

CONTENTS

Unit No.	Title	Page No.
SEMESTER - I		
1.	Shares and Mutual Funds	01
2.	Permutations and Combinations	16
3.	Linear Programming Problems	47
4.	Measures of Central Tendency	59
5.	Measures of Dispersion	85
6.	Elementary Probability Theory	99
7.	Statistical Decision Theory	117



UNIT I**Unit -1****SHARES AND MUTUAL FUNDS****Unit Structure :**

- 1.0 Objectives
- 1.1 Introduction
 - 1.1.1 Bonus Shares
 - 1.1.2 Splitting of shares
- 1.2 Mutual Funds
- 1.3 Systematic Investment Plan (SIP)

1.0 OBJECTIVES:

After going through this chapter you will be able to:

- Define a share, face value, market value, dividend, equity shares preferential shares, bonus shares.
- Understand the concept of Mutual fund.
- Calculate Net Income after considering entry load, dividend, change in Net Asset Value (N.A.V) and exit load.
- Understand the Systematic Investment Plan (S.I.P).

1.1 INTRODUCTION

In day-to-day life we hear about shares, share market etc. Here we will see, exactly what these terms deal with .

When a group of persons plan to establish a company, they form a company under the companies Act 1956. Now this company is an established company. The people who establish this company are called promoters of the company. These promoters can now raise a certain amount of capital to start (run) the company. They divide this required amount into small parts called shares.

A share is the smallest unit of capital of a company. Stock is the American term for share. Usually a share is of value Rs. 100 /- or Rs.50/- or Rs. 10 /-or Rs. 5/- or Rs. 2/- or Rs./- 1 . This value is called the face value of the share. These shares are sold to the public. (usually face value is Rs. 10/- , unless otherwise specified). This sale is called the Initial Public Offer (IPO) of the company.

The company issues share certificates to the persons from whom it accepts the money to raise the capital. Persons who have paid money to form the capital are called share holders. Now-a- days the shares are not in the form of paper, but in the electronic dematerialised (Demat) form, hence the allotment of shares is done directly in the demat account, without a certificate.

Face value or nominal value or Par value is the value printed on the share certificate. Since shares exist in electronic demat form, we can say that the face value is the value stated in the I.P.O. subscription form.

The shareholders enjoy the profits (if any) of the company, after providing for the taxes and other reserve funds. This is called as **dividend**.

Types of shares ; Mainly the shares are of two types **i) Preference shares** and **ii) Equity shares** or common shares or ordinary shares .

i) Preference shares : These shares have a priority over the equity shares. From the profits made by a company, a dividend at a fixed rate is paid to them first, before distributing any profit amount to the equity shareholders. Also, if and when the company is closed down then while returning of the capital, these shareholders get a preference. Again, preference shares are mainly of two types:

- a) Cumulative Preference shares:** In case of loss or inadequate profit , The preference shareholders are not paid their fixed rate of dividend , then the dividend is accumulated in the subsequent years to these shareholders & is paid preferentially whenever possible .
- b) Non-cumulative Preference shares:** As in the case of cumulative preference shares, here the unpaid dividends do not accumulate.

ii) Equity shares : These are the shares for whom the dividend and the return of capital is paid after paying the preference shareholders. In case of equity shares , the rate of dividend is not fixed and it is decided by the Board of Directors .

Share Market

Shareholders are allowed to buy or sell shares like commodities. Selling or buying a share for a price higher than its face value is legal. The share prices are allowed to be subject to the market forces of demand and supply and thus the prices at which shares are traded can be above or below the face value.

The place at which the shares are bought and sold is called a share market or stock Exchange and the price at which a share is traded is called its Market Price (MP) or the Market value. If the market price of a share is same as its face value, then the share is said to be traded at Par.

If M.P. is greater than face value of a share , then the share is said to be available at a premium or above par and is called premium share or above par share.

If M.P. is lower than face value of a share , then the share is said to be available at a discount or below par & the share is called a discount share or below par share .

The purchase and sale of shares can take place through brokerage firms and depository Participant (DP). e.g. Sharekhan.com, Kotak Securities ltd , ICICI direct.com etc. They charge a commission on the purchase and sale of shares, which is called as a brokerage. The brokerage is charged as a percentage of the M.P. of the share. Normally it is below 1%.

1.1.1 Bonus Shares

Sometimes, when a company's free reserves are high, it may choose to capitalize a part of it by converting it into shares. This is done by issuing bonus shares to existing shareholders. These bonus shares are issued free of cost. The ratio of bonus shares to the existing shares is fixed.

Getting bonus shares increases the number of shares of shareholders. But since this applies to all the shareholders in a fixed ratio, hence the percentage of a shareholder's ownership of the company remains same as before.

Now, we will study some examples based on the above concepts:

Example 1

Mr. Prashant invested Rs. 75,375/- to purchase equity shares of a company at market price of Rs. 250 /- through a brokerage firm, charging 0.5% brokerage. The face value of a share is Rs. 10/-. How many shares did Mr. Prashant purchase?

Solution: Brokerage per share = $250 \times \frac{0.5}{100} = 1.25$

$\therefore$ cost of purchasing one share = $250 + 1.25 = 251.25$

$\therefore$ Number of shares purchased = $\frac{75375}{251.25} = 300$

Example 2

Mr. Sandeep received Rs. 4,30,272 /- after selling shares of a company at market price of Rs. 720 /- through Sharekhan Ltd., with brokerage 0.4%. Find the number of shares he sold.

Solution: Brokerage per share = $720 \times \frac{0.4}{100} = 2.88$

selling price of a share = $720 - 2.88 = 717.12$

$\therefore$ Number of shares sold = $\frac{430272}{717.12} = 600$

Example 3

Ashus Beauty World ' has issued 60,000 shares of par value of Rs. 10/- each. The company declared a total dividend of Rs. 72,000 /- . Find the rate of dividend paid by the company.

Solution: Face value of 60,000 shares = 60,000 x 10 = 6,00,000

$$\text{Rate of Dividend} = \frac{\text{Total Dividend}}{\text{Face value of 60,000 share}} \times 100$$

$$= \frac{72000}{600000} \times 100 = 12$$

∴ The rate of dividend paid by the company is 12%

Example 4

The capital of ABC Company consists of Rs. 15 lakhs in 6 % cumulative preference shares of Rs. 100 each and Rs. 30 lakhs in equity shares of Rs.10/- each. The dividends on cumulative preference shares for earlier year was not paid . This year , the company has to distribute profit of Rs . 3 lakh after keeping 20 % as reserve fund. Find the percentage rate of dividend paid to the equity shareholders.

$$\text{Solution: Reserve fund} = \frac{20}{100} \times 300000 = \text{Rs. } 60,000/-$$

$$\text{Profit to be distributed} = 3,00,000 - 60,000 = 2,40,000$$

Annual dividend for 6 % cumulative preference shareholders

$$= \frac{6}{100} \times 1500000 = 90,000$$

This needs to be paid for 2 years (last year & current year) as the preference shares are cumulative & last year's dividend was not paid .

$$\begin{aligned} \therefore \text{Total Dividend paid to Preference shareholders} \\ = 2 \times 90000 = 1,80,000/- \end{aligned}$$

Now , dividend to be distributed to the equity shareholders

$$= 2,40,000 - 1,80,000 = \text{Rs. } 60,000/-$$

$$\therefore \text{Rate of dividend} = \frac{60,000}{30,00,000} \times 100 = 2$$

∴ The rate of dividend to the equity shareholders is 2 %

Example 5

Mr. Dinesh bought some shares of a company which had a face value of Rs.100 /-. The company declared a dividend of 15 % but Mr. Dinesh's rate of return on investment was only 12% . At what market price did he purchase the shares ? There was no brokerage involved.

Solution:

$$\text{Dividend on one share} = \frac{\text{Rate of Dividend}}{100} \times \text{face value of one share}$$

$$= 15 \times 100 = \text{Rs. } 15/-$$

$$\therefore \text{Rate of Return on investment} = \frac{\text{Dividend on one share}}{\text{purchase price of 1 share}} \times 100$$

$$\therefore 12 = \frac{15}{\text{purchase of 1 share}} \times 100$$

$$\therefore \text{purchase price of 1 share} = \frac{15}{12} \times 100 = 125$$

Example 6: Comparison of two stocks

Mr. Subu invested Rs. 20,000 /- in Rs. 100/- shares of company A at the rate of Rs. 125/- per share . He received 10 % dividend on these shares. Mr. Subu also invested Rs. 24,000/- in Rs. 10/- shares of company B at Rs.12/- per share and he received 15 % dividend. Which investment is more beneficial?

Solution : Company A

$$\text{Rate of return} = \frac{\text{Dividend on one share}}{\text{purchase price of 1 share}} \times 100 = \frac{10}{125} \times 100 = 8\%$$

Company B

$$\text{Rate of return} = \frac{1.5}{12} \times 100 = 12.5 \%$$

Investment in company B is more profitable .

Example 7

Ms. Ashma Mehta bought 300 shares of a company of face value Rs. 100 /- each at a market price of Rs. 240 /- each . After receiving a dividend at 8 % , she sold the shares at Rs . 256 /- each. Find her rate of return on investment. There was no brokerage involved.

Solution: Difference in the market price = 256-240= 16

$$\begin{aligned} \text{Dividend on 1 share} &= \frac{\text{Rate of dividend}}{100} \times \text{face value of 1 share} \\ &= \frac{8}{100} \times 100 = 8 \end{aligned}$$

Rate of Return on Investment

$$\begin{aligned} &= \frac{(\text{Price change}) + (\text{Dividend on 1 share})}{\text{purchase price of 1 share}} \times 100 = \frac{16+8}{240} \times 100 \\ &= \frac{2400}{240} = 10 \end{aligned}$$

$\therefore$ The rate of return on investment was 10 % .

1.1.2 Splitting of shares:

Sometimes companies split the face value of a share & break it up into smaller units . For e.g. a Rs. 100 /- share can be split into 10 shares each of face value Rs. 10 /- or a Rs. 10/- share can be split into two shares of face value Rs. 5/- each . Usually this does not affect a shareholder's wealth . However , it can make selling of a part of the holdings easier.

Example 8

Mr. Joshi purchased 30 shares of Rs. 10/- each of Medi computers Ltd. on 20th Jan. 2007, at Rs. 36/- per share. On 3rd April 2007, the company decided to split their shares from the face value of Rs. 10/- per share to Rs. 2/- per share. On 4th April 2007, the market value of each share was Rs. 8/- per share. Find Mr. Joshi's gain or loss, if he was to sell the shares on 4th April 2007? (No brokerage was involved in the transaction).

Solution: On 20th Jan 2007

$$\text{Purchase cost of 30 shares} = 30 \times 36 = 1080/-$$

On 3rd April 2007, each Rs. 10/- share became 5 shares of Rs. 2/- each.

$$\therefore \text{No. of shares} = 30 \times 5 = 150$$

$$\begin{aligned} \text{On 4th April 2007, market value of 150 shares was @ Rs. 8 each} \\ = 150 \times 8 = 1200 \end{aligned}$$

$$\therefore \text{His gain} = 1200 - 1080 = 120/-$$

Example 9

Rahul purchased 500 shares of Rs. 100 of company A at Rs. 700 /-. After 2 months, he received a dividend of 25 % . After 6 months, he also got one bonus share for every 4 shares held. After 5 months, he sold all his shares at Rs. 610/- each. The brokerage was 2% on both, purchases & sales. Find his percentage return on the investment.

Solution: For purchase:

$$\text{Face value} = \text{Rs. } 100 \text{ /- , No. of shares} = 500, \text{ market price} = \text{Rs. } 700/-$$

$$\text{Dividend} = 25\%, \text{ brokerage} = 2\%$$

$$\text{Purchase price of one share} = 700 + \frac{2}{100} \times 700 = 714$$

$$\therefore \text{Total purchase} = 500 \times 714 = \text{Rs. } 3,57,000/-$$

$$\text{Dividend} = \frac{25}{100} \text{ of } 100 \text{ i.e. Rs. } 25 \text{ /- per share}$$

$$\therefore \text{Total dividend} = 500 \times 25 = \text{Rs. } 12,500 \text{ /-}$$

Now, bonus shares are 1 for every 4 shares .

$$\therefore \text{No. of bonus shares} = \frac{1}{4} \times 500 = 125$$

$$\therefore \text{Total No. of shares} = 500 + 125 = 625$$

For sales,

$$\text{No. of shares} = 625, \text{ market price} = 610 \text{ , Brokerage } 2\%$$

$$\text{Sale price of one share} = 610 - 2\% \text{ of } 610 = 597.8$$

$$\begin{aligned} \therefore \text{Total sale value} &= \text{sale price of one share} \times \text{No. of shares} \\ &= 597.8 \times 625 = \text{Rs. } 3,73,625/- \end{aligned}$$

$$\begin{aligned} \text{Net profit} &= \text{sale value} + \text{Dividend} - \text{purchase value} \\ &= 3,73,625 + 12,500 - 3,57,000 \end{aligned}$$

$$\begin{aligned}
 &= \text{Rs. } 29,125/- \\
 \therefore \% \text{ gain} &= \frac{29,125}{3,57,000} \times 100 = 8.16 \\
 &= 8.16
 \end{aligned}$$

EXERCISE :

1) Mr. Amar invested Rs 1,20,480/- to buy equity shares of a company at market price of Rs . 480 /- at 0.4 % brokerage. Find the No. of shares he purchased.

Ans: 250

2) Aditi invested Rs. 19,890 /- to purchase shares of a company with face value of Rs.10/- each , at market price of Rs. 130/- . She received dividend of 20 % as well Afterwards , she sold these shares at market price of Rs. 180/- . She had to pay brokerage of 2 % for both purchase and sales of shares. Find her net profit.

Ans: No. of shares =150 , sales = 26460 , Dividend = 300 , purchase = 19,890, profit= 6870

3) Amol wants to invest some amount in company A or company B , by purchasing equity shares of face value of Rs. 10 /- each , with market price of R. 360/- and Rs. 470/- respectively . The companies are expected to declare dividends at 20 % and 45% respectively . Advise him on the choice of shares of company.

Ans: company B is a better choice .

4) Find the percentage gain or loss if 200 shares of face value Rs. 10/- were purchased at Rs . 350/- each and sold later at Rs. 352 /- , the brokerage being 0.5 % on each of the transaction .

Ans: -0.43 % i.e. a loss of

43 %

5) Find the number of shares if the total dividend at 8% on the shares with face value Rs.10/- was Rs. 240.

Ans :-

300

1.2 MUTUAL FUNDS

In the previous unit shares, we have studied how one can transact in shares. Now, we will study what are the mutual funds and how they function.

An investor can invest money directly in shares or he can invest his money through mutual funds. Mutual funds are managed by large financial services with a professional team of fund Managers & research experts.

Mutual fund is a pool of money, drawn from investors .The amount collected is invested in different portfolios of securities, by the

fund managers and the profits (returns), proportional to the investment, are passed back to the investors.

At a given time, the total value is divided by the total number of units to get the value of a single unit at a given time. This is called Net Asset Value (NAV).

$$\therefore \text{NAV} = \frac{\text{Net Assets of the scheme}}{\text{Total No. of units outstanding}}$$

$$\text{or } \text{NAV} = \frac{\text{Total Assets} - \text{liabilities}}{\text{Total No. of units outstanding}}$$

There are mixed or hybrid funds which invest in both debt and equity. The offer documents give the guidelines / constraints under which the fund managers would operate. e.g. investment in equity 80 % to 100 %, investment in money markets 0 % to 20 % etc.

In India, the mutual funds are governed by SEBI (Securities and Exchange Board of India). There are different companies, called the 'Fund Houses' (like SBI or Reliance or HDFC) which float different mutual funds. Each such fund is called a 'scheme', e.g. HDFC has a scheme 'HDFC Tax saver' etc.

Like IPO of a company's share, a mutual fund scheme starts by having a N.F.O. (New Fund Offer). Investors can invest by purchasing Units of the mutual funds. Usually a unit is of Rs. 10/-. A share is the smallest unit of a company's capital, whereas in mutual funds, even a fraction of a unit can be purchased after the N.F.O.

Let us study the following example to understand this concept:

Example 10

A mutual fund's scheme shows the following on 01/01/2007

Total value of securities (Equity, Bonds etc.)	Rs. 1500 crores
Cash	Rs. 100 crores
Liabilities	Rs. 200 crores
Total No. of units outstanding	Rs. 100 crores

$$\therefore \text{NAV} = \frac{\text{Rs. 1500 crores} + \text{Rs. 100 crores} - \text{Rs. 200 crores}}{100 \text{ crores}}$$

$$= \text{Rs. } \frac{1400 \text{ crores}}{100 \text{ crores}} = \text{Rs. } 14 \text{ per unit.}$$

The NAV of a mutual Fund scheme is calculated and disclosed to the public for every working day. The NAV changes daily. Investors try to invest when NAV is low and sell the units and get profits when the NAV is high.

Most mutual fund schemes are not traded at stock market. Thus, investor purchases as well as sells the units to AMC i.e. Asset management company, This sale is called redemption of units.

Basically funds are of two types :-

1) close ended funds 2) open ended funds .

1) Close ended mutual funds :- These are offered with a fixed date of maturity and can be purchased from mutual fund companies during a specific period . The investor can get the amount after expiry date of the fund . If an investor wants to exit before the maturity date , he can sell the units on the stock exchange at a discount or through a buy-back option by the fund .

2) Open ended funds : These have no fixed date of maturity and the units can be sold or repurchased at any time .The no. of units & its capital changes daily .

Entry load & Exit load : Some mutual fund schemes collect a charge when investors purchase or redeem units . These are usually percentage of NAV . The charge levied while purchasing a unit is called the entry load & the charge collected on redemption is called exit load .

Usually , the debt funds have not loads . When there are no charges while purchasing or selling of units , these funds are called No Load Funds .

Mutual Funds can be broadly categorised into two types : 'Dividend ' funds which offer a dividend and 'Growth ' funds which do not offer a dividend .

In mutual funds , the dividend given has no direct relation to the profit earned . The mutual fund invests the money in different shares that may or may not give a dividend at different times & different rates . The fund manager may at any arbitrary point , decide to give a part of the units' value back to the investors . This is called dividend .

For a growth fund , the NAV does not come down due to dividends . It moves up or down purely on the basis of the gains or losses of the securities that the fund has invested in.

For a growth fund , the gains per unit are purely from the difference between the redemption price and the purchase price i.e. the total gain is purely the capital gain . For a dividend fund , the total gain is the addition of the capital gain & the dividend .

Capital gain = Amount received after redemption - Amount invested .

$$\text{Rate of Return} = \frac{\text{Change in NAV} + \text{Dividend}}{\text{NAV at the beginning of the period}} \times 100$$

(This is for a given period) .

$$\text{Annualised rate of Return} = \text{Rate of Return} \times \frac{365}{n}$$

where n is the number of days .

Some important Terms :

i) Assets :- It refers to market value of investment of M.F. in government securities , bonds etc. , its receivables , accrued income & other assets .

ii) liabilities :- It includes all expenses like accrued expenses , payables and other liabilities for the M.F. scheme .

iii) Net Assets :- Total Assets - liabilities

iv) The valuation Date is the date on which NAV is calculated .

Example 11

Mr. Deore invested Rs. 25,000/- to purchase 2,500 units of ICICI MF - B plan on 4th April 2007 . He decided to sell the units on 14th Nov. 2007 at NAV of Rs. 16.4 /-. The exit load was 2.5 % . Find his profit (Calculations are upto 2 decimal points)

Solution :

No. of units = 2500 , purchase cost = Rs. 25,000/-

NAV on the date of sale = RS. 16.4/- , exit load = 2.5% = of 16.4 = 0.41

$$\therefore \text{selling price of 1 unit} = 16.4 - 0.41 = 15.99$$

$$\therefore \text{sale value} = 2500 \times 15.99 \\ = \text{Rs. } 39,975/-$$

$$\therefore \text{Profit} = 39,975 - 25,000 \\ = \text{Rs. } 14,975 .$$

Example 12

Ragini invested Rs. 94,070/- in mutual Fund when NAV was Rs. 460 /- with entry load of 2.25 % . She received a dividend of Rs. 5/- per unit . She, later sold all units of fund with an exit load of 0.5 % . If her gain was Rs. 1654/-, find NAV at which she sold the units . (Calculations are upto 2 decimal points)

Solution : purchase price of one unit = $460 + 2.25\% \text{ of } 460$
 $= 460 + 10.35 = 470.35$

$$\text{No. of units purchased} = \frac{94,070}{470.35} = 200$$

$$\text{Total dividend} = 200 \times 5 = 1000$$

$$\text{Gain} = \text{Profit} + \text{Dividend}$$

$$\therefore 1654 = \text{Profit} + 1000$$

$$\therefore \text{Profit} = 1654 - 1000 = 654$$

While selling let NAV of one unit be y

$$\begin{aligned}\therefore \text{sale price of one unit} &= \text{NAV} - \text{exit load} \\ &= y - 0.5\% \text{ of } y \\ &= 0.995 y\end{aligned}$$

$$\therefore \text{sale price of 200 units} = 200 \times 0.995 y = 199 y$$

Also, profit = Total sale - Total purchase

$$654 = 199y - 94,070$$

$$\therefore 199y = 654 + 94,070$$

$$\therefore 199y = 94724$$

$$\therefore y = 476$$

$$\therefore \text{NAV at which she sold units} = \text{Rs. } 476/-$$

Example 13

If a mutual fund had NAV of Rs. 28/- at the beginning of the year and Rs. 38/- at the end of the year, find the absolute change and the percentage change in NAV during the year.

Solution : NAV at the beginning = Rs. 28/-

NAV at the end = Rs. 38/-

$$\therefore \text{Absolute change in NAV} = 38 - 28 = \text{Rs. } 10/-$$

$$\% \text{ change} = \frac{\text{Absolute change}}{\text{NAV at the beginning}} \times 100 = \frac{10}{28} \times 100 = 35.71 \%$$

Example 14

If NAV was Rs. 72/- at the end of the year, with 12.5 % increase during the year, find NAV at the beginning of the year.

Solution : Let 'x' be the NAV at the beginning of the year.

$$\therefore \text{Absolute change in NAV} = 12.5 \% \text{ of } x = \frac{12.5}{100} \times x = 0.125 x$$

$$\therefore \text{NAV at the end of the year} = x + 0.125 x = 1.125 x$$

$$\therefore 1.125 x = 72$$

$$\begin{aligned}\therefore x &= \frac{72}{1.125} \\ &= 64\end{aligned}$$

$$\therefore \text{NAV's initial value was Rs. } 64/-$$

Example 15

Rohit purchased some units in open end equity fund at Rs. 16/- . The fund distributed interim dividend of Rs. 5/- per unit, and the NAV of the fund at the end of the year was Rs. 25/- . Find the total percentage return . (Calculations are upto 2 decimal points)

Solution : Total gain = change in NAV + Dividend

$$= (25 - 16) + 5$$

$$= 9 + 5$$

$$= 14$$

$$\therefore \text{Total percentage gain} = \frac{\text{Total gain}}{\text{NAV at the beginning}} \times 100$$

$$= \frac{14}{16} \times 100 = 87.5 \%$$

Example 16

Mr. Hosur purchased some units in open- end fund at Rs. 30/- and its NAV after 18 months was Rs. 45/- . Find the annualised change in NAV as a percentage .

Solution : change in NAV for 18 months = 45-30 = Rs. 15/-

$$\begin{aligned} \therefore \text{annualised change} &= \frac{\text{change in NAV}}{\text{NAV at beginning}} \times \frac{12}{\text{No. of months}} \times 100 \\ &= \frac{15}{30} \times \frac{12}{18} \times 100 \\ &= 33.33 \% \end{aligned}$$

Check your progress :

1) Mr. Kamble purchased 586.909 units of Kotak cash plus retail Growth on 1st June 2007 when the NAV was RS. 20.4461. Its NAV as on 3rd December, 2007 was Rs. 21.1960/- . The fund has neither entry load nor an exit load. Find the amount invested on 1st June 2007 and the value of Mr. Kamble's investment on 3rd December 2007 .

Ans . 12,000 , 12440.12 .

2) Ms . Kannan purchased 113.151 units of 'FT India Prima Plus' on 9th April 2007 and redeemed all the units on 7th Aug 2007 when the NAV was Rs. 35.5573 . The entry load was 2.25 % and the exit load was 1 % . If she gained Rs. 483.11 , find the NAV on 9th April 2007 . (Calculations are upto 2 decimal points)

Ans . 30.2514

3) Mr. Pandit invested Rs. 10,000/- in Birla Sunlife Equity Fund-Dividend plan ' on 10/07/2007 , when the NAV was Rs. 78.04 ,and redeemed all the units on 12/11/2007 when the NAV was Rs. 84.54 . In the meanwhile , on 31/08/2007 , she had received a dividend @ Rs. 10 per unit . Find her total gain and the rate of return considering loads as follows:

The entry load was 2.25 % and the exit load was 0.5 % The number of units were calculated correct upto 3 decimal places.

Ans . Total gain = Rs . 1794.46, Rate of return = 17.94%

4) Given the following information , calculate NAV of the mutual fund :-

No. of units =15000

Market value of investments in Govt . securities = Rs. 20 lakhs

Market value of investments in corporate Bonds = Rs. 25 lakhs

Other Assets of the fund = Rs. 15 lakhs

Liabilities of the fund = 6 lakhs

Ans . Rs. 360/- .

5) Mumtaz purchased 1200 units of TATA BIG Bond- G Rs. 12,000 /- on 14th April 2007 . She sold her units on 9th Dec 2007 at NAV of

Rs. 15.36/- . The short term gain tax (STGT) was 10% of the profit .
Find her net profit . (Calculations upto 2 decimal points)

Ans . profit = 6432 , STGT = 643.2 , Net profit = 5788.8
(profit- STGT) .

1.3 SYSTEMATIC INVESTMENT PLAN (SIP)

In SIP an investor invests a fixed amount (e.g. say 1000/-) every month on a fixed date (e.g. 1st of every month) . In general the minimum amount is Rs. 1000/- per month , in diversified equity schemes . It can be even Rs. 500/- as well in ELSS schemes . If this is done for many months , then each time units are purchased at a different NAV . Over a period of few months, an investor gets the benefit of a phenomenon called 'Rupee cost Averaging' .

Rupee-cost- averaging :- If NAV increases , the no. of units decreases & if NAV decreases , the no. of units purchased increases . Thus on the whole , it lowers the average cost of units because indirectly ,the investor buys more units when NAV prices are low & he buys less units , when NAV prices are high . It is called Rupee-cost-Averaging .

Consider the following example :-

Mr. Shaikh keeps Rs. 5000/- on 3rd of every month for 4 months as follows :-(Calculations are correct to 2 points of decimal)

Month	Amount (in Rs.)	NAV	No. of units he gets
1	5000	109.48	$5000/109.48=45.67$
2	5000	112.36	$5000/112.36=44.50$
3	5000	108.14	$5000/108.14=46.24$
4	5000	105.62	$5000/105.62=47.34$
Total	20,000		183.75

$\therefore$ Avg price of units = $20,000 / 183.75 = 108.84$

If Mr. Shaikh would have invested the entire amount of Rs. 20,000/- On 3rd of first month only , with NAV Rs. 109.48/- , the no. of units purchased would have been $20,000/ 109.48 = 182.68$

Thus he gained more units and average price of units also was Rs.108.84 instead of Rs.109.48 which was NAV on 3rd of the first month

If SIP is followed for a long period of time , it can create wealth to meet a person's future needs like housing , higher education etc .

Now , we will study the following examples to understand SIP .

Example 17

Mr. Patil invested in a SIP of a M.F. , a fixed sum of Rs. 10,000/- on 5th of every month , for 4 months . The NAV on these dates were Rs. 34.26 , 46.12 , 39.34 and 41.85 . The entry load was 2.25 % through out the period . Find the average price , including the entry load , using the Rupee-cost-Averaging method .How does it compare with the Arithmetic mean of the prices ? (Calculations are correct to 4 digits decimal)

Solution :

Month	NAV	Entry load = 2.25%	Total price	No.ofunits=1000/ Total price
1	34.2600	0.7708	35.0308	285.4627
2	46.1200	1.0377	47.1577	212.0544
3	39.3400	0.8851	40.2252	248.6006
4	41.8500	0.9141	42.7916	233.6906
TOTAL			165.2053	979.8083

By using Rupee-cost-Averaging method :-

$$\text{Avg Price} = \frac{\text{Total amount}}{\text{Total No. of units}}$$

$$= \frac{40,000}{979.8083} = 40.8243$$

$$\text{A.M. of price} = \frac{\text{Total price}}{4} = \frac{165.2053}{4} = 41.3013$$

∴ Avg. price using Rupee-cost- Averaging method is less than A.M. of prices .

Example 18

Mr. Desai invested Rs. 5000/- on 1st of every month for 5 months in a SIP of a M.F. with NAV's as 48.15 , 52.83 ,41.28,35.44 & 32.65 respectively .There was no entry load charged . Find the average price , Mr. Desai paid using the Rupee-cost-Averaging method . After 6 months ,he sold all his units , when NAV was Rs. 51.64 with contingent deferred sales charge (CDSC) as 2.25 % . Find his net gain. (Calculations are correct to 2 digits decimal)

Solution : consider the following table :-

Month	Amount (in Rs.)	NAV	No. of units
1	5000	48.15	5000/48.15=103.84
2	5000	52.83	5000/52.83=94.64
3	5000	41.28	5000/41.28=121.12
4	5000	35.44	5000/35.44=141.08
5	5000	32.65	5000/32.65=153.14
TOTAL	25000		613.82

$$\text{Avg . price of units} = \frac{25000}{613.820} = 40.73$$

For selling :

selling price of one unit = $51.64 - 2.25\% \text{ of } 51.64 = 50.48$

Total sales = $50.48 \times 613.82 = 30,991.77$

$\therefore$ Net gain = $30,991.77 - 25,000 = \text{Rs. } 5991.77$.

Check your progress :

1) Mr. Thomas started a SIP in 'HDFC long term advantage Fund ' . On the 10th July , Aug and Sept 2007 he invested Rs. 1000/- each at the NAVs Rs. 44.100, Rs. 43.761/-, s. 45.455 respectively . The entry load was 2.25% . Find his average acquisition cost per unit upto 3 decimal places . (Calculations are up to 3 decimal points).

Ans. Rs. 45.427/- .

2) Maneeshad Rs. 20,000/- on 2nd of every month for 5 onths in a SIP of a mutual fund , with NAVs as Rs. 53.12 , Rs. 56.26 , Rs. 48.86 ,Rs.50.44 and Rs. 54.62 respectively . The entry load was 2.25 % throughout this period .Find average price , including the entry load , using the Rupee-cost -Averaging method and compare it with Arithmetic mean of prices .
(Calculate up to 2 decimal points)

Ans . 53.70 , 53.84 .



Unit - 2

PERMUTATIONS AND COMBINATIONS

Unit Structures :

- 2.1 Introduction
- 2.2 Prerequisites and Terminology
- 2.3 Permutation
- 2.4 Formula to compute permutations of n different objects
- 2.5 Permutations of n objects not all different
- 2.6 Combination

2.1 INTRODUCTION

We often hear people saying that ‘probably it will rain today’; ‘it is likely that India will win the match against Australia on this pitch’; ‘the chance of passing the CET is 10% only’ and so on. The words ‘probably’, ‘by chance’ or ‘likely’ are statistical terms but are very commonly used by all of us. It is very natural that people are interested to know about the possibility that something happens. People interested in sports are eager to know, the possibility of their team to win a game; political activists want to be sure their chances of winning an election, meteorological department would like to know about the weather, an economists may want to know the chance that sales will increase if the price of a commodity is decreased etc. Also it becomes necessary sometimes to know in how many different ways a particular event may happen.

All these calculations can be broadly classified into two types namely *permutations* and *combinations*.

2.2 PREREQUISITES AND TERMINOLOGY

Before going into the detailed study of these two methods, we shall discuss some of the prerequisites which are useful in understanding the concepts and formulae related to them.

Factorial:

The factorial of a natural number n is defined as the product of all numbers from 1 to n . It is denoted as $n!$

For example:

Factorial of 3 i.e. $3! = 3 \times 2 \times 1 = 6$ (This is read as 3 factorial equal to 6)

Similarly, $5! = 5 \times 4 \times 3 \times 2 \times 1 = 120$

In general, $n! = n(n-1)(n-2)(n-3)\dots 3 \times 2 \times 1$

This formula gives a recursive relation: $n! = n(n-1)!$

For example:

$$6! = 6 \times 5 \times 4 \times 3 \times 2 \times 1 = 6 \times 5!$$

Remark: We define $0! = 1$

Fundamental Principal of Counting:

If there are m ways doing one thing and n ways of doing another thing then the total number of ways of doing both the things together is mn .

2.3 PERMUTATION

A *permutation* of n objects is an arrangement of some of these (or all) objects in a definite order.

The order in which the arrangement is done is important in permutations.

Example 1: In how many different ways can three friends Mitesh, Ritesh and Paresh stand for a group photograph?

Ans: Let us denote these friends by their first alphabet M, R and P. The three friends can be arranged as shown below:

M R P M P R R M P R P M P M R P R M

The number of ways is 6.

One should not be satisfied with this answer. The question that should come to our mind is how did we arrange them so? Well, if we observe the arrangement again, it can be seen that first M's place was fixed and the remaining two were arranged. This step was repeated again for R and P. In terms of permutation what we did was arranging 3 objects amongst themselves. The next question should be what if there are 10 friends? Can you write down their different arrangements explicitly as above?

2.4 FORMULA TO COMPUTE PERMUTATIONS OF N DIFFERENT OBJECTS

${}^n P_r$: The number of ways of arranging r objects out of n objects is denoted by ${}^n P_r$ and is calculated by the formula: ${}^n P_r = \frac{n!}{(n-r)!}$

For e.g: ${}^4 P_2 = \frac{4!}{(4-2)!} = \frac{4 \times 3 \times 2 \times 1}{2 \times 1} = 12$

Remark: The number of ways of arranging all n objects is thus

$${}^nP_n = \frac{n!}{(n-n)!} = \frac{n!}{0!} = n!$$

Now if we go back to the first example, where we had to arrange 3 friends, then the number of ways of their arrangement using the above formula is $3! = 6$, which is the same answer what we had got by arranging them explicitly.

Example 2: In how many ways can 6 people be photographed, if only 4 can be seated at a time?

Ans: Here $n = 6$ and $r = 4$.

$$\therefore \text{no. of ways of arranging 4 out of 6 people} = {}^6P_4 = \frac{6!}{(6-4)!} =$$

$$\frac{6 \times 5 \times 4 \times 3 \times 2 \times 1}{2 \times 1} = 720$$

Example 3: If ${}^nP_5 = 42 {}^nP_3$ then find the value of n .

$$\text{Ans: } {}^nP_5 = \frac{n!}{(n-5)!} \quad \text{and} \quad {}^nP_3 = \frac{n!}{(n-3)!}$$

$$\text{Since } {}^nP_5 = 42 {}^nP_3$$

$$\therefore \frac{n!}{(n-5)!} = 42 \frac{n!}{(n-3)!}$$

$$\therefore \frac{(n-3)!}{(n-5)!} = 42$$

$$\text{i.e. } \frac{(n-3)(n-4)(n-5)!}{(n-5)!} = 42$$

$$\therefore (n-3)(n-4) = 42$$

$$\therefore n^2 - 7n + 12 = 42$$

$$\therefore n^2 - 7n - 30 = 0$$

$$\therefore (n-10)(n+3) = 0$$

$$\Rightarrow n = 10, -3$$

$$\text{As } n \neq -3 \quad \therefore n = 10$$

Example 4: Show that ${}^nP_r = n \cdot {}^{n-1}P_{r-1}$

$$\begin{aligned} \text{Ans: Consider R.H.S.} &= n \cdot {}^{n-1}P_{r-1} = n \cdot \frac{(n-1)!}{[(n-1)-(r-1)]!} \\ &= \frac{n(n-1)!}{(n-1-r+1)!} = \frac{n!}{(n-r)!} = {}^nP_r = \text{L.H.S.} \end{aligned}$$

Example 5: In how many different ways can a 4 digit number be formed from the numbers 1, 2, 3, ..., 9, with no digit being repeated?

Ans: Since no repetition is allowed, the number of ways of forming a four digit number from the given 9 digits is

$${}^9P_4 = \frac{9!}{(9-4)!} = \frac{9 \times 8 \times 7 \times 6 \times 5!}{5!} = 3024.$$

Example 6: In how many different ways can the letters of the word “MATHS” be arranged if no letter is to be repeated in the same word?

Ans: The word “MATHS” consists of four letter M, A, T, H and S.

The required number of ways is arranging 4 objects all at a time.

Thus, the no. of different ways = $4! = 24$.

Example 7: Eight candidates are to appear for an interview in a company. In how many ways can the HR manager schedule the candidates for their interview? What if two of the eight candidates are not be interviewed?

Ans: The number of ways of arranging 8 candidates is $8! = 40320$
If two out the eight are not be arranged, it means to arrange the remaining six out of 8.

This can be done in 8P_6 ways.

$$\begin{aligned} \therefore \text{the number of ways of scheduling the candidates now} &= \frac{8!}{(8-6)!} \\ &= \frac{8!}{2!} = \frac{40320}{2} = 20160 \end{aligned}$$

Example 8: Find the number of ways in which 3 books on Economics, 4 books on Mathematics and 2 books on Law can be arranged on a book shelf so that books of the same subject are together. Find the number of arrangements if books are to be arranged at random.

Ans: We can consider this problem of arranging three blocks **E** (books on Economics), **M** (books on Mathematics) and **L** (books on Law), which can be done in $3!$ ways.

Now,

3 books in Economics can be arranged amongst themselves in $3!$ ways,
4 books in Mathematics can be arranged amongst themselves in $4!$ ways,
and 2 books in Law can be arranged amongst themselves in $2!$ Ways

Thus, by the fundamental principle of counting,

the total number of arrangements = $3! \times 3! \times 4! \times 2! = 1728$ ways.

If there is no condition on the position of any book of any subject we may consider this of arranging $3 + 4 + 2 = 9$ books, which can be done in $9!$ Ways

i.e. 362880 ways.

Example 9: In how many ways can 3 boys and 4 girls be seated for a group photograph if (i) no two boys sit together, (ii) no two girls sit together, (iii) all boys sit together

Ans: (i) *no two boys sit together*

4 girls can be arranged in $4! = 24$ number of ways

In each such arrangement there are five places (marked as X) where boys can be seated so that no two boys sit together as shown below:

X **G** X **G** X **G** X **G** X

Now, 3 boys can be seated in these 5 places in 5P_3 ways = **60 ways**

Thus, the total number of arrangements = $24 \times 60 = 1440$.

(ii) *no two girls sit together:*

3 boys can be arranged in $3!$ number of ways

In each such arrangement there are four places (marked as X) where girls can be seated so that no two girls sit together as shown below:

X **B** X **B** X **B** X

Now, 4 girls can be seated in these 4 places in 4P_4 ways = **4! ways**

Thus, the total number of arrangements = $3! \times 4! = 144$.

(iii) *all boys sit together*

All boys can be considered as one block and the remaining 4 girls as four blocks. The required arrangement is of these $1 + 4 = 5$ blocks, which can be done in **5! ways**.

The 3 boys can be arranged amongst themselves in **3! ways**.

Thus the total number of ways in which 3 boys and 4 girls can be seated for a group photograph such that all boys are seated together is = $5! \times 3! = 720$.

Example 10: In how many different ways can the letters of the word “JOGESHWARI” be arranged such that (i) there is no restriction, (ii) the word starts with ‘A’, (iii) the word ends with ‘W’, (iv) the word begins with ‘A’ and ends with ‘W’, (v) the vowels are together, (vi) the letters W, A, R are never together.

Ans: The given word “JOGESHWARI” consists of 10 distinct letters, of which 4 are vowels (A, E, I, O) and 6 are consonants (G, H, J, R, S, W)

(i) If no restriction is there then the total number of arrangements = $10! = 362880$

(ii) *The word starts with ‘A’*

Since the first place is fixed for the 10 letter word, it remains to arrange the remaining 9 letters.

Thus, the total number of ways = $9! = 362880$

(iii) *The word ends with ‘W’*

This is similar to the above problem

(iv) *the word begins with 'A' and ends with 'W'*

Since two letters A and W are fixed, it remains to arrange the remaining 8 letters which can be done in $8! = 40320$ ways.

(v) *the vowels are together*

4 vowels can be considered as one block and the remaining 6 consonants as 6 blocks.

Thus, the problem now reduces to arranging these 7 blocks, which can be done in $7! = 5040$ ways.

The 4 vowels can be arranged amongst themselves in $4! = 24$ ways.

Thus by fundamental principle of counting,

the total number of ways = $5040 \times 24 = 120960$.

(vi) *the letters W, A, R are never together*

The required number of ways = total number of all arrangements – number of ways of arrangements where W, A, R are together.

Now, the total number of ways of arranging 10 letters = $10!$

No of ways of arranging the letters W, A, R together = $8! \times 3!$ (from the above case (v))

Thus the required number of ways = $10! - (8! \times 3!) = 3628800 - 241920$
 $= 3386880$.

2.5 Permutations of n objects not all different

Consider a permutation of 2 white and 4 black hats of the same type. Then the permutation $W_1W_2B_1B_2B_3B_4$ is same as $W_2W_1B_1B_2B_3B_4$ as we cannot differentiate between the two white hats. Thus, in permutations where certain numbers of objects are similar, we have to remove the duplications of same permutations.

Let n_1 objects be of one kind, n_2 objects be of second kind, ..., n_k objects be of k^{th} kind then the number of distinct permutations of all objects taken together i.e. $n = n_1 + n_2 + \dots + n_k$ is given by: $\frac{n!}{n_1!n_2!n_3!\dots n_k!}$

Example 11: A College Librarian Mrs. Akalpita orders 15 books of which 5 books are of Business Law, 5 books of Business Mathematics, 3 books on Principles of Management and 2 books on F.H.S. In how many different ways can she arrange them on the shelf?

Ans: Given $n = 15$, $n_1 = 5$, $n_2 = 5$, $n_3 = 3$ and $n_4 = 2$

The number of ways of arranging the 20 books on the shelf

$$= \frac{15!}{5!5!3!2!} = 7567560$$

Example 12: In how many different ways can the letters of the word “MALAYALAM” be arranged horizontally?

Ans: In the given word “MALAYALAM” has $n = 9$ letters of which M, A and L are repeating.

Here ‘M’ is repeated 2 times, $\therefore n_1 = 2$

‘A’ is repeated 4 times, $n_2 = 4$

‘L’ is repeated 2 times, $n_3 = 2$

$\therefore$ the number of permutations = $\frac{9!}{2!4!2!} = 3780$.

Example 13: Find the number of permutations of the letters of the word “TOMMORROW” such that (i) no two M’s are together, (ii) all the O’s are not together.

Ans: The given word “TOMMORROW” has 9 letters of which ‘O’, ‘M’ and ‘R’ are repeated.

Now, ‘O’ is repeated 3 times, $n_1 = 3$

‘M’ is repeated 2 times, $n_2 = 2$

‘R’ is repeated 2 times, $n_3 = 2$

$\therefore$ number of permutations of letters of the word “TOMMORROW”

$$= \frac{9!}{3!2!2!} = 15120 \quad \dots (1)$$

(i) *no two M’s are together*

We assume the two M’s as one block and the remaining 7 letters as remaining 7 blocks.

The number of ways of arranging these $7+1 = 8$ blocks is $8!$, in which the letters R and O are repeated 2 and 3 times respectively.

Hence, the number of permutations in which both the M’s are together is

$$= \frac{8!}{3!2!} = 3360. \quad \dots (2)$$

From (1) and (2), the number of distinct permutations in which no two M’s are together = $15120 - 3360 = 11760$.

(ii) *all the O’s are not together*

We consider the three O’s as one block and the remaining 6 letters as one block.

The number of ways of arranging the $6 + 1 = 7$ blocks is $7!$, in which the letters M and R are repeated 2 times each.

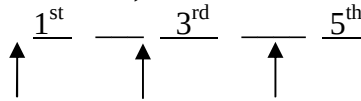
Hence, the number of permutations in which all the O’s are together is

$$= \frac{7!}{2!2!} = 1260 \quad \dots (3)$$

From (1) and (3), the number of distinct permutations in which all three O's are not together = $15120 - 1260 = 13860$

Example 14: Find the number of permutations of the letters of the word “VOWEL” such that the vowels occupy the odd places in the arrangement.

Ans: The given word “VOWEL” has 2 vowels ‘E’, ‘O’ and 3 consonants. There are two odd places viz. 1st, 3rd and 4th.



Two vowels can be placed in three places in 3P_2 ways. For each such way the 3 consonants can be placed in the remaining three places in $3!$ ways. By Fundamental principal of counting, the total number of arrangements is given by ${}^3P_2 \times 3! = 6 \times 6 = 36$ ways.

Exercise

- Find the number of ways in which 4 boys can be seated for a group photograph.
- In how many ways can 3 boys and 2 girls be seated for a photograph.
- In how many ways can 7 books be arranged on a book shelf?
- Find the number of ways of making 3 people sit on 3 chairs?
- Find the number of ways of making 5 people sit on 3 chairs?
- In how many ways can 3 boys and 2 girls be seated for a photograph in two lines with the first line of boys and second line of girls?
- 20 students of a class are seated in three lines for a group photograph. The first line has 8 chairs; second line has 3 chairs and third 5 chairs. In how many ways can this be done?
- In how many ways can 6 people be selected for 3 posts in a company?
- In how many ways can 10 students be selected for Student's Council which consists of only 4 members?
- Find the number of ways in which a 5 digit number be formed from the numbers 1, 2, 3,, 9 if (i) no digit is repeated, (ii) repetition is allowed.
- Find the number of ways of forming a 4 digit even number from the digits 1, 2, 3,, 9 if no digit is to be repeated?
- In how many ways can 3-digit odd number be formed from the digits 1, 2, 3,, 9 if (i) no digit is repeated, (ii) repetition is allowed.
- Find the number of ways of arranging the letters of the word (i) MASK, (ii) MOTHER, (iii) RATION, (iv) YES, (v) BHARTI, (vi) GREATFUL.
- Find the value of n from the following:
 - ${}^nP_5 = 6 {}^nP_3$
 - ${}^nP_6 = 56 {}^nP_4$
 - ${}^nP_7 = 12 {}^nP_5$
- Show that ${}^{n+1}P_{r+1} = (n+1) {}^nP_r$.

16. Find the number of ways in which 2 books on Mathematics, 3 books on Law and 2 books on Economics can be arranged on a shelf so that books of the same subject are together. Also find the number of arrangements if books are to be arranged at random.
17. Find the number of ways in which 4 books on Physics, 3 books on Chemistry and 2 books on Biology can be arranged on a shelf so that (i) books of the same subject are together; (ii) No two books on Biology are together; (iii) No two books on Chemistry are together and (iv) No two books on Physics are together.
18. In how many ways can 4 boys and two girls be seated if (i) no two girls sit together, (ii) no two boys sit together and (iii) both the girls sit together?
19. In how many ways can 3 Africans and 3 Americans be seated so that (i) atleast two Americans always sit together and (ii) exactly two Africans sit together?
20. In how many distinct ways can the letters of the word "CHEMISTRY" be arranged such that (i) there is no restriction; (ii) the word begins with a vowel; (iii) the word begins and ends with a vowel; (iv) the vowels are together and (v) the letters T, R and Y are never together.
21. In how many distinct ways can the letters of the word "ANDHERI" be arranged such that (i) the word begins with A; (ii) the word begins with A and ends with R; (iii) the word begins with vowel and ends with vowel and (iv) the letters N, R and I are never together?
22. A College Librarian Mrs. Parita orders 12 books of which 4 are of Maths, 3 are of English Literature, 3 books on Sociology and 4 books on Philosophy. In how many different ways can the books be arranged?
23. In how many distinct ways can the letters of the word "DISMISS" ?
24. In how many distinct ways can the letters of the word "STATISTICS" such that (i) no two I's are together; (ii) no two T's are together ?
25. Find the number of permutations of the letters of the word "VIRAR" such that the vowels occupy the even places in the arrangement.

2.6 COMBINATION

A combination of n objects is an arrangement of some of these (or all) objects where the order of arrangement is not considered.

For e.g.: A combination considers the arrangements 'ab' and 'ba', of two letters 'a' and 'b' as the same.

nC_r : The selection (or rejection) of r objects out of n objects is denoted by nC_r and is calculated by the formula: ${}^nC_r = \frac{n!}{r!(n-r)!}$

For example: ${}^5C_2 = \frac{5!}{2!(5-2)!} = \frac{5!}{2!3!} = \frac{5 \times 4 \times 3 \times 2 \times 1}{2 \times 3 \times 2 \times 1} = 10$

We know that the factorial formula can be written recursively. Hence we can write the numerator $n!$ as $n(n-1)(n-2)\dots(n-r+1)(n-r)!$. Due to this recursive formula, the calculation of nC_r becomes easier as show below:

$$\text{Now, } {}^nC_r = \frac{n(n-1)(n-2)\dots(n-r+1)(n-r)!}{r!(n-r)!} = \frac{n(n-1)(n-2)\dots(n-r+1)}{r!}$$

For example: ${}^{10}C_4 = \frac{10 \times 9 \times 8 \times 7}{4 \times 3 \times 2 \times 1} = 210$. This method simplifies and speeds up the calculations

Results (without proof):

1. ${}^nC_0 = {}^nC_n = 1$
2. ${}^nC_r = {}^nC_{n-r}$
3. ${}^nC_r + {}^nC_{r-1} = {}^{n+1}C_r$

Exercise

Evaluate the following:

- | | | | | |
|---------------|------------------|---------------|---------------|----------------|
| (1) 8C_1 | (2) ${}^{15}C_0$ | (3) 6C_6 | (4) 8C_3 | (5) 8C_5 |
| (6) 8C_2 | (7) 5C_3 | (8) 6C_3 | (9) 6C_4 | (10) 6C_2 |

Example 15: In how many ways can 2 students be selected for a student's committee out of 7 students?

Ans: Here $n=7$ and $r=2$

$$\therefore \text{the no. of ways of selecting 2 out of 7 students} = {}^7C_2 = \frac{7 \times 6}{2 \times 1} = 21.$$

Example 16: A question paper has 8 questions and only 5 questions are to be attempted. In how many ways a student can select any 5 questions?

Ans: Here $n=8$, $r=5$

$$\therefore \text{the required number of ways} = {}^8C_5 = {}^8C_3 = \frac{8 \times 7 \times 6}{3 \times 2 \times 1} = 56$$

Example 17: In how many ways can 2 boys and 2 girls be selected from a group of 6 boys and 5 girls?

Ans: Here there are two calculations to be made which are interdependent.

(i) 2 boys out of 6 boys can be selected in 6C_2 ways and

(ii) 2 girls out of 5 girls can be selected in 5C_2 ways.

$$\begin{aligned} \text{The total number of ways of such a selection} &= {}^6C_2 \times {}^5C_2 = \frac{6 \times 5}{2} \times \frac{5 \times 4}{2} \\ &= 150. \end{aligned}$$

Example 18: A book shelf has 10 books of which 6 are of Accounts and remaining of Management. In how many ways can a person select 3 books on Accounts and 1 book on Management?

Ans: There are 6 books of Accounts and $10 - 6 = 4$ books on Management.

3 books out of 6 books on Accounts can be selected in ${}^6C_3 = 20$ ways.

1 book out of 4 books on Management can be selected in ${}^4C_1 = 4$ ways.

By fundamental principal of counting,
the total number of ways is $20 \times 4 = 80$.

Example 19: A Committee of 6 people is to be formed from a Staff of 4 Managers, 6 Officers and 2 peons. Find the number of distinct committees in which there are: (i) 2 persons from each category; (ii) no peons; (iii) exactly 2 Managers and (iv) atleast 2 Managers.

Ans:

(i) 2 persons from each category can be selected in

$${}^4C_2 \times {}^6C_2 \times {}^2C_2 = 90 \text{ ways.}$$

(ii) No peons to be selected means the selection of 6 people is to be done from remaining $4 + 6 = 10$ people.

This can be done in ${}^{10}C_6 = {}^{10}C_4 = \frac{10 \times 9 \times 8 \times 7}{4 \times 3 \times 2 \times 1} = 210$ ways.

(iii) Exactly 2 Managers out of 4 can be selected in ${}^4C_2 = 6$ ways.

The remaining 4 persons in the committee are to be selected from the remaining $6 + 2 = 8$ persons, which can be done in ${}^8C_4 = \frac{8 \times 7 \times 6 \times 5}{4 \times 3 \times 2 \times 1} = 70$

ways.

The total number of ways = $6 \times 70 = 420$

(iv) Atleast two managers can be selected in the following ways:

	Managers	Others
Selection I	2 out of 4	4 out of 8
Selection II	3 out of 4	3 out of 8
Selection III	4 out of 4	2 out of 8

Selection I can be done in ${}^4C_2 \times {}^8C_4 = 420$ ways

Selection II can be done in ${}^4C_3 \times {}^8C_3 = 192$ ways

Selection II can be done in ${}^4C_4 \times {}^8C_2 = 28$ ways

Thus, the total number of ways = $420 + 192 + 28 = 640$.

Exercise

1. In how many ways can 6 books be selected out of 10 books?
2. In how many ways can 4 boys be selected out of 7 boys?
3. A question paper contains two sections. Section I consists of 4 questions of which two are to be attempted and Section II consists of 5 questions of which 3 are to be attempted. Find the number of ways of attempting the questions in the paper.
4. There are 6 bulbs of which 3 are defective are to be put in two sockets in a room. Find the number of ways in which the room is lighted.
5. There are 8 books on History and 4 books on Geography. In how many ways can 4 books on History and 3 books on Geography be selected?

6. A box contains 4 white and 5 black balls. How many selections of 2 balls can be made so that (i) one ball of each color is selected; (ii) both balls are white; (iii) no white ball is selected.
7. A box contains 6 blue, 4 green and 2 white balls. How many selections of three balls can be made so that (i) one ball of each color is selected; (ii) atleast one white ball is selected and (iii) no white ball is selected.
8. 10 candidates appear for an interview. The selection committee has time to interview only 6 candidates. Find the number of ways of selecting the candidates.
9. 10 candidates appear for an interview. 2 candidates are disqualified as they did not bring necessary documents and the selection panel has time to interview only 5 people. In how many ways can this be done?
10. A committee consisting of 4 men and 3 women is to be formed from 7 men and 6 women. Find the number of selections in which (i) a particular man is selected; (ii) a particular women is not selected; (iii) a particular man is selected and a particular is rejected.
11. A College has 8 professors and 5 lecturers. In how many ways can a committee of 5 teachers be formed such that it consists of (i) two professors, (ii) two lecturers, (iii) atleast two professors, (iv) atleast two lecturers, (v) atleast two professors and lecturers and (vi) no lecturer.
12. A Cultural Committee of 8 persons from 6 men and 5 women is to be constituted in a Society. Find the number of distinct committees if it should consist of (i) atleast 3 women, (ii) exactly 2 women and (iii) atleast 4 men.
13. A question paper has three sections with each consisting of 4 questions. A student has to attempt a total of 5 questions with atleast one from each section. Find the number of ways in which this can be done.
14. A case is under discussion in front of 5 judges. Find the number of ways in which the judgment is given with a majority.
15. The Mumbai Royal Cricket Club has 22 players of which 3 are wicket keepers, 6 are fast bowlers, 2 are spinners and 3 all rounder's. In how many ways can a team be formed if it should include (i) one wicket keeper, 2 fast bowlers and one all rounder; (ii) 2 fast bowlers, 1 spinner and 1 wicket keeper; (iii) atleast one spinner, (iv) atleast one spinner and one all rounder.



UNIT II**Unit-3****LINEAR PROGRAMMING PROBLEMS****Unit Structure :**

- 3.0 Objective
- 3.1 Introduction
- 3.2 Meaning of Linear Programming Problem
- 3.3 Formulation of Linear Programming Problem
- 3.4 Sketching of Graphs
 - 3.4.1 Graph of a linear equation
 - 3.4.2 Graphs of Linear Inequalities
- 3.5 Solution of L.P.P. by graphical method

3.0 OBJECTIVE

From this chapter students should learn Introduction, meaning of linear programming problem, formulation of linear programming problem, some examples on formulation, Sketching of graphs of linear equation and linear inequalities, and solving of linear programming problem graphically.

3.1 INTRODUCTION

Planning is the heart and the soul of any project, be it a business empire or a simple task required to be done by an individual. However, we will be discussing planning w.r.t. production houses here.

Every organization, uses labour, machine, money, materials, time etc. These are called resources. As one cannot have an unlimited supply of resources, there is always an upper limit on these resources. Therefore, the management has to plan carefully and systematically to use these resources, so as to get the maximum profit at a minimum cost. This is Basic Principle of running any business successfully. Such a planning is called “Programming the strategies”. This is done by writing a management problem as a mathematical model and then solve it scientifically.

Thus, a programming problem consists of Business problems where one faces several limitations causing restriction. One has to remain within the frame – work of these restriction and optimize (maximize or minimize, as the case may be) his goals. The strategies of doing it successfully, is called solving a programming problem.

We will be learning Linear Programming in the chapter, which is the most widely used technique in Production Planning.

3.2 MEANING OF LINEAR PROGRAMMING PROBLEM

As the name suggest, a Linear Programming problem is the problem of maximizing or minimizing a linear function, subject to linear constraints.

Consider a general Programming Problem with a certain goal. Obviously, there are restrictions also.

- i. If the restrictions, when expressed mathematically, are in the form of Linear inequalities, the programming problem is called a Linear Programming problem (L.P.P)
The restrictions are also called constraints.
- ii. If the constraints do not have more than two variables, the L.P.P. can be solved graphically.
- iii. The goals when written mathematically are called objective function.

We will be solving only those linear programming problems having only one objective to be achieved at a time. We will be required to optimize (i.e. maximize or minimize) this objective function.

For example: if the variable is the profit, then we would like to maximize it, but if the variable is the cost or expenditure, we would naturally wish to minimize it. Hence, solving a Linear Programming Problem means, optimizing the given objective function within the given constraints.

Our first job will be to transform the management problems into appropriate mathematical modules.

3.3 FORMULATION OF LINEAR PROGRAMMING PROBLEM

The easiest way to learn this concept is with the help of examples. We begin with the following:

EXAMPLE 1:

A manufacturer produces two types of toys for children, Flutes and drums, each of which must be processed through two machines A and B. The maximum availability of machine A and B per day are 12 and 18 hours respectively. The manufacturing a Flutes requires 4 hours in machine A and 3 hours in machine B, whereas a drum require 2 hours of machine A and 6 hours of machine B, if the profit for Flute is Rs. 20 and per drum is Rs. 50, formulate the problem to maximize the profit.

Solution: Let us suppose that, the manufacture produces x Flutes and y drums per day. Then tabulating the given data as follows, we observe that, if x Flutes and y drums are manufactured per day, he will require $4x + 2y$ hours of machine A and $3x + 6y$ hours of machine B.

Machine	Flutes (x)	Drum (y)	Maximum available
A	4	2	12
B	3	6	18

Since the availability of the machine A and B are not more than 12 and 18 hours respectively, we must have $4x + 2y \leq 12$ and $3x + 6y \leq 18$.

Note that he may not be able to complete a flute or a drum in a day therefore, x or y need not be integers, but x and y can never be negative, since nobody can manufacture a negative number of production. Hence, the condition $x \geq 0, y \geq 0$ will have to be taken

Thus, the problem will be written mathematically as

$$4x + 2y \leq 12$$

$$3x + 6y \leq 18 \text{ and}$$

$$x \geq 0, y \geq 0$$

These inequalities are the constraints on the problem.

Now to find the objective function.

Since the profit per Flute and drum is Rs. 20 and Rs. 50 respectively, the objective function would be

Profit $z = 20x + 50y$, which is to be maximized

Thus Maximize $z = 20x + 50y$

Subject to $4x + 2y \leq 12$

$$3x + 6y \leq 18$$

$$x \geq 0, y \geq 0$$

EXAMPLE 2:

Three different kinds of food A, B and C are to be considered to form a weekly diet. The minimum weekly requirements for fats, carbohydrates and proteins are 12, 30 and 20 units respectively. One Kg. of food A has 2, 16 and 4 units respectively of these ingredients and one Kg. of food B has 6, 4 and 3 units respectively whereas one Kg. of food C has 1, 5 and 7 kgs of these ingredients. If the cost per kg. of food A is Rs. 75, per kg. and that of food B is Rs. 80 and per kg. of food C is Rs. 60, construct the problem to minimize the cost.

Solution: If x kg. of food A, y kg. of food B and z kg. of food C are to be considered for weekly diet, then the data can be represented by the following tabular form.

	A (x)	B (y)	C (z)	Minimum requirements
Fats	2	6	1	12
Carbohydrates	16	4	5	30
Proteins	4	3	7	20

The constraints can be written as

$$2x + 6y + z \geq 12$$

$$16x + 4y + 5z \geq 30$$

$$4x + 3y + 7z \geq 20,$$

$$x \geq 0, y \geq 0$$

Since the cost per kg. of the food A, B and C are given to be Rs 75, Rs. 80 Rs. 60 respectively, the objective function would be:

Cost $C = 75x + 80y + 60z$, which is to be minimized under the given constraints.

Thus Minimize $C = 75x + 80y + 60z$

Subject to $2x + 6y + z \geq 12$
 $16x + 4y + 5z \geq 30$
 $4x + 3y + 7z \geq 20,$
 $x \geq 0, y \geq 0$

EXAMPLE 3:

Two types of food packets A and B are available. Each contain vitamins N_1 and N_2 . A person need 4 decigrams of N_1 and 12 decigrams of N_2 per day. Food packet A contain 2 decigram of vitamin N_1 and 4 decigram of vitamin N_2 . Food packet B contain 1 decigrams of vitamin N_1 and 4 decigrams of vitamin N_2 . Food packed A and B cost Rs. 15 and Rs. 10 respectively. Formulate L.P.P. which will minimize the cost.

Solution: Let x = no. of packet of food A. y = no. of packet of food B.

Food packet A cost Rs. 15 and food packet B cost Rs. 10.

$\therefore$ Objective function i.e. cost function

$Z = 15x + 10y$ subject to

Side constraints

	Food packed A	Food packet B	Requirement ($\geq$)
Vitamin N_1	2	1	4
Vitamin N_2	4	4	12

x and y are no. of packet of food A and B respectively.

$\therefore x \geq 0, y \geq 0$

Mathematical form

Min $Z = 15x + 10y$

Subject to $2x + y \geq 4$

$$4x + 4y \geq 12$$

$$x \geq 0, y \geq 0$$

EXAMPLE 4:

A machine is used for producing two products A and B. Product A is produced by using 4 units of chemical salt and 2 units of chemical mixture. Product B is produced by using 2 unit chemical salt and 3 units of chemical mixture. Only 100 units of chemical salt and 1500 units of chemical mixture are available. The profit on product A is Rs. 30 and on B is Rs. 20 per unit. Formulate this L.P.P. to maximize the profit.

Solution: Let x = no. of unit of product A be produce.
 y = no. of unit of product B be produce.

Profit on product A is Rs. 30 per unit of chemical mixture are available.

	Product A	Product B	Availability
Chemical salt	4	2	1000
Chemical mixture	2	3	1500

x and y are numbers of units $\therefore x \geq 0, y \geq 0$.

Mathematical form

$$\text{Max } z = 30x + 20y$$

$$\text{Subject to } 4x + 2y \leq 1000 \quad (\text{chemical salt})$$

$$2x + 3y \leq 1500 \quad (\text{chemical mixture})$$

$$x \geq 0, y \geq 0.$$

3.4 SKETCHING OF GRAPHS:

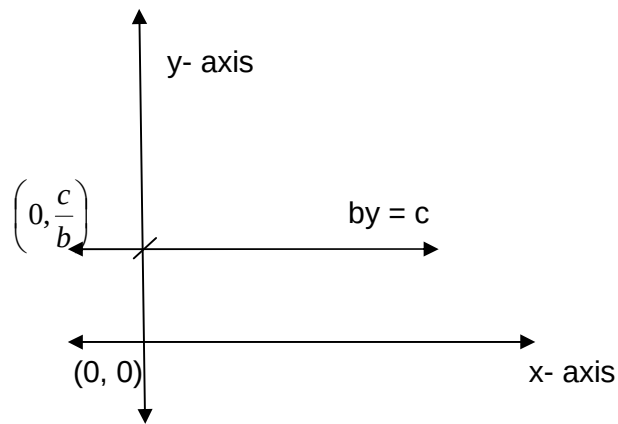
Before we begin to solve the problem of L.P.P. graphically, we shall first see how to sketch the graph of

1. Linear equation
2. Linear inequality

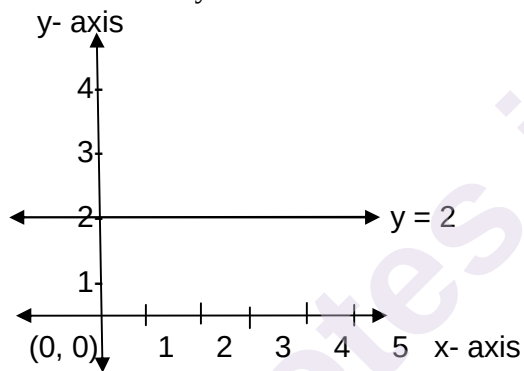
3.4.1. Graph of a linear equation:

A linear equation $ax + by = c$ in two variables x and y where a , b , and c are constants. Here a and b are constant coefficient of x and y respectively not all zero simultaneously, the graph of the linear equation $ax + by = c$ represents a straight line in xy plane, intercepting the x -axis and y -axis. There are four possibilities.

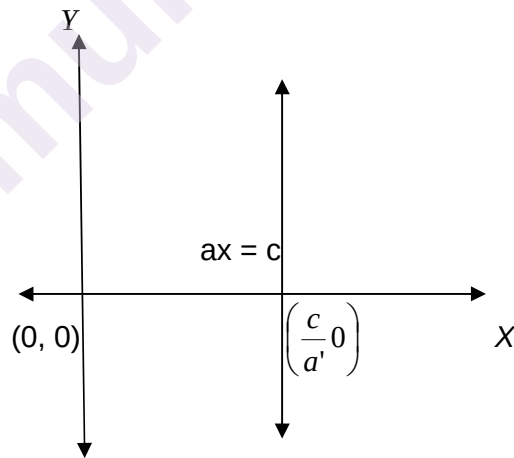
Case (i): If $a = 0$, the linear equation reduces to $by = c$. The graph of this equation is a straight line parallel to x -axis and intercepting y -axis at c/b .



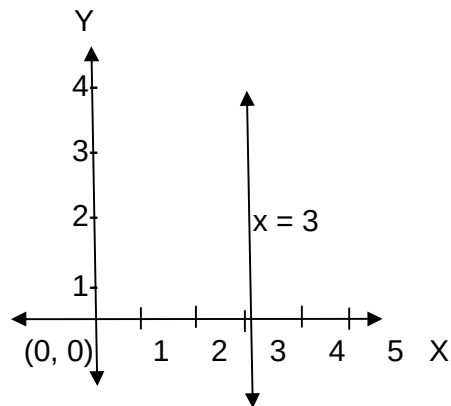
Particularly,
 Sketching of Graph of $3y = 6$
 $\therefore y = 2$



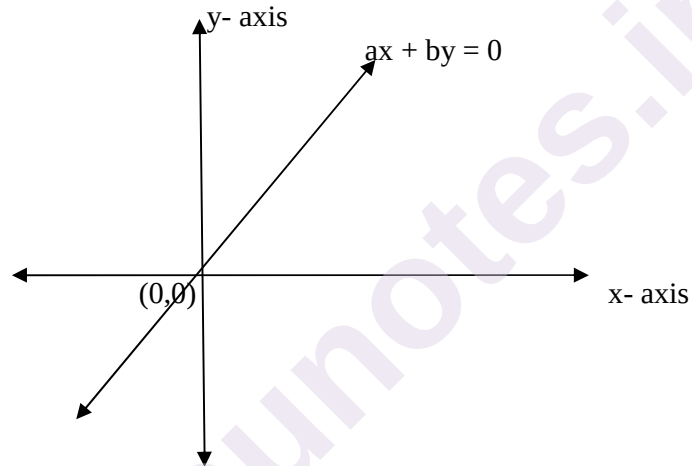
Case (ii): If $a \neq 0$, $b = 0$ and $c \neq 0$. Linear equation reduces to $ax = c$. The graph of this equation is a straight line parallel to y-axis intercepting x-axis at c/a



Particularly, Sketching of Graph of $9x = 27$
 $\therefore x = 3$

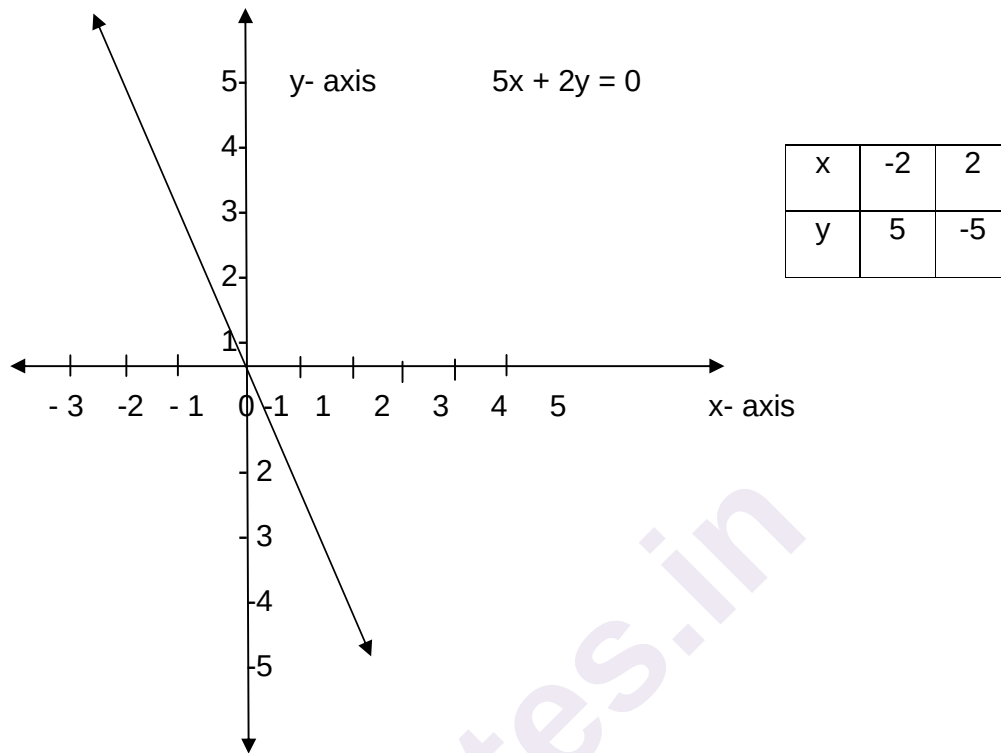


Case (iii): If $a = 0$, $b = 0$, $c = 0$, the equation reduces to $ax + by = 0$. The graph of this equation is a straight line passing through the origin as shown in the figure.

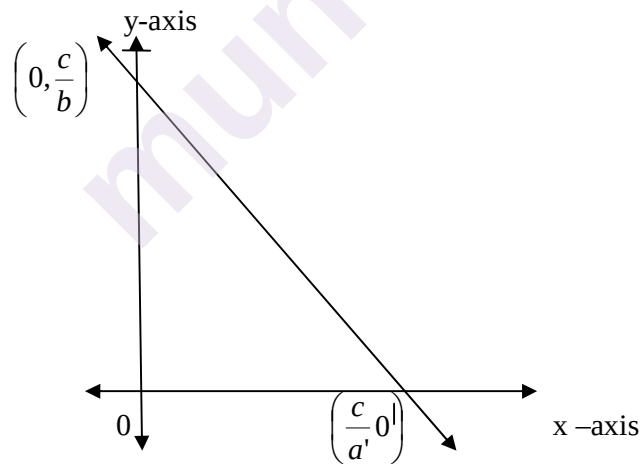


Particularly, Sketching of Graph of $5x + 2y = 0$.
 x - intercept and y - intercept are '0'

$\therefore$ Graph of equation $5x + 2y = 0$ is a straight line passing through the origin.

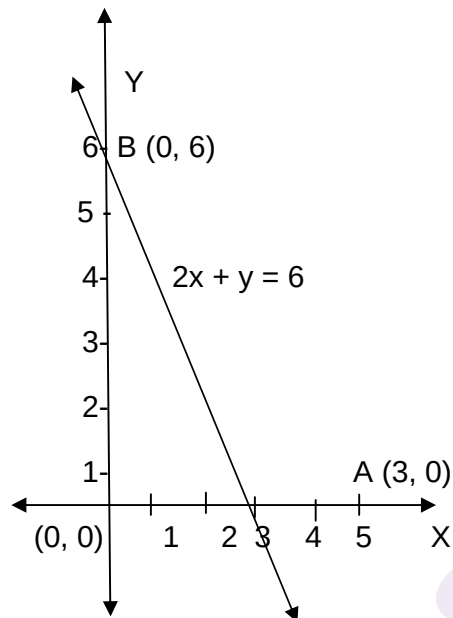


Case (iv): If $a \neq 0$, $b \neq 0$, $c \neq 0$, the equation reduces to $ax + by = c$. The graph of this equation is shown below.



Particularly, Sketching of Graph of $2x + y = 6$

x	y	points
0	6	(0,6)
3	0	(3,0)



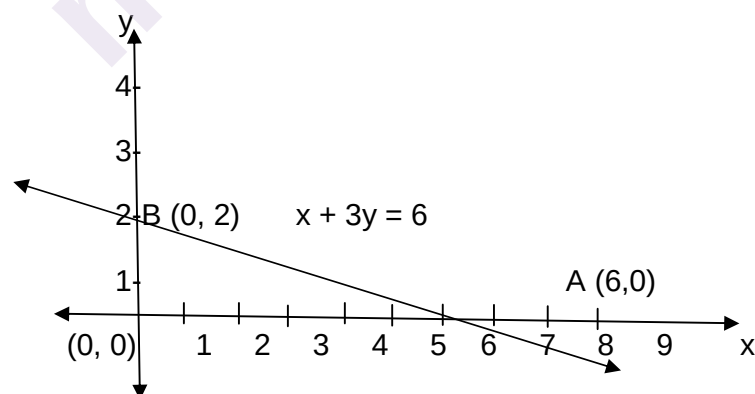
Exercise

Draw the graph of the following linear equations:

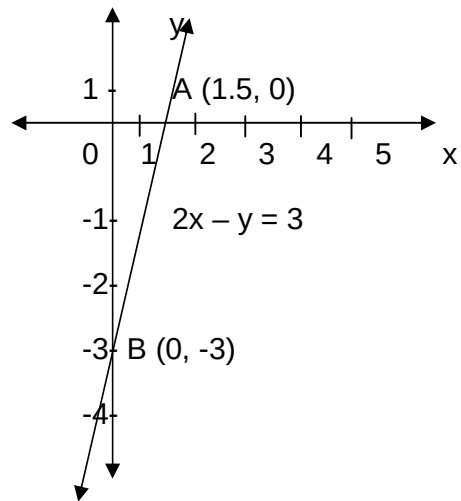
1. $x + 3y = 6$;
2. $2x - y = 3$;
3. $3x - 5y = 8$;
4. $x = 5$;
5. $x = -3$;
6. $y = -2$.
- 7.

Answers :

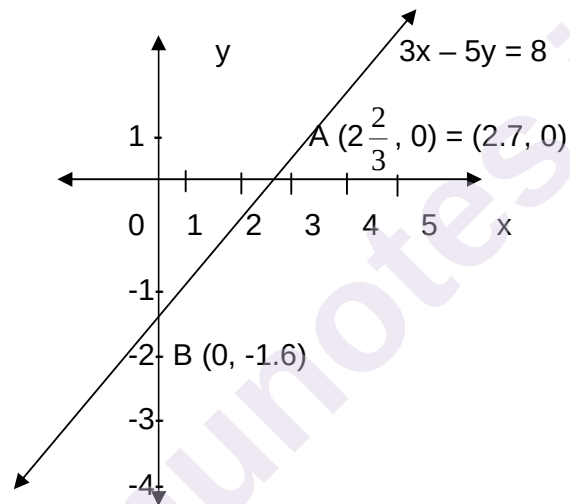
1



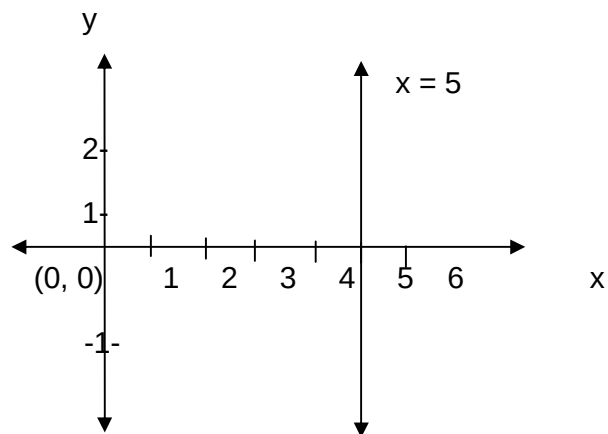
2.

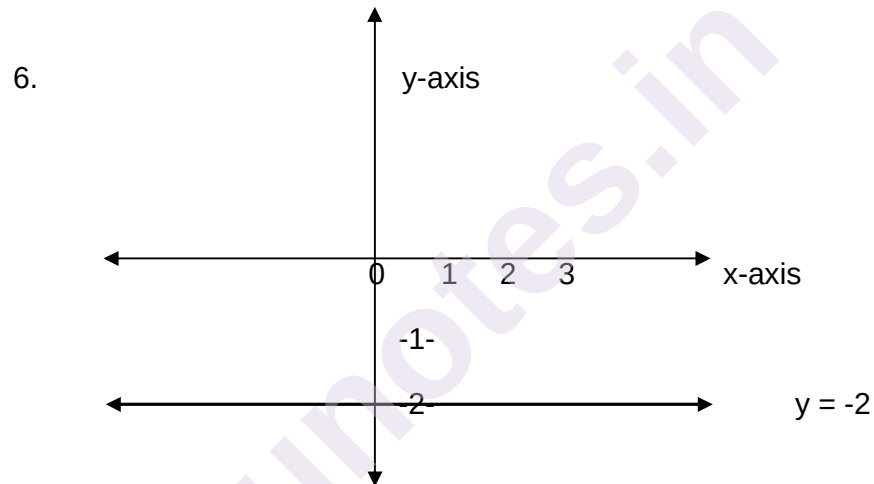
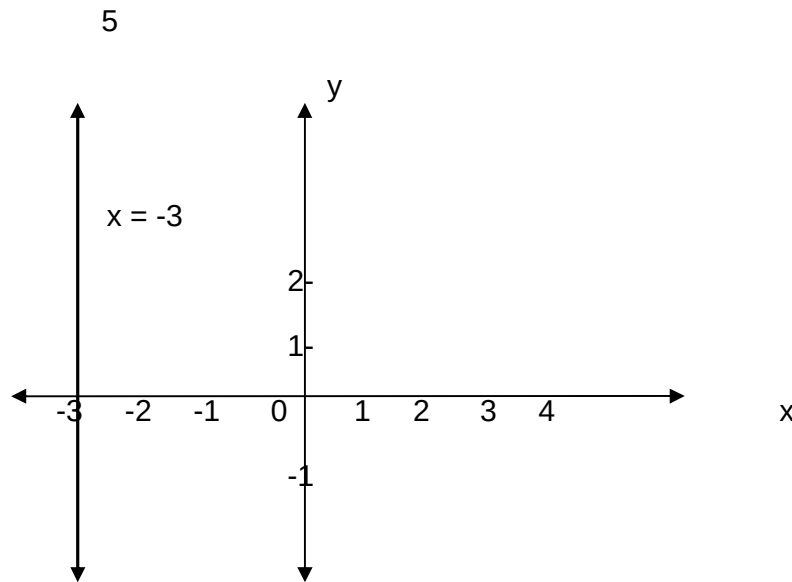


3.



4.





3.4.2. Graphs of Linear Inequalities:

$ax + by + c$ is called a linear expression. We have seen that if $ax + by + c = 0$, then it is a linear equation and it represents a straight line, if a, b, c are real numbers and a, b are both not zero.

Linear inequalities are of four types:

1. $ax + by + c \leq 0$,
2. $ax + by + c < 0$,
3. $ax + by + c \geq 0$,
4. $ax + by + c > 0$

These inequalities represent region of the plane, if a, b, c are real numbers and a, b are not both zero. We are interested in knowing what these regions are.

First draw the line $ax + by + c = 0$ on the graph paper. We know that the plane is now divided into three mutually disjoint set viz. the line itself and the two half planes one on each side of the line. Let us denote by P_1 , the plane which contain the origin and by P_2 the half plane which does not contain the origin.

Now consider the origin $(0, 0)$. Putting $x = 0$ and $y = 0$ in the linear expression $ax + by + c$, we get c . We consider two cases.

- i. $c < 0$ and
- ii. $c > 0$.
- iii. $c = 0$
- iv.

Case (I): if $c < 0$, then the half plane P_1 which contain the origin represents the inequality $ax + by + c < 0$. Naturally, the half plane P_2 which does not contain the origin represents the inequality $ax + by + c > 0$

Further, the inequality $ax + by + c \leq 0$ is represented graphically by the union of the half plane P_1 and the line $ax + by + c = 0$. Whereas, the inequality $ax + by + c \geq 0$ is represented graphically by the union of the half plane P_2 and the line $ax + by + c = 0$.

Case (II): If $c > 0$, then the half plane P_1 which contain the origin represent the in inequality $ax + by + c > 0$. Naturally, the half plane P_2 which does not contain the origin represents the inequality

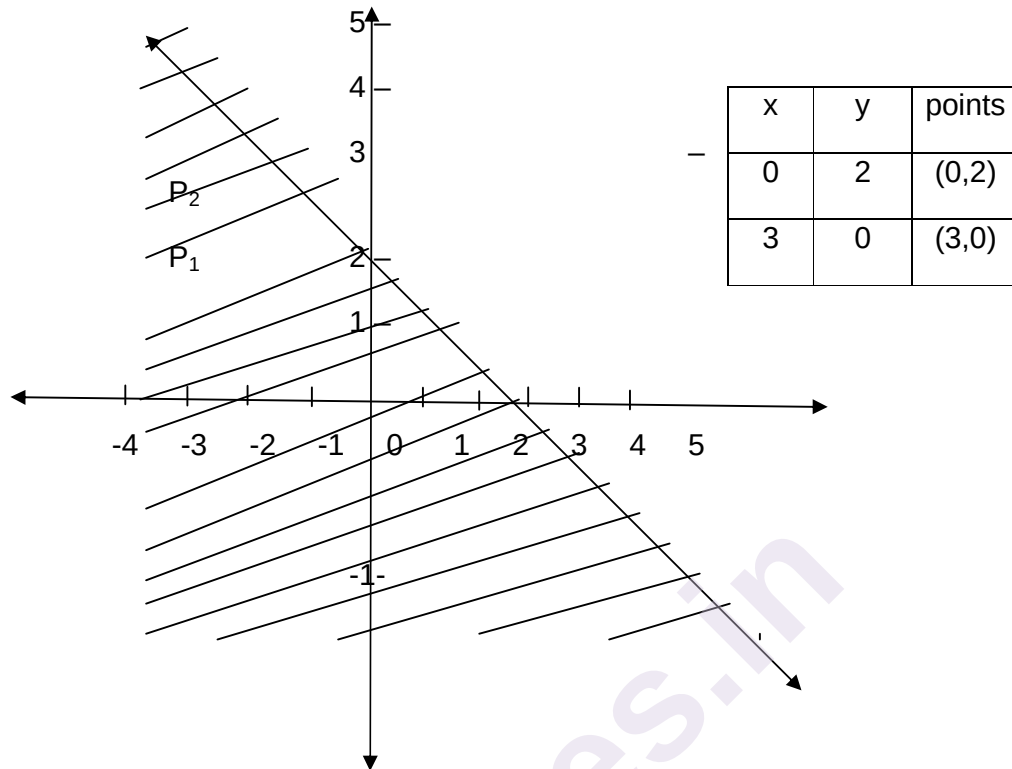
$$ax + by + c < 0$$

Further, the inequality $ax + by + c \leq 0$ is represented by the union of the half plane P_1 and the line $ax + by + c = 0$. Whereas, the inequality $ax + by + c \geq 0$ is represented graphically by the union of the half plane P_2 and the line $ax + by + c = 0$.

EXAMPLE 1: Represent the inequality $2x + 3y \leq 6$ graphically.

Solution: Draw the line $2x + 3y = 6$ on the graph paper. Consider the origin $(0, 0)$. Putting $x = 0$ and $y = 0$ in the inequality, left hand side we get $0 < 6$. Thus $(0, 0)$ satisfies the inequality $2x + 3y \leq 6$.

Hence shaded area including the line will be $2x + 3y \leq 6$



EXAMPLE 2:

Represent the inequality $2x + y \leq 5$ graphically.

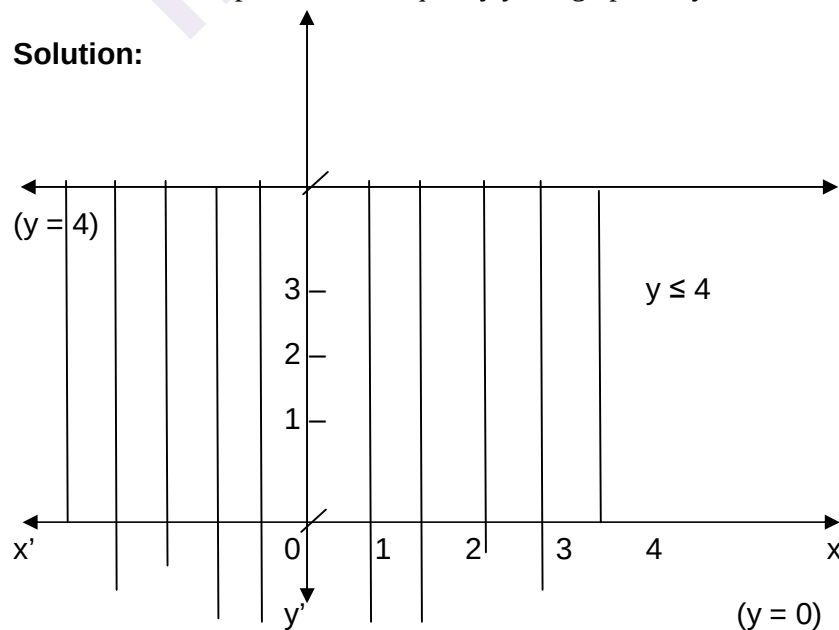
Solution:

Draw the line $2x + y = 5$ on the graph paper. Consider the origin $O(0, 0)$. Putting $x = 0$ and $y = 0$ in the left hand side we get $0 > 5$. But we want $2x + y \geq 5$.

Hence the shaded area will be away from origin above the line.

EXAMPLE 3. Represent the inequality $y \leq 4$ graphically.

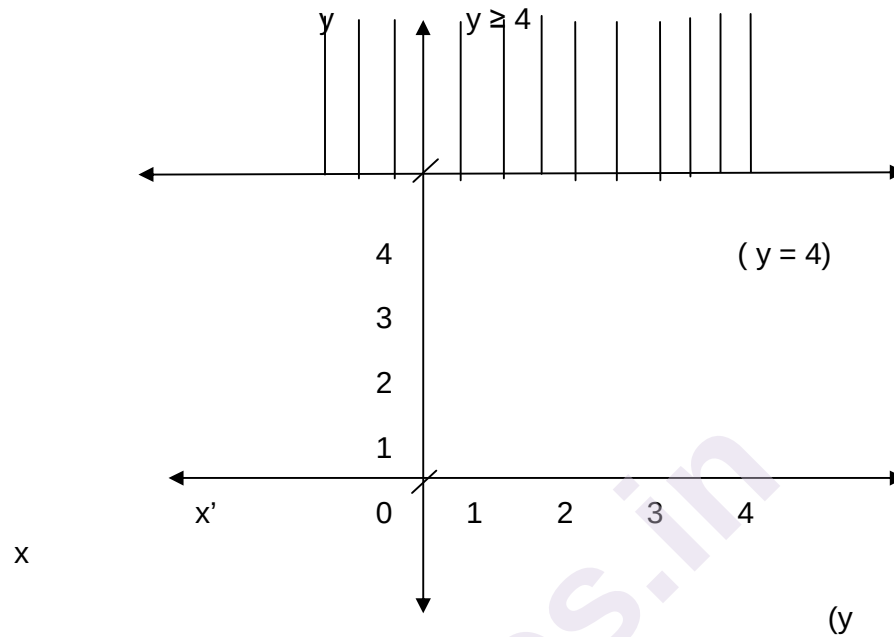
Solution:



($y = 0$) satisfies the inequality.

$\therefore$ Region represented by $y \leq 4$ is towards origin. Below the line $y = 4$.

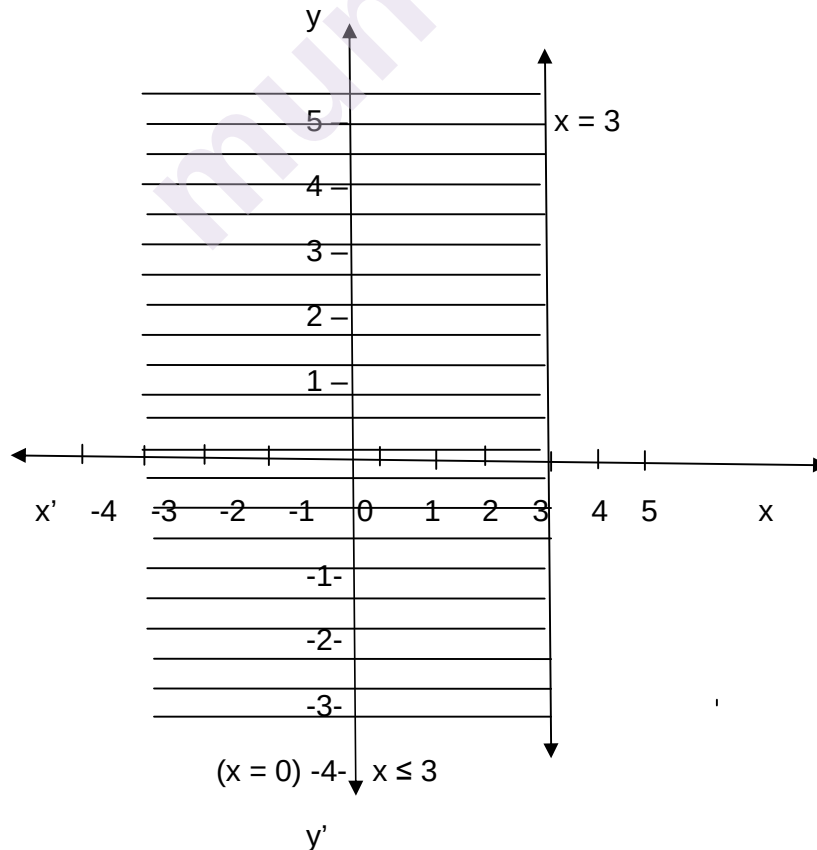
EXAMPLE 4: Represent the inequality $y \geq 4$ graphically.



$= 0$)

$y = 0$ does not satisfy the inequality. $\therefore$ Region represented by $y \geq 4$ is above the line $y = 4$

EXAMPLE 5: Represent the inequality $x \leq 3$

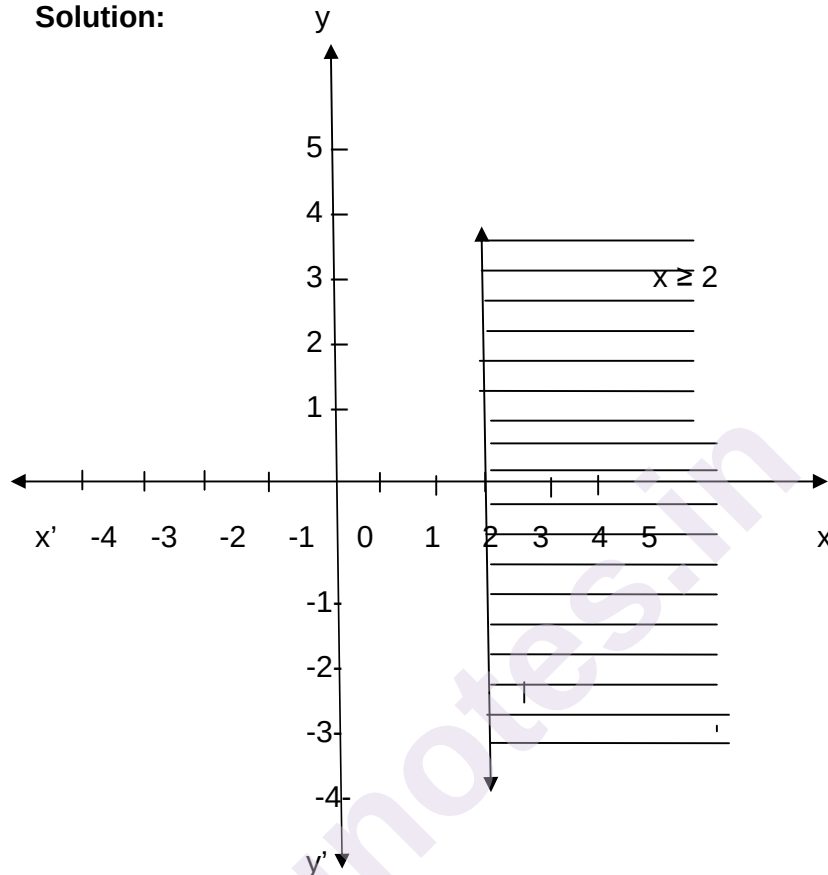


O (0, 0), $x = 0$ satisfies the inequality

$\therefore$ Region is to be left side of the line $x = 3$

EXAMPLE 5: Represent the inequality $x \geq 2$ graphically

Solution:

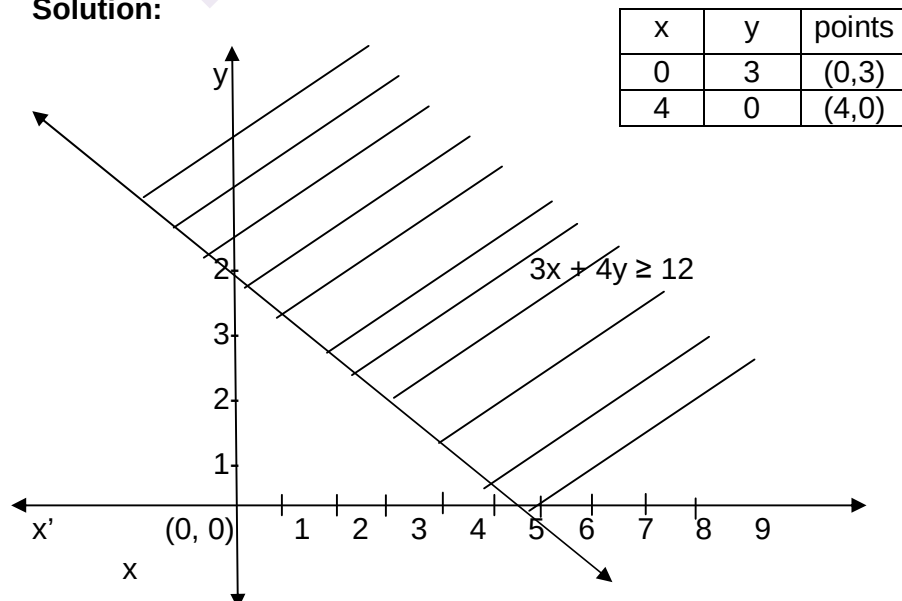


$x = 0$ does not satisfy the inequality $x \geq 2$

$\therefore$ Region is to be right side of the line $x = 2$

EXAMPLE 6: Represent the inequality $3x + 4y \geq 12$ graphically.

Solution:



First draw graph of equation $3x + 4y = 12$.

Origin $O(0, 0)$ does not satisfy the inequality $3x + 4y \geq 2$.

$\therefore$ Region represented by the inequality $3x + 4y \geq$ does not have origin as a point.

Exercise

Sketch the graph of the following linear inequalities.

- 1) $x + 3y \leq 6$
- 2) $x + 3y \leq 6$
- 3) $2x + 5y \geq 10$
- 4) $2x + 5y \geq 10$
- 5) $x \leq 2$
- 6) $y \geq 1$
- 7) $x \geq 0$
- 8) $y \geq 0$

3.5 SOLUTION OF L.P.P. BY GRAPHICAL METHOD

In the earlier section we studied how to formulate L.P.P. However, we need to find solution for the L.P.P. i.e. we have to find the values of the variable which will optimize (Maximize or minimize) the objective function and satisfy all the inequality constraints as well as non – negativity restrictions.

Let us revise some concept we studied in the earlier section in linear inequalities. These concepts will be useful to find solution of L.P.P. by graphical method.

We studied how to find the points in two dimensional co – ordinate geometry, which will satisfy the given linear inequality. If this feasible region is bounded, it is in the form of polyhedron. The set of points of feasible region, which is polyhedron is called polyhedral set. A polyhedral set is called a convex set if a line joining any two points from the set lies entirely in the set. Hence, the feasible region is the convex set if it is a polyhedral set and the line joining any two points from polyhedral set entirely lie in the polyhedral set (feasible region). In the terminology of L.P.P. we can say that feasible region is the set of points, which satisfies the inequality constraints as well as non – negative restriction. Such as set of points is called feasible solution of L.P.P. To find optimal solution of L.P.P. the following result is extremely useful.

Result:

If the convex set of feasible solution is a convex polyhedron set, then at least one of the extreme points giving an optimal solution.

To find solution to L.P.P. graphically by using these results these are the steps in the method.

Following steps are involved in solving L.P.P. graphically.

The steps involved in this method are:

- 1) Draw the graph for the inequality restriction.
- 2) Indicate the area of feasible solution (feasible region)
- 3) Determine the co-ordinates of all points at the corners (points) of all the feasible region.
- 4) Find the value of objective function corresponding to all the solution points determined in (3).
- 5) Determine the feasible solution which optimizes the value of the objective function.

Example1: Solve graphically the L.P. Problem.

Maximize $z = 9x + 12y$,

subject to $2x + 3y \leq 18$,

$2x + y \leq 10$,

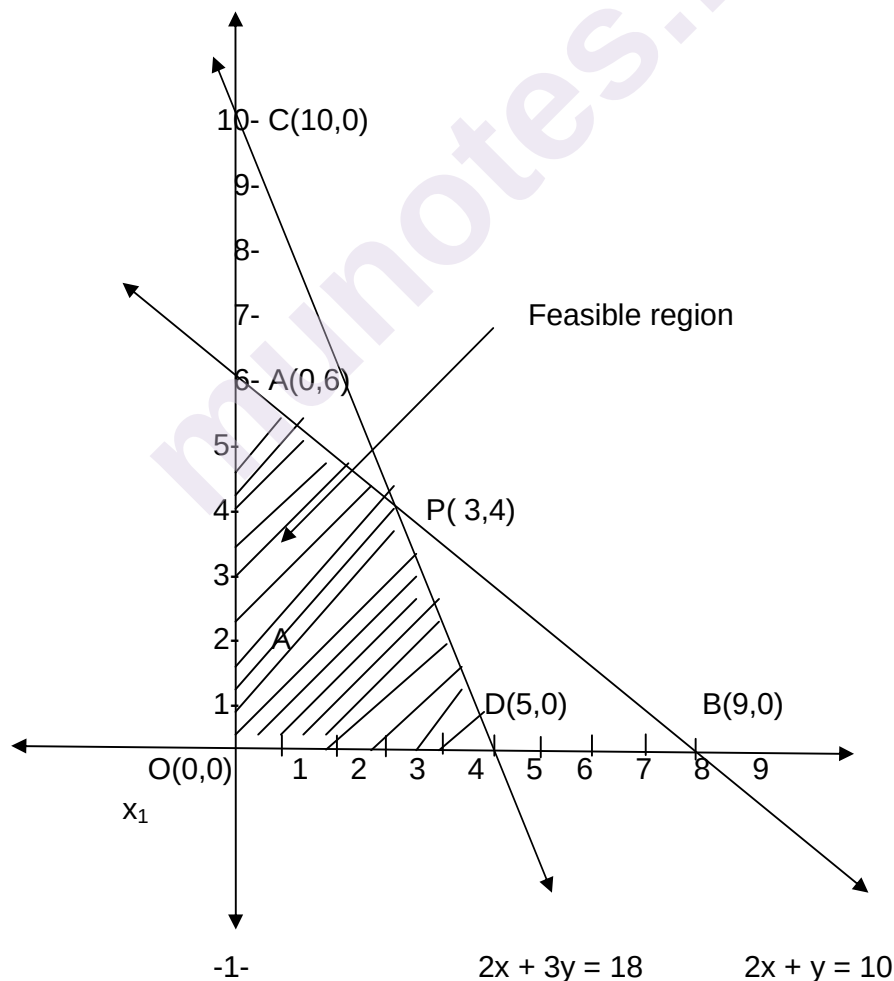
$x \geq 0, y \geq 0$

Solution: For the line $2x + 3y \leq 18$

x	y	Points
0	6	(0,6)
9	0	(9,0)

For the line $2x + y \leq 10$

x	y	Points
0	10	(0,10)
5	0	(5,0)



We first find the region, each point of which satisfies all the constraints by drawing graphs of all the constraints. This is called the region of

feasible solution. The region of feasible solution is the quadrilateral OAPD. The corner point of this region are O, A, P and D.

We find the co-ordinates of P by solving the following equations simultaneously.

$$\begin{array}{rcl} 2x + 3y & = & 18, \\ 2x + y & = & 10, \\ \hline 2y & = & 8 \\ y & = & 4 \end{array}$$

Substituting value of y in $2x + y = 10$, we get $2x + 4 = 10$. $\therefore x = 3$

So the point of intersecting = P (3, 4)

Objective function is $Z = 9x + 12y$,

To check the value of Z at end points of shaded feasible region O, A, P and D

At O (0, 0), $Z = 9(0) + 12(0) = 0$

At A(0, 6), $Z = 9(0) + 12(6) = 72$

At P(3, 4), $Z = 9(3) + 12(4) = 75$

At D (5, 0), $Z = 9(5) + 12(0) = 45$

So, at P (3, 4), Z is Maximized and the maximum value is 75 for $x = 3$ and $y = 4$.

Example2: Solve graphically the L.P. Problem.

Maximize $Z = 5x_1 + 3x_2$

subject to the constraints $3x_1 + 5x_2 \leq 15$,

$5x_1 + 2x_2 \leq 10$,

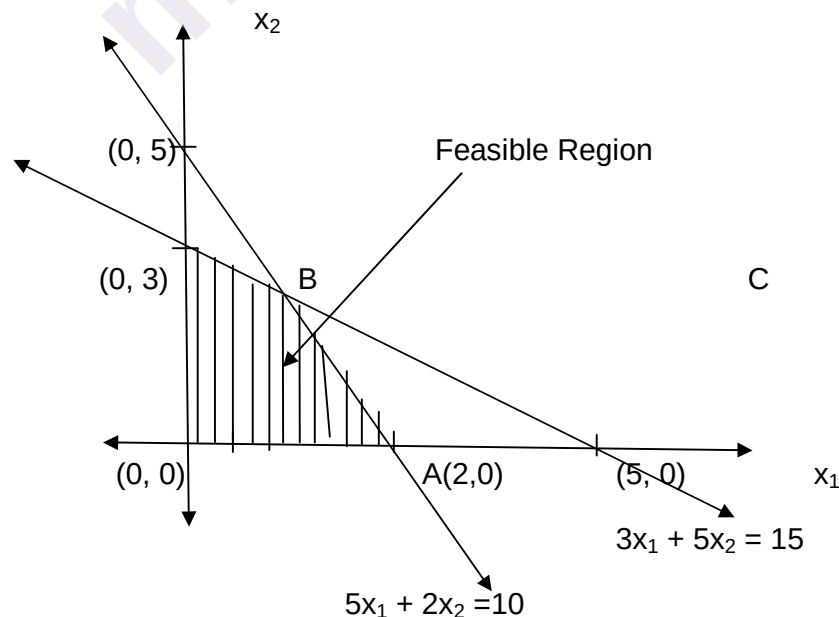
$x_1 \geq 0, x_2 \geq 0$.

Solution: For the line $3x_1 + 5x_2 \leq 15$
10

For the line $5x_1 + 2x_2 \leq$

x_1	x_2	Points
0	3	(3,0)
5	0	(5,0)

x_1	x_2	Points
0	5	(0,5)
2	0	(2,0)



We first find the region, each point of which satisfies all the constraints by drawing graphs of all the constraints. This is called the region of feasible solution. The region of feasible solution is the quadrilateral OACB. The corner point of this region are O, A, B and C.

We find the co-ordinates of C by solving the following equations simultaneously.

$$\begin{array}{rcl}
 3x_1 + 5x_2 & = & 15 \quad \dots\dots\dots(1) \times 5 \\
 \text{and } 5x_1 + 2x_2 & = & 10 \quad \dots\dots\dots(2) \times 3 \\
 15x_1 + 25x_2 & = & 75 \\
 15x_1 + 6x_2 & = & 30 \\
 \hline
 19x_2 & = & 45 \\
 x_2 & = & \frac{45}{19}
 \end{array}$$

substituting x_2 value in any one of the equation, we get $x_1 = \frac{20}{19}$

The objective function is $Z = 5x_1 + 3x_2$.

We find the value of Z at O, A, B, C.

At O (0,0), $Z = 5 \times 0 + 3 \times 0 = 0$

At A (2,0), $Z = 5 \times 2 + 3 \times 0 = 10$

At C $\left(\frac{20}{19}, \frac{45}{19}\right)$, $Z = 5 \times \frac{20}{19} + 3 \times \frac{45}{19} = \frac{235}{19}$

Z has maximum value $\frac{235}{19}$ at C $\left(\frac{20}{19}, \frac{45}{19}\right)$.

i.e. $x_1 = \frac{20}{19}$, $x_2 = \frac{45}{19}$

Example3:

Maximize $Z = 4x_1 - 2x_2$.

Subject to $x_1 + 3x_2 \leq 6$

$x_1 + x_2 \leq 2$

$x_1 \geq 0, x_2 \geq 0$.

Solve the above L.P. problem graphically.

Solution:

Consider the given condition as equality

$x_1 + 3x_2 = 6$ (I)

$x_1 + x_2 = 2$ (II)

For point of intersection, solve the equations simultaneously

$$\begin{array}{rcl}
 x_1 + 3x_2 & = & 6 \\
 x_1 + x_2 & = & 2 \\
 \hline
 2x_2 & = & 4
 \end{array}$$

$\therefore x_2 = 1$

Substituting $x_2 = 1$ in (ii), $x_1 = 3$

So the point of intersecting = P (3, 1)

To draw the straight line representing (i), (ii) consider

x_1	x_2	Points
0	2	(0,2)
6	0	(2,0)

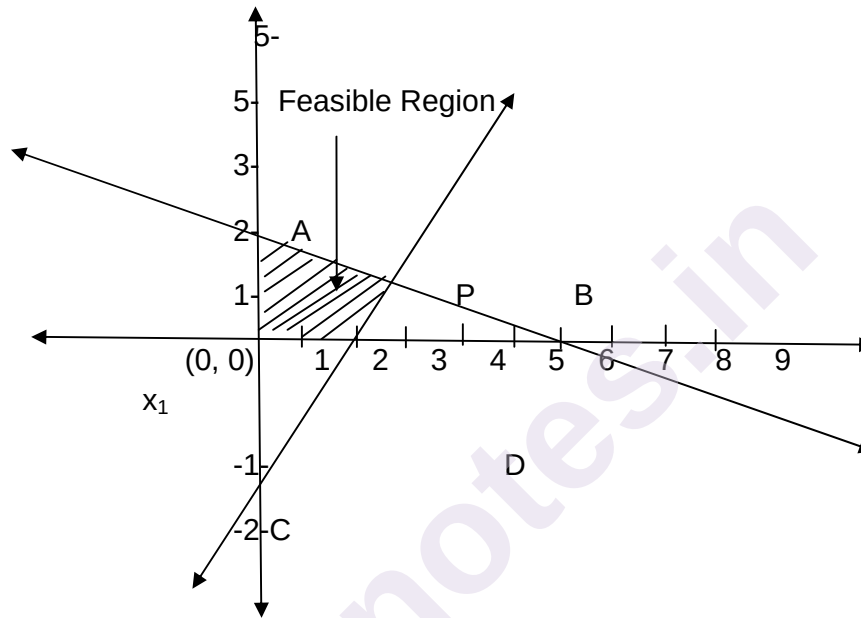
x_1	x_2	Points
0	-2	(0,-2)
2	0	(2,0)

(i)

(ii)

By plotting the point A (0,2), B(6, 0), C(0, -2) and D(2, 0) and joining we get straight line AB and CD representing equation (i) and (ii) and the feasible region is shaded as

$$3x + 4y \geq 12$$

 y_2 

Note : To get P, the line CD is extended.

To check the value of Z at end points of shaded feasible region A, P and D

$$\text{At } (0, 0), Z = 4(0) - 2(0) = 0$$

$$\text{At } (0, 2), Z = 4x_1 - 2x_2 = 4(0) - 2(2) = -4$$

$$\text{At } P(3, 1), Z = 4(3) - 2(1) = 10$$

$$\text{At } D(2, 0), Z = 4(2) - 2(0) = 8$$

So, at P (3, 1), Z is Maximized and the maximum value is 10 for $x_1 = 3$ and $x_2 = 1$

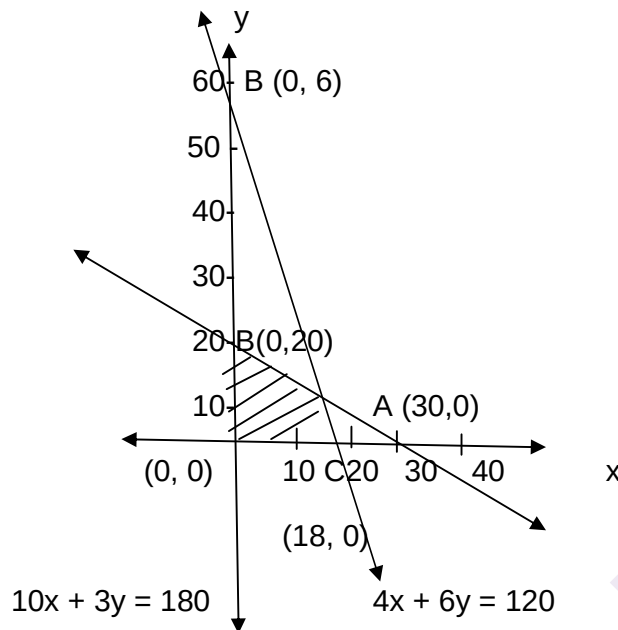
Example 4:

Maximize $z = 800x + 100y$,
 subject to $4x + 6y \leq 120$,
 $10x + 3y \leq 0, y \geq 0$

Solution : Now we solve this problem by the graphical method as follows: First we draw the graph of the solution set of the linear inequalities.

$$4x + 6y \leq 120, 10x + 180, x \geq 0, y \geq 0$$

At each point of this region, we can find the values of the objective function $z = 800x + 100y$.



Hence this region is called a feasible region. We want to find that point at which the objective function $z = 800x + 100y$ has a maximum value. For this purpose, we use the following result.

If the feasible region is a polygonal one, then the maximum and minimum value of the objective function lie at some vertices of the region.

Now we draw the line AB and CD whose equations are $4x + 6y = 120$ and $10x + 3y = 180$ respectively. For this purpose, we find their point of intersection with the co-ordinate axes. It is convenient to form a table for their co-ordinates as follows:

For the line
 $4x + 6y = 120$

x	y	points
30	0	A(30,0)
0	20	B(0,20)

For the line
 $10x + 3y = 180$

x	y	points
18	0	C(18,0)
0	60	D(0,60)

Shade the feasible region OCPBO by horizontal lines. Its vertices are O (0, 0), C (18, 0), B (0, 20) and P which is the point of intersection of the line AB and CD. Hence to get the co-ordinates of P, we solve their equation:

$$4x + 6y = 120 \quad \dots(1)$$

$$10x + 3y = 180 \quad \dots(2)$$

Multiplying equation (2) by 2 and subtracting equation (1) from it,

$$\begin{array}{rcl}
 20x + 6y & = & 360 \\
 4x + 6y & = & 120 \\
 \hline
 16x & = & 240
 \end{array}$$

$$\begin{aligned}
 x &= 15 \\
 \therefore \text{from (2)} \\
 150 + 3y &= 180 \\
 \therefore 3y &= 30 \therefore y = 10 \\
 \therefore P \text{ is } (15, 10).
 \end{aligned}$$

We find the values of the objective function at these vertices:

$$z(O) = 800 \times 0 + 100 \times 0 = 0$$

$$z(P) = 800 \times 15 + 100 \times 10 = 13000$$

$$z(C) = 800(18) + 100 \times 0 = 14400$$

$$z(B) = 800 \times 0 + 100 \times 20 = 2000$$

$\therefore z$ has the maximum value 14400 at the point C (18, 0).

$\therefore$ the manufacturer should produce 18 scooters and 0 bicycles, in order to have maximum profit of Rs. 14400.

Example5:

$$\text{Minimize } z = 40x + 37y$$

$$\text{Subject to constraints: } 10x + 3y \geq 180$$

$$2x + 3y \geq 60$$

$$x \geq 0, y \geq 0$$

Solution:

We first draw the line AB and CD whose equation are $10x + 3y = 180$ and $2x + 3y = 60$.

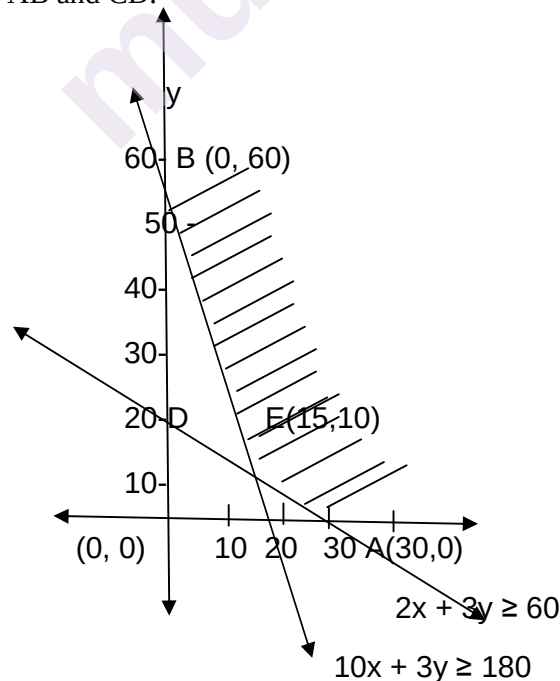
For the line $10x + 3y = 180$

For the line $2x + 3y = 60$

x	y	Points
18	0	(18,0)
0	60	(0,60)

x	y	Points
30	0	(30,0)
0	20	(0,20)

The feasible region is shaded in the figure. Its vertices are C (30, 0), B (0, 60) and E, where E is the point of intersection of the lines AB and CD.



For the point E, we solve the two equations simultaneously.

$$\begin{array}{rcl}
 10x + 3y & = & 180 \\
 2x + 3y & = & 60 \\
 \hline
 8x & = & 120 \\
 x & = & 15
 \end{array}$$

$\therefore 2x + 3y = 60$ gives $30 + 3y = 60$,

i.e., $3y = 30$, i.e. $y = 10$

$\therefore E$ is $(15, 10)$

The values of the objective function $z = 40x + 37y$ at these vertices are:

Vertex (x, y)	$z = 40x + 37y$
C (30, 0)	$z(C) = 40 \times 30 + 37 \times 0 = 1200$
B (0, 60)	$z(B) = 40 \times 0 + 37 \times 60 = 2220$
E (15, 10)	$z(E) = 40 \times 15 + 37 \times 10 = 600 + 370 = 970$.

$\therefore z$ has minimum value 970 at the point E (15, 10), where $x = 15$ and $y = 10$.

Example6: Minimize $z = 4x + 2y$

Subject to constraints: $x + 3y \geq 3$, $2x + y \geq 2$, $x \geq 0$, $y \geq 0$

Solution:

We first draw the line AB and CD.

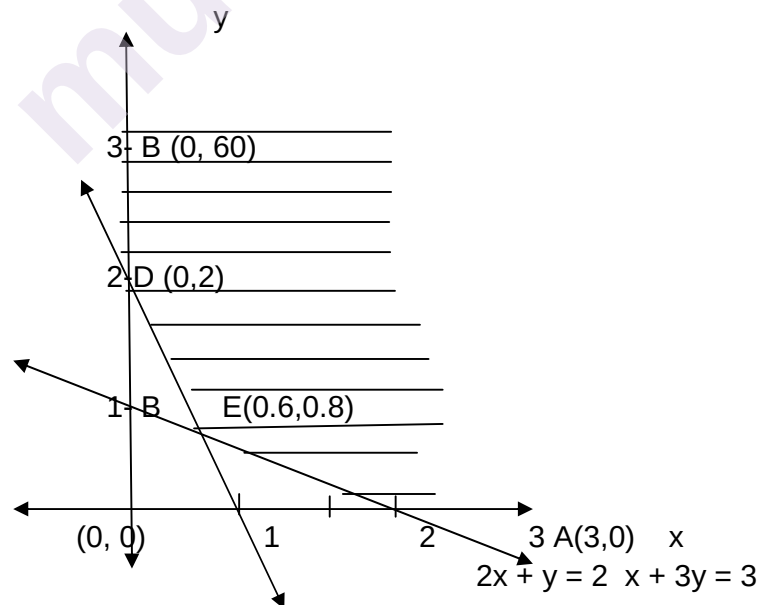
For the line $x + 3y = 3$

For the line $2x + y = 2$

x	y	Point
3	0	A(3,0)
0	1	B(0,1)

x	y	Point
1	0	C(1,0)
0	2	D(0,2)

The feasible region is shaded in the figure. Its vertices are A (3, 0), D (0, 2) and E, where E is the point of intersection of the lines AB and CD.



To find the point E, we solve the two equations (1) and (2) simultaneously.

$$\begin{array}{rcl}
 2x + 6y & = & 6 \\
 2x + y & = & 2 \\
 \hline
 5y & = & 4 \\
 x & = & \frac{4}{5} = 0.8
 \end{array}$$

Substituting $y = 0.8$ in (1), we get,

$$\begin{array}{rcl}
 \bullet \bullet & x + 3 \times 0.8 & = 3 \\
 \bullet \bullet & x + 2.4 & = 3 \\
 \bullet \bullet & x & = 3 - 2.4 = 0.6
 \end{array}$$

Thus E is (0.6, 0.8)

The values of the objective function $z = 4x + 2y$ at the vertices are calculated below:

Vertex (x, y)	$z = 4x + 2y$
C (3, 0)	$z (A) = 4 \times 3 + 2 \times 0 = 12$
E (0.6, 0.8)	$z (E) = 4 \times 0.6 + 2 \times 0.8 = 2.4 + 1.6 = 4$
D(0, 2)	$z (D) = 4 \times 0 + 2 \times 2 = 4$

We can see the minimum value of z is 4, at two vertices E(0.6, 0.8) and D(0, 2). Thus z is minimum at any point on the line segment ED.

Exercise

Maximize:

- 1) $z = 5x + 10y$, subject to
 $5x + 8y \leq 40, 3x + y \leq 12, x \geq 0, y \geq 0$
- 2) $z = 7x + 6y$, subject to
 $2x + 3y \leq 13, x + y \leq 5, x \geq 0, y \geq 0$
- 3) $z = 5x + 3y$, subject to
 $2x + y \leq 27, 3x + 2y \leq 48, x \geq 0, y \geq 0$
- 4) $z = 20x + 30y$, subject to
 $x + y \leq 6, 3x + y \leq 12, x \geq 0, y \geq 0$
- 5) $z = 6x + 7y$, subject to
 $2x + 3y \leq 12, 2x + y \leq 8, x \geq 0, y \geq 0$
- 6) $z = 30x + 20y$, subject to
 $2x + y \leq 20, x + 3y \leq 15, x \geq 0, y \geq 0$
- 7) $z = 20x + 25y$, subject to
 $5x + 2y \leq 50, x + y \leq 12, x \geq 0, y \geq 0$
- 8) $z = 90x + 130y$, subject to
 $2x + 3y \leq 18, 2x + y \leq 12, x \geq 0, y \geq 0$

Minimize:

- 1) $z = x + 4y$, subject to
 $x + 3y \geq 3, 2x + y \geq 2, x \geq 0, y \geq 0$
- 2) $z = 5x + 2y$, subject to
 $10x + 2y \geq 20, 5x + 5y \geq 30, x \geq 0, y \geq 0$
- 3) $z = 9x + 10y$, subject to
 $x + 2y \geq 30, 3x + y \geq 30, x \geq 0, y \geq 0$
- 4) $z = 50x + 55y$, subject to

$$x + 3y \geq 30, 2x + y \geq 20, x \geq 0, y \geq 0$$

5) $z = 80x + 90y$, subject to

$$6x + 5y \geq 300, 2x + 3y \geq 120, x \geq 0, y \geq 0$$

6) $z = 13x + 15y$, subject to

$$3x + 4y \geq 360, 2x + y \geq 100, x \geq 0, y \geq 0$$

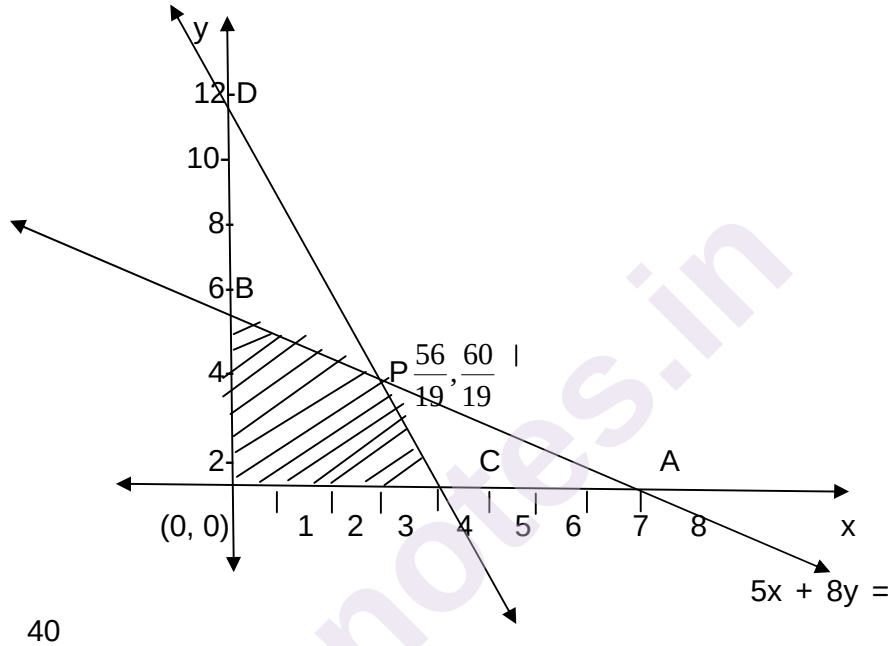
7) $z = 12x + 20y$, subject to

$$x + y \geq 7, 5x + 2y \geq 20, x \geq 0, y \geq 0$$

Answer:

Maximize:

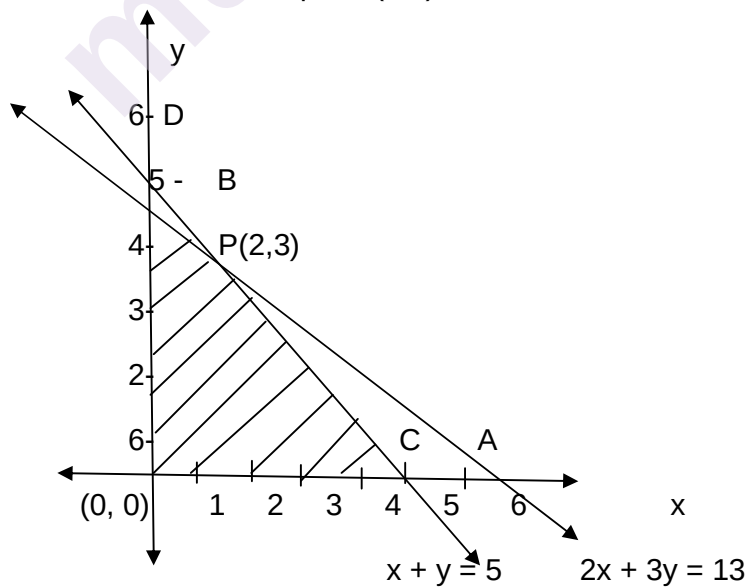
1) Max. value is 50 at the point (0,5)



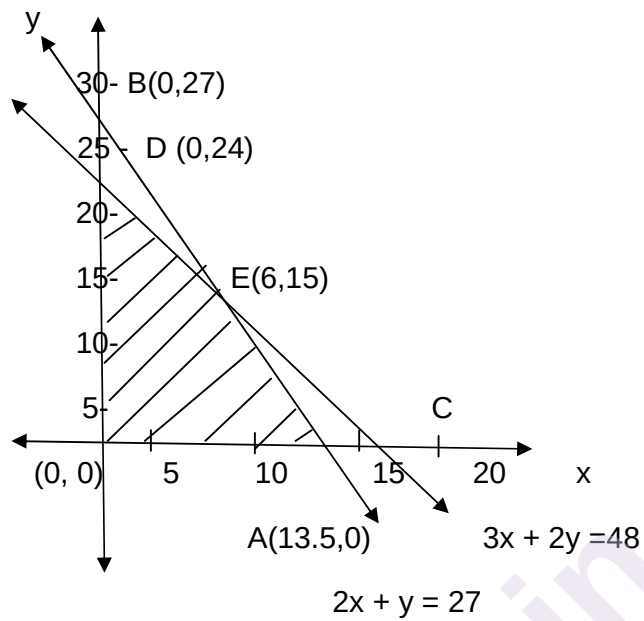
40

$$3x + y = 12$$

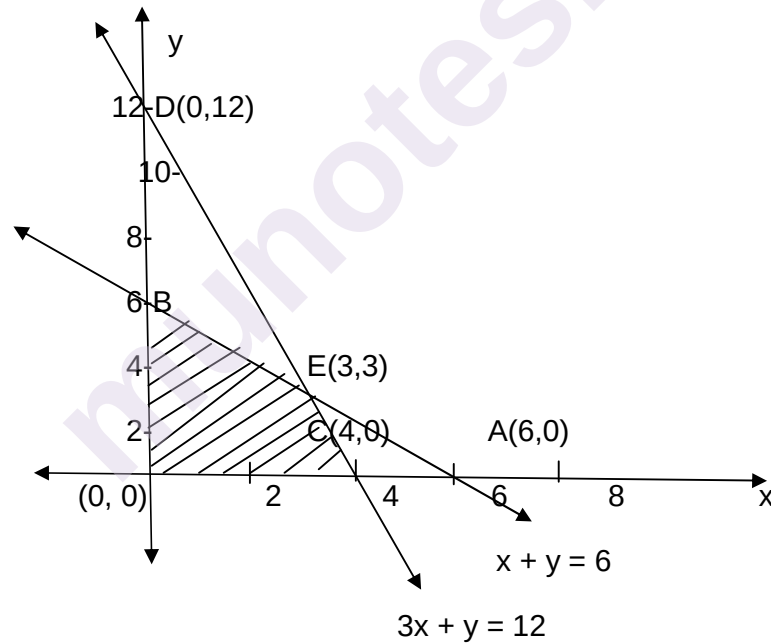
2) Max. value is 35 at the point (5,0)



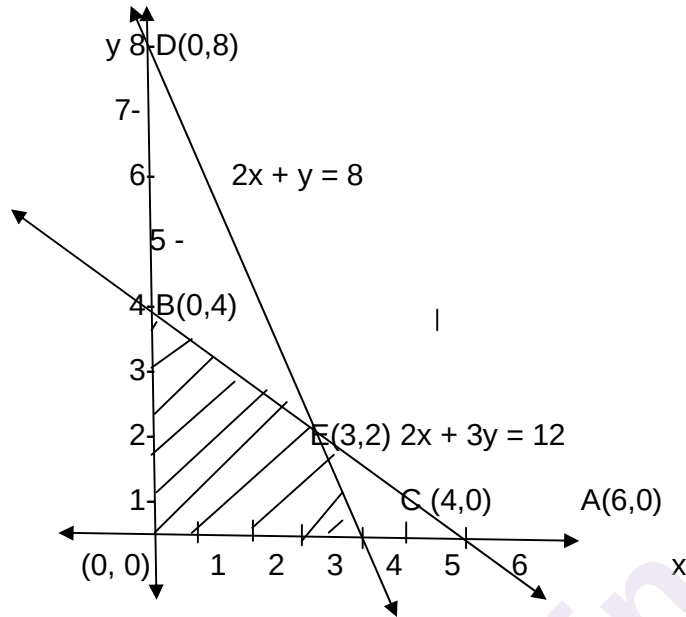
3) Max. value is 75 at the point. (6,15)



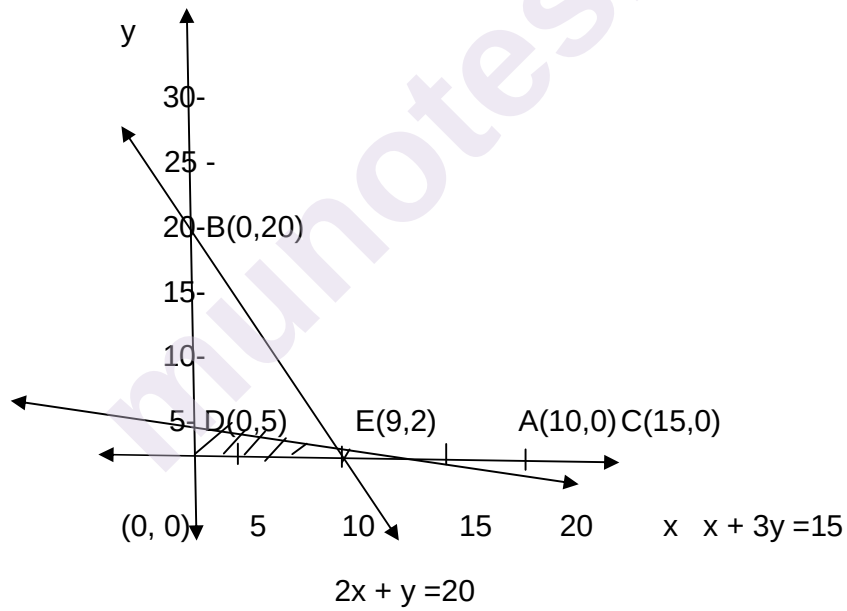
4) Max. value is 180 at the point (0,6)



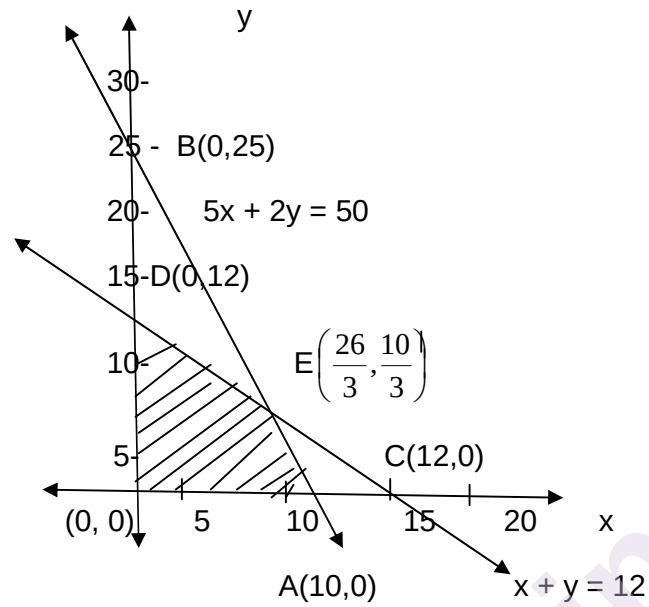
5) Max. value is 32 at the point (3,2)



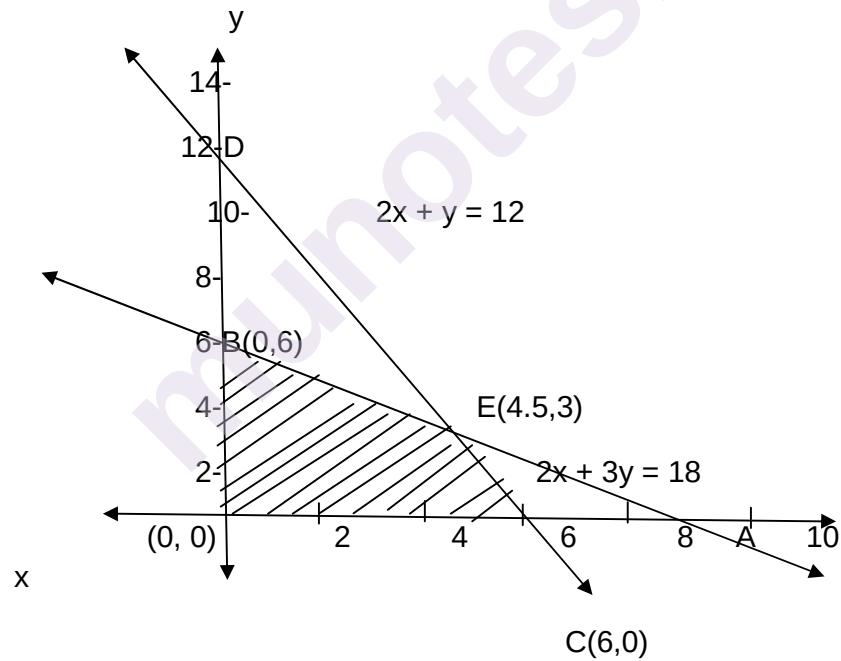
6) Max. value is 310 at the point (9,2)



7) Max. value is 300 at the point (0,12)

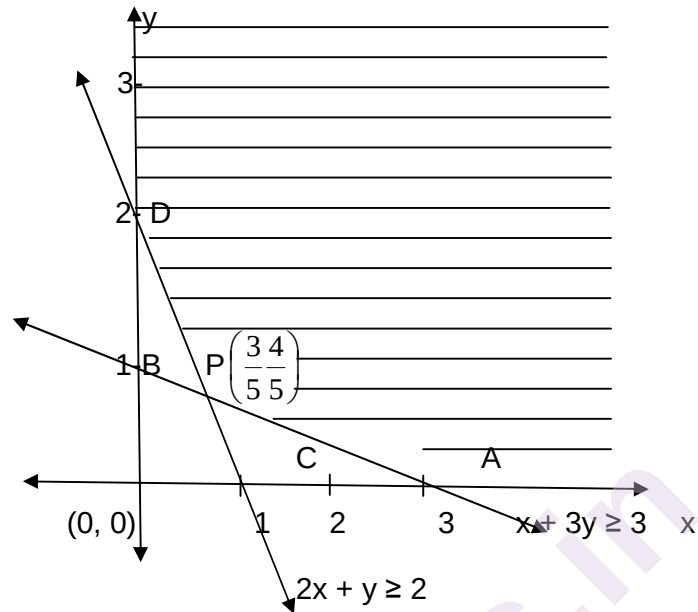


8) Max. value is 795 the point (4.5, 3)

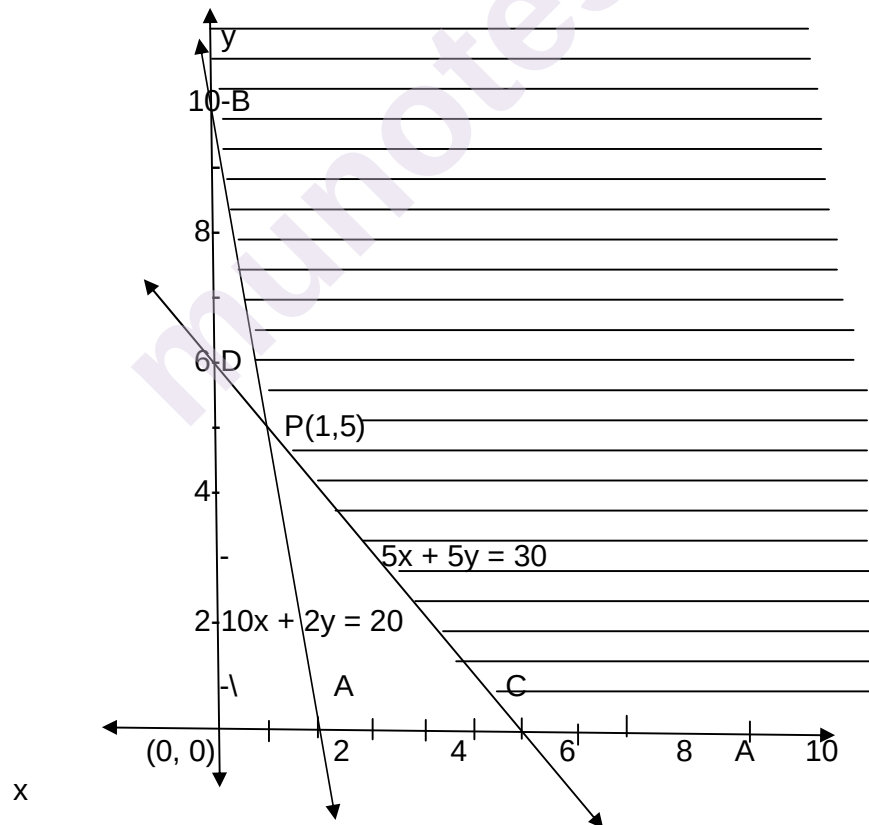


Minimize:

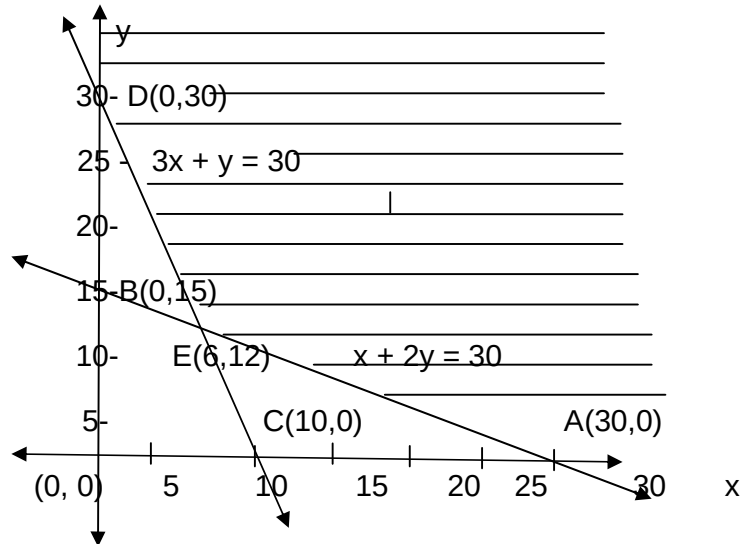
- 1) Min. value is 3 at the point (3,0)



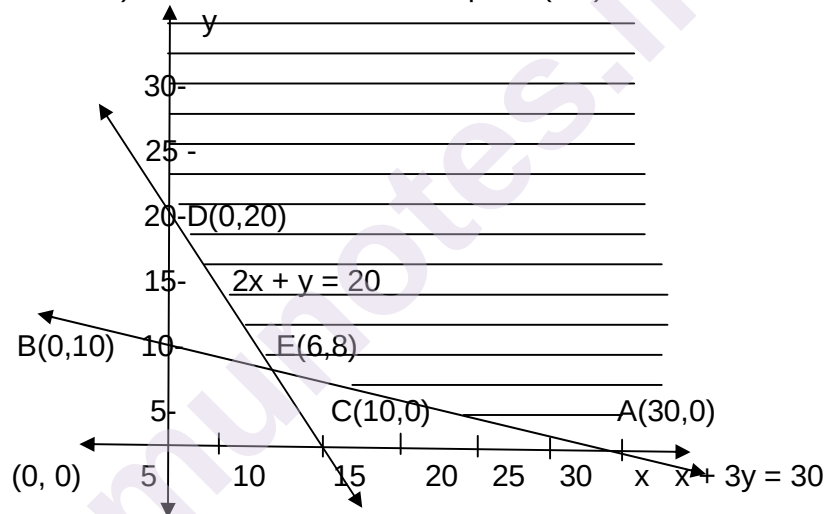
- 2) Min. value is 15 at the point (1,5)



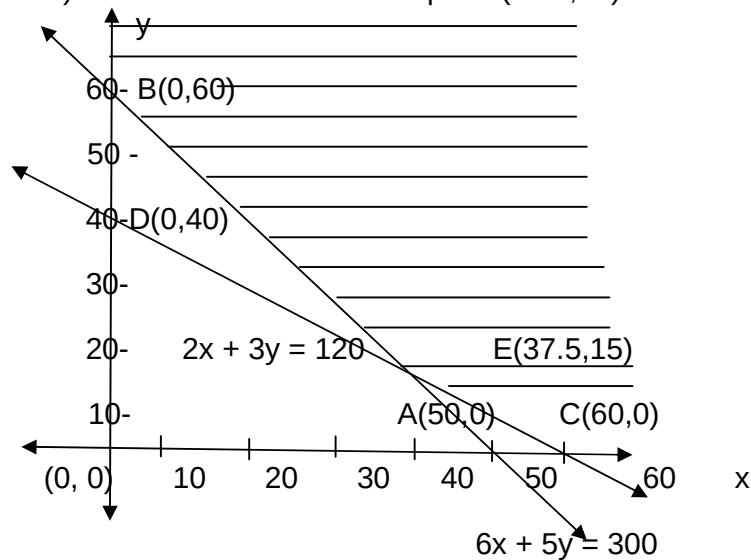
3) Min. value is 174 at the point (6,12)



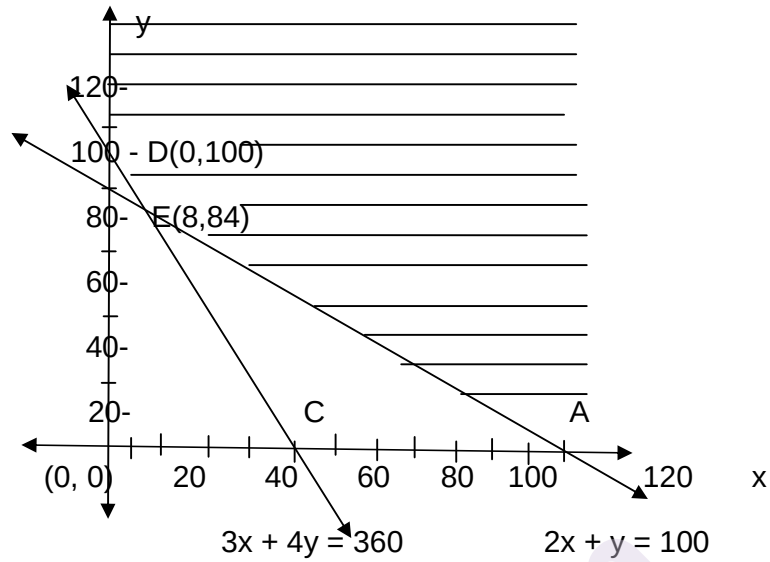
4) Min. value is 740 at the point (6,8)



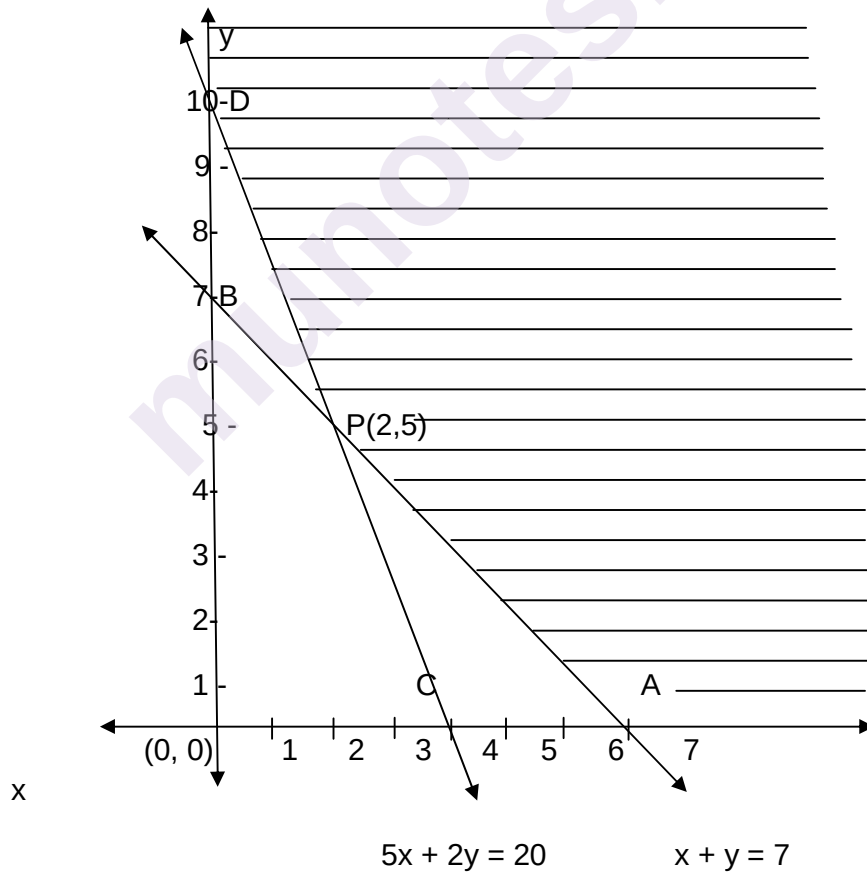
5) Min. value is 4350 at the point (37.5,15)



6) Min. value is 1364 at the point (8,84)



7) Min. value is 84 at the point (7,0)



UNIT III

Unit-4

MEASURES OF CENTRAL TENDENCY

Unit Structures :

- 4.0 Objectives
- 4.1 Introduction
- 4.2 Requisites for an Ideal Average
- 4.3 Types of Averages
- 4.4 Arithmetic Mean (A.M)
- 4.5 Weighted Arithmetic Mean
- 4.6 Combined Mean
- 4.7 Median
- 4.8 Quartiles
- 4.9 Mode
- 4.10 Graphical Location of Mode

4.0 OBJECTIVES

After reading this chapter you will be able to:

- 1) Define an average.
- 2) Understand different types of averages.
- 3) Calculate Arithmetic mean and weighted Arithmetic mean.
- 4) Calculate Median and Quartiles.
- 5) Calculate Mode by formula.
- 6) Graphical location of Median and Mode.

4.1 INTRODUCTION

In the earlier chapter we have discussed the Diagrams and Graphs which are used for presentation of data. Diagrams furnish only approximate information. Graphs are more precise and accurate than diagrams. But they cannot be effectively used for further statistical analysis. To study the characteristics in detail, the data must be further analysed.

There is a tendency in almost every statistical data that most of the values concentrate at the centre which is referred as 'central tendency'. The typical values which measure the central tendency are called measures of central tendency. Measures of central tendency are commonly

known as 'Averages.' They are also known as first order measures. Averages always lie between the lowest and the lightest observation.

The purpose for computing an average value for a set of observations is to obtain a single value which is representative of all the items and which the mind can grasp simply and quickly. The single value is the point or location around which the individual items cluster.

The word average is very commonly used in day to day conversation. For example, we often talk of average boy in a class, average height of students, average income, etc. In statistics, the term average has a different meaning. Averages are very much useful for describing the distribution in concise manner. The averages are extremely useful for comparative study of different distributions.

4.2 REQUISITES FOR AN IDEAL AVERAGE

The following are the characteristics which must be satisfied by an ideal average:

- (i) It should be rigidly defined.
- (ii) It should be easy to calculate and easy to understand.
- (iii) It should be based on all observations.
- (iv) It should be suitable for further mathematical treatment.
- (v) It should be least affected by the fluctuations of sampling.
- (vi) It should not be affected by extreme values.

4.3 TYPES OF AVERAGES

The following are the five types of averages which are commonly used in practice.

1. Arithmetic Mean or Mean (A.M)
2. Median
3. Mode
4. Geometric Mean (G.M)
5. Harmonic Mean (H.M)

Of these, arithmetic mean, geometric mean and harmonic mean are called mathematical averages; median and mode are called positional averages. Here we shall discuss only the first three of them in detail, one by one.

4.4 ARITHMETIC MEAN (A.M)

(i) **For Simple or Ungrouped data:**(where frequencies are not given)
Arithmetic Mean is defined as the sum of all the observations divided by the total number of observations in the data and is denoted by $\bar{x}$, which is read as 'x-bar'

i.e. $\bar{x} = \frac{\text{sum of all observations}}{\text{number of observations}}$

In general, if $x_1, x_2, \dots, x_n$ are the n observations of variable x , then the arithmetic mean is given by

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

If we denote the sum $x_1 + x_2 + \dots + x_n$ as $\sum x$, then

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

Note: The symbol $\sum$ is the Greek letter *capital sigma* and is used in Mathematics to denote the sum of the values.

Steps: (i) Add together all the values of the variable x and obtain the total, i.e., $\sum x$.
(ii) Divide this total by the number of observations.

Example 1:

Find the arithmetic mean for the following data representing marks in six subjects at the H.S.C examination of a student.

The marks are 74, 89, 93, 68, 85 and 76.

Solution: $n = 6$

$$\bar{x} = \frac{\sum x}{n} \Rightarrow \bar{x} = \frac{74 + 89 + 93 + 68 + 85 + 76}{6} = \frac{485}{6}$$

$$\bar{x} = 80.83$$

Example 2: Find arithmetic mean for the following data .

425 , 408 , 441 , 435 , 418

Solution: $n = 5$ and $\bar{x} = \frac{\sum x}{n}$

$$\therefore \bar{x} = \frac{425 + 408 + 441 + 435 + 418}{5} = \frac{2127}{5}$$

$$\therefore \bar{x} = 425.4$$

(ii) For Grouped data (or) discrete data:
(values and frequencies are given)

If $x_1, x_2, \dots, x_n$ are the values of the variable x with corresponding frequencies $f_1, f_2, \dots, f_n$ then the arithmetic mean of x is given by

$$\bar{x} = \frac{f_1 x_1 + f_2 x_2 + \dots + f_n x_n}{f_1 + f_2 + \dots + f_n}$$

Or

$$\bar{x} = \frac{\sum fx}{\sum f}$$

Where $\sum f = f_1 + f_2 + \dots + f_n$ = sum of the frequenciesIf we denote $N = \sum f$, then $\bar{x} = \frac{\sum fx}{N}$ **Steps :** (1) Multiply the frequency of each row with the variable x and obtain the total $\sum fx$.(2) Divide the total $\sum fx$ by the total frequency.**Example 3:** Calculate arithmetic mean for the following data .

Age in years:	11	12	13	14	15	16	17
No.ofStudents:	7	10	16	12	8	11	5

Solution:

Age in years (x)	No.of Students (f)	fx
11	7	77
12	10	120
13	16	208
14	12	168
15	8	120
16	11	176
17	5	85
Total	$N = 69$	$\sum fx = 954$

$$\sum f = N = 69, \sum fx = 954$$

$$\bar{x} = \frac{\sum fx}{N} = 13.83 \text{ years.}$$

Example 4: Calculate the average bonus paid per member from the following data :

Bonus (in Rs.)	40	50	60	70	80	90	100
No. of persons	2	5	7	6	4	8	3

Solution :

Bonus (in Rs) x	No. of persons (f)	fx
40	2	80
50	5	250
60	7	420
70	6	420

80	4	320
90	8	720
100	3	300
Total	N= 35	$\Sigma fx = 2510$

Now, $N = 35$ and $\Sigma fx = 2510$

$$\bar{x} = \frac{\Sigma fx}{N} = \frac{2510}{35} = \text{Rs. } 71.71$$

(iii) **Grouped data** (class intervals and frequencies are given)

Formula : $\bar{x} = \frac{\Sigma fx}{N}$ where $N = \Sigma f$ and x = Mid-point of the class intervals

- Steps :** (1) Obtain the mid-point of each class interval
 (Mid-point = $\frac{\text{lower limit} + \text{upper limit}}{2}$)
 (2) Multiply these mid-points by the respective frequency of each class interval and obtain the total Σfx .
 (3) Divide the total obtained by step (2) by the total frequency N .

Example 5:

Find the arithmetic mean for the following data representing marks of 60 students.

Marks	10-20	20-30	30-40	40-50	50-60	60-70	70-80
No.ofStudents	8	15	13	10	7	4	3

Solution:

Marks	No. of students (f)	Mid-point (x)	fx
10-20	8	15	120
20-30	15	25	375
30-40	13	35	455
40-50	10	45	450
50-60	7	55	385
60-70	4	65	260
70-80	3	75	225
Total	N = 60		$\Sigma fx = 2270$

$$\bar{x} = \frac{\Sigma fx}{N} = \frac{2270}{60} = 37.83$$

$\therefore$ The average marks are 37.83

Example 6:

Calculate the arithmetic mean of heights of 80 Students for the following data.

Heights in cms	130-134	135-139	140-144	145-149	150-154	155-159	160-164
No. of Students	7	11	15	21	16	6	4

Solution:

The class intervals are of inclusive type. We first make them continuous by finding the class boundaries. 0.5 is added to the upper class limits and 0.5 is subtracted from the lower class limits to obtain the class boundaries as

$$130-0.5= 129.5 \text{ and } 134+0.5=134.5$$

$$135-0.5= 134.5 \text{ and } 139+0.5= 139.5 \text{ and so on.}$$

Heights (in cms)	Class boundaries	No. of Students(f)	Mid-point(x)	fx
130-134	129.5-134.5	7	132	524
135-139	134.5-139.5	11	137	1507
140-144	139.5- 144.5	15	142	2130
145-149	144.5- 149.5	21	147	3087
150-154	149.5-154.5	16	152	2432
155- 159	154.5-159.5	6	157	942
160-164	159.5-164.5	4	162	648
Total		N = 80		$\Sigma fx = 1270$

$$N = \Sigma f = 80 \text{ and } \Sigma fx = 11270$$

$$\therefore \bar{X} = \frac{\Sigma fx}{N} = \frac{11270}{80} = 140.88$$

$\therefore$ The average height of students = 140.88 cm.

4.5 WEIGHTED ARITHMETIC MEAN

One of the limitations of the arithmetic mean discussed above is that it gives equal importance to all the items. But these are cases where the relative importance of the different items is not the same. In these cases, weights are assigned to different items according to their importance. The term ‘weight’ stands for the relative importance of the different items.

If $x_1, x_2, \dots, x_n$ are the n values of the variable x with the corresponding weights $w_1, w_2, \dots, w_n$, then the weighted mean is given by

$$x_w = \frac{w_1x_1 + w_2x_2 + \dots + w_nx_n}{w_1 + w_2 + \dots + w_n} = \frac{\sum wx}{\sum w}$$

Where x_w = Weighted Arithmetic Mean and $\sum w$ = Sum of the weights.

Steps:

- (1) Multiply the weights w by the variable x and obtain the total $\sum wx$.
- (2) Divide the total $\sum wx$ by the sum of the weights $\sum w$.

Example 7:

Find the weighted mean for the following data.

x	28	25	20	32	40
W	3	6	4	5	8

Solution:

x	W	xw
28	3	84
25	6	150
20	4	80
32	5	160
40	8	320
Total	26	794

$$\therefore x_w = \frac{\sum wx}{\sum w} = \frac{794}{26} = 30.54 \text{ units}$$

Example 8:

A candidate obtained the following marks in percentages in an examination. English 64, Mathematics 93, Economics 72, Accountancy 85 and Statistics 79. The weights of these subjects are 2, 3, 3, 4, 1 respectively. Find the candidate's weighted mean.

Solution:

Subject	Percentage of Marks (x)	Weights (w)	wx
English	64	2	128
Mathematics	93	3	279
Economics	72	3	216
Accountancy	85	4	340
Statistics	79	1	79
Total		13	1042

$$\sum w = 13, \sum wx = 1042$$

$$x_w = \frac{\Sigma wx}{\Sigma w} = \frac{1042}{13} = 80.15 \quad \therefore \text{Weighted Mean Marks} = 80.15$$

4.6 COMBINED MEAN

If there are two groups containing n_1 and n_2 observations with means $\bar{x}_1$ and $\bar{x}_2$ respectively, then the combined arithmetic mean of two groups is given by

$$\bar{x} = \frac{n_1 x_1 + n_2 x_2}{n_1 + n_2}$$

The above formula can be generalized for more than two groups. If $n_1, n_2, \dots, n_k$ are sizes of k groups with means $x_1, x_2, \dots, x_k$ respectively then the mean $\bar{x}$ of the combined group is given by

$$\bar{x} = \frac{n_1 x_1 + n_2 x_2 + \dots + n_k x_k}{n_1 + n_2 + \dots + n_k}$$

Example 9:

If average salaries of two groups of employees are Rs . 1500 and Rs . 2200 and there are 80 and 70 employees in the two groups. Find the mean of the combined group.

Solution:

Given :	Group I	Group II
	$n_1 = 80$	$n_2 = 70$
	$\bar{x}_1 = 1500$	$\bar{x}_2 = 2200$

$$\begin{aligned}
 \therefore \bar{x} &= \frac{n_1 x_1 + n_2 x_2}{n_1 + n_2} \\
 &= \frac{80(1500) + 70(2200)}{80 + 70} \\
 &= \frac{120000 + 154000}{150} \\
 &= \frac{274000}{150} \\
 &= 1826.67
 \end{aligned}$$

$\therefore$ The average monthly salary of the combined group of 150 employees is Rs . 1826 .67

Example 10:

The mean weight of a group of 50 workers is 58 kgs. The second group consists of 60 workers with average weight 62 kgs. and there are 90 workers in the third group with average weight 56 kgs. Find the average weight of the combined group.

Solution:

Given	Group I	Group II	Group III
No. of workers	$n_1 = 50$	$n_2 = 60$	$n_3 = 90$
Mean weight	$\bar{x}_1 = 58$	$\bar{x}_2 = 62$	$\bar{x}_3 = 56$

$$\begin{aligned}
 \therefore \bar{x} &= \frac{n_1 \bar{x}_1 + n_2 \bar{x}_2 + n_3 \bar{x}_3}{n_1 + n_2 + n_3} \\
 &= \frac{50(58) + 60(62) + 90(56)}{50 + 60 + 90} \\
 &= \frac{11660}{200} = 58.3
 \end{aligned}$$

$\therefore$ The average weight of the combined group of 200 workers is 58.3 kgs

Merits of Arithmetic Mean .

- (1) It is rigidly defined .
- (2) It is easy to understand and easy to calculate .
- (3) It is based on each and every observation of the series .
- (4) It is capable for further mathematical treatment .
- (5) It is least affected by sampling fluctuations.

Demerits of Arithmetic Mean .

- (1) It is very much affected by extreme observations.
- (2) It can not be used in case of open end classes.
- (3) It can not be determined by inspection nor it can be located graphically.
- (4) It can not be obtained if a single observation is missing .
- (5) It is a value which may not be present in the data .

Exercise 5.1

1. Find the arithmetic mean for the following sets of observations.

- (i) 48, 55, 83, 65, 38, 74, 58
- (ii) 154, 138, 165, 172, 160, 145, 157, 185
- (iii) 2254, 2357, 2241, 2012, 2125

Ans: (i) $\bar{x} = 60.14$ (ii) $\bar{x} = 159.5$ (iii) $\bar{x} = 2197.8$

2. Calculate the mean for the following distribution:

$x:$	15	17	19	21	23	25
$f:$	6	11	8	15	5	4

Ans: $\bar{x} = 19.57$

3. Find the mean for the following data.

Size of Shoe:	5	6	7	8	9	10
No. of pairs	22	35	28	42	15	12

Ans: $\bar{x} = 7.19$

4. Calculate the arithmetic mean for the following data

Daily Wages in Rs:	40-80	80-20	120-60	160-00	200-40	240-80	280-20
No. of workers	8	15	13	19	12	14	10

Ans: $\bar{x}=181.32$

5. Calculate the mean for the following data representing monthly salary of a group of employees.

Salary in Rs:	500-1000	1000-1500	1500-2000	2000-2500	2500-3000	3000-3500
No. of persons:	7	11	13	8	5	4

Ans: $\bar{x}=1802.08$

6. Find the mean for the following data.

Weekly wages in Rs:	50-99	100-149	150-199	200-249	250-299	300-349
No. of workers	12	17	15	20	8	7

Ans: $\bar{x}=184.63$

7. Find the arithmetic mean for the following data

Age in yrs:	10-19	20-29	30-39	40-49	50-59	60-69	70-79
No. of persons:	6	12	19	23	15	8	6

Ans: $\bar{x}=43.15$

8. Find the weighted mean for the following data.

x:	25	22	15	30	18
w:	6	4	3	8	2

Ans: $\bar{x}_w=24.30$

9. A Student Scores 44 in Test I, 32 in Test II, 27 in Test III and 38 in Test IV, These are to be weighted with weights 4,3,3 and 4 respectively. Find the average score.

Ans: $\bar{x}_w=36.07$

10. The average daily wages for 120 workers in a factory are Rs. 78. The average wage for 80 male workers out of them is Rs.92 Find the average wage for the remaining female workers.

Ans: Rs.50

11. There are three groups in a class of 200 Students. The first group contains 80 Students with average marks 65, the second group

consists of 70 Students with average marks 74. Find the average marks of the Students from the third group if the average for the entire class is 68.

Ans: 64.4

12. Find the combined arithmetic mean for the following data:

	Group I	Group II	Group III
Mean	125	141	131
No.of observations.	150	100	50

Ans: 131.67

4.7 MEDIAN

The median by definition refers to the middle value in a distribution. Median is the value of the variable which divides the distribution into two equal parts. The 50% observations lie below the value of the median and 50% observations lie above it. Median is called a positional average. Median is denoted by M.

(i) For Simple data or ungrouped data:

Median is defined as the value of the middle item of a series when the observations have been arranged in ascending or descending order of magnitude.

Steps: 1) Arrange the data in ascending or descending order of magnitude. (Both arrangements would give the same answer).

2) When n is odd:

Median = value of $\left(\frac{n+1}{2}\right)^{\text{th}}$ term.

3) When n is even:

Median = Arithmetic Mean of value of $\left(\frac{n}{2}\right)$ & $\left(\frac{n}{2} + 1\right)^{\text{th}}$ terms.

i.e, adding the two middle values and divided by two.
where n = number of observations.

Example 11

Find the median for the following set of observations
65, 38, 79, 85, 54, 47, 72

Solution: Arrange the values in ascending order:

38, 47, 54, 65, 72, 79, 85

n = 7 (odd number)

The middle observation is 65

$\therefore$ Median = 65 units.

using formula:

$$\text{Median} = \text{value of } \left(\frac{n+1}{2} \right)^{\text{th}} \text{ term.}$$

$$= \text{value of } \left(\frac{7+1}{2} \right)^{\text{th}} \text{ term.}$$

$$= 4^{\text{th}} \text{ term}$$

$$\text{Median} = 65$$

Example 12: Find the median for the following data.

25, 98, 67, 18, 45, 83, 76, 35

Solution: Arrange the values in ascending order

18, 25, 35, 45, 67, 76, 83, 98

$n = 8$ (even number)

The pair 45, 67 can be considered as the middle pair.

$$\therefore \text{Median} = \text{A.M of the pair} = \frac{45+67}{2}$$

$$\text{Median} = 56$$

Using formula :

$$\text{Median} = \text{A. M. of values of } \left(\frac{n}{2} \right)^{\text{th}}, \left(\frac{n+1}{2} \right)^{\text{th}} \text{ terms}$$

$$= \text{A. M. of } \left(\frac{8}{2} \right)^{\text{th}}, \left(\frac{8+1}{2} \right)^{\text{th}} \text{ terms}$$

$$= \text{A. M. of } (4, 4+1)^{\text{th}} \text{ terms}$$

$$= \text{A. M. of } (4, 5)^{\text{th}} \text{ terms}$$

$$= \text{A. M. of } 45 \text{ and } 67$$

$$= \frac{45 + 67}{2}$$

$$\text{Median} = 56$$

(ii) For ungrouped frequency distribution:

Steps:

- 1) Arrange the data in ascending or descending order of magnitude with respective frequencies .
- 2) Find the cumulative frequency (c. f) less than type .
- 3) Find $N/2$, N = total frequency.
- 4) See the c. f column either equal or greater than $N/2$ and determine the value of the variable corresponding to it . That gives the value of Median .

Example 13: From the following data find the value of Median.

Income in Rs:	1500	2500	3000	3500	4500
No. of persons	12	8	15	6	5

Solution:

Income (in Rs.) (x)	No. of persons (f)	c.f less than type
1500	12	12
2500	8	20
3000	15	35
3500	6	41
4500	5	46
	N= 46	

$$\frac{N}{2} = \frac{46}{2} = 23$$

In c.f column, we get 35 as the first cumulative frequency greater than 23.
The value of x corresponding to the c.f. 35 is 3000.

$\therefore$ Median = Rs. 3000

Example : 14

Find the median for the following data .

x :	25	20	30	40	35	15
f :	19	14	23	12	20	8

Solution : Arrange the values in ascending order .

x	f	c.f less than type
15	8	8
20	14	22
25	19	41
30	23	64
35	20	84
40	12	96
	N= 96	

$$\frac{N}{2} = \frac{96}{2} = 48$$

In c.f column, greater than 48 is 64.

$\therefore$ The corresponding value of x to the c.f is 30.

Median = 30 units.

(iii) For grouped data.

Steps:

- 1) Find the c.f less than type
- 2) Find $N/2$, N = total frequency.
- 3) See the c.f column just greater than $N/2$.
- 4) The Corresponding class interval is called the Median class.
- 5) To find Median, use the following formula.

$$M = l_1 + \left[\frac{N/2 - cf}{f} \right] (l_2 - l_1)$$

Where

l_1 = lower class-bounding of the median class.

l_2 = upper class-bounding of the median class.

f = frequency of the median class.

$c.f$ = cumulative frequency of the class interval preceding the median class.

Example: 15

The following data relate to the number of patients visiting a government hospital daily:

No. of patients :	1000-1200	1200-1400	1400-1600	1600-1800	1800-2000	2000-2200
No. of days :	15	21	24	18	12	10

Solution :

No. of patients	No. of days	c.f. less than type
1000-1200	15	15
1200-1400	21	36
1400-1600	24	60
1600-1800	18	78
1800-2000	12	90
2000-2200	10	100
	N=100	

$$N/2 = 100/2 = 50$$

See the c.f column greater than 50 is 60. The corresponding class interval 1400-1600 is the median class.

$$\therefore l_1 = 1400; l_2 = 1600; f = 24; c.f = 36$$

$$\therefore M = l_1 + \left[\frac{N/2 - cf}{f} \right] (l_2 - l_1) = 1400 + \left[\frac{50 - 36}{24} \right] (1600 - 1400)$$

$$= 1400 + [14] (200) = 1400 + 116.67$$

$M = 1516.67$ patients

Example 16: Find the median for the following distribution:

Weights (in kgs)	30-34	35-39	40-44	45-49	50-54	55-59	60-64
No. of students	4	7	15	21	18	10	5

Solution : Convert the given class intervals into exclusive type

Weights	No. of students	c.f. less than type
29.5-34.5	4	4
34.5-39.5	7	11
39.5-44.5	15	26
44.5-49.5	21	47
49.5-54.5	18	65
54.5-59.5	10	75
59.5-64.5	5	80
	N=80	

$$N/2 = 80/2 = 40$$

Median class is 44.5-49.5

$$l_1 = 44.5 ; l_2 = 49.5 , f = 21, c.f. = 26$$

$$\begin{aligned}
 M &= l_1 + \left[\frac{N/2 - cf}{f} \right] (l_2 - l_1) \\
 &= 44.5 + \left[\frac{40 - 26}{21} \right] (49.5 - 44.5) \\
 &= 44.5 + [14/21] (5) \\
 &= 44.5 + 3.33 \\
 M &= 47.83 \text{ units.}
 \end{aligned}$$

Merits of Median :

1. It is rigidly defined .
2. It is easy to understand and easy to calculate .
3. It is not affected by extreme observations as it is a positional average.
4. It can be calculated , even if the extreme values are not known .
5. It can be located by mere inspection and can also be located graphically.
6. It is the only average to be used while dealing with qualitative characteristics which can not be measured numerically .

Demerits of Median :

1. It is not a good representative in many cases.
2. It is not based on all observations.
3. It is not capable of further mathematical treatment .
4. It is affected by sampling fluctuations.

5. For continuous data case, the formula is obtained on the assumption of uniform distribution of frequencies over the class intervals. This assumption may not be true.

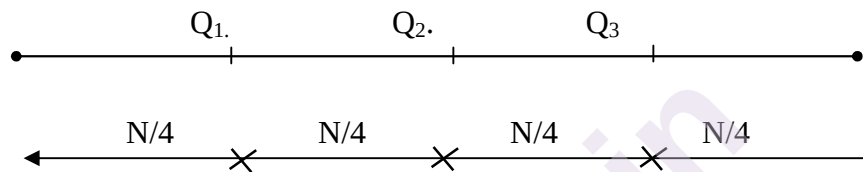
4.8 QUARTILES

When a distribution is divided into four equal parts, each point of division is called as quartile and each part is of 25% (one-fourth) of the total observations. There are three partition values such as Q_1 , Q_2 , and Q_3 .

Q_1 is called first quartile or lower quartile.

Q_3 is called third quartile or upper quartile.

Q_2 is called second quartile which coincides with median. Therefore Q_2 is nothing but Median.



The Steps involved for computing the quartiles is basically the same as that of computing median.

To find Q_1 : (First Quartile)

- Steps:**
- 1) Find the c.f less than type.
 - 2) Find $N/4$, N =total frequency.
 - 3) See the c.f. column just greater than $N/4$.
 - 4) The corresponding class interval is called the quartile class.
 - 5) To find Q_1 , use the following formula:

$$Q_1 = l_1 + \left[\frac{\frac{N}{4} - cf}{f} \right] (l_2 - l_1)$$

Where l_1 = lower limit of the quartile class

l_2 = upper limit of the quartile class.

f = frequency of the quartile class.

c.f = cumulative frequency of the class interval preceding the quartile class.

To find Q_3 : (Third Quartile)

- Steps:**
- 1) Find the c.f less than type.
 - 2) Find $3N/4$, N =total frequency.
 - 3) See the c.f column just greater than $3N/4$
 - 4) The corresponding class interval is called the quartile class.
 - 5) To find Q_3 , use the following formula.

$$Q_3 = l_1 + \left[\frac{\frac{3N}{4} - cf}{f} \right] (l_2 - l_1)$$

Where l_1 = lower limit of the quartile class.

l_2 = upper limit of the quartile class.

f = frequency of the quartile class.

$c.f.$ = cumulative frequency of the class interval preceding the quartile class.

To find Q_2 (Second Quartile)

Q_2 = Median (Discussed above)

Example 17: Find the three quartiles for the following data.

Commission (in Rs):	100-120	120-140	140-160	160-180	180-200	200-220	220-240
No. of Salesmen:	13	38	45	56	27	17	12

Solution:

Commission (in Rs.)	No. of Salesmen	c.f. less than type
100-120	13	13
120-140	38	51
140-160	45	96
160-180	56	152
180-200	27	179
200-220	17	196
220-240	12	208
	n = 208	

To find Q_1 :

$$\frac{N}{4} = \frac{208}{4} = 52$$

See the c.f. column just greater than 52.

Quartile class: 140-160

$$l_1 = 140, l_2 = 160, f = 45, c.f = 51$$

$$Q_1 = l_1 + \frac{[N/4 - c.f.](l_2 - l_1)}{f}$$

$$= 140 + \frac{[52 - 51](160 - 140)}{45}$$

$$= 140 + [1/45](20)$$

$$= 140 + 0.44$$

$$Q_1 = 140.44$$

To find Q_2 = Median :

$$N/2 = 208/2 = 104$$

See the c.f. column just greater than 104.

Median class = 160-180

$$l_1 = 160; l_2 = 180; f = 56; c.f = 96$$

$$\begin{aligned}
 M &= l_1 + \frac{[N/2 - c.f.]}{f} (l_2 - l_1) \\
 &= 160 + \frac{[104 - 96]}{56} (180 - 160) \\
 &= 160 + [8/56] (20) \\
 &= 160 + 2.86 \\
 Q_2 &= 162.86
 \end{aligned}$$

To find Q_3 :

$$3N/4 = 3(208/4) = 156$$

See the $c.f$ column just greater than 156.

Quartile class = 180-200

$l_1 = 180$; $l_2 = 200$; $f = 27$; $c.f = 152$

$$\begin{aligned}
 Q_3 &= l_1 + \frac{[3N/4 - c.f.]}{f} (l_2 - l_1) \\
 &= 180 + \frac{[156 - 152]}{27} (200 - 180) \\
 &= 180 + [4/27] (20) \\
 &= 180 + 2.96
 \end{aligned}$$

$$Q_3 = 182.96$$

$$\begin{aligned}
 \therefore Q_1 &= 140.44 \\
 Q_2 &= M = 162.86 \\
 Q_3 &= 182.96
 \end{aligned}$$

Graphical location of Quartiles :

The median and the quartiles obtained graphically from the $c.f$ curve less than type curve as follows:

First draw a cumulative frequency curve less than type.

To locate $Q_2 = \text{Median}$:

Locate $n/2$ on the Y – axis and from it draw a perpendicular on the $c.f$ curve. From the point where it meets the $c.f$ curve draw another perpendicular on the X- axis and the point where it meets the X- axis is called the median .

To locate Q_1 and Q_3 :

Locate $n/4$ and $3N/4$ on the Y- axis and proceed as above to obtain Q_1 and Q_3 respectively on the X- axis.

Note that the values obtained from the graph are approximate figures.
They do not represent exact figures.

Example 18:

Draw a 'less than' cumulative frequency curve for the following data and hence locate the three quartiles graphically.

Age in yrs.	0-10	10-20	20-30	30-40	40-50	50-60	60-70	70-80
No. of persons	15	13	25	22	25	10	5	5

Solution :

Age (in yrs.)	No. of persons.	c.f. less than type
0-10	15	15
10-20	13	28
20-30	25	53
30-40	22	75
40-50	25	100
50-60	10	110
60-70	5	115
70-80	5	120
	N=120	

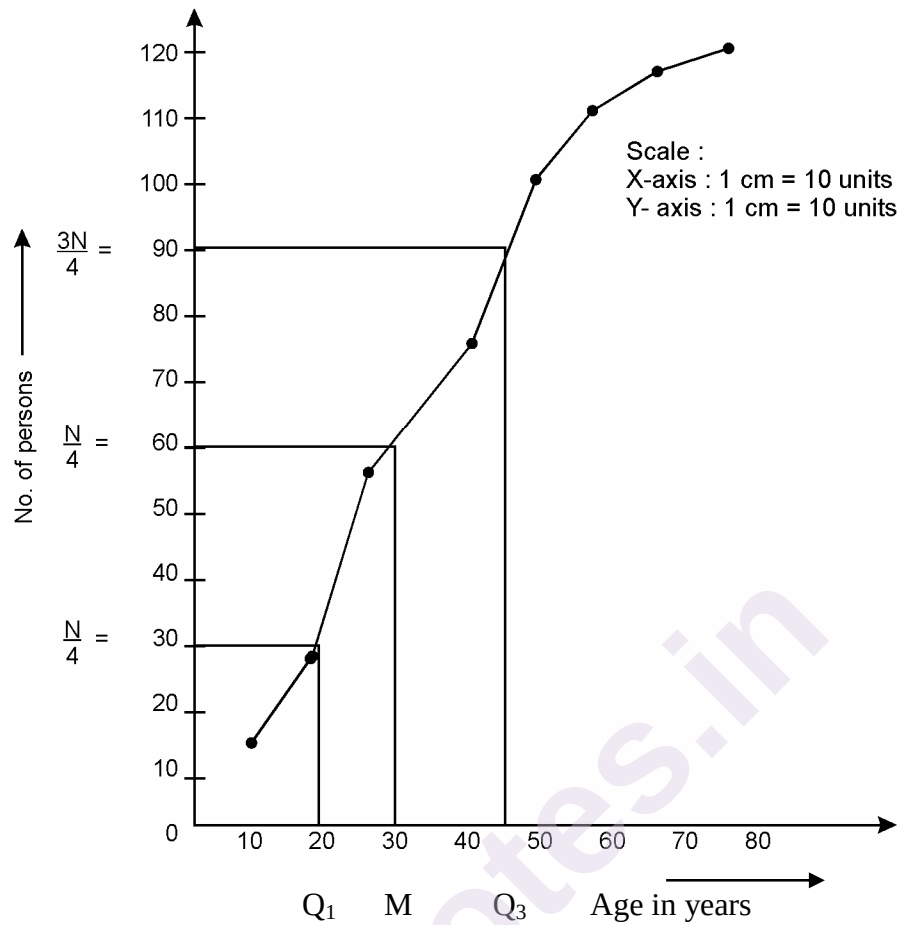
X-axis (upper limit)	Y-axis c.f. < type
10	15
20	28
30	53
40	75
50	100
60	110
70	115
80	120

$$N/4 = 120/4 = 30$$

$$N/2 = 120/2 = 60$$

$$3N/4 = 3 (120/4)$$

$$3N/4 = 90$$



From the graph : $Q_1 = 21$ approximately
 $M = 33$ approximately
 $Q_3 = 46$ approximately

Exercise 5.2

1. Find the median for the following data

(i) 50, 28, 35, 98, 75, 44, 58

(ii) 16, 22, 10, 12, 30, 37, 28, 40, 15, 20

Ans: (i) $M=50$ (ii) $M= 21$

2. Calculate the median for the following distribution :

$x:$	25	30	35	40	45	50	55
$f:$	8	14	23	28	10	6	4

Ans: $M= 40$

3. Find the median for the following data.

Weekly Wages (in Rs)	50-100	100-150	150-200	200-250	250-300	300-350
No. of Workers:	20	18	32	18	12	15

Ans: $M= 180.47$

4. Find the median height for the following distribution:

Height in cms:	150-154	154-158	158-162	162-166	166-170	170-174
No. of Students:	6	15	23	20	12	10

Ans $M=161.83$

5. Find the three quartiles for the following distribution:

Income in Rs:	0-500	500-1000	1000-1500	1500-2000	2000-2500	2500-3000	3000-3500
No. of families	3	5	10	18	15	9	6

Ans: $Q_1=1425$, $M=1916.67$, $Q_3=2450$

6. Find the median and the two quartiles for the following data. Also locate them graphically.

Rainfall (in cms):	15-20	20-25	25-30	30-35	35-40	40-45	45-50
No. of years:	3	7	10	16	12	8	4

Ans: $Q_1=27.5$, $M=33.13$, $Q_3=38.75$

6. The following is the data representing profits in thousands of Rs. of some companies. Find the quartiles and hence locate them graphically.

Profit:	70-80	80-90	90-100	100-110	110-120	120-130
No. of Companies	11	15	19	23	22	10

Ans: $Q_1=89.33$, $M=102.17$, $Q_3=113.18$

4.9 MODE

(i) For Raw data:

Mode is the value which occurs most frequently, in a set of observations. It is a value which is repeated maximum number of times and is denoted by Z .

Example 19: Find mode for the following data.

64, 38, 35, 68, 35, 94, 42, 35, 52, 35

Solution:

As the number 35 is repeated maximum number of times that is 4 times.

$\therefore$ Mode = 35 units.

(ii) For ungrouped frequency distribution:

Mode is the value of the variable corresponding to the highest frequency.

Example 20: Calculate the mode for the following data.

Size of Shoe:	5	6	7	8	9	10
No.of Pairs:	38	43	48	56	25	22

Solution: Here the highest frequency is 56 against the size 8.

$\therefore$ Modal size = 7

(iii) For Grouped data:

In a Continuous distribution first the modal class is determined. The class interval corresponding to the highest frequency is called modal class.

Formula:

$$Z = l_1 + \left[\frac{f_1 - f_0}{(f_1 - f_0) + (f_1 - f_2)} \right] (l_2 - l_1)$$

Where l_1 = lower boundary of the modal class

l_2 = upper boundary of the modal class

f_1 = frequency of the modal class

f_0 = frequency of the class interval immediately preceding the modal class.

f_2 = frequency of the class interval immediately succeeding the modal class.

Example 21

Find the mode for the following data.

Daily wages(in Rs):	20-40	40-60	60-80	80-100	100-120	120-140	140-160
No.of employees:	21	28	35	40	24	18	10

Solution:

Daily wages (in Rs.)	No. of employees
20- 40	21
40-60	28
60-80	35 f_0
80- 100	40 f_1
100-120	24 f_2
120-140	18
140- 160	10

The maximum frequency is 40.

$\therefore$ Modal class is 80-100.

$\therefore l_1 = 80; l_2 = 100; f_0 = 35; f_1 = 40, f_2 = 24$

$$\begin{aligned}
 Z &= l_1 + \left[\frac{f_1 - f_0}{(f_1 - f_0) + (f_1 - f_2)} \right] (l_2 - l_1) \\
 &= 80 + \left[\frac{40 - 35}{(40 - 35) + (40 - 24)} \right] (100 - 80) \\
 &= 80 + \left[\frac{5}{5 + 16} \right] \times 20 = 80 + 4.76
 \end{aligned}$$

$$Z = 84.76$$

Modal daily wages = Rs. 84.76

Example: 21

Find the mode for the following data.

Marks:	10-19	20-29	30-39	40-49	50-59	60-69	70-79
No of Students:	8	22	31	44	15	13	10

Solution:

Marks	Class boundaries	No. of Students
10-19	9.5-19.5	8
20-29	19.5-29.5	22
30-39	29.5-39.5	31 f_0
40-49	39.5-49.5	44 f_1
50-59	49.5-59.5	15 f_2
60-69	59.5-69.5	13
70-79	69.5-79.5	10

The maximum frequency is 44.

$\therefore$ The modal class is 39.5-49.5

$$l_1 = 39.5 \quad f_0 = 31$$

$$l_2 = 49.5 \quad f_1 = 44, \quad f_2 = 15$$

$$\begin{aligned}
 Z &= l_1 + \left[\frac{f_1 - f_0}{(f_1 - f_0) + (f_1 - f_2)} \right] (l_2 - l_1) \\
 &= 39.5 + \left[\frac{44 - 31}{(44 - 31) + (44 - 15)} \right] (49.5 - 39.5) = 39.5 + \left[\frac{13}{13 + 29} \right] \times 10 \\
 &= 39.5 + 3.095 \\
 &= 42.595
 \end{aligned}$$

$$Z = 42.60$$

Modal Marks = 42.60

4.10 GRAPHICAL LOCATION OF MODE

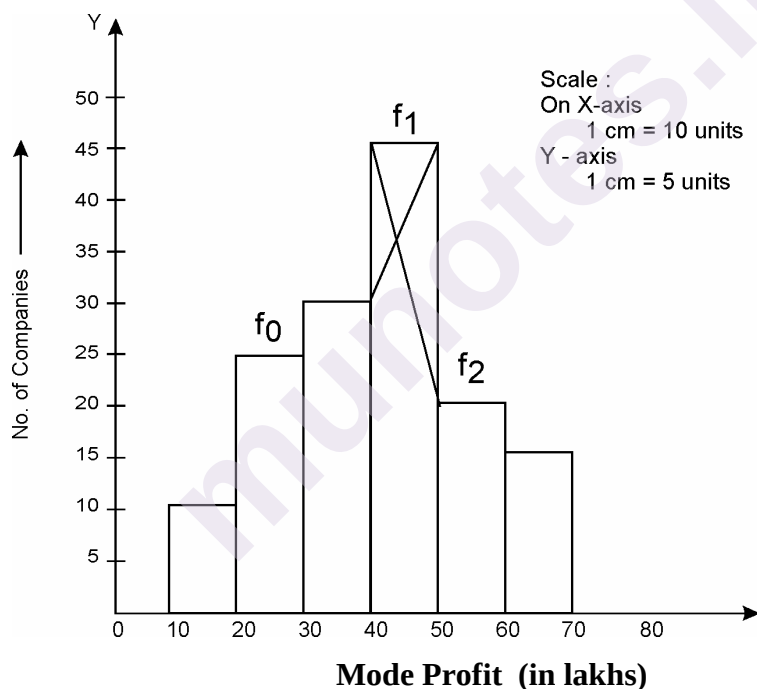
Mode can be obtained from a histogram as follows . The method can be applied for class intervals of equal length having a unique maximum frequency .

In the histogram, the rectangle with the maximum height represents the modal class. The right upper corner of this rectangle is connected with right upper corner of the preceding rectangle by a straight line. Similarly, the left upper corner of the maximum height of the rectangle is connected with left upper corner of the succeeding rectangle by a straight line. These two straight lines are intersecting at a point. From the point of intersection, a perpendicular is drawn on x-axis, the foot of which gives the value of Mode.

Example 23:

Draw a histogram for the following data and hence locate mode graphically.

Profit (in lakhs):	10-20	20-30	30-40	40-50	50-60	60-70
No. of Companies	12	25	31	45	20	15



Modal profits = Rs. 44 lakhs.

Merits and Demerits of Mode .

Merits :

1. It is easy to understand and easy to calculate.
2. It is unaffected by the presence of extreme values.
3. It can be obtained graphically from a histogram.
4. It can be calculated from frequency distribution with open-end classes.
5. It is not necessary to know all the items. Only the point of maximum concentration is required.

Demerits :

- (1) It is not rigidly defined .
- (2) It is not based on all observations.
- (3) It is affected by sampling fluctuations.
- (4) It is not suitable for further mathematical treatment .

Exercise 4.3

(1) Find the mode for the following data .

(i) 85, 40, 55, 35, 42, 67, 75, 63, 35, 10, 35

(ii) 250, 300, 450, 300, 290, 410, 350, 300

Ans: (i) $Z = 35$ (ii) $Z = 300$

(2) Find the mode for the following data.

$x:$	15	18	20	22	24
$f:$	8	6	13	18	10

Ans: $Z = 22$

(3) Calculate the modal wages for the following distribution.

Wages in Rs:	50-65	65-80	80-95	95-110	110-125	125-140
No. of workers	4	7	15	28	11	5

Ans: $Z = 101.5$

(4) Find mode for the following data.

Life in hrs:	200-400	400-600	600-800	800-1000	1000-1200	1200-1400	1400-1600
No. of bulbs:	15	22	40	65	38	26	10

Ans: $z = 896.15$

(5) Find the mode for the following data and hence locate mode graphically.

Salary in Rs:	1000-1500	1500-2000	2000-2500	2500-3000	3000-3500	3500-4000
No. of persons:	23	35	50	30	18	12

Ans: $z = 2214.29$

(6) Find the mode for the following data, also locate mode graphically.

Class Interval :	100-150	150-200	200-250	250-300	300-350	350-400	400-450
Frequency:	15	29	31	42	56	35	17

Ans: $z = 320$

(7) Find Mean Median and mode for the following data.

Income (in Rs.):	1000-2000	2000-3000	3000-4000	4000-5000	5000-6000	6000-7000
No. of employees:	15	26	45	32	12	10

Ans: $\bar{x}$ = 3714.29; M = 3644.44; Z = 3593.75

(8) Prove that median lies between mean and mode from the following data:

Weekly (wages in Rs.)	100-200	200-300	300-400	400-500	500-600	600-700	700-800
No. of workers.	18	23	32	41	35	20	15

Ans: $\bar{x}$ = 443.48; M = 446.34; Z = 460

(9) Calculate mean median and mode for the following data.

Height (in.cms)	125-130	130-135	135-140	140-145	145-150	150-155
No. of children	6	13	19	26	15	7

Ans: $\bar{x}$ = 140.52; M = 140.96; Z = 141.94

(10) Find arithmetic mean median and mode for the following data.

Age in yrs:	10-20	20-30	30-40	40-50	50-60	60-70	70-80
No. of persons:	12	18	23	28	20	16	7

Ans: $\bar{x}$ = 43.23; M = 43.21; Z = 43.85



UNIT III

Unit-5

MEASURES OF DISPERSION

Unit Structures :

- 5.0 Objectives
- 5.1 Introduction
- 5.2 Range
- 5.3 Quartile Deviation
- 5.4 Mean Deviation
- 5.5 Standard Deviation

5.0 OBJECTIVES

After reading this chapter you will be able to:

- Compute different types of deviations like; Range, Quartile Deviation, Standard Deviation and Mean Deviation.
- Compute relative measures of deviation like; coefficient of Range/Quartile Deviation/Mean Deviation/Variation.
- Compute combined standard deviation.

5.1 INTRODUCTION

Why Do We Need to Study the Measure of Dispersion?

The common averages or measures of central tendency indicate the general magnitude of the data and give us a single value which represents the data, but they do not tell us the degree of *spread out* or the extent of variability in individual items in a distribution. This can be known by certain other measures called *measures of dispersion*.

We will discuss the most commonly used statistical measures showing the degree and the characters of variations in data.

Dispersion in particular helps in finding out the variability of the data i.e., the extent of dispersal or scatter of individual items in a given distribution. Such dispersal may be known with reference to the central values or the common averages such as mean, median and mode or a standard value; or with reference to other values in the distribution. The need for such a measure arises because mean or even median and mode may be the same in two or more distribution but the composition of individual items in the series may vary widely. We give an example to illustrate this.

Region	Rainfall (cms)				
East	106	106	106	106	106
West	90	98	106	114	122
North	120	130	84	126	70

In the above example note that the average rainfall in all the three regions is the same i.e. 106 cms however it can be easily seen that it would be wrong to conclude that all the three regions have the same climatic pattern. This is so because in East all the values are equal to the average; whereas in the western region they are centered around their mean, and in north they are widely scattered. It may thus be misleading to describe a situation simply with the aid of an average value.

In measuring dispersion, it is necessary to know the **amount** of variation and the **degree** of variation. The former is designated as absolute measures of dispersion and expressed in the denomination of original variates such as inches, cms, tons, kilograms etc while the latter is designated as related measures of dispersion. We use the absolute measures of dispersion when we compare two sets of data with the same units and the same average type. If the two sets of data do not have the same units then we cannot use the absolute measures and we use the relative measures of dispersion.

Absolute measures can be divided into **positional** measures based on some items of the series such as (I) Range, (ii) Quartile deviation or semi-inter quartile range and those which are based on all items in series such as (I) Mean deviation, (ii) Standard deviation. The relative measure in each of the above cases are called the coefficients of the respective measures. For purposes of comparison between two or more series with varying size or number of items, varying central values or units of calculation, only relative measures can be used.

5.2 RANGE

Range is the simplest measure of dispersion.

When the data are arranged in an array the difference between the largest and the smallest values in the group is called the **Range**.

Symbolically: Absolute Range = $L - S$, [where L is the largest value and S is the smallest value]

$$\text{Relative Range} = \frac{\text{Absolute Range}}{\text{sum of the two extremes}}$$

Amongst all the methods of studying dispersion range is the simplest to calculate and to understand but it is not used generally because of the following reasons:

- 1) Since it is based on the smallest and the largest values of the distribution, it is unduly influenced by two unusual values at either end. On this account, range is usually not used to describe a sample having one or a few unusual values at one or the other end.
- 2) It is not affected by the values of various items comprised in the distribution. Thus, it cannot give any information about the general characters of the distribution within the two extreme observations.

For example, let us consider the following three series:

Series: A

6 46 46 46 46 46 46 46

Series: B

6 6 6 6 46 46 46 46

Series: C

6 10 15 25 30 32 40 46

It can be noted that in all three series the range is the same, i.e. 40, however the distributions are not alike: the averages in each case is also quite different. It is because range is not sensitive to the values of individual items included in the distribution. It thus cannot be depended upon to give any guidance for determining the dispersion of the values within a distribution.

5.3 QUARTILE DEVIATION

The dependence of the range on extreme items can be avoided by adopting this measure. Quartiles together with the median are the points that divide the whole series of observations into approximately four equal parts so that quartile measures give a rough idea of the distribution on either side of the average.

Since under most circumstances, the central half of the distribution tends to be fairly typical, quartile range $Q_3 - Q_1$ affords a convenient measure of absolute variation. The lowest quarter of the data (upto Q_1) and the highest quarter (beyond Q_3) are ignored. The remainder, the middle half of the data above Q_1 and below Q_3 or $(Q_3 - Q_1)$ are considered. This, when divided by 2, gives the semi-interquartile range or quartile deviation.

$$\text{Q.D.} = \frac{Q_3 - Q_1}{2}$$

In a symmetrical distribution, when Q_3 (75%) plus Q_1 (25%) is halved, the value reached would give Median, i.e., the mid-point of 75% and 25%. Thus, with semi-interquartile range 50% of data is distributed around the median. It gives the expected range between which 50% of all data should lie. The Quartile Deviation is an absolute measure of dispersion. The corresponding relative measure of dispersion is

$$\text{Coefficient of Q.D.} = \frac{Q_3 - Q_1}{Q_3 + Q_1}$$

For the following Frequency Distribution we show how to calculate the Quartile Deviation and the coefficient of Quartile Deviation.

x	f	c.f.
10	6	6
15	17	23
20	29	52
25	38	90

30	25	115
35	14	129
40	9	138
90	1	139

It may be noted that:

Corresponding to Q_1 , the c.f. = $\frac{n+1}{4}$

Corresponding to Q_2 , the c.f. = $\frac{2(n+1)}{4}$

Corresponding to Q_3 , the c.f. = $\frac{3(n+1)}{4}$

Thus from the given data :

Q_3 = Value of the variates corresponding to c.f. $\frac{3 \times 140}{4} = 105$ which corresponds to 30

Q_2 = Value of the variates corresponding to c.f. $\frac{2 \times 140}{4} = 70$ which corresponds to 25

Q_1 = Value of the variates corresponding to c.f. $\frac{140}{4} = 35$, which corresponds to 20

Quartile deviation or Q.D. = $\frac{Q_3 - Q_1}{2} = \frac{30 - 20}{2} = 5$

And Co-efficient of Quartile Deviation

$$= \frac{Q_3 - Q_1}{Q_3 + Q_1} = \frac{30 - 20}{30 + 20} = \frac{10}{50} = 0.2$$

Example

Find the Quartile Deviation of the daily wages (in Rs.) of 11 workers given as follows: 125, 75, 80, 50, 60, 40, 50, 100, 85, 90, 45.

Solution: Arranging the data in ascending order we have the wages of the 11 workers as follows:

40, 45, 50, 50, 60, 75, 80, 85, 90, 100, 125

Since the number of observations is odd (11), the 1st Quartile is given by:

$$Q_1 = (11 + 1)/4 = 3^{\text{rd}} \text{ observation} = 50.$$

... (1)

$$Q_3 = 3(11 + 1)/4 = 9^{\text{th}} \text{ observation} = 90.$$

... (2)

$$Q.D. = \frac{Q_3 - Q_1}{2} = (90 - 50)/2 = 20 \quad \dots$$

from (1) and (2)

Example

The following data gives the weight of 60 students in a class. Find the range of the weights of central 50% students.

Weight in kg	30 – 35	35 – 40	40 – 45	45 – 50	50 – 55	55 – 60
No. of students	4	16	12	8	10	5

Solution:

To find range of the weight's of central 50 % students means to find the inter quartile range. For that we require Q_1 and Q_3 . The column of less than cf is introduced as follows:

Weight in kg	30 – 35	35 – 40 ↑	40 – 45	45 – 50	50 – 55 ↑	55 – 60
No. of students	4	16	12	8	10	6
cf	4	20	32	40	50	56

 Q_1 class Q_3 class

To find Q_1 : $N = 56$. Thus $m = N/4 = 14$.

The cf just greater than 14 is 20, so 35 – 40 is the 1st quartile class and $l_1 = 35$, $l_2 = 40$, $i = 40 - 35 = 5$, $f = 16$ and $pcf = 4$.

$$\therefore Q_1 = l_1 + \left[\frac{m - pcf}{f} \times i \right] = 35 + \left[\frac{14 - 4}{16} \times 5 \right] = 35 + 3.125 = 38.125$$

To find Q_3 : $N = 56$. Thus $m = 3N/4 = 42$.

The cf just greater than 42 is 50, so 50 – 55 is the 3rd quartile class and $l_1 = 50$, $l_2 = 55$, $i = 55 - 50 = 5$, $f = 10$ and $pcf = 40$.

$$\therefore Q_3 = l_1 + \left[\frac{m - pcf}{f} \times i \right] = 50 + \left[\frac{42 - 40}{10} \times 5 \right] = 50 + 1 = 51$$

$$\therefore \text{inter quartile range} = Q_3 - Q_1 = 51 - 38.125 = 12.875 \text{ kg}$$

Thus, the range of weight for the central 50% students = 12.875kg

Example

Find the semi – inter quartile range and its coefficient for the following data:

Size of shoe	0	1	2	3	4	5	6	7	8	9	10
No. of boys	7	10	15	11	18	10	16	5	12	6	2

Solution: The less than $c f$ are computed and the table is completed as follows:

Size of shoe	0	1	2	3	4	5	6	7	8	9	10
No. of boys	7	10	15	11	18	10	16	5	12	6	2
$c f$	7	17	32	43	61	71	87	92	104	110	112

Here $N = 112$.

To find Q_1 : $m = N/4 = 28$.

The first $c f$ just greater than 28 is 32, so the 1st quartile is $Q_1 = 2$

To find Q_3 : $m = 3N/4 = 84$.

The first $c f$ just greater than 84 is 87, so the 3rd quartile is $Q_3 = 6$

$$\therefore \text{the semi inter quartile range i.e. Q.D.} = \frac{Q_3 - Q_1}{2} = \frac{6 - 2}{2} = 2$$

$$\text{Coefficient of Q.D.} = \frac{Q_3 - Q_1}{Q_3 + Q_1} = \frac{6 - 2}{6 + 2} = 0.5$$

5.4 MEAN DEVIATION

Both the range and quartile deviation do not show the scatterness around an average and as such do not give a clear idea of the dispersion of the distribution, these measures also exclude some data and consequently do not give a complete picture based on the entire data.. To study the deviations around the average we now introduce two more measures of dispersion the **mean deviation** and the **standard deviation**. The mean deviation is also called as the **average deviation**.

The essence of average deviation lies in the concept of dispersion, which is the average amount of scatter of individual items from either the mean or the median ignoring the algebraic signs.

This measure takes into account the whole data. When it is calculated by averaging the deviations of the individual items from their arithmetic mean, taking all deviation to be positive, the measures is called mean deviation. It may be pointed that we are concerned with the distance of the individual items from their averages and not with their position, which may be either above or below the average.

Suppose that we have sample of six observations 0,2,3,4,4, and 5. The mean of these observations is: $\frac{0+2+3+4+4+5}{6} = 3$

Now we obtain the deviation of each observation from the mean of 3. For the first observation, for example, this gives a deviation of $0 - 3 = -3$. Similarly, the other observations are : -1, 0, + 1, + 1 and + 2. It may seem that a good way to measure dispersion would be to take the

mean of six deviations. But this gives 0, since, $-3-1+0+1+1+2$. For practice in using symbols, we give symbolic definition of the mean deviations. To avoid the cancelling off of the positive deviations with the negative deviations we take only the magnitude (absolute value) of each of these deviations .

$$\text{So in the above case the M.D.} = \frac{\sum |x - x_i|}{n} = \frac{3+1+0+1+1+2}{6} = \frac{8}{6}$$

For a frequency distribution the calculation of M.D. is done as follows

$$\text{M.D.} = \frac{\sum f |d|}{\sum f} \text{ where } f \text{ is the frequency of the observation and } d = |x_i - \bar{x}| \text{ where } \bar{x} \text{ is the mean.}$$

$$\text{The coefficient of M.D.} = \frac{\text{M.D.}}{\text{Mean}}$$

We now take an example to calculate the Mean deviation and the coefficient of mean deviation for a frequency distribution.

Example

Find mean deviation from mean for the following distribution:

X	10	15	20	25	30	25	40	90
f	6	17	29	38	25	14	9	1

Solution

Calculation of mean deviation from mean :

x	f	f . x	dx = $ x - 25.43 $	f . dx
10	6	60	15.43	92.58
15	17	255	10.43	177.31
20	29	580	5.43	157.47
25	38	950	0.43	16.34
30	25	750	4.57	114.25
35	14	490	9.57	133.98
40	9	360	14.57	131.13
90	1	90	64.57	64.57
139	3535			$\sum f \cdot dx = 887.63$

$$\text{In the above example the mean is } = \bar{x} = \frac{\sum f \cdot x}{n} = \frac{3535}{139} = 25.43 \text{ } n = \sum f$$

$$\text{And mean Deviation from Mean} = \frac{\sum (f \cdot |dx|)}{n} = \frac{887.63}{139} = 6.386 \text{ } n = \sum f$$

And Coefficient of mean deviation is $= \frac{6.386}{25.43} = 0.25$

The mean deviation tells us that some values were above the mean and some below. On the average, the deviation of all values combined was 6.386. Co-efficient of mean deviation or 0.25.

In case of continuous frequency distribution we take the mid value of each class as the value of x and proceed to do the calculations in the same manner as above.

Example

Find the mean deviation from mean and its coefficient for the following data giving the rainfall in cm in different areas in Maharashtra: 105, 90, 102, 67, 71, 52, 80, 30, 70 and 48.

Solution: Since we have to compute M.D. from mean, we first prepare the table for finding mean and then introduce columns of absolute deviations from the mean.

$$\bar{x} = \frac{105 + 90 + 102 + 67 + 71 + 52 + 80 + 30 + 70 + 48}{10} = \frac{715}{10} = 71.5$$

x	$d = x - \bar{x} $
105	33.5
90	18.5
102	30.5
67	4.5
71	0.5
52	19.5
80	8.5
30	41.5
70	1.5
48	23.5
$\Sigma x = 715$	$\Sigma d = 182$

Now,

$$\text{M.D. from mean} = \frac{\Sigma d}{n}$$

$$\delta \bar{x} = \frac{182}{10}$$

$$\therefore \delta \bar{x} = 18.2$$

Coefficient of M.D. from mean

$$= \frac{\delta \bar{x}}{\bar{x}} = \frac{18.2}{71.5}$$

Example

The marks obtained by 10 students in a test are given below. Find the M.D. from median and its relative measure.

Marks: 15 10 10 03 06 04 11 17 13 05

Solution: The marks of 10 students are arranged in ascending order and its median is found. The column of absolute deviations from median is introduced and its sum is computed. Using the formula mentioned above, M.D. from median and its coefficient is calculated.

x	$d = x - M $
03	5
04	4
05	3
06	2
10	2
10	2
11	3
13	5
15	7
17	9
Total	$\Sigma d = 42$

Since $N = 10$,

Median = A.M. of 5th and 6th observation

$$\therefore M = \frac{6+10}{2} = 8$$

$$\text{M.D. from median} = \frac{\Sigma d}{n} = \frac{42}{10}$$

$$\therefore \delta M = 4.2$$

$$\text{Coefficient of M.D.} = \frac{\delta M}{M} = \frac{4.2}{8} = 0.525$$

Example

On the Mumbai – Nashik highway the number of accidents per day in 6 months are given below. Find the mean deviation and coefficient of M.D.

No. of accidents	0	1	2	3	4	5	6	7	8	9	10
No. of days	26	32	41	12	22	10	05	01	06	15	10

No. of accidents	No. of days (f)	cf	$d = x - M $	$f.d$
0	26	26	2	52
1	32	58	1	32
2	41	99	0	0
3	12	111	1	12
4	22	133	2	44
5	10	143	3	30
6	05	148	4	20
7	01	149	5	05
8	06	155	6	36
9	15	170	7	105
10	10	180	8	80
Total	$N = 180$	-	-	$\Sigma fd = 416$

Solution:

Now, $N = 180 \therefore m = N/2 = 180/2 = 90$.

The cf just greater than 90 is 99. The corresponding observation is 2. $\therefore M = 2$

$$\text{M.D. from median} = \delta M = \frac{\Sigma fd}{N} = \frac{416}{180} = 2.31$$

$$\text{Coefficient of M.D.} = \frac{\delta M}{M} = \frac{2.31}{2} = 1.15$$

Example

The following data gives the wages of 200 workers in a factory with minimum wages Rs. 60 and maximum wages as Rs. 200. Find the mean deviation and compute its relative measure.

Wages less than	80	100	120	140	160	180	200
No of workers	30	45	77	98	128	172	200

Solution: The data is given with less than cf , we first convert them to frequencies then find the median and follow the steps to compute M.D. as mentioned above.

Wages in Rs	No. of workers (f)	cf	x	$d = x - 141.33 $	fd
60 – 80	30	30	70	71.33	2139.9
80 – 100	15	45	90	51.33	769.95
100 – 120	32	77	110	31.33	1002.56
120 – 140	21	98	130	11.33	237.93
140 – 160	30	128	150	8.67	260.1
160 – 180	44	172	170	28.67	1261.48
180 – 200	28	200	190	48.67	1362.76
Total	$N = 200$	-	-	-	$\Sigma fd = 7034.68$

Now, $N = 200 \therefore m = N/2 = 200/2 = 100$.

The median class is 140 – 160. $\therefore l_1 = 140, l_2 = 160, i = 160 - 140 = 20, f = 30$ and $pcf = 98$

$\therefore M$

=

$$l_1 + \left[\frac{m - pcf}{f} \times i \right] = 140 + \left[\frac{100 - 98}{30} \times 20 \right] = 140 + 1.33 = 141.33$$

$$\text{M.D. from median} = \delta M = \frac{\Sigma fd}{N} = \frac{7034.68}{200} = 35.1734$$

$$\text{Coefficient of M.D.} = \frac{\delta M}{M} = \frac{35.1734}{141.33} = 0.25$$

5.5 STANDARD DEVIATION

As we have seen range is unstable, quartile deviation excludes half the data arbitrarily and mean deviation neglects algebraic signs of the deviations, a measure of dispersion that does not suffer from any of these defects and is at the same time useful in statistic work is **standard deviation**. In 1893 Karl Pearson first introduced the concept. It is considered as one of the best measures of dispersion as it satisfies the requisites of a good measure of dispersion. The standard deviation measures the absolute dispersion or variability of a distribution. The greater the amount of variability or dispersion greater is the value of standard deviation. In common language a small value of standard deviation means greater uniformity of the data and homogeneity of the distribution. It is due to this reason that standard deviation is considered as a good indicator of the representativeness of the mean.

It is represented by σ (read as ‘sigma’); σ^2 i.e., the square of the standard deviation is called **variance**. Here, each deviation is squared.

The measure is calculated as the average of deviations from arithmetic mean. To avoid positive and negative signs, the deviations are squared. Further, squaring gives added weight to extreme measures, which is a desirable feature for some types of data. It is a square root of arithmetic mean of the squared deviations of individual items from their arithmetic mean.

The mean of squared deviation, i.e., the square of standard deviation is known as variance. Standard deviation is one of the most important measures of variation used in Statistics. Let us see how to compute the measure in different situation.

$$\text{s.d} = \sigma = \sqrt{\frac{\sum(x - \bar{x})^2}{n}}$$

For a frequency distribution standard deviation is

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

We will now take an example of a frequency distribution and calculate the standard deviation.

Example

From the following frequency distribution, find the standard deviation using the formula for grouped data:

Class interval	Frequency
10 – 20	9
20 – 30	18
30 – 40	31
40 – 50	17
50 – 60	16
60 - 70	9
Total	100

Interval	Mid-point X	Frequency f	fx	dx	dx ²	f dx ²
10 – 20	15	9	135	- 24	576	5184
20 – 30	25	18	450	-14	196	3528
30 – 40	35	31	1085	- 4	16	496
40 – 50	45	17	765	6	36	612
50 – 60	55	16	880	16	256	4096
60 - 70	65	9	585	26	676	6084
Total		100	3900			20000

$$\text{Standard Deviation } \sigma = \sqrt{\frac{\sum f dx^2}{\sum f}} = \sqrt{\frac{20000}{100}} = \sqrt{200} = 14.1$$

The same problem can also be solved with the step deviation method which is useful when the numbers are large.

Taking assumed mean, $A = 35$, $dx = \frac{x-35}{c}$, where $c = 10$

Interval	Mid-point X	Frequency f	$dx = \frac{x-35}{c}$	dx^2	f dx	f dx^2
10 – 20	15	9	-2	4	-18	36
20 – 30	25	18	-1	1	-18	18
30 – 40	35	31	0	0	0	0
40 – 50	45	17	1	1	17	17
50 – 60	55	16	2	4	32	64
60 – 70	65	9	3	9	27	81
Total		100		Total	40	216

$$\begin{aligned} \text{S.D.} = \sigma &= \sqrt{\frac{\sum f \cdot dx^2}{\sum f} - \left(\frac{\sum f \cdot dx}{\sum f} \right)^2 \times c} = \sqrt{\frac{216}{100} - \left(\frac{40}{100} \right)^2 \times 10} \\ &= \sqrt{2 \times 10} = 14.1421 \end{aligned}$$

Relative measure of standard deviation or coefficient of variation

$$CV = \frac{S.D.}{\text{mean}} \times 100 = \frac{14.14}{39} \times 100 = 36.26$$

Example

The marks of internal assessment obtained by FYBMS students in a college are given below. Find the mean marks and standard deviation.

22 30 36 12 15 25 18 10 33 29

Solution: We first sum all the observations and find the mean. Then the differences of the observations from the mean are computed and squared. The positive square root average of sum of square of the differences is the required standard deviation.

x	$d = x - \bar{x}$	d^2
22	-1	1
30	7	49
36	13	169
12	-11	121
15	-8	64
25	2	4
18	-5	25
10	-13	169
33	10	100
29	6	36
$\Sigma x = 230$	-	$\Sigma d^2 = 738$

$$\text{I. } \bar{x} = \frac{\Sigma x}{n} = \frac{230}{10} = 23.$$

$$\begin{aligned} \text{II. } \sigma &= \sqrt{\frac{\Sigma d^2}{n}} = \sqrt{\frac{738}{10}} \\ &= \sqrt{73.8} = 8.59 \end{aligned}$$

$$\therefore \sigma = 8.59$$

Example

Find the standard deviation for the following data:

03 12 17 29 10 05 18 14 12 20

Solution: We find the sum of the observations and the sum of its squares. Using formula 2.2, S.D. is computed as follows:

x	x^2
3	9
12	144
17	289
29	841
10	100
05	25
18	324
14	196
12	144
20	400
$\Sigma x = 140$	$\Sigma x^2 = 2472$

The formula 2.2 gives the S.D. as below:

$$\sigma = \sqrt{\frac{\Sigma x^2}{n} - \left(\frac{\Sigma x}{n}\right)^2} = \sqrt{\frac{2472}{10} - \left(\frac{140}{10}\right)^2}$$

$$= \sqrt{247.2 - 196} = \sqrt{51.2}$$

Example

Compute the standard deviation for the following data

x	100	102	104	106	108	110	112
f	5	11	7	9	13	10	12

Solution: *Short-cut method:*

In problems where the value of x is large (consequently its square also will be very large to compute), we use the short-cut method. In this method, a fixed number x_0 (which is usually the central value among x) is subtracted from each observation. This difference is denoted as $u = x - x_0$. Now the columns of fu and fu^2 are computed and the S.D. is

calculated by the formula: $\sigma = \sqrt{\frac{\Sigma fu^2}{N} - \left(\frac{\Sigma fu}{N}\right)^2}$. One can observe that

this formula is similar to that mentioned in 2.3. This formula is called as change of Origin formula.

In this problem we assume $x_0 = 106$. The table of calculations is as follows:

x	$u = x - 106$	f	fu	fu^2
100	-6	5	-30	180
102	-4	11	-44	176
104	-2	7	-14	28
106	0	9	0	0
108	2	13	26	52
110	4	10	40	160
112	6	12	72	432
Total	-	$N = 67$	$\Sigma fu = 50$	$\Sigma fu^2 = 1028$

In this table, the column of fu^2 is computed by multiplying the entries of the columns fu and u .

From the table we have: $\Sigma f.u^2 = 1028$, $\Sigma fu = 50$ and $N = 67$.

$$\therefore \sigma = \sqrt{\frac{\Sigma f.u^2}{N} - \left(\frac{\Sigma fu}{N}\right)^2} = \sqrt{\frac{1028}{67} - \left(\frac{50}{67}\right)^2}$$

$$\therefore \sigma = \sqrt{15.34 - 0.56} = \sqrt{14.78}$$

$$\therefore \sigma = 3.84$$

Some important points to be noted regarding Standard deviation:

The standard deviation being an algebraic quantity, it possesses the following important characteristics :

1. It is rigidly defined
2. It is based on all the observations, i.e. the value of the standard deviation will change if any one of the observations is changed.
3. In the case of the value which lies close to the mean, the deviations are small and therefore variance and standard deviation are also small. Variance and standard deviation would thus be zero when all the values are equal.
4. If the same amount is added to or subtracted from all the values, the mean shall increase or decrease by the same amount; also deviations from the mean in the case of each value would remain unchanged and hence variance and standard deviation shall remain unchanged
5. In case a number of samples are drawn from the same population, it may be observed that standard deviation is least affected from sample to sample as compared to other three measures of dispersion.

Limitations: Standard deviation lays down the limits of variability by which the individual observation in a distribution will vary from the mean. In other words, Mean ± 1 standard deviation will indicate the range within which a given percentage of values of the total are likely to fall i.e., nearly 68.27% will lie within mean ± 1 standard deviation, 95.45% within mean ± 2 standard deviation and 99.73% within mean ± 3 standard deviation.

The point may be illustrated by taking an example of distribution of weight of 1000 school students with a mean height of 40 Kgs and standard deviation of 6 Kgs. If the groups of students is a normal one, about two thirds (68.26) of the students would fall within ± 1 standard deviation from the mean. Thus 683 students would weigh between 34 and 46 i.e., 40 ± 6 Kg. Further, when 2 standard deviations are added and subtracted from the mean, the total population covered would be 95.44% and in the case of 3 standard deviation it would cover 99.73%.



UNIT IV

Unit-6

ELEMENTARY PROBABILITY THEORY

Unit Structures :

- 6.0 Objectives
- 6.1 Introduction
- 6.2 Types of Events
- 6.3 Algebra of Events
- 6.4 Solved Examples
- 6.5 Counting Principle and Combination
- 6.6 Random Variable and Expected Value
- 6.7 Normal Distribution

6.0 OBJECTIVES

- To understand the uncertainty (chance) involved in the unpredictable events.
- To find the probability (numeric value of the uncertainty) and various rules of probability to measure the uncertainty.
- To find expected value (Mean) and variance in random experiments.
- Use of normal distribution to find proportions and percentage of the observations referred to certain continuous variables.

6.1 INTRODUCTION

In our day-to-day life conversation we hear the statements like Most probably it will rain today. Or a sales manager makes the claim the sales will cross Rs.500 cores.

Both these statements show that the claims are subject to uncertainty and cannot be predicted in advance with 100% guaranty.

Probability measures the certainty in such type unpredicted events. The origin of probability lies in Gambling or the games of choices such as tossing a fair coin, throwing a cubic die or removing a card from a pack of playing cards. Today probability plays an important role in the field of Economics, Finance, and Medicine etc. for making inferences and predictions.

To understand the concept of probability and learn the methods of calculating the probabilities, we should first define understand some basic terms and concepts related to probability.

Random Experiment: Any act or trial in which we are not sure about the result is called as the random experiment.

e.g.. Tossing a fair coin. Throwing a cubic die. Removing a number in the game of Housie.

Outcome: The possible result of the random experiment is called an outcome.

e.g. When we toss a coin, there are two possible outcomes Head(H) and Tail(T).

or when we throw a cubic die the possible outcomes of the no of dots on the uppermost face are 1,2,3,4,5,or6.

Sample Space: The collection of all the possible outcomes in the Random experiment is called the sample space. It is denoted by S. The outcomes listed in the sample space are called the sample points. The sample space may be finite, countable infinite or infinite in nature. The no of sample points in the sample space is denoted by $n(S)$.

e.g. When we toss a pair of unbiased coins, the sample space is
 $S = \{HH, HT, TH, TT\}$ $n(S) = 4$

Or when a cubic die is thrown the sample space is
 $S = \{1, 2, 3, 4, 5, 6\}$ $n(S) = 6$

Event: An event is a well-defined subset of the sample space. It is denoted by the letters like A,B,C etc. The no of sample points in the event is denoted by $n(A)$.

e.g. In the experiment of throwing a cubic die when the sample space is
 $S = \{1, 2, 3, 4, 5, 6\}$. We can define the events as follows

Event A: The no of dots appeared is multiple of 3.

$$A = \{3, 6\} \quad n(A) = 2$$

B: The no of dots appeared is a prime number.

$$B = \{2, 3, 5\} \quad n(B) = 3$$

Random Experiment	Event
Tossing a pair of coins $S = \{HH, HT, TH, TT\}, n(S) = 4$	A: Both coins show Head. $A = \{HH\}$ B: At least one coin show Head. $B = \{HT, TH, HH\} \quad n(B) = 3$
Tossing three coins at a time $S = \{HHH, TTT, HHT, HTH, THH, HTT, THT, TTH\}, n(S) = 8$	A: exactly two Head turns up. $A = \{HHT, HTH, THH\}$
Throwing a pair of cubic dice. $S = \{(1,1), (1,2), (1,3), (1,4), (1,5), (1,6), \dots, (5,6), (6,6)\}$ $n(S) = 36$	A: The sum of the dots on the uppermost faces is 6 or 10 $A = \{(1,5), (5,1), (2,4), (4,2), (3,3), (4,6), (6,4), (5,5)\} \quad n(A) = 8$ B: The sum of the dots on the uppermost faces is divisible by 4. $B = \{(1,3), (3,1), (2,2), (2,6), (6,2), (3,5), (5,3), (4,4), (6,6)\} \quad n(B) = 9$
Selecting a two digit number $S = \{10, 11, 12, \dots, 99\}$ $n(S) = 100$	A: The number is a perfect square. $A = \{16, 25, 36, 49, 64, 81\} \quad n(A) = 6$ B: The number is > 99 . $B = \{ \} \quad n(B) = 0$

6.2 TYPE OF EVENTS

Simple event- The event containing only one sample point is called a simple event.

e.g Tossing a pair of coins $S = \{HH, HT, TH, TT\}$

Now the event defined as

A: Both coins show Head.

Is a simple event.

Null event: It is the event containing no sample point in it is called a null or impossible event. It is the impossible happening and is denoted by ' Φ '. e.g in the experiment of throwing a cubic die when the sample space is

$S = \{1, 2, 3, 4, 5, 6\}$. We define the event

Event A: The no of dots appeared is a two digit number

$A = \{ \}$ i.e. $A = \Phi$ and $n(A) = 0$.

6.3 ALGEBRA OF EVENTS

Union of two events A and B ($A \cup B$): union $A \cup B$ is the event containing all the sample points in A and B together.

$A \cup B = \{\text{Elements in either A or B or in both A and B together}\}$

$= \{x \text{ such that } x \in A \text{ or } x \in B\}$ or $\in A \cup B$ both

e.g. When $A = \{1, 2, 3\}$ and $B = \{2, 3, 5\}$

$A \cup B = \{1, 2, 3, 5\}$

Intersection of two events A and B ($A \cap B$): For the events A and B defined on the sample space S associated with the random experiment

E, intersection of A and B is the event containing all the sample points common to A and B both.

$$A \cap B = \{\text{Elements in A and B both}\} \\ = \{x \text{ such that } x \in A \text{ and } x \in B\}$$

e.g. When $A = \{1, 2, 3\}$ and $B = \{2, 3, 5\}$

$$A \cap B = \{2, 3\}$$

Mutually exclusive events: The two events A and B are said to be mutually exclusive or disjoint if they have no common element in them. i.e. their intersection is an empty set.

Disjoint events cannot occur simultaneously.

e.g. When $A = \{1, 2, 3\}$ and $B = \{4, 5\}$

Mutually exclusive and exhaustive events: The two events A and B are said to be mutually exclusive and exhaustive if they are disjoint and their union is S.

e.g. when a cubic die is thrown the sample space is

$$S = \{1, 2, 3, 4, 5, 6\}$$

Now if the events A and B defined on S have the sample points as follows

$$A = \{1, 2, 3\} \text{ and } B = \{4, 5, 6\} \text{ then } A \cap B = \Phi \text{ and } A \cup B = S$$

Hence A and B are mutually exclusive and exhaustive.

Complement of an event A: Let A be any event defined on the sample space S, then its complement A' is the event containing all the sample points in S which are not in A.

$$A' = \{\text{elements in S which are not in A}\}$$

$$A = \{x \text{ s.t. } x \in S \text{ but } x \notin A\}$$

e.g. When $S = \{1, 2, 3, 4, 5, 6\}$ and $A = \{2, 3\}$

$$A' = \{1, 4, 5, 6\}.$$

Probability of an event A: (Classical definition) Suppose S is the sample space associated with the random experiment E, and A is any event defined on the sample space S, then its probability $P(A)$, is defined as

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

In other words probability of A is the proportion of A in S.

Example: A cubic die is thrown up find the probability that, the no of dots appeared is a prime number.

Solution: When a cubic die is thrown the sample space is

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

Now we define the event A as,

A: the no of dots appeared is a prime number.

$$A = \{2, 3, 5\} \quad n(A) = 3$$

Using the above formula, we get

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

From the above we can note that,

Probability of event A always lies between 0 and 1 i.e. $0 \leq P(A) \leq 1$.

$P(\text{Sample space}) = P(S) = 1$ and $P(\Phi) = 0$.

Also, when events A and B are such that $A \subset B$ then, $P(A) \leq P(B)$.

6.4 SOLVED EXAMPLES

Example 1:

An unbiased die is thrown, find the probability that, i) the no of dots is less than 3 ii) the no of dots is divisible by 3.

Solution: when a cubic die is thrown the sample space is

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

(i) Let A denote the event that no of dots < 3 . $A = \{1, 2\}$ i.e. $n(A) = 2$

$$\therefore P(A) = \frac{n(A)}{n(S)} = \frac{2}{6} = 0.33$$

(ii) Now let B denote the event of no of dots is divisible by 3.

$$B: \{3, 6\} \text{ and } n(B) = 2$$

$$\therefore P(B) = \frac{n(B)}{n(S)} = \frac{2}{6} = 0.33$$

Example 2:

Three unbiased coins are tossed at a time. Find the probability that, (a) exactly two Head turns up and (b) at most two Head turns up.

Solution: When three coins are tossed up at a time the sample is

$$S = \{HHH, HHT, HTH, THH, HTT, THT, TTH, TTT\} \quad n(S) = 8.$$

Now to find the required probability, we define the events as follows,

Event A: Exactly two Head turns up.

$$A = \{HHT, THH, HTH\} \quad n(A) = 3$$

$$\therefore P(A) = \frac{n(A)}{n(S)} = \frac{3}{8}$$

Event B: At most two Head turns up.

$$B: \{HHT, HTH, THH, HTT, THT, TTH, TTT\} \quad n(B) = 7.$$

$$\therefore P(B) = \frac{n(B)}{n(S)} = \frac{7}{8}.$$

Example 3:

A pair of fair dice is rolled. Write down the sample space and find the probability that, a) the sum of dots on the uppermost face is 6 or 10, b) the sum of dots is multiple of 4 and c) the sum of the dots is < 6 .

Solution: When a pair of dice is rolled, the sample space is

$$S = \{(1, 1), (1, 2), (1, 3), (1, 4), (1, 5), (1, 6) \\ (2, 1), \dots, (2, 6) \\ \dots, (5, 6), (6, 6)\} \quad n(S) = 36.$$

To find the probability we define the events,

a) event A: the sum of dots on the uppermost face is 6 or 10.

$$A = \{(1,5), (5,1), (2,4), (4,2), (3,3), (4,6), (6,4), (5,5)\} \quad n(A) = 8.$$

$$\therefore P(A) = \frac{n(A)}{n(S)} = \frac{8}{36}$$

b) event B: The sum of the dots on the uppermost faces is divisible by 4.

$$B: \{(1,3), (3,1), (2,2), (2,6), (6,2), (3,5), (5,3), (4,4), (6,6)\} \quad n(B) = 9.$$

$$\therefore P(B) = \frac{n(B)}{n(S)} = \frac{9}{36} = 0.25.$$

c) event C: the sum of the dots is < 6 .

$$C: \{(1,1), (1,2), (1,3), (1,4), (2,1), (2,2), (2,3), (3,1)\} \quad n(C) = 8.$$

$$\therefore P(C) = \frac{n(C)}{n(S)} = \frac{8}{36}.$$

6.5 COUNTING PRINCIPLE AND COMBINATION

In some experiments like selecting cards, balls or players we cannot list out the complete sample space but count the no of sample points in the given sample space. To count the no of points in the sample space in these experiments, we define the concepts of combination. First, we state the **Counting principle** (Fundamental Principle of Mathematics), as follows

Counting principle: If A and B are two different things can be independently performed in m and n different possible ways then, by counting principle

(a) both A and B together can be performed in m.n possible ways.

(b) any one of them i.e. A or B can be performed in m+n possible ways.

Combination: Combination of r different things from n things e.g. selecting 3 balls from 5 balls, or 4 students from the group of 10 students of a class. It is calculation by the formula,

$${}^nC_r = \frac{n!}{r!(n-r)!} \quad \text{where } n! = n(n-1)(n-2)\dots 3 \times 2 \times 1.$$

$$(e.g. 5! = 5 \times 4 \times 3 \times 2 \times 1 = 120)$$

Illustrations-

(i) We can select a group of 3 students from 5 students in 5C_3 ways and

$${}^5C_3 = \frac{5!}{2!(5-2)!} = \frac{5!}{2!3!} = \frac{5 \times 4 \times 3 \times 2 \times 1}{2 \times 1 \times 3 \times 2 \times 1} = 10.$$

(ii) A student can select 4 different questions from 6 independently in 6C_4 ways.

$${}^6C_4 = \frac{6!}{2!(6-4)!} = \frac{6!}{2!4!} = \frac{6 \times 5 \times 4 \times 3 \times 2 \times 1}{2 \times 1 \times 4 \times 3 \times 2 \times 1} = 15.$$

- (iii) A box contains 4 red and 6 green balls then by counting principle,
 a) 2 red **and** 3 green together can be drawn in ${}^4C_2 \times {}^6C_3$ ways.
 b) 3 red **or** 3 green can be drawn in ${}^4C_3 + {}^6C_3$ ways.

Example 4:

Three cards are drawn from the pack of 52 playing cards. Find the probability that (a) all three are spade cards, (b) all three are of same suit, and (c) there are two Kings and one Queen.

Solution: When three cards are drawn from the pack of 52 cards, no of points in the sample is

$$n(S) = {}^{52}C_3 = \frac{52!}{3!(52-3)!} = \frac{51 \times 50 \times 49}{3 \times 2 \times 1} = 22100.$$

Now, we define the event,

a) A: all three are spade cards.

There are 13 spade cards, so 3 of them can be drawn in ${}^{13}C_3$

$$n(A) = {}^{13}C_3 = \frac{13!}{3!10!} = \frac{13 \times 12 \times 11}{3 \times 2 \times 1} = 286$$

$$\therefore P(A) = \frac{n(A)}{n(S)} = \frac{286}{22100} = 0.0129.$$

a) event B: all three are of same suit

There are four suits (Club, Diamond, Heart & Spade) of 13 cards each. Three from them can be drawn in ${}^{13}C_3$ ways each. Now by the counting principle,

$$n(B) = {}^{13}C_3 + {}^{13}C_3 + {}^{13}C_3 + {}^{13}C_3 = 4 \times 286 = 1144. \quad \textbf{(refer to illustrations above)}$$

$$\text{Hence } P(B) = \frac{1144}{22100} = 0.051.$$

b) event C: there are two Kings and one Queen.

There are 4 kings and 4 Queens in the pack. So 2 kings and 1 Queen can be drawn in 4C_2 and 4C_1 respectively.

Therefore by counting principle,

$$n(C) = {}^4C_2 \times {}^4C_1 = 6 \times 4 = 24.$$

$$\text{Where, } {}^4C_2 = \frac{4!}{2!2!} = \frac{4 \times 3 \times 2 \times 1}{2 \times 1 \times 2 \times 1} = 6 \quad \& \quad {}^4C_1 = \frac{4!}{1!3!} = \frac{4 \times 3 \times 2}{3 \times 2 \times 1} = 4.$$

$$\therefore P(C) = \frac{n(C)}{n(S)} = \frac{24}{22100}$$

Example 5:

A box contains 5 Red and 4 Green balls. Two balls are drawn at random from the box, find the probability that i) Both are of same color ii) only red balls are drawn.

Solution: In the box there 9(4+5) balls in total, so 2 of them can be drawn in 9C_2 ways.

$$n(S) = {}^9C_2 = \frac{9!}{2!7!} = \frac{9 \times 8}{2 \times 1} = 36.$$

(i) We define the event A: Both are of same color

$$n(A) = {}^4C_2 + {}^5C_2 = \frac{4!}{2!2!} + \frac{5!}{2!3!} = 6 + 10 = 16.$$

$$\therefore P(A) = \frac{n(A)}{n(S)} = \frac{16}{36} = 0.44$$

Hence probability of same colour is 0.44

(ii) Let us define event B: only red balls are drawn

Now both the balls will be drawn from 4 red balls only

$$n(B) = {}^4C_2 = 6$$

$$\therefore P(B) = \frac{n(B)}{n(S)} = \frac{6}{36} = 0.16$$

Hence probability of both red balls is 0.16.

Exercise

1. Define the terms, i) Sample space. ii) An event. iii) Mutually exclusive events
2. Define the probability of an event. Also state the properties of the probability of event.
3. An unbiased coin is tossed three times, write down sample points w.r.t. following events: (a) Head occur only two times, (b) Head occur at least 2 times and (c) There are more Heads than Tail.
4. A pair of fair dice is rolled, write down the sample points w.r.t. following events: The sum of the no of dots appearing on the uppermost faces is (i) 7 or 11 (ii) multiple of 3 (iii) a perfect square.
5. A pair of coins is tossed at a time find the probability that,
6. Both the coins show Head. ii) No coin shows Head. iii) Only one Head turns up.
7. A cubic die is thrown, find the probability that the no of dots appeared is (a) A prime number, (b) A number multiple of 2.
8. A box contains 20 tickets numbered 1-20. A ticket is drawn at random from the box, find the probability that, i) the ticket bears a number < 5. ii) the number on the ticket is divisible by 4. iii) it is a cube of a natural number.
9. A card is drawn from the pack of 52 playing cards find the probability that, (a) The card is a king card (b) It is a face card.
10. If two fair dice are thrown, what is the probability the sum of the no of dots on the dice is, a) greater than 8. b) between 5 and 8.
11. Three unbiased coins are tossed at a time find the probability that, (a) exactly one Head turns up. (b) At most 2 Heads turn up. (c) All 3 coins show Heads.
12. Two cards are drawn from the pack of 52 playing cards find the probability that, (a) Both the cards are of same suits. (b) Both are Ace cards. (c) Both are of different colour.

13. A pair of dice is rolled, write down the sample space and find the probability that, the sum of the no of dots appearing on the uppermost faces is, (a) 6 or 10. (b) multiple of 5. (c) a perfect square.
14. Four unbiased coins are tossed at a time find the probability that, (i) exactly 2 Heads turn up. (ii) at most 3 Heads turn up. (iii) at least 3 coins show Heads.
15. Three cards are drawn from the pack of 52 playing cards find the probability that, (i) all 3 cards are of same colour. (ii) Two cards are face cards. (iii) Only face cards are drawn.
16. Two boxes identical in size and shape respectively contain 3 red, 4 blue and 5 red, 2 blue balls. One ball is drawn at random from each box. What is the probability that both the balls are of same colour.
17. A committee of 4 is to be formed from 3 Professors and 7 students in a college. Find the probability that it includes, a) only 2 Professors. b) there are at least 3 students.
18. A box contains 6 red, 4 green and 3 white balls. Two balls are drawn at random, find the probability that, a) both are of same colour. b) no white ball is drawn. c) the balls are of different colour.
19. Given $P(A)=0.6$, $P(B)=0.5$, and $P(A \cap B)=0.4$. Find $P(A \cup B)$; $P(A/B)$; $P(B/A)$.
20. For two events A and B; $P(A)=\frac{2}{5}$, $P(B')=\frac{1}{3}$, $P(A \cup B)=\frac{5}{6}$. Find $P(A \cap B)$; $P(A/B)$; $P(\text{only } A)$; $P(\text{Only one})$.
21. For two mutually exclusive events A and B, $P(A)=0.7$ and $P(B)=0.5$, find $P(A \cap B)$ and $P(A/B)$
22. For the independent events A and B, $P(A)=\frac{1}{2}$, $P(B)=\frac{2}{5}$. Find $P(A \cap B)$; $P(A \cup B)$; $P(\text{only } B)$.
23. One of the two purses contains 4 Gold coins and 5 Silver coins, another purse contains 3 Gold and 6 Silver coins. A coin is drawn at random from one of the purses, find the probability that it is a silver coin.
24. Two students A and B are solving a problem on Mathematics independently. Their chances of solving the problem are $\frac{1}{2}$ and $\frac{1}{3}$ respectively. Find the probability that, i) the problem is solved. ii) it is solved by only one of them.
25. A government contractor applies for 2 tenders of supplying breakfast supply and lunch box supply. His chances of getting the contracts are 0.6 and 0.5 respectively. Find the probability that, (i) he will get either of the contract (ii) he will get only one contract.

6.6 RANDOM VARIABLE AND EXPECTED VALUE

Random variable is a real valued function defined on the sample space.

Suppose S is the sample space associated with the random experiment E , then to every sample point in S we can assign a real number denoted by a variable X called as a random variable on S . e.g. When we toss a coin three times the sample is

$S = \{HHH, HHT, HTH, THH, HTT, THT, TTH, TTT\}$, now if we define a variable

X : No of tosses showing Heads.

Then X takes values 0,1,2,3 where

$\{X=0\} \Leftrightarrow \{TTT\}; \quad \{X=1\} \Leftrightarrow \{HTT, THT, TTH\};$

$\{X=3\} \Leftrightarrow \{HHH\} \quad \{X=2\} \Leftrightarrow \{HHT, HTH, THH\}.$

Hence to each sample points in S we have assigned a real number, which uniquely determine the sample point.

The variable X is called as the random variable defined on the sample space S .

We can also find the probabilities of values 0,1,2,3 of the r.v. X as follows

$P(\{X=0\}) = P(\{TTT\}) = 1/8, P(\{X=1\}) = P(\{HTT, THT, TTH\}) = 3/8,$

$P(\{X=3\}) = P(\{HHH\}) = 1/8, P(\{X=2\}) = P(\{HHT, HTH, THH\}) = 3/8.$

Now we can express these probabilities in the form of a table,

$X:$	0	1	2	3
$P(x)$	1/8	3/8	3/8	1/8

X	$P(x)$
0	1/8
1	3/8
2	3/8
3	1/8
Total	1

This is called as the probability distribution of random variable X .

In general probability distribution of X satisfies the following conditions;

- (i) all $p(x)$ are positive. i.e. $p(x) \geq 0$
- (ii) $\sum p(x) = 1$ for all x .

A **random variable X** defined on the sample space S may be finite or infinite, at the same time it may take only countable values (without decimal) such variables are called as discrete random variables. On the other hand some variables like height, weight, income do take the fractional values also and called as the continuous random variables.

Expected value of X , $E(X)$:

Suppose a random variable X defined on sample space S takes values $x_1, x_2, x_3, \dots, x_n$. with respective probabilities $p_1, p_2, p_3, \dots, p_n$; $P(X = x_1) = p_1$, it's expected value is defined as,

$$E(X) = \sum x_i \cdot p(x_i)$$

Expected value is also called as the mean of X .

Variance of X ; V(X): For the random variable X, variance is defined as,

$$\begin{aligned} V(X) &= E(X - E(X))^2 \\ &= E(X^2) - [E(X)]^2 \\ &= \sum x^2 \cdot p(x) - (\sum x \cdot p(x))^2 \end{aligned}$$

Root of variance is called as standard deviation S.D.

Solved examples:

Example 6:

A discrete random variable X has the following probability distribution.

x:	-2	-1	0	1	2
p(x)	k	0.2	2k	2k	0.1

Find k. Also find the expected value of random variable X.

Solution: Since X is a random variable with given p(x), it must satisfy the conditions of a probability distribution.

$$\sum p(x) = 1 \Rightarrow 5k + 0.3 = 1 \Rightarrow k = 0.7/5 = 0.14$$

p(x) values are 0.14; 0.2; 0.28; 0.28; 0.1

Now we calculate the expected value by the formula,

$$E(X) = \sum x \cdot p(x) = 0 \text{ ----- from the table.}$$

x	p(x)	xp(x)
-2	0.14	-2 x 0.14 = -0.28
-1	0.2	-0.2
0	0.28	0
1	0.28	0.28
2	0.1	0.2
Total	1	0 = $\sum x \cdot p(x)$

Example 7:

A random variable follows the probability distribution given below,

X	0	1	2	3	4
p(x)	0.12	0.23	0.35	0.20	0.10

Obtain the expected value and variance of X.

Solution: The expected value and variance are given by the formula,

$$E(X) = \sum x \cdot p(x) \text{ and } V(X) = \sum x^2 \cdot p(x) - (\sum x \cdot p(x))^2$$

x	p(x)	xp(x)	x ² .p(x)
0	0.12	0x0.12=0	0x0=0
1	0.23	0.23	0.23
2	0.35	0.70	1.40
3	0.20	0.60	1.80
4	0.10	0.40	1.60
Total	1.00	1.93 = $\sum x \cdot p(x)$	4.03 = $\sum x^2 \cdot p(x)$

Now from the table,

$$E(X) = \sum x \cdot p(x) = 1.93.$$

$$\begin{aligned} V(X) &= \sum x^2 \cdot p(x) - (\sum x \cdot p(x))^2 \\ &= 4.03 - (1.93)^2 \end{aligned}$$

$$V(X) = 0.35.$$

Hence, Mean $E(X) = 1.93$ units and $V(X) = 0.35$ units.

Example 8:

Find mean and variance of the random variable X whose probability distribution is given by

X:	-2	-1	0	1	2
P(x)	1/16	1/8	5/8	1/8	1/16.

Solution: For the random variable X, we have

From the table we get,

$$E(X) = \sum x.p(x) = 0.$$

$$\text{And } V(X) = \sum x^2.p(x) - (\sum x.p(x))^2$$

$$= \frac{12}{16} - 0 = \frac{12}{16}$$

x	p(x)	xp(x)	x ² .p(x)
-2	1/16	-2/16	-2x(-2/16) = 4/16
-1	1/8	-1/8	1/8
0	5/8	0	0
1	1/8	1/8	1/8
2	1/16	2/16	4/16
Total	1.00	0 = $\sum x.p(x)$	12/16 = $\sum x^2.p(x)$

Example 9:

A uniform die is thrown find the expected value of the random variable X denoting the no on the uppermost face.

Solution: When a uniform die is thrown the random variable

X: the no on the uppermost face, takes the possible values 1,2,3,4,5 or 6.

With the same probability of occurrence.

Therefore we can find the mean or expected value of X by using the formula,

$$E(X) = \sum x.p(x)$$

$$= \frac{21}{6} = 3.5 \text{ ----- from the}$$

table.

Hence the mean of X is 3.5.

x	p(x)	xp(x)
1	1/6	1/6
2	1/6	2/6
3	1/6	3/6
4	1/6	4/6
5	1/6	5/6
6	1/6	6/6
Total	6/6 = 1	21/6 = $\sum x.p(x)$

EXERCISE II

1. A random variable X has the following probability distribution:

X:	-2	-1	0	1	2	3
P(x)	0.1	k	0.2	2k	0.3	k

Find the value of k. Find the expected value and variance of x.

2. A random variable X has the following probability distribution:

X:	0	1	2	3	4	5
P(x)	0.1	0.1	0.2	0.3	0.2	0.1

Find the expected value and variance of x.

3. An unbiased coin is tossed four times. Find the expected value and variance of the random variable defined as number of Heads.

6.7 NORMAL DISTRIBUTION

Normal distribution deals with the calculation of probabilities for a continuous random variable like Height of players, Marks of students, or Wages of workers. We define the normal distribution as follows.

A continuous random variable X is said follow a normal distribution with parameters μ and σ , written as $X \sim N(\mu, \sigma^2)$ if it's probability function is given by

$$p(x) = \frac{1}{\sigma\sqrt{2\pi}} e^{-\frac{1}{2}\left(\frac{x-\mu}{\sigma}\right)^2} \text{ where } x, \mu \in \mathbb{R}, \sigma > 0$$

$$= 0 \quad \text{otherwise.}$$

Here the constants are $\mu = \text{Mean}(X)$; $\sigma = \text{S.D.}(X)$

$\pi = 3.142$ and $e = 2.718$ (approx).

Before we learn to calculate the probabilities on normal distribution, we state the characteristics of the normal distribution stated below.

Characteristics of the normal distribution

i) The graph of normal distribution is a bell shaped curve.
 ii) The area under the curve reads the probabilities of normal distribution hence total area is 1 (one).

iii) The curve is symmetric about it's mean μ . Hence,
 Area on l.h.s. of $\mu = \text{Area on r.h.s. of } \mu = 0.5$. Since area reads probability,

$$P(X < \mu) = P(X > \mu) = 0.5 = 50\%.$$

iv) Hence mean μ divides the curve into two equal parts so it is also the median.

The curve has it's maximum height at $x = \mu$, therefore it the mode of the distribution.

v) Hence for normal distribution **Mean = Median = Mode = μ**

For the probability calculations, we define the variable

$$Z = \frac{x - \text{Mean}}{\text{S.D.}} = \frac{x - \mu}{\sigma}$$

Mean (Z) = 0 and S.D.(Z) = 1. Z is called a standard normal variable (s.n.v.)

Also $P(X) = P(Z) = \text{Area}(Z)$. The area (probability) values of z are tabulated.

vi) The lower (Q_1) and upper quartiles (Q_3) are equidistant from the mean

$$\mu \text{ i.e. } \mu - Q_1 = Q_3 - \mu \Rightarrow \mu = \frac{Q_1 + Q_3}{2}$$

vii) The mean deviation (M.D.) of normal distribution is $\frac{4}{5} \sigma$

viii) The quartile deviation (Q.D.) of normal distribution is 0.67σ .

Area under the normal curve between

- | | |
|-------------------------------------|------------------------------------|
| (i) $\mu \pm \sigma$ is 68.27% | (ii) $\mu \pm 2 \sigma$ is 95.45% |
| (iii) $\mu \pm 3 \sigma$ is 99.73%. | (iv) $\mu \pm 1.645 \sigma$ is 90% |
| (v) $\mu \pm 1.96 \sigma$ is 95% | (vi) $\mu \pm 0.67 \sigma$ is 50%. |

Solved examples:

Example 10:

A continuous random variable X follows a normal distribution with mean 50 and S.D. of 10. Find the following probabilities for X ,

- a) $P(X \geq 55)$ b) $P(45 \leq X \leq 60)$ c) $P(X \leq 45)$.

Solution: For the normal variable X , we have

Mean (X) = $\mu = 50$ and standard deviation = $\sigma = 10$.

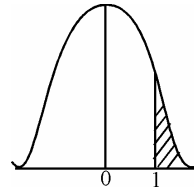
$\therefore X \sim N(\mu, \sigma^2) = N(50, 10^2)$.

We define the variable $Z = \frac{x - \mu}{\sigma} = \frac{x - 50}{10}$

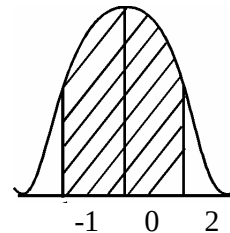
Mean (Z) = 0 and S.D. (Z) = 1. Z is called a standard normal variable (s.n.v.)

Also $P(X) = P(Z) = \text{Area}(Z)$.

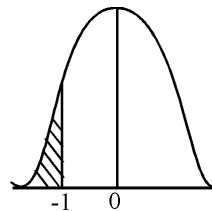
$$\begin{aligned} \text{a) } P(X \geq 55) &= P\left(\frac{x - 50}{10} \geq \frac{55 - 50}{10}\right) \\ &= P(Z \geq 1) \\ &= \text{Area r.h.s. of } +1 \\ &= 0.5 - \text{Area from } 0 \text{ to } 1 \\ &= 0.5 - 0.3413 = 0.1587. \end{aligned}$$



$$\begin{aligned} \text{b) } P(45 \leq X \leq 60) &= P\left(\frac{45 - 50}{10} \leq \frac{x - 50}{10} \leq \frac{60 - 50}{10}\right) \\ &= P(-1 \leq Z \leq 2) \\ &= \text{Area between } -1 \text{ \& } +2 \\ &= \text{Area from } -1 \text{ to } 0 + \text{Area from } 0 \text{ to } 2. \\ &= 0.3413 + 0.4772 = 0.8185. \end{aligned}$$



$$\begin{aligned} \text{c) } P(X \leq 45) &= P\left(\frac{x - 50}{10} \leq \frac{45 - 50}{10}\right) \\ &= P(Z \leq -1) \\ &= \text{Area on l.h.s. of } -1. \\ &= 0.5 - \text{Area from } -1 \text{ to } 0. \\ &= 0.5 - 0.3413 = 0.1587. \end{aligned}$$



Example 11:

The marks of 150 students in the class is said to follow a normal with mean 60 and S.D. of 10. Find, the expected no of students scoring marks below 45. Percentage of students scoring marks between 55 and 70.

Solution: Let X : Marks of students;

Mean(X) = 60 and S.D.(X) = $\sigma = 10$.

X has normal distribution with $\mu = 60$ and $\sigma = 10$.

i.e. $X \sim N(\mu, \sigma^2) = N(60, 10^2)$

We define $Z = \frac{X - \mu}{\sigma} = \frac{X - 60}{10}$.

$\therefore$ Mean (Z) = 0 and S.D.(Z) = 1. Z is called a standard normal variable (s.n.v.)

Also $P(X) = P(Z) = \text{Area}(Z)$.

Now, to find the expected no of students scoring marks below 45, we find

$P(\text{marks less than 45})$

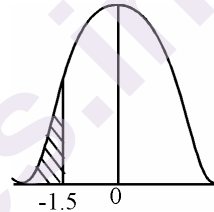
$$= P(X \leq 45) = P\left(\frac{X - 60}{10} \leq \frac{45 - 60}{10}\right)$$

$$= P(Z \leq -1.5)$$

= Area on l.h.s. of -1.5.

$$= 0.5 - \text{Area from } -1.5 \text{ to } 0.$$

$$= 0.5 - 0.4332 = 0.0668 = 6.68\%.$$



Expected no of students = $6.68\%(150) = 10$.

Similarly, to find percentage of students scoring marks between 55 and 70.

Consider, $P(\text{marks between 55 and 70})$

$$= P(55 \leq X \leq 70)$$

$$= P\left(\frac{55 - 60}{10} \leq \frac{X - 60}{10} \leq \frac{70 - 60}{10}\right)$$

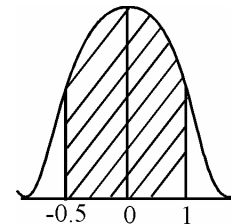
$$= P(-0.5 \leq Z \leq 1)$$

= Area between -0.5 & +1

= Area from -0.5 to 0 + Area from 0 to 1.

$$= 0.1915 + 0.3413 = 0.5328 = 53.28\%.$$

$\therefore$ 53.28% students have scored marks between 55 and 70.

**Example 12:**

The height of 250 soldiers in a military camp confirms a normal distribution with mean height of 155cms. and S.D. of 20cms. Find the proportion of soldiers with height above 170 cms. Also find the height of the shortest soldier in the group of tallest 20% soldiers.

Solution: Let r.v. X denotes the height of a soldier.

Mean(X) = 155 and S.D.(X) = $\sigma = 20$.

$\therefore$ X has normal distribution with $\mu = 155$ and $\sigma = 20$.

i.e. $X \sim N(\mu, \sigma^2) = N(155, 20^2)$

We define $Z = \frac{x - \mu}{\sigma} = \frac{x - 155}{20}$.

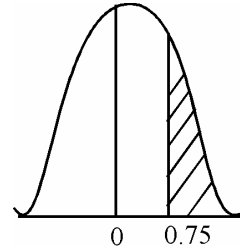
$\therefore$ Mean (Z) = 0 and S.D.(Z) = 1. Z is called a standard normal variable (s.n.v.)

Also $P(X) = P(Z) = \text{Area}(Z)$.

Now, to find the proportion of soldiers with height above 170 cms, we find,

$P(\text{soldier's height is above 170 cms})$

$$\begin{aligned} P(X \geq 170) &= P\left(\frac{x-155}{20} \geq \frac{170-155}{20}\right) \\ &= P(Z \geq 0.75) \\ &= \text{Area r.h.s. of 0.75} \\ &= 0.5 - \text{Area from 0 to 0.75.} \\ &= 0.5 - 0.2734 = 0.2766. \end{aligned}$$



$\therefore$ Proportion of soldiers with height above 170 cms is $(0.2766 \times 250) : 250$
i.e. 69 : 250.

Now, let the height of the shortest soldier in the group of tallest 20% soldiers be h .

$P(\text{height less than } h) = 20\% = 0.2$ i.e. $P(X \leq h) = 0.2$

Consider, $P(X \leq h) = P\left(\frac{x-155}{20} \leq \frac{h-155}{20} = t \text{ say}\right)$ i.e. $P(Z \leq t) = 0.2$.

Area on l.h.s. of $t = 0.2$ (t is less than 0 i.e. negative.) (since area on l.h.s. < 0.5)

area from t to 0 = 0.3.

Now from the normal area table, area from 0 to 0.84 is 0.3.

Hence, $t = -0.84$ ($t = \frac{h-155}{20} = -0.84$ i.e. $h = 155 + 20(-0.84) = 155 -$

$16.8 = 138.2$.)

Therefore, the height of the shortest soldier in the group of tallest 20% soldiers is 138 cms.

Example 13:

The daily wages of 300 workers in a factory are normally distributed with the average wages of Rs.2500 and S.D. of wages equals to Rs.500. Find the percentage of workers earning wages between Rs.3000 and Rs.4000. Also find the wages of the lowest paid worker in the group of highest paid 30% workers.

Solution: Let r.v. X denotes the wages of a worker.

Mean(X) = 2500 and S.D.(X) = σ = 500.

$\therefore$ X has normal distribution with $\mu = 2500$ and $\sigma = 500$.

i.e. $X \sim N(\mu, \sigma^2) = N(2500, 500^2)$

We define $Z = \frac{x - \mu}{\sigma} = \frac{x - 2500}{500}$

$\therefore$ Mean (Z) = 0 and S.D.(Z) = 1. Z is called a standard normal variable (s.n.v.)

Also $P(X) = P(Z) = \text{Area}(Z)$.

Now, to find the percentage of workers earning wages between Rs 3000 and Rs 4000.

Consider, $P(\text{wages between Rs.3000 and Rs 4000})$

$$= P(3000 \leq X \leq 4000)$$

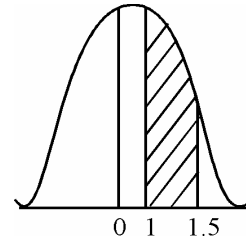
$$= P\left(\frac{3000 - 2500}{500} \leq \frac{x - 2500}{500} \leq \frac{4000 - 2500}{500}\right)$$

$$= P(1 \leq Z \leq 1.5)$$

$$= \text{Area between } +1 \text{ \& } +1.5$$

$$= \text{Area from 0 to 1.5} - \text{Area from 0 to 1.}$$

$$= 0.4332 - 0.3413 = 0.0919 = 9.19\%.$$



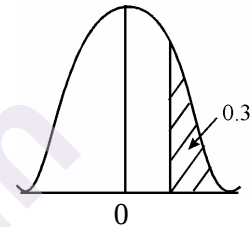
$\therefore$ 9.91% workers are earning wages between Rs 3000 and Rs 4000.

Now, let the wages of the lowest paid worker in the group of highest paid 30% workers be h .

$$P(\text{wages greater than } h) = 30\% = 0.3 \quad \text{i.e. } P(X \geq h) = 0.3.$$

$$\text{Consider, } P(X \geq h) = P\left(\frac{x - 2500}{500} \geq \frac{h - 2500}{500} = t \text{ say}\right)$$

$$\text{i.e. } P(Z \geq t) = 0.3.$$



$$t = 0.2$$

$\therefore$ Area on r.h.s. of $t = 0.3$ (t is greater than 0 i.e. positive. (since area on r.h.s. < 0.5))

area from 0 to $t = 0.2$.

Now from the normal area table, area from 0 to 0.52 is 0.2.

$$\text{Hence, } t = 0.52 \Rightarrow \frac{h - 2500}{500} = 0.52 \Rightarrow h = 2500 + 500(0.52) = 2760.$$

Therefore, the wages of the lowest paid worker in the group of highest paid 30% workers are Rs. 2760.

EXERCISE III

1. Define a normal variable. State the properties of normal distribution.

2. What is mean by a standard normal variable. What are the mean and standard of a standard normal variable.

3. A continuous random variable X follows a normal distribution with mean 50 and S.D. of 10. Find the following probabilities for X ,

$$\text{a) } P(X \geq 55) \quad \text{b) } P(45 \leq X \leq 60) \quad \text{c) } P(X \leq 45).$$

Given, Area under the normal curve,

From 0 to 1 is 0.3413.

From 0 to 2 is 0.4772.

4. The marks of 150 students in the class is said to follow a normal with mean 60 and S.D. of 10. Find, the expected no of students scoring marks below 45. Percentage of students scoring marks between 55 and 70.

Given, Area under the normal curve,
 From 0 to 0.5 is 0.1915.
 From 0 to 1 is 0.3413.
 From 0 to 1.5 is 0.4332.

5. The height of 250 soldiers in a military camp confirms a normal distribution with mean height of 155cms. and S.D. of 20cms. Find the proportion of soldiers with height above 170 cms. Also find the height of the shortest soldier in the group of tallest 20% soldiers.

Given, Area under the normal curve,
 From 0 to 1.5 is 0.4332.
 From 0 to 0.84 is 0.3.

6. The daily wages of 300 workers in a factory are normally distributed with the average wages of Rs.2500 and S.D. of wages equals to Rs.500. Find the percentage of workers earning wages between Rs.3000 and Rs.4000. Also find the wages of the highest paid worker in the group of lowest paid 30% workers.

Given, Area under the normal curve,
 From 0 to 1 is 0.3413.
 From 0 to 1.5 is 0.4332.
 From 0 to 0.52 is 0.2.

7. A normal distribution has mean $\mu = 15$ and $\sigma = 5$. Find the following probabilities.

$$P(X \geq 20) \quad P(10 \leq X \leq 17.5) \quad P(X \leq 12).$$

Given, Area under the normal curve,
 From 0 to 0.4 is 0.1554.
 From 0 to 0.5 is 0.1915.
 From 0 to 1 is 0.3413.

8. The weights of 450 students in a school are normally distributed with the average weight of 50 kg. and S.D. 5 kg. Find the percentage of students with weight:

- i) less than 45 kg. ii) Between 40 and 47 kg.

Given, Area under the normal curve,
 From 0 to 0.4 is 0.1554. From 0 to 0.5 is 0.1915. From 0 to 1 is 0.3413.



UNIT V

Unit - 7

STATISTICAL DECISION THEORY

Unit Structure:

- 7.0 Objectives
- 7.1 Introduction

7.0 OBJECTIVES

After going through this chapter you will be able to understand:

- The decision making situation.
- The decision making criteria to arrive at optimum decision.
- The decision tree technique for multi-stage decision making.

7.1 INTRODUCTION

The decision theory may be applied to problems whether it involves financial management or a plant assembly line, whether the time span is ten years, five years or one day and whether it is in public or private sector. In each of such decision-making problems, there are certain common elements which are called ingredients of decision problems. These ingredients are:

1. Alternative Courses of Action: The process of decision-making involves the selection of a single act from among some set of alternative acts. Decision is needed in a problem situation where two or more alternative courses of action are available, and where only one of these actions can be taken. Obviously, if there is only one course of action available, no decision is required since that action must be taken in order to solve the problem. For the sake of simplicity the possible actions are symbolised by $a_1, a_2, a_3, \dots$ etc. The totality of all possible actions is called action space denoted by A . If there are only three possible actions, we write $A = \text{action space} = (a_1, a_2, a_3)$. The decision procedure involves selecting among the alternatives a single course of action that can be actually carried out. If such a course of action is selected that cannot be carried out in the existing situation and circumstance, it will then amount to waste of time and resources. Quite often the objective of precision is to select an act which will accomplish some predesignated purpose. The decision taken may be regarded as satisfactory or not depending upon whether it has helped in the attainment of that objective.

2. Uncertainty: In all decision problems "uncertainty" is found to be a common element. When the outcome of some action is not known in advance, the outcome is said to be uncertain. When there are many possible outcomes of an event (also called *states of nature*) one cannot predict with certainty what will happen—it is only in terms of probabilities we may be able to talk. The term "state of nature" does not mean nature in the ordinary sense of the word, it is a general term which is used to encompass all those factors beyond the control of the decision-maker that affect the outcome of his decision. The various states of nature (outcomes) are symbolised by $\theta_1, \theta_2, \theta_3 \dots$ etc. Totality of all outcomes is called nature space or state space symbolised by Ω . If an action leads to three outcomes θ_1, θ_2 and θ_3 then we write : $\Omega = (\theta_1, \theta_2, \theta_3)$.

For example, if a product is marketed it may be highly appreciated (outcome θ_1), it may not appeal to the customers (outcome θ_2) or it may be liked by a certain fraction of the customers say, 25% (outcome θ_3).

It must be pointed out that sometimes a distinction is made between decision-making under "risk" and decision-making under "uncertainty". When the state of nature is unknown, but objective or empirical data is available so that the decision-maker can use these data to assign probabilities to the various states of nature the procedure is generally referred to as decision-making under 'risk.' When the state of nature is unknown and there is no objective information on which probabilities can be based, the procedure is referred to as decision-making under 'uncertainty'. It may, however, be noted that even when no objective information is available, the decision-maker may, in Bayesian decision theory, assign subjective probability to the states of nature to help in taking a decision. Once probabilities are assigned, regardless of the manner in which they were obtained, the decision procedure that follows is exactly the same. Hence for practical purposes risk and uncertainty are essentially the same and decision-making under both circumstances will be referred to as decision-making under uncertainty.

3. Pay Offs: In order to evaluate each possible course of action, the result of each event with each course of action has a *value (or pay off)* placed upon it. A number of consequences result from each action under different conditions, the conditions being various states of nature, the consequences will be more in number. In practical situations, particularly in business and economic problems, consequences can be expressed in terms of money and utility.

The consequences may be evaluated in several ways such as:

- (i) in terms of profit,
- (ii) in terms of cost.
- (iii) in terms of opportunity loss: (The opportunity loss is defined as the difference between the highest possible profit for an event and the actual profit obtained for the actual action taken.)
- (iv) units of satisfaction or utility.

4. Decision criteria: The decision-maker must-determine how to select the best course of action. In most decision problems the expected monetary value is used as a decision criterion. When consequences are evaluated in terms of profit, they are called payoffs. A payoff table is prepared and it shows the relation between all possible states of nature, all possible actions and the values associated with the consequences. A specimen of payoff table is given below;

GENERAL FORMAT OF A PAYOFF TABLE

Acts	a_1	a_2	a_3		a_k		a_m
States of nature							
θ_1	p_{11}	p_{12}	p_{13}		p_{1k}		p_{1m}
θ_2	p_{21}	p_{22}	p_{23}		p_{2k}		p_{2m}
θ_3	p_{31}	p_{32}	p_{33}		p_{3k}		p_{3m}
$\vdots$	$\vdots$	$\vdots$	$\vdots$	$\vdots$	$\vdots$	$\vdots$	$\vdots$
θ_i	p_{i1}	p_{i2}	p_{i3}				
$\vdots$	$\vdots$	$\vdots$	$\vdots$	$\vdots$	$\vdots$	$\vdots$	$\vdots$
θ_n	p_{n1}	p_{n2}	p_{n3}		p_{nk}		p_{nm}

In the above table, the column heading designate the various actions out of which the decision-maker may choose while the row heading shows the admissible states of nature under which the decision-maker has to take decision .

A payoff may be thought of as a conditional value or conditional profit (loss). It is conditional value in the sense that is associated with each course of action there is a certain profit (or loss), given that a specific state of nature has occurred. A payoff table thus contains all conditional values of all possible combinations and states of nature. The payoff table indicates that there is no single act which is best for all the states of nature. Therefore, in order to make an optimal decision some criterion and additional information are necessary.

The calculation of payoff depends on the problem. Very often it is a relatively easy matter and sometimes a bit algebraic reasoning is required.

With the payoff table, the decision-maker may be able to reach the optimal solution of a problem, if he has a knowledge of what event is going to occur. Since there is uncertainty about occurrence of events, a decision-maker must make some prediction or forecast usually in terms of probability of occurrence of events. With the probabilities of occurrence assigned, the last step of statistical decision theory is to analyse these probabilities by calculating expected payoff (EP) of expected monetary

value for each course of action. The decision criterion here is to choose the Optimal Act (OA), the act that yields the highest EP.

An alternative decision criterion of statistical decision theory is what is called *expected opportunity loss, EOL*. This criterion also leads to the same result as obtained from expected profits (EP).

Calculations of EOL are the same as those for EP except for the fact that we have to use conditional opportunity loss (COL) instead of payoffs. It may be pointed out that COL of the optimal acts is zero, COL of any act other than OA is positive and is the difference between the payoff of OA and the act taken.

If we replace the payoffs by their corresponding opportunity losses we get a new table called loss table.

Example 1:

A baker produces a certain type of special pastry at a total average cost of Rs. 3 and sells it at a price of Rs. 5. This pastry is produced over the weekend and is sold during the following week; such pastry being produced but not sold during a week's time is totally spoiled and has to be thrown away. According to past experience the weekly demand for these pastries is never less than 78 or greater than 80. You are required to formulate action space, payoff table and opportunity loss table.

Solution: It is clear from the problem given that the manufacturer will not produce less than 78 or more than 80 pastries. Thus, there are three courses of action open to him:

a_1 = produce 78 pastries

a_2 = produce 79 pastries

a_3 = produce 80 pastries

thus the action space or $A = (a_1, a_2, a_3)$

The state of nature is the weekly demand for pastries. There are three possible states of nature, i.e.,

θ_1 = demand is 78 pastries

θ_2 = demand is 79 pastries

θ_3 = demand is 80 pastries

Hence the state space $\Omega = (\theta_1, \theta_2, \theta_3)$

The uncertainty element in the problem is the weekly demand. The bakery profit conditioned by the weekly demand. Cell values of payoff table are computed as follows:

Demand x Price - Production x Cost

P_{11} = payoff when action a_1 is taken but the state of nature is θ_1
 = Rs. $[5 \times 78 - 3 \times 78] = \text{Rs. } 156.$

P_{12} = payoff when action a_2 is taken but the state of nature is θ_1
 = Rs. $[5 \times 78 - 3 \times 79] = \text{Rs. } 153.$

P_{13} = payoff when action a_3 is taken but the state of nature is θ_1
 = Rs. $[5 \times 78 - 3 \times 80] = \text{Rs. } 150$.

Similarly P_{21} = payoff when action a_1 is taken and the state of nature is θ_2
 = Rs. $[5 \times 78 - 3 \times 78] = \text{Rs. } 156$.

P_{22} = Rs. $[5 \times 79 - 3 \times 79] = \text{Rs. } 158$.

P_{23} = Rs. $[5 \times 79 - 3 \times 80] = \text{Rs. } 155$.

Also, P_{31} = payoff when action a_1 is taken and the state of nature is θ_3

P_{31} = Rs. $[5 \times 78 - 3 \times 78] = \text{Rs. } 156$.

P_{32} = Rs. $[5 \times 79 - 3 \times 79] = \text{Rs. } 158$.

P_{33} = Rs. $[5 \times 80 - 3 \times 80] = \text{Rs. } 160$

An EMV indicates the average profit that would be gained if a particular alternative were selected in many similar decision-making situations. Since decisions are often made on a one time basis, the decision criterion is to choose the alternative course of action that maximizes the expected monetary value.

These values are tabulated below:

PAYOFF TABLE

Action <i>State of nature</i>	$a_1 = 78$	$a_2 = 79$	$a_3 = 80$
$\theta_1 = 78$	156	153	150
$\theta_2 = 79$	156	158	155
$\theta_3 = 80$	156	158	160

To calculate opportunity loss, we first determine maximum payoff in each state of nature.

In first state of nature θ_1 , maximum payoff = Rs. 156

In second state of nature θ_2 , maximum payoff = Rs. 158

In third state of nature θ_3 , maximum payoff = Rs. 160

$L_{11} = 156 - 156 = 0$, $L_{12} = 156 - 153 = 3$, $L_{13} = 156 - 150 = 6$

$L_{21} = 158 - 156 = 2$, $L_{22} = 158 - 158 = 0$, $L_{23} = 158 - 155 = 3$

$L_{31} = 160 - 156 = 4$, $L_{32} = 160 - 158 = 2$, $L_{33} = 160 - 160 = 0$.

The loss table corresponding to payoff table is given below.

Action State of nature	a_1	a_2	a_3
θ_1	0	3	6
θ_2	2	0	3
θ_3	4	2	0

Example 2: Suppose in a process of producing bulbs, the number of producing defective bulbs is constant but unknown. Consider a lot of 100 bulbs which is either to be sold at Rs. 5 each giving a double money back

guarantee for each defective item or to be junked at a cost of Rs. 100 for the lot. Construct action space, state space, and payoff table.

Solution: Two possible actions are open to the manufacturer, namely,
 a_1 = Junk the lot at a loss of Rs. 100 for the lot,
 a_2 = Sell the lot at Rs. 5 each giving double money back guarantee for each defective item.

Thus $A = (a_1 a_2)$ = action space

The number of defective bulbs in the lot designates the status of nature. The number of defective bulbs can be 0, 1, 2, 3. 100. i.e. there are 101 possible states of nature. Let O_i denote that there are "i" defective bulbs in the lot.

Thus state space or $\Omega = \{O_0, O_1, O_2, \dots, O_{100}\}$

Since there are 101 states of nature, it is very difficult to work out all possible payoff. Suppose the manufacturer takes action a_1 , then he will lose Rs. 100.

$$P_{i1} = -100 \text{ for all } i = 0, 1, 2, \dots, 100$$

Suppose the manufacturer takes action a_2 and the state of nature is that, there are defective bulbs in the lot then the payoff is given by

$$P_{i2} = 100 \times 5 - i \times 10 = 500 - 10i \quad \text{For all } i = 0, 1, 2, \dots, 100.$$

Thus his payoffs are

$$\begin{array}{l} P_{i1} = -100 \quad \} \quad \text{for all } i = 0, 1, 2, \dots, 100 \\ P_{i2} = 500 - 10i \quad \} \quad \text{for all } i = 0, 1, 2, \dots, 100 \end{array}$$

I (i) Maximax or Hurwicz Decision criterion: The maximax or Hurwicz decision criterion is a criterion of super optimism. It is based upon the idea that we do get some favourable or lucky breaks. Since, nature can be good to us, the decision maker should select that state of nature which will yield him the highest payoff for the selected strategy. Thus, the maximax criterion attempts to maximise the maximum gain i.e. maximax chooses the act that is the "best of the best". The procedure involved in the criterion is to look at the various payoffs for each strategy and select the highest amount. Therefore, the maximum of these maximum payoffs is selected. This payoff is referred to as a maximax (maximum of maximums).

i) Example 3 : Consider the following payoff table :

Strategies	S_1	S_2	S_3
States of Nature			
N_1	24	17	10
N_2	24	27	20
N_3	24	27	30
Maximum gains	24	27	(30)
Minimum gains	(24)	17	10

For maximax criterion, first of all maximum payoff of each strategy is identified. Maximum payoff for strategy S_1 is 24, for S_2 is 27 and for S_3 , it is 30. Next, we must determine the largest payoff of these maximum payoffs which is 30. Hence, based on maximax criterion the best strategy would be S_3 .

(ii) Maximin or Wald decision criterion: The maximin criterion is a criterion of pessimism. Wald suggested that the decision maker should always be pessimistic or conservative, resulting in a maximin criterion. This maximin payoff approach means that the decision maker, under continually adverse circumstances, should select the strategy that will give him as large a minimum payoff as possible i.e. the criterion attempts to maximise the minimum gains or it tries to "pick the best of the worst". Under maximin criterion, minimum payoff of each strategy is selected. In the example given in (i) part, minimum payoffs for S_1 is 24, for S_2 it is 17, and for S_3 , the minimum payoff is 10. Once the minimum payoff corresponding to each strategy has been selected, the next step is to select the maximum of these minimum payoffs (maximum of minimums) which is 24 in our above example. Hence, strategy S_1 is the maximin gain strategy; In essence, the worst state of nature that could happen would give a payoff of 24.

(iii) Minimax Regret Criterion: Suppose that the manager is concerned about how the decision he makes might be viewed in the future after the state of nature is known with certainty. In such situation, the minimax regret criterion can be used. To employ this criterion, one must transform the payoff matrix into a regret matrix by replacing every payoff in a row of the payoff matrix with the difference obtained by subtracting the payoff from the row's maximum payoff. This regret or opportunity loss represents the amount of profit that a person lost because he did not select the most profitable act. *When the minimax criterion is used, the decision maker expects the worst event to materialise and so he selects the act that will give the minimum of the maximum opportunity losses.*

Example 4:

Given the following pay-off function for each act a_1 and a_2 $Qa_1 = -30 + 50x$
 $Qa_2 = -90 + 20x$

- (i) What is the break-even value of x ?
- (ii) If $x = 10$, which is the better act?
- (iii) If $x = 10$, what is the regret of the poor strategy?
- (iv) If $x = -5$ which is the better act?
- (v) If $x = -5$ what is the regret of the poorer strategy?

Solution: Equating Qa_1 and Qa_2

$$-30 + 50x = -90 + 20x$$

$$\text{or } x = -2$$

$\therefore -2$ is the break even point.

(ii) Substituting $x = 10$ in Qa_1 and Qa_2

$$Qa_1 = 470 \quad Qa_2 = 110$$

Hence a_1 is a better strategy.

(iii) Regret with $x = 10$ is equal to $470 - 110 = 360$

This happens when strategy a_2 is adopted.

(iv) for $x = -5$ $Qa_1 = -280$

$Qa_2 = -190$

a_2 is a better strategy.

(v) Regret in (iv) = $-190 - (-280)$

= $-190 + 280$

= 90

Example 5:

A food products company is contemplating the introduction of a revolutionary new product with new packaging to replace the existing product at much higher price (S_1) or a moderate change in the composition of the existing product with a new packaging at a small increase in price (S_2) or a small change in the composition of the existing except the word 'New' with a negligible increase in price (S_3). The three possible states of nature of events are (i) high increase in sales (N_1), (ii) no change in sales (N_2), and (iii) decrease in sales (N_3). The marketing department of the company worked out the payoffs in terms of yearly net profits for each of the strategies for these events (expected sales) This is represented in the following table:

State of nature	Pay offs (in Rs.)		
Strategies	N_1	N_2	N_3
S_1	7,00,000	3,00,000	1,50,000
S_2	5,00,000	4,50,000	0
S_3	3,00,000	3,00,000	3,00,000

Which strategy should the executive concerned choose on the basis of

(a) Maximin Criterion,

(b) Maximax Criterion,

(c) Minimax Regret Criterion,

(d) Laplace Criterion ?

Solution. a) *Maximin Criterion*. When this criterion is adopted that course of action is selected which maximises the minimum payoffs.

Strategy	Minimum Payoffs Rs
S_1	1,50,000
S_2	0
S_3	3,00,000

The executive would choose strategy S_3

(b) *Maximax Criterion*. In this criterion we select that strategy which gives the maximum payoffs.

Strategy	Maximum Payoffs
S_1	7,00,000
S_2	5,00,000
S_3	3,00,000

The executive should choose S_1

(c) *Minimax Regret Criterion*. When this criterion is adopted three steps are necessary :

- (i) to determine the opportunity loss for each strategy by subtracting from maximum pay-off of a state of nature, the payoffs of all the strategies.
- (ii) to determine the maximum opportunity loss for each strategy.
- (iii) to select the strategy which minimises the maximum of the loss.

OPPORTUNITY LOSS TABLE

	N_1	N_2	N_3
S_1	$7,00,000 - 7,00,000 = 0$	$4,50,000 - 3,00,000 = 1,50,000$	$3,00,000 - 1,50,000 = 1,50,000$
S_2	$7,00,000 - 5,00,000 = 2,00,000$	$4,50,000 - 4,50,000 = 0$	$3,00,000 - 0 = 3,00,000$
S_3	$7,00,000 - 3,00,000 = 4,00,000$	$4,50,000 - 3,00,000 = 1,50,000$	$3,00,000 - 3,00,000 = 0$

Maximum opportunity loss

S_1	$= 1,50,000$
S_2	$= 3,00,000$
S_3	$= 4,00,000$

The executive should choose strategy S_1 for it minimises the maximum opportunity loss.

Laplace criterion. This involves three things:

- (i) assigning equal opportunity to each state of nature.
- (ii) calculating the expected monetary value E_1MV .
- (iii) selecting that strategy whose E_1MV is maximum.

E_1MV TABLE

$$\begin{aligned}
 S_1 : \quad & \frac{1}{3} (7,00,000 + 3,00,000 + 1,50,000) = \frac{11,50,000}{3} = 3,83,333.33 \\
 S_2 : \quad & \frac{1}{3} (5,00,000 + 4,50,000 + 0) = \frac{9,50,000}{3} = 3,16,666.67 \\
 S_3 : \quad & \frac{1}{3} (3,00,000 + 3,00,000 + 3,00,000) = \frac{9,00,000}{3} = 3,00,000
 \end{aligned}$$

Since the E_2MV is highest for strategy 1, hence the executive should select strategy S_1 .

Example 6:

A management is faced with the problem of choosing one of three products for manufacturing. The potential demand for each product may

turn out to be good, moderate or poor. The probabilities for each of the state of nature were estimated as follows :

Product	Nature of Demand		
	Good	Moderate	poor
X	0.70	0.20	0.10
Y	0.50	0.30	0.20
Z	0.40	0.50	0.10

The estimated profit or loss under the three states may be taken as

X	30,000	20,000	10,000
Y	60,000	30,000	20,000
Z	40,000	10,000	-15,000(loss)

Prepare the expected value table and advise the management about the choice of product.

Computation of EV for various Acts

States of Nature	Expected payoff (in Rs. '0000) for various acts								
	X			Y			Z		
	x_{1j}	p_{1j}	$x_{1j}p_{1j}$	x_{2j}	p_{2j}	$x_{2j}p_{2j}$	x_{3j}	p_{3j}	$x_{3j}p_{3j}$
Good	30	0.7	21	60	0.5	30	40	0.4	16
Moderate	20	0.2	04	30	0.3	09	10	0.5	05
Poor	10	0.1	01	20	0.2	04	-15	0.1	-1.5
Expected Monetary value			26			43			19.5

Since the expected value is highest for second course of action, the management is advised to produce Y.

Example 7:

A TV dealer finds that the procurement cost of a TV is Rs. 20 and the cost of a unit shortage is Rs. 50. For one particular model of TV the probability distribution of weekly sales is as follows:

Weekly Sales	0	1	2	3	4	5	6
Probability	0.10	0.10	0.20	0.20	0.20	0.15	0.05

How many units per week should the dealer buy? Also find E.V.P.I.

Solution : The cost matrix is constructed below. Also derived from it are the expected costs for the various strategies.

Demand Prob.		Strategies (buy so many)						
		0	1	2	3	4	5	6
0	0.10	0	20	40	60	80	100	120
1	0.10	50	20	40	60	80	100	120
2	0.20	100	70	40	60	80	100	120
3	0.20	150	120	90	60	80	100	120
4	0.20	200	170	140	110	80	100	120
5	0.15	250	220	190	160	130	100	120
6	0.05	300	270	240	210	180	150	120
Expected		147.5	122.5	102.5	92.5	92.5	102.5	120

Thus either buy 3 or 4.

Example 8: Following are the records of demand of an item for the past 300 days.

θ = Demand in units	No. of days	Prob (θ)
10,000	18	0.06
11,000	90	0.30
12,000	120	0.40
13,000	60	0.20
14,000	<u>12</u>	<u>0.04</u>
	300	1.00

(i) What is the expected demand?

(ii) It costs Rs. 15 to make an item which sells for Rs. 20 normally but at the end of the day surplus has to be disposed at Rs. 10 per item. What is the, optimum output?

Solution: Expected demand = $10 \times 0.06 = 0.60$.
 $= 11 \times 0.30 = 3.30$
 $= 12 \times 0.40 = 4.80$
 $= 13 \times 0.20 = 2.60$
 $= 14 \times 0.04 = \underline{0.56}$
 11.86

Expected demand = 11,860 (Answer)

The daily pay-off matrix is constructed below:

		Strategies				
P(θ)	states of nature	10	11	12	13	14
0.06	10	50	45	40	35	30
0.30	11	50	55	50	45	40
0.40	12	50	55	60	55	50
0.20	13	50	55	60	65	60
0.04	14	50	55	60	65	70
Expected Pay off		50	54.4	55.8	53.2	48.6

Demand of 12,000 is the optimum with an expected payoff of $55.8 \times 1000 = 55,800$ rupees.;

Example 9: The sales manager of Beta Co. is highly experienced in the fad market. He is sure that the sales of JUMBO (during the period it has special appeal) will not be less than 25,000 units. Plant capacity limits total production to a maximum of 80,000 units during JUMBO's, brief life. According to the sales manager, there are 2 chances in 5 for a sales volume of 50,000 units. The probability that it will be more than 50,000 is four times the probability that it will be less than 50,000. If sales exceed 50,000 units volumes of 60,000 and 80,000 are equally likely. A 70,000 unit volume is 4 times as likely as either. It costs Rs. 30 to produce a unit of JUMBO whereas its selling price is estimated at Rs. 50 per unit. Initial investment is estimated at Rs. 8,00,000. Should the venture of production be undertaken?

Solution:

$$\text{Prob. (50,000)} = \frac{2}{5} = 0.40$$

$$\text{Prob. (less than or more than 50,000)} = 0.60$$

$$\text{Prob. (less than 50,000): Prob. (More than 50,000): } 1:4$$

$$\text{Thus Prob. (less than 50,000)} = 0.60 \times \frac{1}{1+4} = 0.12$$

$$\text{Prob. (More than 50,000)} = 0.60 \times \frac{4}{1+4} = 0.48$$

$$\text{Prob. (60,000): Prob. (70,000): Prob. (80,000):: } 1:4:1$$

$$\text{Thus Prob (60,000)} = 0.48 \times \frac{1}{1+4+1} = 0.08$$

$$\therefore \text{ Prob (80,000)} = 0.08$$

$$\text{and Prob 70,000} = \frac{4}{1+4+1} \times 0.48 = 0.32$$

Now we have complete probability distribution of sales except that for sales less than 50,000 we have a summarised probability of 0.12. The payoff matrix is compiled below. We used 25,000 in place of less than 50,000 as the worst contingency.

Pay-off Table

		(Rs.in'000)				
Demand	Prob	< 50,000 (take it 25,000)	50,000	60,000	70,000	80,000
< 50,000 (worst 25,000)	0.12	500	-250	-550	-850	-1150
50,000	0.40	500	1,000	700	400	100
60,000	0.08	500	1,000	1200	900	600
70,000	0.32	500	1,000	1200	1400	1100
80,000	0.08	500	1,000	1200	1400	1600
Expected Payoff		500	850	790		

Expected Payoff is, thus, Rs. 850,000. This exceeds 800,000 the initial investment; therefore, the venture ought to be undertaken.

2 Decision Trees

A Decision Tree diagram is a graphical representation of various alternatives and the sequence of events in a decision problem. In constructing a decision tree, there are certain conventions to be followed. The tree is constructed starting from left and moving towards right. The square box denotes a decision point at which the available strategies are considered. The circle O represents the chance node or event, the various states of nature or outcomes emanate from this chance event. At each decision or chance node, there can be one or more branches represented by a straight line. In figure below, there are two branches at the decision node, each branch represents a strategy. Similarly, at each of the two chance nodes, there are three branches. Any branch that is not followed by another decision or chance node is called a *terminal branch*.

Let us illustrate the decision tree with an example.

Example 10: A manufacturer of toys is interested to know whether he should launch a deluxe model or a popular model of a toy. If the deluxe model is launched, the probabilities that the market will be good, fair or poor are given by 0.3, 0.4 and 0.3 respectively with payoffs Rs. 1,40,000, Rs. 70,000 and Rs. (-10,000). If the popular model is introduced, the corresponding probabilities are given by 0.4, 0.3 and 0.3 with respective payoffs Rs. 1,50,000, Rs. 80,000 and Rs. (-15,000). The problem is to decide which model should be launched. The decision tree for the given problem is drawn below:

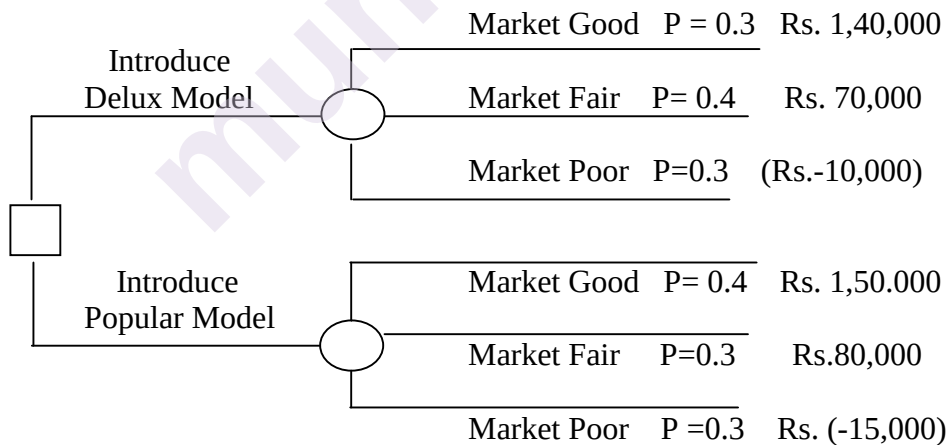


Fig. 1

Thus the decision tree shows that structure of the decision problem. To carry out the decision tree analysis, conditional payoffs are estimated for every combination of actions and events (i.e., for every path through the tree). These payoffs can be either positive or negative. Also the probabilities for each event must be assessed by the decision maker. To analyse the decision tree begin at the end of the tree and work backward. For each set of events branches, the *Expected Monetary Value (EMV)* is

calculated and for each set of decision branches, the one with the highest EMV is selected. This EMV now represents the EMV for the decision point from which this set of event branches emanates. Following the same procedure we move back on the decision tree to the next decision point. This technique of analysing tree is called the roll-back technique. Thus, there are two rules concerning roll-back technique:

- (i) If branches emanate from a circle, the total expected pay off may be calculated by summing the expected value of all the branches.
- (ii) If branches emanate from a square, we calculate the total expected benefit for each branch emanating from that square and let the total expected pay-off be equal to the value of the branch with the highest expected benefit.

Let us analyse the tree given above by the roll back technique. Here point 1 is decision point and C & D are the chance nodes.

$$\begin{aligned}\text{EMV (at C)} &= .3 \times \text{Rs. } 1,40,000 + .4 \times \text{Rs. } 70,000 + .3 \times (-\text{Rs. } 10,000) \\ &= \text{Rs. } 42,000 + \text{Rs. } 28,000 - \text{Rs. } 3,000 = \text{Rs. } 67,000\end{aligned}$$

$$\begin{aligned}\text{EMV (at D)} &= .4 \times \text{Rs. } 1,50,000 + .3 \times \text{Rs. } 80,000 + .3 \times (-\text{Rs. } 15,000) \\ &= \text{Rs. } 60,000 + \text{Rs. } 24,000 - \text{Rs. } 4,500 = \text{Rs. } 79,500\end{aligned}$$

The decision at point 1 is to introduce the popular model since it results in the highest EMV

Example 11: Mr. X is trying to decide whether to travel to Sri Lanka from Delhi to negotiate the sale of a shipment of china novelties. He holds the novelties stock and is fairly confident, but by no means sure that if he makes the trip, he will sell the novelties at price that will give him profit of Rs. 30,000. He puts the probability of obtaining the order at 0.6. If he does not make the trip, he will certainly not get the order.

If the novelties are not sold in Sri Lanka there is an Indian customer who will certainly buy them at a price that leaves him a profit of Rs. 15,000 and his offer will be open at least till Mr. X returns from Sri Lanka. Mr. X estimates the expenses of trip to Sri Lanka at Rs. 2,500. He is however, concerned that his absence, even for only three days, may lead to production inefficiencies in the factory. These could cause him to miss the deadline on another contract, with the consequence that a late penalty of Rs. 10,000 will be invoked. Mr. X assesses the probability missing the deadline under these circumstances at 0.4. Further, he believes that in his absence there will be a lower standard of house-keeping in the factory, and the raw material and labour costs on the other contract will rise by about Rs. 2,000 above the budgeted figure.

Draw an appropriate decision tree for Mr. X's problem and using EMV as the appropriate criterion for decision, find the appropriate initial decision.

Solution:

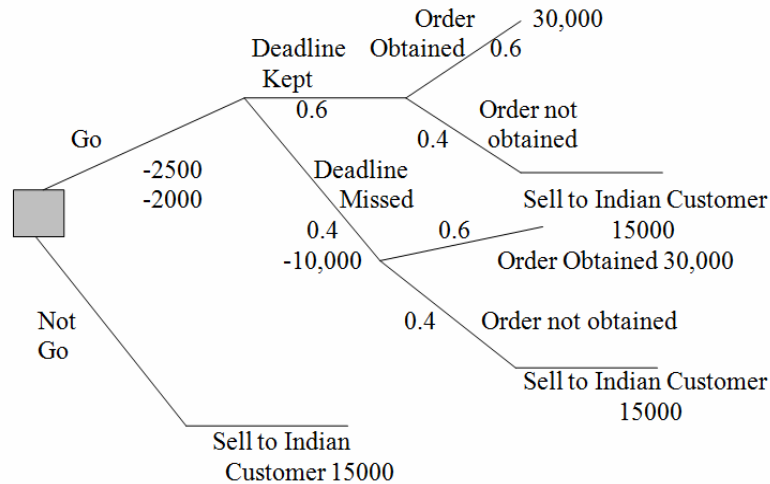


Fig. 2

The decision tree is drawn in the above figure. Calculations of EMV proceed from right to left.

<i>Go</i>	<i>Not Go</i>
$30,000 \times 0.6 = 18,000$	$15,000 \times 1 = 15,000$
$15,000 \times 0.4 = \underline{6000}$	
$\underline{24000}$	
$24,000 \times 0.6 = 14,400$	
plus	
$24,000 \times 0.4 = 9600$	
minus	
$10,000 \times 0.4 = \underline{4000}$	
$\underline{20,000}$	
minus	
$\underline{4,500}$	
$15,500$	

Hence Mr. X should proceed to Sri Lanka

Steps in Decision Tree Analysis

In a decision analysis, the decision maker has usually to proceed through the following six steps:

1. Define the problem in structured terms. First of all, the factors relevant to the solution should be determined. Then probability distributions that are appropriate to describe future behaviour of those factors estimated. The financial data concerning conditional outcomes is collected.

2. Model the decision process. A decision tree that illustrates all alternatives in the problem is constructed. The entire decision process is presented schematically and in an organised step-by-step fashion.

3. Apply the appropriate probability values and financial data. To each of the branches and sub-branches of the decision tree the appropriate probability values and financial data are applied. This would enable to distinguish between the probability value and conditional monetary value associated with each outcome.

4. "Solve" the decision tree. Using the methodology mentioned above proceed to locate that particular branch of the tree that has the largest expected value or that maximises the decision criterion.

5. Perform sensitivity analysis. Determine how the solution reacts in changes in inputs. Changing probability values and conditional financial values allows the decision maker to test both the magnitude and the direction of the reaction.

6. List the underlying assumptions. The accounting cost finding and other assumptions used to arrive at a solution should be explained. This would also enable others to know what risks they are taking when they use the results of your decision tree analysis. The limits under which the results obtained will be valid and not valid should be clearly specified.

Advantages of Decision Tree Approach

The decision tree analysis as a tool of decision-making is important because of the following:

1. Decision trees are of great help in complicated kinds of decision problems. However, in a simple problem there is no advantage of constructing a decision tree.
2. It structures the decision process making managers, approach decision making in an orderly sequential fashion.
3. It requires the decision-maker to examine all possible outcomes desirable and undesirable.
4. It communicates clearly the decision-making process to others.
5. It allows a group to discuss alternatives by focusing on each financial figure, probability value and underlying assumption one at a time. Thus, group can move in orderly steps towards a consensus decision instead of debating a decision in its entirety.
6. It can be used with a computer so that many different sets of consumptions can be simulated and their effects on the final outcomes observed.

Example 12: M/s J. Bloggs & Co. is currently working with a process, which, after paying for materials, labour etc. brings a profit of Rs. 10,000. The following alternatives are made available to the company.

- (i) The company can conduct research (R_1) which is expected to cost Rs. 10,000 and having 90% probability of success, the company gets a gross income of Rs. 25,000. (ii) The company can conduct research

- (R₂) expected to cost Rs. 5,000 and having a probability of 60% success. If successful, the gross income will be Rs. 25,000.
- (iii) The company can pay Rs. 6,000 as royalty of a new process which will bring a gross income of 20,000.
- (iv) The company continues the current process.

Because of limited resources, it is assumed, that only one of the two types of research can be carried out at a time. Which alternative should be accepted by the company?

Solution: The decision tree which represents the possible courses of action is depicted in figure 4. Point 1 is a 'decision box' located 'now' on the time scale. The four possibilities arising here are shown. Upon failure of a particular research, say, R₁, there are again 3 original alternatives to be sorted out, that of R₁ being excluded. If R₂ fails after failure of R₁, the company is left with only two choices, i.e., either to pay royalty or continue the existing process.

Branch Current

Net Return = Rs. 10,000

Branch Licence

Net Return = 20,000 - 6,000 = Rs. 14,000.

Branch R₁ First

Value at Point 3

Current Rs. 10,000

Licence Rs. 14,000

Value at point C: Expected gross profit
 $= 25,000 \times 0.6 + 14,000 \times 0.4$
 $= 20,600.$

Value at Point 2

Licence Rs. 14,000

Current Rs. 10,000

R₂ Rs. 20,600 - Rs. 5,000 = Rs. 15,600

Value at Point A Rs. 25,000 $\times$ 0.9 + 0.1 $\times$ 15,600 = Rs. 24,060

Value at Branch R₁ Rs. 24,060 - Rs. 10,000 = Rs. 14,060

Branch R₂ First

Value at point 5

Current Rs. 10,000

Licence Rs. 14,000

Value at Point at D: Expected gross profit
 $= 0.9 \times 25,000 + 0.1 \times 14,000 = \text{Rs. } 23,900$

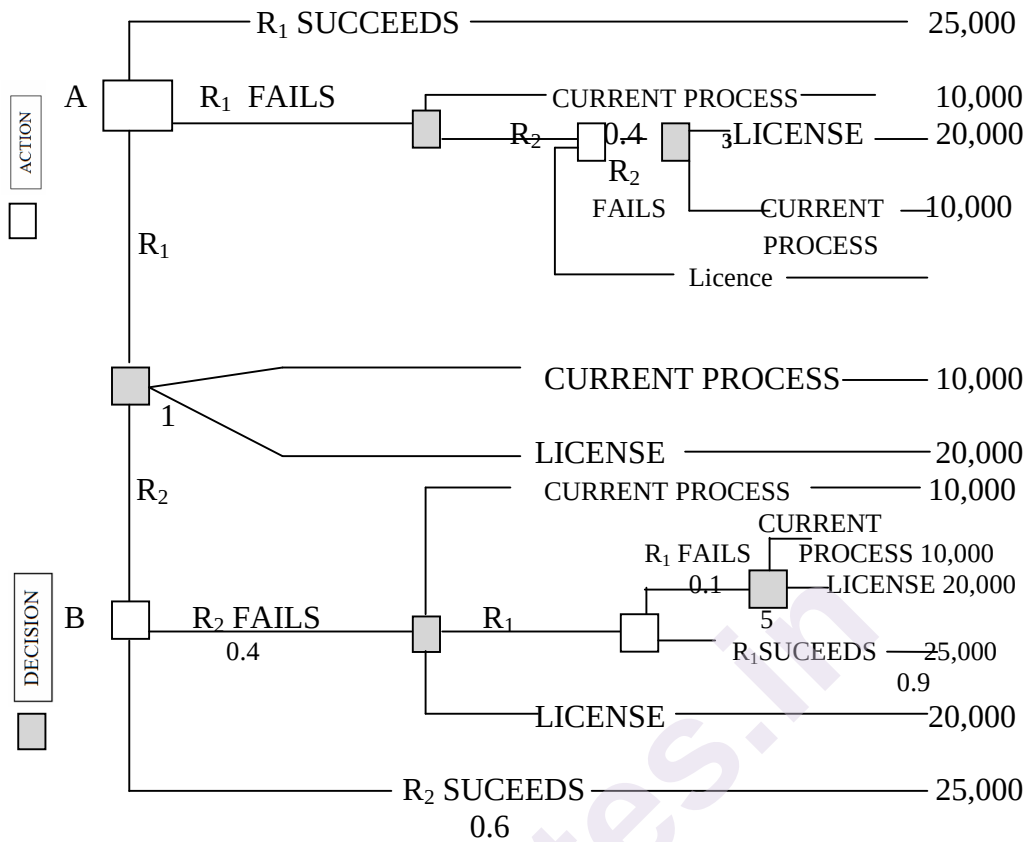


Fig 3

Value at point 4

Current

Rs. 10,000

Licence

Rs. 14,000

R_1

= Rs. 23,900 - Rs. 10,00 = Rs. 13,900

Value at Point B

= $0.4 \times 14,000 + 0.6 \times 25,000$ = Rs. 20,600

Value of Branch R_2 = Rs. 20,600 - Rs. 5,000 = Rs. 15,600

Thus R_2 , followed by licence upon former's failure is the best course of action.

EXERCISES

1. The research department of consumer products division has recommended the marketing department to launch a soap with 3 different perfumes. The marketing manager has to decide the type of perfume to launch under the following estimated payoff for various levels of sales.

Type of perfume	Estimated sales (units)		
	25,000	15,000	10,000
I	200	25	20
II	30	20	10
III	50	30	5

Find the best decision using (i) Maximax, (ii) Maximin, (iii) Minimax Regret and (iv) Laplace criteria.

Ans: (i) I, (ii) I, (iii) I, (iv) I

2. A consumer product company is examining the introduction of a new product with new packaging or replace the existing product at much higher price (S_1) or moderate change in the composition of the existing product with new packaging at a small increase in price (S_2) or a small change in the composition of the existing product except the word 'new' with the negligible increase in price (S_3). The three possible states of nature are E_1 : increase in sales, E_2 : no change in sales, E_3 : decrease in sales. The marketing department has worked out the following payoffs in terms of profits. Which strategy should be considered under (i) Minimax Regret (ii) optimistic (iii) equi-probability conditions?

States of Nature	Strategies		
	S_1	S_2	S_3
E_1	30,000	40,000	25,000
E_2	50,000	45,000	10,000
E_3	40,000	40,000	40,000

Ans: (i) S_2 , (ii) S_1 , (iii) S_2

3. Construct a payoff matrix for the following situations and find the best decision using (i) Maximin, (ii) Maximax, (iii) Laplace criteria.

Product		
	Fixed Cost Rs.	Variable cost unit in Rs.
X	250	12
Y	350	10
Z	500	5

The likely demand (units) of products

Poor demand 300

Moderate demand 700

High demand 1000

Selling price of each product is Rs. 25.

Ans: (i) Z, (ii) Z, (iii) Z

4. The management of a company is faced with the problem of choosing one of the three products A, B, C for manufacturing. The demand for these products can be good, moderate or poor. The probabilities of each state of nature are estimated as follows :

Product	Nature of demand		
	Good	Moderate	Poor
A	0.7	0.2	0.1
B	0.4	0.5	0.1
C	0.5	0.3	0.2

The estimated profit/loss under the three states of nature are given as follows:

Product	Profit in 10,000 Rs.		
	Good	Moderate	Poor
A	30	20	10
B	40	10	-15
C	60	30	20

Using EMV criterion, advice the management the optimum decision.

Ans: Select product C

5. You are given, the following payoff table for three acts A_1 , A_2 , A_3 and the states of nature S_1 , S_2 , S_3 .

States of nature	A_1	A_2	A_3
E_1	25	-10	-125
E_2	400	440	400
E_3	650	740	750

The probabilities of the states of nature are 0.1, 0.7, 0.2 respectively. Find the optimum decision using EMV criterion.

Ans: Optimum decision at A_2

6. An investor is given the following investment alternatives and percentage rate of return.

State of Nature	Course	of	Action
	Regular Share	Risky Share	Property
Good	7%	-10%	-12%
Better	10%	12%	18%
Best	15%	18%	30%

Over past 100 days, 50 days have market in better condition and for 20 days market conditions are best. Using this information state optimum investment strategy for the investment. Use EMV criterion.

Ans: Invest in property.

7. A retailer purchases grapes every morning for Rs. 50/- a case and sells for Rs. 80/- a case. Unsold cases at the end of the day are donated to old age homes. Past sales have ranged from 15 to 18 cases per day. The following are the details of demand and probability.

Cases demanded :	15	16	17	18
Probability :	0.1	0.2	0.4	0.3

The retailer wants to know how many cases should he buy to maximize his profit. Use EMV criterion.

Ans: 17 cases

8. Following is the pay-off matrix corresponding to four states of nature S_1, S_2, S_3, S_4 and four courses of action A_1, A_2, A_3, A_4 .

State of Nature	Course of Action				Probability of State
	A_1	A_2	A_3	A_4	
S_1	50	400	-50	0	0.15
S_2	300	0	200	300	0.45
S_3	-150	100	0	300	0.25
S_4	50	0	100	0	0.15

- (i) Calculate expected pay off and find best course of action using EMV.
(ii) Calculate EOL for each course of action hence find best action using EOL.

Ans: (i) A_4 , (ii) A_4

9. Following is pay-off table corresponding to four acts A_1, A_2, A_3, A_4 and four states of nature E_1, E_2, E_3, E_4 with the probability of the events of this table $P(E_1) = 0.20, P(E_2) = 0.15, P(E_3) = 0.40, P(E_4) = 0.25$. Calculate the expected pay off and expected opportunity loss and suggest best course of action.

Acts	Events			
	E_1	E_2	E_3	E_4
A_1	40	50	200	0
A_2	300	200	0	100
A_3	50	100	40	200
A_4	300	0	100	50

Ans: A_2

10. The florist shop promises its customers delivery within three hours on all orders. All flowers are purchased the previous day and delivered to

florist by 8.00 a.m. the next morning. Demand distribution of roses is as follows.

Dozens of roses	7	8	9	10
Probability	0.1	0.2	0.4	0.3

Florist purchases roses for Rs. 10.00 per dozen and sells them at Rs. 30 per dozen. All unsold roses are distributed in a local hospital free of cost. How many dozens of roses should the florist order each evening?

Ans: 9 dozens of roses

11. The probability of the demand for tourist cars for hiring on any day at ABC travels and tours are as under:

No. of cars demanded	0	1	2	3	4
Probability	0.1	0.2	0.3	0.2	0.2

The cars have a fixed cost of Rs. 900 each day to keep and the daily hire charges are Rs. 2000. (i) If the tour owner company owns 4 cars what is its daily expectation? (ii) Use EMV and EOL criterion to suggest how many cars the company should keep.

Ans: (i) Rs. 800/-, (ii) 2 cars.

12. Unique home appliances finds that the cost of holding a cooking ware in stock for a month is Rs. 200. Customer who cannot obtain a cooking ware immediately tends to go to other dealers and he estimates that for every customer who cannot get immediate delivery he loses an average of Rs. 500. The probabilities of a demand of 0, 1, 2, 3, 4, 5 cooking ware in a month are 0.05, 0.1, 0.2, 0.3, 0.2, 0.15 respectively. Determine the optimum stock level of cooking wares. Using EMV criterion.

Ans: 4 Cooking wares

13. A person has the choice of running a hot snack stall or an ice cream stall at a certain holiday hotel during coming season. If weather is cool and rainy he can expect to make a profit of Rs. 1,50,000 and if it is warm he can expect to make profit of Rs. 40,000 by running hot snack stall. If he runs ice-cream and cold drink shop he can make a profit of Rs. 1,60,000 if the weather is warm and only Rs. 30,000 if whether is cool and rainy. The odds in favour of warm weather are 2 : 3 and that of having cool and rainy are 3 : 2. Use EMV to find Best f Action:

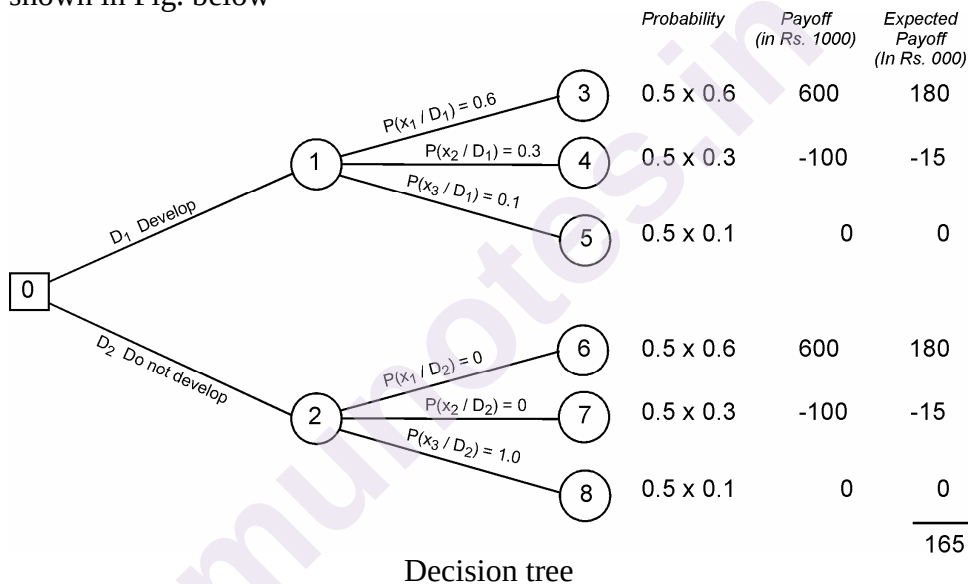
Ans: Hot snack stall.

14. You are given the following estimates concerning a Research and Development programme:

Decision Di	Probability of decision Di given research R $P(D_i \setminus R)$	Outcome number	Probability of outcome X_i given D_i $P(X_i \setminus D_i)$	Payoff value of outcome, X_i (Rs '000)
Develop	0.5	1	0.6	600
		2	0.3	-100
		3	0.1	0
Do not develop	0.5	1	0.0	600
		2	0.0	-100
		3	1.0	0

Construct and evaluate the decision tree diagram for the above data. Show your working for evaluation.

The decision tree of the given problem along with necessary calculation is shown in Fig. below

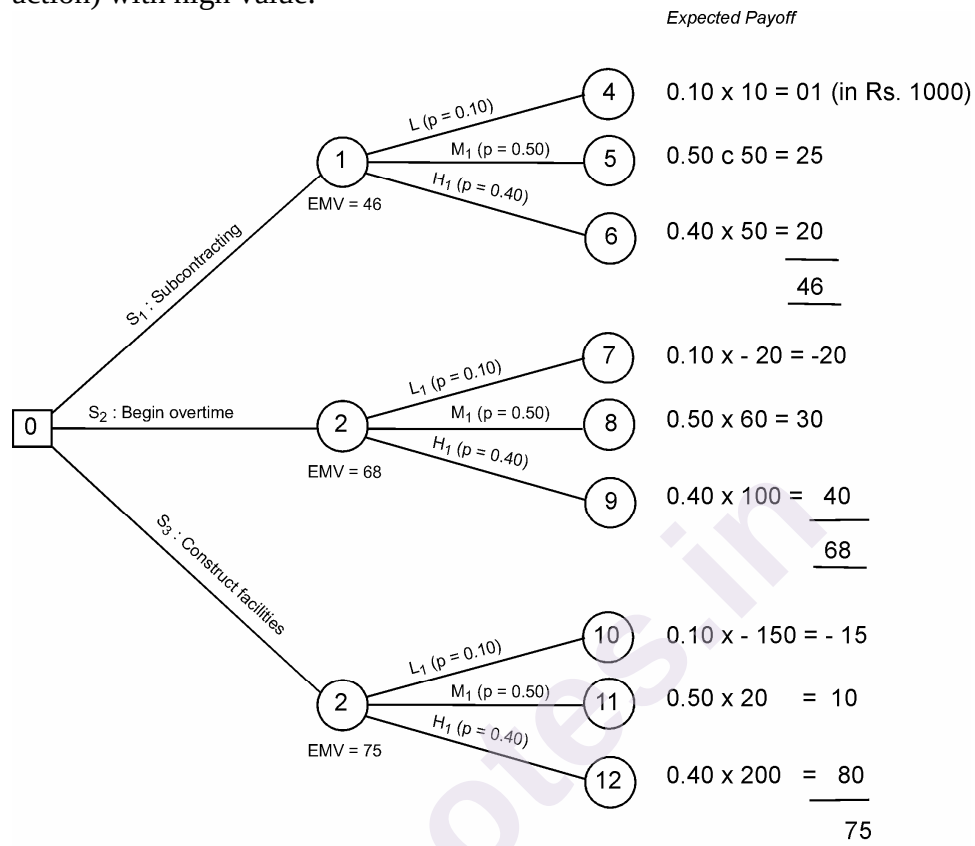


15. A glass factory specializing in crystal is developing a substantial backlog and the firm's management is considering three courses of sub-contracting (S1), being overtime production (S2), and construct new facilities (S3). The correct choice depends largely upon future demand which may be low, medium, or high. By consensus, management ranks the respective probability as 0.10, 0.50 and 0.40. A cost analysis reveals effect upon the profits that is shown in the table below:

Show this decision situation in the form of a decision tree and indicate the most preferred decision and corresponding expected value.

A decision tree which represents possible courses of action and nature are shown in the Fig. In order to analyse the tree we start working backward from the end branches.

The most preferred decision at the decision node 0 is found by calculating expected value of each decision branch and selecting the path (course of action) with high value.



Decision tree

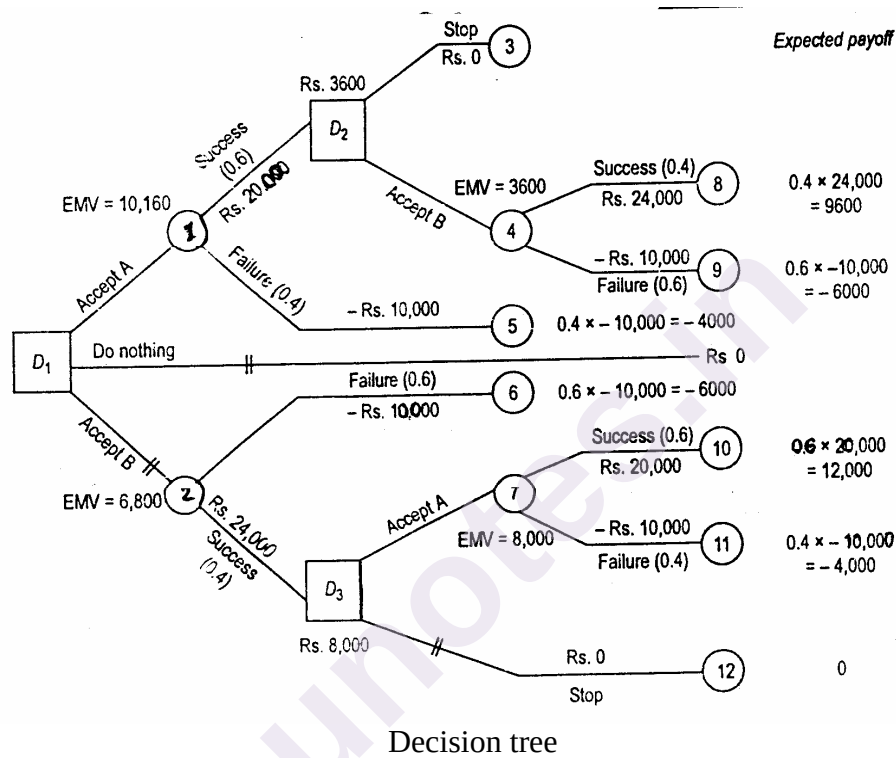
	Decision Point	Outcome	Probability	Conditional value (Rs.)	Expected value
D ₃	(i) Accept A	Success	0.6	20,000	12,000
		Failure		-10,000	-4,000
	(ii) Stop	-	-	-	Rs. 8,000 0
D ₂	(i) Accept B	Success	0.4	24,000	9,600
		Failure	0.6	-10,000	-6,000
	(ii) Stop	-	-	-	Rs. 3,600 0
D ₁	(i) Accept A	Success	0.6	20,000+3,600	14,160
		Failure	0.4	-10,000	-4,000
					Rs. 10,160
	(ii) Accept B	Success	0.4	24,000+8,000	12,800
		Failure	0.6	-10,000	-6,000
	(ii) Do nothing	-	-	-	Rs. 6,800 0

Since node 3 has the highest EMV, therefore, the decision at node 0 will be to choose the course of action S₃, i.e., construct new facilities.

16. A person wants to invest in two independent investment schemes: A and B, but he can undertake only at a time due to certain constraints. He can choose A first and then stop, or if A is not successful then B or vice-versa. The probability of success of A is 0.6, while for B it is 0.4. The

investment in both the schemes requires an initial capital outlay of Rs. 10,000 and both return nothing if the venture is unsuccessful. Successful completion of A will return Rs. 20,000 (over cost) and successful completion of B will return Rs. 24,000 (over cost). Draw decision tree and determine the best strategy.

The decision tree corresponding to the given information is depicted in the Fig.



Since EMV = Rs 10,160 at node 1 is the highest the best strategy at node D1 is to accept course of Action A first and if A is successful then Accept B.



QUESTION PAPER PATTERN**MARKS:- 100****TIME:- 3 HRS****N.B :**

- 1) All questions are compulsory
- 2) All question carry equal marks
- 3) Figures to the right indicate marks to a sub-question.
- 4) Graphs paper will be supplied on request.
- 5) Use of non-programmable calculator is allowed.

SECTION-I

- Q.1 Attempt any four of the following
(a) 5 marks (b) 5 marks (c) 5 marks
(d) 5 marks (e) 5 marks 20 marks
- Q.2 Attempt any four of the following
(a) 5 marks (b) 5 marks (c) 5 marks
(d) 5 marks (e) 5 marks 20 marks

SECTION-II

- Q.3 Attempt any four of the following
(a) 5 marks (b) 5 marks (c) 5 marks
(d) 5 marks (e) 5 marks 20 marks
- Q.4 Attempt any four of the following
(a) 5 marks (b) 5 marks (c) 5 marks
(d) 5 marks (e) 5 marks 20 marks
- Q.5 Attempt any four of the following
(a) 5 marks (b) 5 marks (c) 5 marks
(d) 5 marks (e) 5 marks 20 marks

