	1/20	3/20	a	$2a$	1/20

   Find a and obtain the c. d. f. of X.

6. A fair coin is tossed 4 times. Let X denote the number of heads obtained. Identify the probability distribution of X and state the formula for p. m. f. of X.

7. Find the probability of the number of successes in two tosses of a die, where success is defined as (i)number greater than 4 (ii) six appears in at least one toss.

8. A random variable X has the following probability distribution.

x	1	2	3	4	5	6	7
$P(x)$	k	$2k$	$2k$	$3k$	k^2	$2k^2$	$7k^2+k$

   Determine (i) k, (ii) $P(X < 3)$, (iii) $P(X > 6)$, (iv) $P(0 < X < 3)$.

9. The following is the c. d. f. of a. r. v. X.

x	-3	-2	-1	0	1	2	3	4
$F(x)$	0.1	0.3	0.5	0.65	0.75	0.85	0.9	1

   Find the probability distribution of X and $P(-1 \leq X \leq 2)$.

10. Find the expected value and variance of the r. v. X if its probability distribution is as follows.

   (i)

x	1	2	3
$P(X=x)$	1/5	2/5	2/5

(ii)

x	-1	0	1
$P(X=x)$	1/5	2/5	2/5

(iii)

x	1	2	3	...	n
$P(X=x)$	$1/n$	$1/n$	$1/n$	...	$1/n$

(iv)

x	0	1	2	3	4	5
$P(X=x)$	1/32	5/32	10/32	10/32	5/32	1/32

11. A player tosses two coins. He wins Rs. 10 if 2 heads appear, Rs. 5 if 1 head appears, and Rs. 2 if no head appears. Find the expected value and variance of winning amount.

12. Let the p. m. f. of the r. v. X be

$$P(x) = \begin{cases} \dfrac{3-x}{10} & \text{for } x = -1, 0, 1, 2. \\ 0 & \text{otherwise} \end{cases}$$

Calculate $E(X)$ and $\text{Var}(X)$.

13. Suppose error involved in making a certain measurement is a continuous r. v. X with p. d. f.

$$f(x) = \begin{cases} k(4-x^2) & \text{for } -2 \le x \le 2. \\ 0 & \text{otherwise} \end{cases}$$

Compute (i) $P(X > 0)$, (ii) $P(-1 < x < 1)$, (iii) $P(X < 0.5 \text{ or } X > 0.5)$

14. The p. d. f. of the r. v. X is given by

$$f(x) = \begin{cases} \dfrac{1}{2a} & \text{for } 0 < x < 2a. \\ 0 & \text{otherwise} \end{cases}$$

Show that $P(X < a/2) = P(X > 3a/2)$.

15. Determine k if the p. d. f. of the r. v. is

$$f(x) = \begin{cases} ke^{-\theta x} & \text{for } 0 \le x < \infty. \\ 0 & \text{otherwise} \end{cases}$$

Find $P(X > \dfrac{1}{\theta})$ and detemine M if $P(0 < X < M) = \dfrac{1}{2}$

16. The p. d. f. of the r. v. X is given by

$$f(x) = \begin{cases} \dfrac{k}{\sqrt{x}} & \text{for } 0 < x < 4. \\ 0 & \text{otherwise} \end{cases}$$

Determine k, the c. d. f. of X, and hence find $P(X \le 2)$ and $P(x \ge 1)$.

17. Let X denote the reaction temperature in celcius of a certain chemical process. Let X have the p. d. f.

$$f(x) = \begin{cases} \dfrac{1}{10} & \text{for } -5 \le x < 5. \\ 0 & \text{otherwise} \end{cases}$$

Compute $P(X < 0)$.

PART - II

1. Let X ~ B(10,0.2). Find (i) P(X = 1) (ii) P(X ≥ 1) (iii) P(X ≤ 8).

2. Let X~B(n,p) (i) If n = 10 and E(X)= 5, find p and Var(X). (ii) If E(X)= 5 and Var(X) = 2.5, find n and p.

3. If a fair coin is tossed 4 times, find the probability that it shows (i) 3 heads, (ii) head in the first 2 tosses and tail in last 2 tosses.

4. The probability that a bomb will hit the target is 0.8. Find the probability that, out of 5 bombs, exactly 2 will miss the target.

5. The probability that a lamp in the classroom will burn is 0.3. 3 lamps are fitted in the classroom. The classroom is unusable if the number of lamps burning in it is less than 2. Find the probability that the classroom can not be used on a random occasion.

6. A large chain retailer purchases an electric device from the manufacturer. The manufacturer indicates that the defective rate of the device is 10%. The inspector of the retailer randomly selects 4 items from a shipment. Find the probability that the inspector finds at most one defective item in the 4 selected items.

7. The probability that a component will survive a check test is 0.6. Find the probability that exactly 2 of the next 4 components tested survive.

8. An examination consists of 5 multiple choice questions, in each of which the candidate has to decide which one of 4 suggested answers is correct. A completely unprepared student guesses each answer completely randomly. Find the probability that this student gets 4 or more correct answers.

9. The probability that a machine will produce all bolts in a production run within the specification is 0.9. A sample of 3 machines is taken at random. Calculate the probability that all machines will produce all bolts in a production run within the specification.

10. A computer installation has 3 terminals. The probability that anyone terminal requires attention during a week is 0.1, independent of other terminals. Find the probabilities that (i) 0 (ii) 1 terminal requires attention during a week.

11. In a large school, 80% of the students like mathematics. A visitor asks each of 4 students, selected at random, whether they like mathematics. (i) Calculate the probabilites of obtaining an answer yes from all of the selected students. (ii) Find the probability that the visitor obtains the answer yes from at least 3 students.

12. It is observed that it rains on 10 days out of 30 days. Find the probability that (i) it rains on exactly 3 days of a week. (ii) it rains on at most 2 days of a week.

13. If X follows Poisson distribution such that $P(X = 1) = 0.4$ and $P(X = 2) = 0.2$, find variance of X.

14. If X follows Poisson distribution with parameter m such that $\frac{P(X=x+1)}{P(X=x)} = \frac{m}{x+1}$ find mean and variance of X.

✦ ✦ ✦

ANSWERS

Exercise 1.1

1. ₹ 5760, ₹ 42,240
2. ₹ 7500
3. 22
4. ₹ 50,000
5. ₹ 80,000
6. 20%
7. ₹ 5800, 6%
8. ₹ 1,00,000
9. Cash Sales = ₹ 66,000,
   Credit Sales = ₹ 36,000
10. 11.75%
11, Rs 47,500
12. ₹ 10,200
13. C.P. = ₹ 6400, marked price = ₹10,000
14. ₹ 4,464
15. ₹ 1,140

Exercise 1.2

1. ₹ 10500
2. ₹ 8333.33
3. ₹ 7560, ₹ 7000
4. 5 years.
5. 10 % p. a.
6. T. D. = ₹ 182.60, B. D = ₹ 190.80,
   B. G. = ₹ 8.20
7. 6% p. a.
8. B. D = ₹ 313.12, C. V. = ₹ 6621.88
9. 15th April 1998
10. ₹ 43800
11. 5 months
12. T.D. = ₹ 1200, B.D. = ₹ 1220,
    Amount = ₹ 73200
13. B. G. = ₹ 10, T. D. = 1000
14. r = 20%
15. ₹ 12,000

MISCELLANEOUS EXERCISE - 1

I) 1. b. Del credere agent
2. a. factor
3). c. nominal due date
4. a. The legal due date
5. a. Face Value
6. b. Present worth
7. d. True discount
8. b. List Price
9. b. Invoice price = Net selling price
10. b. True discount

II) 1. Drawee
2. Auctioneer
3. Catalogue/list
4. Commercial Discount
5. higher
6. Bankers Gain
7. Legal due date
8. A broker
9. Trade, Catalogue / list
10. Invoice Price

III) 1. Flase
2. False
3. True
4. False
5. True
6. False
7. False
8. True

9. False
10. True.

IV) 1. ₹ 5125
2. ₹ 75,00,000
3. ₹ 1989, ₹ 28611
4. ₹ 1,00,000
5. 8% p. a.
6. C. P. = ₹ 650
7. Salary = ₹ 8000, r = 10%
8. 100
9. ₹ 6923
10. 5 months
11. ₹ 4000, ₹ 4080, ₹ 2,04,000
12. 13%
13. ₹ 1280
14. ₹ 120, ₹ 4680
15. 16th November
16. ₹ 50,000, ₹ 2500
17. ₹ 18000, ₹ 33,000.

Exercise 2.1

1. (i) ₹ 75,000 (ii) ₹ 60,000
2. ₹ 2430, ₹ 364.5
3. 3.25%, ₹ 1430
4. ₹ 2,25,000
5. ₹ 3,85,714.3
6. 50,000 kg
7. ₹ 71,040
8. ₹ 3,60,000
9. From Company A : ₹ 3600
   From Company B : ₹ 2250
   From Company C: ₹ 3150
10. ₹ 1,40,000
11. ₹ 153.6
12. ₹ 5595.8
13. ₹ 1,90,400
14. ₹ 1,40,000

Exercise 2.2

1. ₹ 21,000
2. ₹ 23,205
3. ₹ 10,50,000
4. ₹ 29,975
5. ₹ 91,120
6. ₹ 1,83,555
7. ₹ 1,00,000
8. ₹ 1,50,000
9. 200% p. a.
10. 3 years
11. ₹ 3,641
12. ₹ 824.16
13. ₹ 2,146.5
14. ₹ 2,400
15. ₹ 31,944
16. ₹ 3,703.70

MISCELLANEOUS EXERCISE - 2

I) 1 c 2. d 3. d 4. d 5. c 6. b
7. c 8. b 9. b 10. d

II)

1. premium
2. life
3. property value
4. policy value
5. annuitant
6. installment
7. payment period
8. perpetuity
9. annuity due
10. immediate annuity R.

III) 1. F 2. T 3. F 4. T 5. F 6. T
7.F 8 T 9 F 10 T.

IV) 1. ₹ 432
2. ₹ 96,000
3. ₹ 18,00,000
4. ₹ 85,000
5. ₹ 1,07,000
6. ₹ 32,800
7. ₹ 26,000
8. 16,000 kg
9. 875
10. ₹ 58,800
11. ₹ 78,000 , ₹ 144
12. ₹ 38,400 , ₹ 17,052
13. ₹ 2,07,400

14. ₹ 13,435
15. ₹ 23,201.85
16. 12 years
17. 5% p.a.
18. ₹ 41,250
19. ₹ 49,740
20. ₹ 4550 , ₹ 1930
21. ₹ 2,00,000
22. ₹ 31,488
23. ₹ 4 years
24. ₹ 21,752.30

Exercise 3.1

1) (i) $y = 2 + 0.75x$
   (ii) $y =$ ₹ 11750
2) (i) $x = y + 6$ (ii) $y = 0.87x - 4.18$
3) (i) -0.8 (ii) $-0.4, -0.4$ (iii) -0.4
4) 12.73
5) (i) $y = 1.16x - 10.4$
   Estimate of $y = 99.8$
   (ii) $x = 0.59y + 26.65$,
   Estimate of $x = 70.9$
6) $y = -1.34x + 18.04$
7) $y = 0.3x + 62.1$
8) $y = 2x + 3, x = 0.5y - 1.5$
9) $y = -0.65x + 11.9$,
   $x = -1.3y + 16.4$
10) (i) $y = 2x - 1$, (ii) $y = 7$
11) (i) $y = 0.63x - 2.8$, (ii) $y = 3.5$
12) $y = 6.6x + 38.36$

Exercise 3.2

1) (i) -0.6 (ii) 31.6 (iii) 53.9
2) (i) $5x + 24y = 3251$
   (ii) $6x + 5x = 1664$ (iii) $-\frac{5}{12}$
3) (i) $y = 0.36x + 34$ (ii) $x = 2.19y - 58.59$
   (iii) 0.8879
4) (i) $y = 3.2x + 58, x = 0.2y - 8$
   (ii) ₹ 106 lakh (iii) ₹ 16 lakh
5) (i) Inconsistent as $\frac{b_{xy} + b_{yx}}{2} \geq \mathrm{r}$
   (ii) Inconsistent as sign of $b_{yx.} b_{xy}$ and r are not same.
   (iii) Inconsistent as sign of $b_{yx.}$ and b_{xy} are not same
   (iv) Consistent.
6) $x = 0.82y + 10.24$
7) (i) $y = 10$ (ii) $c = 23.5$
8) (i) $y = 0.72x + 28.8, x = 0.5y + 40.$
   (ii) $x = 90$
9) $x = 0.14y + 5.528, x = 6.928$
10) $y = 182$
11) i) $x =$ ₹ 64 (ii) $y =$ ₹ 8.7
12) $x = 11.2$ (ii) $y = 15.8$

Exercise 3.3

1) $x = 2, \bar{y} = 8.25, r = 0.6$
2) i) $\bar{x} = 13, \bar{y} = 17$ (ii) $= 0.6$ (iii) $\sigma_y = 4$
3) $\bar{x} = 62.5, r = 0.8$
4) 4
5) i) $r = \frac{-2}{3}$ (ii) $\frac{\sigma_x}{\sigma_y} = 1$
6) 32.5
7) $\bar{x} = 28, \bar{y} = 5.75, r = 0.5$
8) i) $\bar{x} = 4, \bar{y} = 5.2$ (ii) 0.375 (iii) $\frac{8}{3}$
9) i) $\bar{x} = 17$ (ii) $\bar{Y} = 19$ (iii) $b_{yx} = \frac{3}{4}$
   iv) $b_{xy} = \frac{1}{2}$ (v) $r = 0.61$
10) $\bar{x} = 30, \bar{y} = 40$ (ii) $r = \frac{2}{3}$
11) x on y is $10x + 3y - 62 = 0$

(ii) $\bar{x} = 5,\ \bar{y} = 4$ (iii) $r = -0.6$

12) (i) $r = 0.6$ (ii) $\bar{x} = 10,\ \bar{y} = 20$

13) (i) $\bar{x} = 17,\ \bar{y} = 19$, (ii) $b_{yx} = \frac{3}{4}$

(iv) $b_{xy} = \frac{1}{2}$ (v) $r = 0.61$

14) (i) $\bar{x} = 60,\ \bar{y} = 100$ (ii) $y = 137.5kg$

MISCELLANEOUS EXERCISE - 3

I) 1) c 2) a 3) b 4) a 5) b 6) b 7) a 8) c 9) c 10) b 11) b 12) b 13) a 14) b 15) b 16) c 17) a 18) b 19) a 20) b 21) a 22) d 23) b 24) b

II) 1) Negative

2) $y - \bar{y} = b_{xy}(x - \bar{x})$

3) $x - \bar{x} = b_{xy}(y - \bar{y})$

4) 2

5) -1

6) $\frac{c}{d} b_{uv}$

7) $\frac{d}{c} b_{uv}$

8) $2|r|$

9) <1

10) r^2

III) 1) True 2) False 3) True 4) True 5) False 6) False 7) False 8) True 9) True 10) True.

IV)

1) $y = 2.588x + 8.2$ For $x = 1.5$ $y = 12.082$

2) $byx = -\frac{3}{2}$

3) $Y = 8$

4) i) $r = \frac{1}{3}$ (ii) $6_y^2 = \frac{16}{9}$

5) Equation of y on x is $2x + 3y = 6$ and equation of x on y is $5x + 7y - 12 = 0$

6) i) $r = -0.6$, ii) $\bar{x} = 1, \bar{y} = -1$

7) i) $\bar{X} = 4, \bar{Y} = 7$ ii) $r = -0.5$ (iii) 10 (iv) 4

8) $(X - \bar{x}) = \frac{6}{11}(Y - \bar{y})$

9) $Y = 3.75x - 39.$

10) $Y = 0.7x + 105,\ Y = 133$

11) $(i)\bar{x} = 10, \bar{y} = 5$ (ii) $b_{yx} = 0.9$, $b_{xy} = 0.4$

iii) $V(X) = 16$ (iv) $r = 0.6$

12) $r = -0.36$

Exercise 4.1

2. Equation of trend line:

$y_t = 4.182 + (1.036)u$. where $u = t - 1967$.

Year	Trend Value	Year	Trend Value
1962	–1.0000	1967	4.1818
1963	0.0364	1968	5.2182
1964	1.0727	1969	6.2545
1965	2.1091	1970	7.2909
1966	3.1455	1971	8.3273
		1972	9.3636

For the year 1975, trend value = 12.4732 (in ‘000 tones)

3.

Year	Trend Value	Year	Trend Value
1962	–	1968	5.2
1963	–	1969	6.4
1964	1.2	1970	7.6
1965	1.6	1971	–
1966	2.4	1972	–
1967	4		

5. Equation of trend line:

$y_t = 4.2 + (0.4485)u$, where $u = 2t - 3961$

Year	Trend Value	Year	Trend Value
1976	0.1636	1981	4.6485
1977	1.0606	1982	5.5455
1978	1.9576	1983	6.4424
1979	2.8545	1984	7.3394
1980	3.7515	1985	8.2364

For the year 1987, trend value = 10.0305.

6.

Year	Trend Value
1976	–
1977	–
1978	2
1979	2.5
1980	3
1981	3.5
1982	4.25
1983	5.5
1984	7
1985	–

8. Equation of Trend line:

$y_t = 5.6364 + (0.7909)\ u$, where, $u = t - 1981$

Year	Trend Value	Year	Trend Value
1976	1.1618	1982	6.4273
1977	2.4727	1983	7.2182
1978	3.2636	1984	8.0091
1979	4.0545	1985	8.8000
1980	4.4855	1986	9.5909
1981	5.6364		

The trend value for the year 1990 = 12.7545 (in million tonnes.)

9.

Year	Trend Value	Year	Trend Value
1976	–	1982	7.3333
1977	2.6667	1983	6.0000
1978	3.3333	1984	7.6667
1979	4.0000	1985	8.0000
1980	5.3333	1986	–
1981	6.3333		

10. (i)

Year	Trend Value	Year	Trend Value
1962	–	1969	5
1963	–	1970	6
1964	0.8	1971	7
1965	1.4	1972	7.6
1966	2.2	1973	8.2
1967	3	1974	8.8
1968	4	1975	–
		1976	–

MISCELLANEOUS EXERCISE - 4

I) 1. (d) 2. (c) 3. (c) 4. (c) 5. (a) 6. (d) 7. (a) 8. (a) 9. (b) 10. (a)

II) 1. trend 2. seasonal 3. cyclical 4. irregular 5. assume 6. does not assume 7. graphical 8. moving average 9. least square 10. trend

III)

1. F	2. T	3. F	4. F	5. F
6. T	7. F	8. F	9. F	10. F

IV) 2. Euqation of trend line.

$y_t = 6.3333 + (0.6333)\ u$, where $u = t - 1978$

Year	Trend Value	Year	Trend Value
1974	3.8	1979	6.9667
1975	4.4333	1980	7.6000
1976	5.0667	1981	8.2333
1977	5.7	1982	8.8667
1978	6.3333		

3.

Year	Trend Value	Year	Trend Value
1974	–	1979	6.8
1975	–	1980	7
1976	6	1981	–
1977	7	1982	–
1978	7	1986	–

5. Equation of trend line:

$y_t = 3.1667 + (0.2797)u$, where $u = 2t - 3953$.

Year	Trend Value	Year	Trend Value
1971	0.0897	1977	3.4464
1972	0.6492	1978	4.0058
1973	1.2086	1979	4.5653
1974	1.7681	1980	5.1247
1975	2.3275	1981	5.6841
1976	2.8867	1982	6.2436

6.

Year	Trend Value
1971	-
1972	1
1973	1.5
1974	2
1975	2.5
1976	3.5
1977	4
1978	3.75
1979	4
1980	5
1981	-
1982	

8. Equation of trend line:

$yt = 5.1 + (0.4758)u$, where $u = \dfrac{2t - 3965}{5}$

Year	Trend Value	Year	Trend Value
1960	0.8182	1985	5.5758
1965	1.7697	1990	6.5273
1970	2.7212	1995	7.4788
1975	3.6727	2000	8.4303
1980	4.6242	2005	9.3818

Trend value for the year 2010 = 10.3338%

9.

Year	Trend Value
1960	-
1965	2.5
1970	3.5
1975	4
1980	4.75
1985	5.75
1990	6.75
1995	8
2000	-
2005	

11. Equation of trend line:

$y_t = 4.8750 + (0.4702)u$, where $u = 2t - 3961$

Year	Trend Value	Year	Trend Value
1977	1.5833	1981	5.3452
1978	2.5238	1982	6.2857
1979	3.4643	1983	7.2262
1980	4.4048	1984	8.1667

Trend value for the year 1988 = 11.928 (in ten thousands).

12.

Year	Trend Value	Year	Trend Value
1977	–	1981	7.3333
1978	1.3333	1982	6.3333
1979	3.6667	1983	5.6667
1980	7.0000	1984	–

14. Equation of trend line:

$y_t = 5.2222 + (0.667)\,u$, where $u = t - 1979$

Year	Trend Value	Year	Trend Value
1975	2.5556	1980	5.8889
1976	3.2222	1981	6.5556
1977	3.8889	1982	7.2222
1978	4.5556	1983	7.8889
1979	5.2222	1994	–

Trend value for the year 1984 = 8.5557.

15.

Year	Trend Value
1975	-
1976	4.25
1977	4.75
1978	5.5
1979	5.75
1980	5
1981	7
1982	-
1983	

17. Equation of trend line:

$$yt = 4.286 + (-1.571)u, \text{ where } u = \frac{t-1995}{5}$$

Year	Trend Value	Year	Trend Value
1980	9.0000	2000	2.7143
1985	7.4286	2005	1.1429
1990	5.8571	2010	– 0.4286
1995	4.2857		

Trend value for the year 2012 = –1.0554 (per ‘000).

18.

Year	Trend Value	Year	Trend Value
1980	–	2000	2.6667
1985	7.3333	2005	1.3333
1990	5.3333	2010	–
1995	4		

19. Equation of trend line:

$y_t = 12.4 + (0.2848)\,u$, where $u = 2t - 3927$

Year	Trend Value	Year	Trend Value
1959	–0.1636	1964	2.6848
1960	0.4061	1965	3.2545
1961	0.9758	1966	3.8242
1962	1.5455	1967	4.3939
1963	2.1152	1968	4.9636

20.

Year	Trend Value	Year	Trend Value
1959	–	1964	1.8
1960	–	1965	1.6
1961	1.4	1966	3.4
1962	1.4	1967	–
1963	2	1968	–

Exercise 5.1

1. 166.97
2. 138.92
3. 128.81
4. 171.59
5. 127.87
6. 131.26
7. 141.77
8. 74.52
9. 124.29
10. 128.87
11. 156.63
12. $x = 15$
13. $y = 60$

Exercise 5.2

1. $P_{01}(L) = 164.29$, $P_{01}(P) = 202.21$, $P_{01}(D - B) = 183.25$, $P_{01}(M - E) = 179.19$.

2. $P_{o1}(L) = 161.11$, $P_{o1}(P) = 161.82$, $P_{o1}(D–B) = 161.46$, $P_{o1}(Al – E) = 161.42$.

3. $P_{o1}(L) = 96.88$, $P_{o1}(P) = 94.47$, $P_{o1}(D – B) = 95.67$, $P_{o1}(M – E) = 95.63$.

4. $P_{o1}(L) = 130.26$, $P_{o1}(P) = 129.19$, $P_{o1}(D –B) = 129.73$, $P_{o1}(M – E) = 129.64$.

5. $P_{o1}(D–B) = 65$, $P_{o1}(F) = 60$.

6. $P_{o1}(L) = 250$, $P_{o1}(P) = 230$, $P_{o1}(D – B) = 240$, P_{o1} (id – E) = 238.24.

7. $P_{o1}(P) = 168.04$

8. P_{o1} (L) = 250.

9. z = 2.

10. $P_{o1}.(D – B) = 1P_{o1}(F)$

11. $P_{o1}(P) = 8$ and $P_{o1}(L) = 2$ because a.m.>g.m.

Exercise 5.3

1. 135.04 2. 115.78 3. 110.67
4. 77 5. 205.88 6. 169.33
7. 2= 18 8. v = 6 9. ₹ 16000

MISCELLAISEOUS EXERCISE - 5

I) 1. (c) 2. (c) 3. (d) 4. (c) 5. (c) 6. (d) 7. (c) 8. (d) 9. (c) 10. (a) 11. (a) 12. (a) 13. (a) 14. (a) .

II)

1. $$\frac{\sum p_1}{\sum p_0}\times 100$$

2. $$\frac{\sum q_1}{\sum q_0}\times 100$$

3. $$\frac{\sum P_1 q_1}{\sum q_1 q_0}\times 100$$

4. $$\frac{\sum P_1 w}{\sum p_0 w}\times 100$$

5. $$\frac{\sum q_1 w}{\sum q_0 w}\times 100$$

6. $$\frac{\sum P_1 q_1 w}{\sum q_0 q_0 w}\times 100$$

7. $$\frac{\sum P_1 q_0}{\sum p_0 q_0}\times 100$$

8. $$\frac{\sum P_1 q_1}{\sum p_0 q_1}\times 100$$

9. $$\frac{1}{2}\left[\frac{\sum P_1 q_0}{\sum p_0 q_0}+\frac{\sum P_1 q_1}{\sum p_0 q_1}\right]\times 100$$

10. $$\sqrt{\frac{\sum P_1 q_0}{\sum p_0 q_0}\times\frac{\sum P_1 q_1}{\sum p_0 q_1}}$$

11. $$\frac{\sum {}_1 (q_0 + q_1)}{\sum p_0 (q_0 + q_1)}\times 100$$

12. $$\frac{\sum P_1 \sqrt{q_0 q_1}}{\sum p_0 \sqrt{p_0 q_1}}\times 100$$

III) 1.T 2.F 3.F 4.F 5.F
6.F 7.I? 8.F 9.F 10.F

IV)
1. 154.55 2. 135.48
3. 121.43 4. 42
5. $P_{oi}(L) = 10$
6. 19, $P_{o1}(P) = 136.54$ 6. P_{o1} (D – B) = 134.27
7. $P_{ot}(I_{ti}$ E) = 107.14
8. $P_{oi}(W) = 177.10$
9. $P_{01}(L) = 181.82$, $P_{01}(P) = 172.22$
10. $x = 6$
11. $x = 16.6$
12. 10.75
13. P01(D – B) = 20.5, P01(F) = 20
14. $P_{01}(P) = 155.4$

15. P_{01} (L) = 166.67,
P_{01} (P) = 87.5, P_{01} (D – B) = 127.085,
P_{01}(M–E) = 121.42.

16. P_{01} = 153.85,P_{01}
(P) = 87.5.P_{01}(D – B) = 120.68, P_{01}
(M–E) = 117.24

17. E_{poqi} = 228, P_{01}(P) = 131.58

18. 106.42

19. 86.5

20. 295.25, Its. 29,525

21. $X = 15$

22. ₹ 18,900

Exercise 6.1

1. Maximize $z = 30x + 20y$ Subject to
$10x + 6y \leq 60$, $5x + 4y \leq 35$, $x \geq 0$, $y \geq 0$

2. Minimize $z = 3x + 2y$ Subject to
$2x + y \geq 14$, $2x + 3y \geq 22$, $x + y \geq 1$ $x \geq 0$, $y \geq 0$

3. maximize $P = 350x \times 400y$ Subject to
$3x + 2y \leq 120$, $2x + 5y \leq 160$, $x \geq 0$, $y \geq 0$

4. Maximize $z = 10x + 15y$ Subject to
$2x + 3y \leq 36$, $5x + 2y \leq 50$, $2x + 6y \leq 60$
$x \geq 0$, $y \geq 0$

5. Maximize $P = 13.5x + 55y$ Subject to
$x + 2y \leq 10$, $3x + 4y \leq 12$,
$x \geq 0$, $y \geq 0$

6. Maximize $z = 500x + 750y$ Subject to
$2x + 3y \leq 40$, $x + 4y \leq 70$, $x \geq 0$, $y \geq 0$

7. Minimize $z = 4.5x + 3.5y$ Subject to
$4x + 6y \geq 18$, $14x + 12y \geq 28$, $7x + 8y \geq 14$
$x \geq 0$, $y \geq 0$

8. Maximize $z = x_1 + x_2$ Subject to
$x_1/60 + x_2/90 \leq 1$, $5x_1 + 8x_2 \leq 600$, $x \geq 0$, $y \geq 0$

9. Minimize $C = 20x_1 + 6x_2$ Subject to
$x_1 \geq 4$, $x_2 \leq 2$, $x_1 + x_2 \geq 5$, $x_1 \geq 0$, $x_2 \geq 0$

Exercise 6.2

1. $x = 4$, $y = 2$, $z = 60$
2. $x = 0$, $y = 6$, $z = 36$
3. $x = 4.5$, $y = 2.5$, $z = 59$
4. $x = 2$, $y = 3$, $z = 95$
5. $x = 4$, $y = 5$, $z = 37$
6. $x = 0$, $y = 5$, $z = 5$
7. $x = 1.5$, $y = 4$, $z = 52$
8. $x = 2$, $y = 0.5$, $z = 22.5$

MISCELLAISEOUS EXERCISE - 6

I)

1	2	3	4	5	6	7	8	9	10	11	12	13	14	15
a	c	b	c	a	d	c	b	a	b	b	b	a	c	c

II) 1) I　2) III　3) vertex
4) III and IV　5) $y > x$
6) $x \geq 3$, $y \geq 2$,　7) $x \geq 2y$
8) $x \leq 40$

III) 1), 4), 5), 6), 7) are true and 2), 3) are false.

IV) Subject to

1. $x_1 = 4.5$, $x_2 = 3$ Max $z = 40.5$
2. $x = 3$, $y = 18$, Min $z = 48$
3. infinite number of optimum solution on the line $3x + 5y = 10$ between A (45/16, 5/16) and B(0,2)
4. $x_1 = 2$, $x_2 = 1$ Min. $z = 7$.
5. $x = 4$, $y = 3$ Max $z = 25$
6. $x = 10$, $y = 15$ Max $z = 1350$
7. $x = 3$, $y = 18$ Min $z = 48$.
8. Max $z = 140x + 210y$ Subject to
$3x + 3y \leq 36$, $5x + 2y \leq 50$, $2x + 6y \leq 60$

$x, y \geq 0$ where x = No of tables = 3
y = no of chairs = 9.
Max z = maximum profit = ₹ 2310/–

9. Maximize $z = 180x + 220y$ Subject to
$6x + 4y \leq 120$, $3x + 10y \leq 180$
$x \geq 0, y \geq 0$ Ans. $x = 10, y = 15$.

10. Minimize $z = 4x + 6y$ Subject to
$x + 2y \geq 80$, $3x + y \geq 75$, $x \geq 0, y \geq 0$
Ans. $x = 14, y = 33$

11. Maximize $z = 2000x + 3000y$ Subject to
$3x + 3y \leq 36$, $5x + 2y \leq 50$, $2x + 6y \leq 60$,
$x \geq 0, y \geq 0$ Ans. $x = 3, y = 9$

12. Minimize $z = 800x + 640y$ Subject to
$4x + 2y \geq 16$, $12x + 2y \geq 24$
$2x + 6y \geq 18$, $x \geq 0, y \geq 0$
Ans. Minimum cost = ₹ 3680/–
when $x = 3, y = 2$

13. Maximize $z = 75x + 125y$ subject to
$4x + 2y \leq 208$, $2x + 4y \leq 152$, $x \geq 0, y \geq 0$
Ans $x = 44$, $y = 16$

14) Max $z = 3x + 4y$, s.t. $x + y \leq 450$, $2x + y \leq 600$ $x_1 y \geq 0$
Maximum Profit = ₹ 1800/– at (0, 450)

15. Max $z = 20x + 30y$, s.t. $2x + 2y \leq 210$, $3x + 4y \leq 300$ $x_1 y \geq 0$
Maximum Profit = ₹ 2400/– at (30, 60)

Exercise 7.1

1. P → II Q →1V, R → I, S → III Total cost = 99

2. 1 →I, 2→Ill, 3→IV, 4 → II, 5 → V Total cost = 39

3. 1→C, 2→E 3→A, 4→D, 5→B Total cost = 94

4. M1→A, M2 → B, M3 → E, M4 → D, M5 → C Total cost = 12

5. 1 → A, 2 → C, 3 → B, 4 → D Total Profit = 61

6. M1 → A, M2→ B, M3 → E, M4 →D Total cost = 32

Exercise 7.2

1. Optimal sequence is : 4 –1–3 – 2 – 5 – 6
Idle time for turning operation = 1 min
Total elapsed time = 43 minutes Idle time for threading = 6 minutes

2. Optimal sequence is : I — III — II Total
Idle time for department A = 4 days
elapsed time = 23 days Idle time for department B = 8 days

3. Optimal sequence is : 3 –1– 2
Idle time for date entry operation = 140 min.
Total elapsed time = 430 minutes Idle time for filing = 80 min.

4. Optimal sequence is : B – D – C – E – A
Idle time for machine X = 4 hrs
Total elapsed time = 60 hrs. Idle time for machine Y = 6 hrs.

5. Optimal sequence is : VII – I – IV – V – III – II – VI Idle time for machine.
Idle time for machine B = 13. Total elapsed time = 91 units.

6.

1) Optimal sequence is : V – III – II – VI – VII – IV – I OR III – V – II – VI – VII – IV – I Total elapsed time = 61 hrs Idle time for machine B = 38 hrs. Idle time for machine A = 7 hrs. Idle time for machine C = 15 hrs.

2) Total elapsed time = 40
OR

2 – 5 – 4 – 3 – 1 Idle time for machine A = 8 hrs. Idle time for machine C = 12 hrs machine B is 25 hrs.

7. Optimal sequence is :
   Total elapsed time = 51 hrs. Idle time for machine A = 19 hrs. Idle time for machine B = 31 hrs. 1–4–5–2–3.
   Ideal time for machine C = 9 hrs.

MISCELLANEOUS EXERCISE - 7

I)

1	2	3	4	5	6	7	8	9	10
c	b	c	d	b	d	a	a	d	c

11	12	13	14
a	a	b	b

II)

1. Number of rows is not equal to the number of columns.
2. Balanced
3. Square
4. Square
5. Zero
6. Total elapsed time
7. Idle time
8. Maximum
9. Multiple
10. A–D–B–C

III) 1. False 2. True 3. False 4. True 5. False 6. True 7. True 8. False 9. True 10. False

PART - I

IV)

1. A → I, B →III, C →II, D → IV;
   Minimum Cost = 37
2. A →II, B →III, C → V, D →I, E → IV;
   man hrs Minimum Cost = 525 kms
3. A → V, B →II, C → IV, D → III, E →I;
   Maximum Sale = 65
4. P→IV, Q → III, R → V, S →I, X →II;
5. A → IV, B → III, C → II, D → V, E →I
   A → IV, B→ III, C →V, D → II E → I;
6. E_1 →1, E_2 →IV, E_3 → II, E_4 → V, E_S → III
   Total Expenditure = 20.

PART - II

1. Optimal sequence: 3–4–5–7–2–6–1;
   Idle time for cutting = 2 hrs;
   Total elapsed time = 46 hrs; Idle time for sewing = 4 hrs.
2. Optimal sequence: II–IV–V–III–1;
   Idle time for cutting = 4 hrs;
   Total elapsed time – 21 hrs Idle time for sewing 3 hrs
3. Optimal sequence: III–V–II–VI–I–IV–VII;
   Idle time for cutting = 3 hrs;
   Total elapsed time 55 hrs Idle time for sewing 9 hrs
4. Optimal sequence: 3–2–5–4–1;
   Idle time for machine A = 18 hrs;
   dle time for machine C = 38
   Total elapsed time 102 hrs Idel time for machine B = 62 hrs.
5. Optimal sequence is : 3–1–4–2;
   Idle time for machine P = 12 min.; Idle time for machine R = 31 min.
   Total elapsed time = 74 min. Idle time for machine Q=51 min

Exercise 8.1

1. {0, 2, 4, 6}
2. {0, 1, 2}
3. (i) Probability distribution. All entries are positive and add to 1. (ii) Not a probability distribution. Probability cannot be negative.

(iii) Probability distribution. All entries are positive and add to 1. (iv) Probability distribution. All entries are positive and add to 1. (v) Not a probability distribution. Probabilities do not add to 1. (vi) Not a probability distribution.Probabilites do not add to 1.

4. (i)

x	0	1	2
$P(x)$	1/4	1/2	1/4

(ii)

x	0	1	2	3
$P(x)$	1/8	3/8	3/8	1/8

(iii)

x	0	1	2	3	4
$P(x)$	1/16	1/4	3/8	1/4	1/16

5.

x	0	1	2
$P(x)$	4/9	4/9	1/9

6.

x	0	1	2
P(x)	$(4/5)^4$	$4(4/5)^3$ $(1/5)$	$6(4/5)^2$ $(1/5)^2$
z	3	4	
P(x)	$4/5(1/5)^3$	$(1/5)^4$	

7.

x	0	1	2
$P(x)$	9/16	3/8	1/16

8. (i) 1/10 (ii) 3/10 (iii)3/10 (iv)1/5 = 0.2

9. −0.05, 2.2475

10. 7/2 = 3.5 m 35/12

11. 1.5

12. 1/3

13. 14/3 = 4.667

14. 35/6 = 5.8333,

15.

x	14	15	16	17
P $(X = x)$	2/15	1/15	2/15	3/15
x	18	19	20	21
P $(X = x)$	1/15	2/15	3/15	1/15

Mean = 17.5333, Variance = 5.12381.

16. 0.7, 0.21

Exercise 8.2

1. (i) p. d. f. (ii) p. d. f.
2. (ii) 2.25/16, (ii) 3/16 (iii) 3/4
3. (ii) 1/256 (iii) 5/16
4. (i) 1/2, 35/64 (ii) 6, (a) 11/32, (b) 1/12
5. (a)1/4 (b) 1/2 (c) 7/16
6. (i) 2/5 (ii) 1/5
7. (i) 1/2 (ii) 11/16 (iii) 81/128
8. (i) $x^2/16$ for $0 < x < 4$ (ii) 1/64, 2.89/16, 1
9. (i) 1/8 (ii) 0, (iii) 1 (iv) 7/8
10. (i) 1/log(3), (ii) 2/log (3),

    (iii) $\dfrac{4[log3-1]}{[log3]^2}$

Exercise 8.3

1. (i) 0.375 (ii)0.3125 (iii)0.6875
2. 5/12
3. $1.3 \times (0.9)^3$
4. (i) $1/13^4$ (ii)$12/13^4$ (iii) $(12/13)^4$
5. (i) $(0.8)^5$ (ii) $(1.8)(0.8)^4$ (iii)$1 - (1.8)(0.8)^4$ (iv) $1 - (0.8)^5$
6. (9/10)4
7. 11/243
8. $(1/2)\,(5/6)^5$
9. (i) 4 , 2.4 (ii) 6, 2.4 (iii) 2/5 (iv) 0.16

Exercise 8.4

1. 0.7357
2. 0.0126
3. 0.8008
4. (i) 0.1465 (ii) 0.2381
5. (i) 0.2231 (ii) 0.1911
6. (i) 0.3678 (ii) 0.6322
7. (i) 0.1754 (ii) 0.3840 (iii) 0.4261

MISCELLANEOUS EXERCISE - 8

I) 1. d 2. a 3. b 4. d 5. c
6. b 7. b 8. a 9. b 10. a

II)

1. Counting
2. Measurement
3. 1
4. 1
5. 0
6. Centre of gravity
7. Distribution function
8. Remains constant / independent
9. Poisson

III) 1. False 2. True 3. True 4. True 5. True
6. False 7. True 8. True 9. True

SOLVED EXAMPLES

PART - I

1. (i) Discrete, {0,1,2,....., 100000}
(ii) Continuous (iii) Continuous
(iv) Discrete {0,1,2,3,4,5}. (v) Continuous.
2. (i) 1/21 (ii) 10/21, 1/7, 6/7
3. (i)0.5 (ii) 0.7 (iii) 0.55 (iv) 0.45
4. Bothe probabilities are 1/2 and therefore are equal.
5. a = 1/4

x	1	2	3	4	5
$P(x)$	1/20	3/20	1/4	1/2	1/20
$F(x)$	1/20	1/5	9/20	10/20	1

6. The distribution of X is binomial with n = 4 and p = 1/2. The formula of its p. m. f. is as follows. $P(X = x) = \binom{4}{x}\frac{1}{2^4}$.

The probability distribution of X is tabulated below.

x	0	1	2	3	4
$P(x)$	1/16	1/4	3/8	1/4	1/16

7. (i)

x	0	1	2
$P(x)$	4/9	4/9	1/9

(ii)

x	0	1	2
$P(x)$	25/36	10/36	1/36

8. (i) 1/10 (ii) 3/10 (iii) 17/100 (iv) 3/10
9.

x	-3	-2	-1	0	1	2	3	4
$F(x)$	0.1	0.3	0.5	0.65	0.75	0.85	0.9	1
$P(x)$	0.1	0.2	0.2	0.15	0.10	0.10	0.05	0.10

10. (i) 11/5, 14/25 (ii) 1/5, 14/25
(iii) $(n + 1)/2$, $(n^2 - 1)/12$ (iv) 2.5, 1.25
11. ₹ 5.5, 8.25
12. 0, 1
13. (i) 1/2 (ii) 2/3 (iii) 1
0, 1
14. (i) 1/2, (ii) 11/16, (iii) 81/128]
15. Both probabilities are 1/4 and hence are equal.
16. $k = \theta,\ 1/e,\ \frac{1}{\theta}\log 2$
17. $k = 1/4,\ F(2) = \frac{\sqrt{2}}{2},\ P(X \geq 1) = 1/2$

18. 1/2

PART - II

1. (i) $2(0.8)^9$ (ii) $1-(0.8)^{10}$ (iii) $1-(8,2)(0.2)^9$
2. (i) $p = 1/2$, $\mathrm{Var}(X) = 2.5$ (ii) (i) $n = 10$, $p = 1/2$
3. (i) 63/256, (ii) 1/1024
4. $45\,\frac{4^8}{5^{10}}$
5. 0.784
6. 0.9477
7. 0.3456
8. $\frac{761}{510}$
9. (i) $(0.998)^8$
10. (i) $(0.9)^{10}$ (ii) $(0.9)^9$
11. (i) $\left(\frac{1}{5^4},\frac{16}{54}\right)\frac{96}{5^4},\frac{256}{5^4},\frac{256}{5^4}$

    (ii) (a) $\frac{608}{5^4}$ (b) $1-\frac{33}{5^4}$
12. (i) $\frac{35\times 8\times 81}{5^7}$ (ii) $1-17\frac{3^6}{5^7}$
13. (i) 0.3687 , (ii) 0.0613

14) m = 2

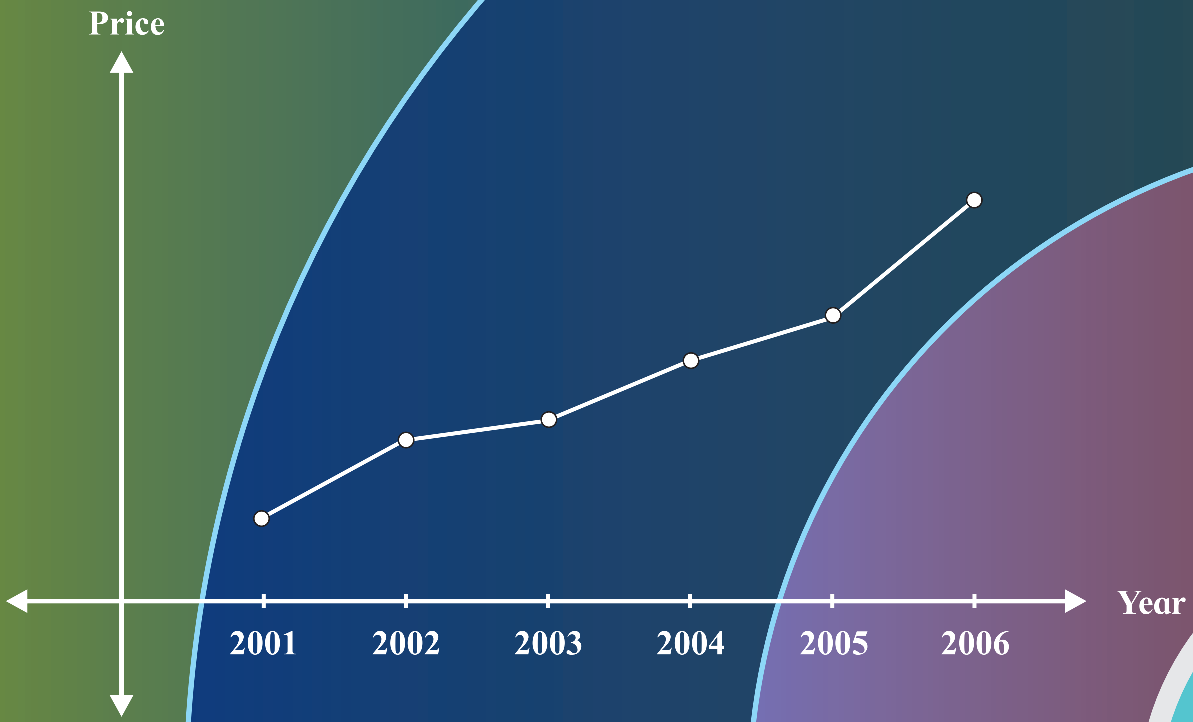
◆ ◆ ◆

NOTES

Machine \ Job	A	B	C	D	E
M1	8	4	5	3	6
M2	5	6	2	7	4

Optimal Sequence

D	B	A	E	C



Maharashtra State Bureau of Textbook
Production and Curriculum Research,
Pune - 411 004
₹ 103.00
इंग्रजी गणित आणि संख्याशास्त्र (वाणिज्य) भाग-२, इ. १२ वी