UNIT I

COMMISSION, BROKERAGE, DISCOUNT AND PARTNERSHIP

OBJECTIVES

Money plays an important role in any transaction. We are going to study some Technical terms involved in business transactions.

1.1 COMMISSION

A producer or manufacturer of goods generally does not sell his goods directly to the ultimate consumer. There are agents who purchase the goods from the manufacturer and sell them to the consumer. In a sense, such agents bring the manufacturers and the consumers together for transaction. The remuneration which an agent gets for his services in the transaction is called *commission*. Most of the business transactions are made through intermediate persons.

1.2 RATE OF COMMISSION

The amount of commission that an agent gets in the transaction depends on the volume of work done or the services rendered by him. His commission is based on the value of the goods bought or sold and is generally fixed on a percentage basis. In some cases, he is paid a commission on the total sales brought by the agent or on different slabs. For example, it may be 5% on the first Rs. 10,000/-, 6% on the next Rs. 5,000/- and so on.

1.3 TYPES OF COMMISSION AGENTS

On account of the different fields of specialization in business activities agents can be differentiated as such as commission agents, brokers, del credere agents, factors, insurance agents, export agents, auctioneers, etc. Agents particularly salesmen are also paid regular salary in addition to commission they earn.

**Commission Agent**: A commission agent is that middle-man who buys or sells goods on behalf of some other person called as principal. He is usually employed by his principal and gets commission as some percentage of the sale value.

**Broker**: A broker is that middle-man who brings together a prospective seller and a prospective buyer and negotiates the sale between them. The
commission that he gets is called brokerage, which may be charged from both the parties. Accordingly, there are stock-brokers, producer, brokers, bullion-brokers, bill-brokers, insurance-brokers depending upon the field of business in which they work.

**Factor:** A factor is an agent who takes possession of the goods and then sells them in his own name. A factor does the transaction without disclosing the name or identity of his principal. He receives payment in his own name and passes receipt for the same.

**Del Credere Agent:** A del credere agent sells goods and at the same time gives a guarantee of the collection of dues from the consumers. In return for this guarantee, he gets an additional commission known as del creders. This commission may be at a flat rate on all the sales or at a higher rate on credit sales only.

**Auctioneer:** A person who wants to sell his goods or property or machinery approaches such an agent. The agent then gives an advertisement in which he gives some description of the goods and the date of sale. The agent sells goods to the highest bidder, i.e. the person who offers to pay the highest amount in the auction. Such an agent is called an auctioneer. Generally, the auctioneer does not disclose the name of the principal.

**Solved Examples:**

**Example 1**
An agent sold Rs. 3,000 worth of articles on 4½% commission basis. Find the commission of the agent.

**Solution:**
The agent sold Rs. 3,000 worth of articles on 4½% commission basis
Agent’s commission = 4½ of Rs. 3,000

\[
\text{Agent’s commission} = \frac{9}{2} \times \frac{1}{100} \times 3000
\]

\[
= Rs. 135
\]

**Example 2**
An agent was paid Rs. 1596 as commission at rate 12% on the sale of bicycles. The selling price of each bicycle was Rs. 950. Find the number of bicycles sold by the agents.

**Solution:**
Let \( \chi \) be the number of bicycle sold by the agent.
Price of one bicycle = Rs. 950
Selling price of \( \chi \) Bicycles = 950 \( \chi \)
\[
@12\% \text{ commission on } 950 x = \frac{12}{100} \times 950 x = 114 x
\]

Agent’s commission
\[
\begin{align*}
114 x &= 1596 \\
114 x &= 1596 \\
x &= \frac{1596}{114} = 14
\end{align*}
\]

Number of Bicycles sold = 14

**Example 3**
An insurance agent gets commission of 20\% on first year premium, 6\% on second and third year’s premium and 4\% on subsequent years premium on an insurance policy of Rs. 40,000. Annual rate of premium being Rs. 30 per thousand. Find the total earning of the agent for which 5 annual premiums have been paid.

**Solution:**
The rate of annual premium is Rs. 30 per thousand.
\[
\text{Annual premium} = \frac{30}{1000} \times 40000 = 1200
\]

\[\therefore \text{Commission for first year} = \frac{20}{100} \times 1200 = 240\]

Commission for second and third year = \(2 \left( \frac{6}{100} \times 1200 \right) = 2 (72) = 144\)

Commission for fourth and fifth year = \(2 \left( \frac{4}{100} \times 1200 \right) = 2 (48) = 96\)

Total Earnings = 240 + 144 + 96 = 480

= Rs. 480

**Example 4**
At 7\% rate of commission, a sales girl, got Rs. 210 on the sale of combs. Find the value of the sale if the price of each comb is Rs. 15 per container. Find the number of containers sold by the sales girl.

**Solution:**
Since the rate of commission is 7\% and she received Rs. 210,

total value of her sales = Rs. = Rs. 210 \(\times\) \(\frac{7}{100}\) = Rs. 3,000.

The price of each container is Rs. 15.

Total number of containers sold = \(\frac{3000}{15}\) = 200
Example 5
Find the commission on total sales worth Rs. 25,000, if the rate of commission is 3% on first Rs. 10,000 and 4½ on sales over Rs. 10,000.

Solution:
Total sales are worth Rs. 25,000.
Commission at 3% on first Rs. 10,000 = Rs. 10,000 × 3 = Rs. 300
Commission at 4½% on sales over Rs. 10,000 = Commission at rate 4½% on Rs. 15000.

\[
\frac{9}{2} \times \frac{1}{100} \times 15,000 = Rs. 675
\]
Total commission = Rs. (300 + 675) = Rs. 975

Example 6
A salesman is allowed 5% commission on the total sales made by him plus a bonus of 1% on the excess of his sale over Rs. 20,000. If the total earning are Rs. 1,450/- on commission alone, find his total earnings?

Solution: Let the total sales be Rs. x.
Then commission at 5% on Rs. x = \[\frac{5}{100} \times x\] .
This is given to be Rs. 1450/-
\[
\frac{x}{20} = 1450
\]
\[
x = 1450 \times 20
\]
\[
x = 29,000
\]
The amount of total sales is = Rs. 29,000/-
Excess of sales over Rs. 20,000 = Rs. (29,000 - 20,000) = Rs. 9,000/-
Bonus at 1% on Rs. 9,000 = Rs. \[\frac{1}{100} \times 9000\] = Rs. 90/-

Example 7
A piece of land was sold for Rs. 19,00,000 through a broker who received 1.25% commission from the seller and 1.75% from the buyer. Find the amount paid by the buyer. Also find the amounts received by the seller and the broker.

Solution: Brokerage paid by the buyer at 1.75% = 1.75% of Rs. 19,00,000 = Rs. \[\frac{1.75}{100} \times 19,00,000\] = Rs. 33,250/-
Total amount paid by the buyer = Cost of the land + Brokerage
= Rs. 19,00,000 + 33,250
= Rs. 19,33,250/-
Brokerage paid by the seller at 1.25% = 1.25% of Rs. 19,00,000 = Rs. \[\frac{1.25}{100} \times 19,00,000\] = Rs. 23,750/-
Total amount received by the seller = Cost of the land - Brokerage
= 19,00,000 - 23,750
Total amount received by the broker = Brokerage from the buyer + Brokerage from the seller
= Rs. (33,250 + 23,750)
= Rs. 57,000/

Example 8
A merchant instructs his agent to buy 1,000 micro tip pens and sell them at 15% above the purchase price. The agent charges 1% commission on the purchase and 3% commission on sales and earns Rs. 534/- as commission. Find the price at which the agent buys the pen.

Solution: Suppose the agent buys each pen for Rs. $x$.
Then cost price of 1000 pens = Rs. 1000 $x$

selling price of 1000 pens at 15% profit.

$$\text{Rs. } \frac{115}{100} \times 1000 \times x = \text{Rs. } 1150 \times x$$

commission at 1% on the purchase.

$$=1\% \text{ of Rs. } 1000 \times x = \text{Rs. } \frac{1}{100} \times 1000 \times x = \text{Rs. } 10 \times x \quad (i)$$

commission at 3% on the sales.

$$=3\% \text{ of Rs. } 1150 \times x = \text{Rs. } \frac{3}{100} \times 1150 \times x = \text{Rs. } \frac{69}{2} x \quad (ii)$$

From (i) & (ii)

Total commission of the agent

$$=\text{Rs. } 10 \times x + \frac{69}{2} \times x = \frac{89}{2} \times x$$

This is given to be Rs. 534/-

$$\frac{89}{2} \times x = 534 \quad x = \frac{532 \times 2}{89} = 12$$

The agent buys each pen for Rs. 12/-

Example 9
A del credere agent charges 3% commission on cash sales and 6% commission on credit sales. If his average commission is 4.3%, find the ratio of cash sales to credit sales.
Solution: Suppose the cash sales are Rs. \( x \) and the credit sales are Rs. \( y \).

Then commission at 3% on cash sales of Rs. \( x \) is \( \frac{3x}{100} \) and

Commission at 6% on credit sales of Rs. \( y \) is \( \frac{6y}{100} \)

Total commission of the agent \( = \frac{3x + 6y}{100} \) \( \text{(i)} \)

Average commission at 4.3% on total sales of Rs. \( (x + y) \)

\( = \frac{4.3(x + y)}{100} = \frac{43(x + y)}{1000} \) \( \text{(ii)} \)

From (i) & (ii)

\( \frac{43x + 43y}{1000} = \frac{3x + 6y}{100} \)

After simplification, \( 43x - 30x = 60y - 43y \)

\( 13x = 17y \)

\( x = \frac{17}{13}y \)

The ratio of cash sales to credit sales is 17:13

Example 10

A salesman is appointed on a fixed monthly salary of Rs. 1,500/- together with a commission at 5% on the sales over Rs. 10,000/- during a month. If his monthly income is Rs. 2,050/-, find his sales during that month.

Solution: Commission at 5% on the sales over Rs. 10,000/-

\[ = \text{monthly income - monthly salary} \]
\[ = \text{Rs. 2050-1500} \]
\[ = \text{Rs. 550/-} \]

the sales over Rs. 10,000/-

\[ = \text{Rs. 550} \times \frac{100}{5} = \text{Rs. 11,000/-} \]
total sales during the month

\[ = \text{Rs. (10,000 + 11,000)} = \text{Rs. 21,000/-} \]

Example 11

A salesman is paid a monthly salary plus a commission based as a percentage on sale. If on the sales of Rs. 20,000/- and Rs.25,000/- in two successive months he received Rs. 1600/- and Rs.1750/- respectively, find his monthly salary and the rate of commission paid on sales.

Solution: Let the monthly salary of the salesman be Rs. \( x \) and \( y \% \) be the rate of commission.
commission in first month at \( y \) % on Rs. 20,000/-

\[
= Rs. \frac{y}{100} \times 20000 = Rs. 200y
\]

For the first month total earnings = salary + commission in the first month

\[
1600 = Rs. (x + 200y) \quad \text{…… (1)}
\]

Also commission in the second month at \( y \) % on Rs. 25,000/-

\[
= Rs. \frac{y}{100} \times 25000 = Rs. 250y
\]

For the second month total earnings = salary + commission in the second month

\[
1750 = Rs. (x + 250y) \quad \text{…… (2)}
\]

Subtracting (1) from (2), we get, \( 50y = 150 \) \( \therefore y = 3 \)

Substituting this value of \( y \) in (1), we get,

\[
x + 200 \times 3 = 1600 \quad \therefore x = 1000
\]

The monthly salary of the salesman is Rs. 1000/- and the rate of commission is 3%.

Example 12
At what price should goods costing Rs. 18,000 be sold through an agent so that after paying her a commission at 4% on sales, a net profit of 20% on cost can be made?

Solution: Let the selling price be Rs. \( x \).

After deducting 4% commission, the net receipt

\[
= 96\% \text{ of } x = \frac{96}{100} x
\]

Now, profit at 20% of Rs. 18,000/-

\[
= \frac{20}{100} \times 18000 = Rs. 3600/-
\]

Net receipt = Cost + Profit

\[
= Rs. (18000 + 3600) = Rs. 21600/-
\]

Thus \( \frac{96}{100} x = 21600 \)

\[
x = \frac{21600 \times 100}{96} = Rs. 22500/-
\]

Hence, the goods should be sold for Rs. 22,500/-

Example 13
A salesman gets 4% commission on the first Rs. 20,000, 7% on the next Rs. 20,000 and 12% on the excess. He also receives an incentive at the
rate of 3% on total sales, if it exceeds Rs. 50,000. Find the total earning of two salesmen with total sales worth of Rs. 44,000 and Rs. 58,000.

Solution:

(i) For the first salesman
Sales is Rs. 44,000 = 20,000 + 20,000 + 4,000
So, commission = 4% of 20,000 + 7% of 20,000 + 12% of 4,000
= 800 + 1400 + 480 = Rs. 2680
As his sales of Rs. 44,000 is less than Rs. 50,000, he will not get any incentive amount.
So, total earning of the first salesman = Rs. 2680

(ii) For the second salesman
Sales is Rs. 58,000 = 20,000 + 20,000 + 18,000
So, commission = 4% of Rs. 20,000 + 7% of Rs. 20,000 + 12% of Rs. 18,000
= 800 + 1,400 + 2,160
= 4360
His sales is Rs. 58,000, which exceeds Rs. 50,000
So, incentive amount = 3% of 58,000 = 1,740
So, total earnings of second salesman = commission + incentive
= 4360 + 1740
= 6,100

Example 14
A merchant gained a net gain of 18.75% on his cost of Rs. 20,000 after giving the agent a commission at 5% on the sale price. Find the sale price.

Solution: Let $x$ be the sale price

\[
\text{Commission} = 5\% \text{ on sale} = \left(\frac{5}{100}\right)x = 0.05x
\]

\[
\text{Gross Profit} = \text{Sale Price} - \text{Cost Price} = x - 20,000
\]

\[
\text{Net Profit} = \text{Gross Profit} - \text{Commission} = (x - 20,000) - 0.05x = 0.95x - 20,000
\]

But Net Profit = 18.75% on Cost Price

\[
= \frac{18.75}{100} \times 20,000
\]

\[
= 3,750
\]

\[
\therefore \quad 0.95x - 20,000 = 3,750
\]

\[
\therefore \quad 0.95x = 23,750
\]

\[
\therefore \quad x = \frac{23,750}{0.95} = 25,000
\]

\[
\therefore \quad \text{The sale price is Rs. 20,000}
\]
Example 15
A merchant employed an agent to buy and sell some article. The agent charged 3% on the purchase value and 2% on the sale value as commission. The purchased value was Rs. 40,000. The merchant after deducting the commission still received a net profit of 19.5% on cost. Find the sale price.

Solution: Let x be the sale price.

Commission on purchase = 3% on purchased price
= \frac{3}{100} \times 40,000
= 1,200

Commission on sales = 2% on sale price
= 0.02x

\therefore \text{Total Commission} = \text{Commission on purchase} + \text{Commission on Sale}
= 1,200 + \frac{2}{100}x

Net Profit = 19.5\% on Purchased Price
= 19.5\% on 40,000 = Rs. 7,800

But \ Net Profit = \text{Sale Price} – \text{Purchase Price} – \text{Total Commission}
7,800 = x – 40,000 – (1,200 + 0.02x)
7,800 = x – 40,000 – (1,200 – 0.02x)
7,800 = x – 0.98x – 41,200
49,000 = 0.98x
\therefore \frac{49000}{0.98} = x
\therefore 50,000 = x
\therefore \text{The sale price was Rs. 50,000}

Example 16
At what price should goods costing Rs.16,000 be sold through an agent so that after paying her a commission at 4\% on sales, a net profit of 20\% on cost can be made?

Solution: Let gross Receipts be Rs. x
After deducting 4\% commission, the net receipt
= 96\% of x = \frac{96}{100}x

Now, profit at 20\% of Rs. 16,000
= \frac{20}{100} \times 16,000 = Rs. 3,200

\therefore \text{Net receipt} = \text{Cost} + \text{Profit}
= Rs. (16,000 + 3,200)
= Rs. 19,200
Thus \( \frac{96}{100} x = 19,200 \)

\[ x = 19,200 \times \frac{100}{96} = \text{Rs. 20,000} \]

Hence, the goods should be sold for Rs. 20,000

Check your progress

1) An agent earns \( 6 \frac{2}{3} \% \) commission on his total sales which are of Rs. 11,520. Find the amount of his commission.  
   **Ans. Rs. 768**

2) An insurance agent gets commission of 25% on first year premium, 7% on second and third years premium and 5% on subsequent years premium on an insurance policy of Rs. 30,000, annual rate of premium being Rs. 40 per thousand. Find the total earning of the agent for which 5 annual premiums have been paid.  
   **Ans. Rs. 588**

3) An insurance agency pays 20% of annual premium as commission to its agent in the first years, 9% in the second and the third year and 7.5% in each subsequent year. A customer insured her car for Rs. 90,000 through an agent and her annual premium was fixed at Rs. 5,600. Find the agent’s total commission amount if the customer has paid annual premium for six year up to now.  
   **Ans. Rs. 3,388**

4) An insurance company pays its agent a commission at 20% on the first year’s premium amount. The rate of commission reduces to 6% for the subsequent years. A customer purchased a policy premium at 5% of the policy amount. Find the agent’s total earning for the premium for the period of seven years.  
   **Ans. Rs. 3,200**

5) A merchant pays 10% commission on total sales and pays del credere at a rate of 3% on credit sales. If cash sales were Rs. 4,500 and credit sales were Rs. 7,000 find the total commission earned by the agent.  
   **Ans. 1,360**

6) A merchant instructs his agent to buy 500 Raincoats and to sell them 20% above the purchase price. The agent charges 1% commission on purchase and 2% commission on the sales and earns Rs. 1,360. Find the price at which the agent buys a Raincoat.  
   **Ans. 80**

7) An agent was instructed to sell 5000 cotton shirts on 2% commission and invest the balance after deducting commission in purchasing kurtas. The commission paid on purchase is \( 3/4 \% \) If the agent earns together Rs. 3550.50 in two transactions, find the price at which a shirt was sold.  
   **Ans. Rs. 26**
8) A del credere agent charges 3% on cash sales and 6% on credit sales. If
the agent earns at an average rate of 4.7% on total sales, find the ratio of
cash sales to the credit sales.  
\textbf{Ans.} 13 : 17

9) A salesman is paid a fixed monthly salary plus commission at a certain
rate on sales. The salesman received Rs. 1,130 and Rs. 1,360 as
remuneration for two successive months and his sales were Rs. 17,100 and
21,700 respectively. Find the fixed monthly salary and the rate of
commission.  
\textbf{Ans.} 275; 5%

10) A del credere agent charges 4% commission on cash sales and 7%
commission on credit sales. If the agent earns at an overall rate of 5.4% of
total sales, find the ratio of cash sales to credit sales.  
\textbf{Ans.} 8 : 7

11) A salesman is paid fixed monthly salary together with commission at a
certain rate on his total sales during the month. The salesman received Rs.
1,256 and Rs. 1,372 as remuneration for two successive months during
which his total sales were Rs. 21,400 and Rs. 24,300 respectively. Find the
fixed monthly salary and the rate of commission.  
\textbf{Ans.} Rs. 400 and 4%

12) An agent was paid Rs. 21,708 as commission at a rate 9% on the sale
of Refrigerators. The selling price of each Refrigerator was 13,400. Find
the number of Refrigerators sold by the agent.  
\textbf{Ans.} 18 Refrigerators

13) A piece of land was sold for Rs. 19,00,000 through a broker who
received 1\textfrac{1}{4} \text{ }\% \text{ commission from the seller and } 1\textfrac{3}{4} \text{ }\% \text{ commission from
the buyer. Find the amount paid the buyer. Find also the amount received
by the seller and the broker}  
\textbf{Ans.} 19,33,250, Rs. 18,76,250, Rs. 57,000

14) A merchant instructs his agent to buy 600 pieces of an article and sell
at 25\% above the purchase price. The agent charges \textfrac{1}{2} \% \text{ commission on
the purchase and } 3\% \text{ commission on the sale and earns Rs. 1,020. Find the
price at which the agent buys an article.}  
\textbf{Ans.} Rs. 40

15) A del credere agent charges 4\% on cash sales and 7\% on credit sales.
If the consolidated rate at which the agent earns is 5\% of total sales find
the ratio of cash sales to the credit sales.  
\textbf{Ans.} 2 : 1

16) A salesman is paid a fixed monthly salary together with commission at
a certain rate on the sales. The salesman received Rs. 1,240 and Rs. 1,184
as remuneration for two successive months in which his total sales were
Rs. 18,500 and Rs. 17,100 respectively. Find the rate of commission and
the fixed monthly salary.  
\textbf{Ans.} 4\% \text{ Rs. 500}
17) A company fixed the rate of commission to its salesman as follows: 4% on the first Rs. 5,000, 5% on the next Rs. 8,000, 7% on the next Rs. 10,000 and 11% on the balance. Company has agreed to pay ¼ % of the total sales as bonus if the sales crossed Rs. 30,000. A salesman of the company secured sales worth Rs. 29,000. Calculate total earning of the salesman.  

**Ans.** Rs. 1,960

18) A merchant asked his agent to sell 250 hats at 2% commission and to invest the balance in purchasing ties. The agent charged 1 1/2% commissions on both the transactions. Find the price for which a hat was sold.  

**Ans.** Rs. 60

19) A merchant asked his agent to sell 200 caps at Rs. 15 per cap and to purchase 400 towels at Rs. 10 per towel. The agent charged 2% on the sales and 1% on the purchase. How much extra money did the merchant have to pay to the agent?  

**Ans.** Rs. 1,110

20) A merchant asked his agent to sell 40 music records at Rs. 25 each and use the money to buy some toys at Rs. 10 each. The agent quoted 15% commission on the sale and 6.25% on the purchase. The merchant instructed the agent to adjust the number of toys such that the merchant would not need to give him any extra money. How many toys would the agent buy?  

**Ans.** 800

21) The income of a salesman remains unchanged through the rate of commission is increased from 4% to 5%. Find the percentage reduction in his sales.  

**Ans.** 20%

22) At what price should goods costing Rs.48,250 be sold through an agent so that after paying him a commission at 3 1/2%, a net gain of 20% on the cost may be made?  

**Ans.** Rs. 60,000

### 1.5 DISCOUNT

During a festival season you must have seen advertisement banners displayed in front of each shops offering discount on the purchases made by the customer. Such advertisements are published in the newspapers also. Each shop owner proclaims to sell his goods at 20%, 15%, 30% discount etc.

Discount means reduction in price or less than marked price. Marked price /printed price/Catalogue price/List price is the price quoted on the article for sale.
Trade Discount is allowed on the list price (L.P)
List Price - Trade Discount = Invoice Price
Invoice Price - Cash Discount = Net selling price
Net selling price – Profit = cost

**Example 17**
Sachin buys a kurta in a Khadi Bhandar marked at Rs. 400. If a discount of 18% is allowed, find the discount.

**Solution:**
Marked price = List price (L.P.) = Rs. 400
Rate of Discount = 18% on L.P.
Discount = 18% on L.P.
Discount = 18% on 400 = $\frac{18}{100} \times 400 = Rs.72$

**Example 18**
Priya purchased a table whose marked price is Rs.7,000. The dealer allowed 6% discount on it. Find (i) discount on Rs.7,000 (ii) amount paid by Priya for the table.

**Solution:**
(i) Discount = 6% on L.P.

$$= 6\% \text{ on } Rs. \ 7000 = \frac{6}{100} \times 7000$$

= Rs. 420

(ii) Amount paid by Priya for the table (or) S.P.

$$= \text{L.P. – Discount} = 7,000 – 420$$

= Rs. 6,580

**Example 19**
A furniture maker sells a wooden table listed Rs.2,500 for Rs.2,200. Find the rate of discount.

**Solution:**
Listed price = Rs. 2,500
S.P = Rs.2,200

.$\therefore$ Discount = L.P. – S.P. = 2,500 – 2,200

= Rs. 300

Rate of discount = $\frac{\text{Discount} \times 100}{\text{L.P.}}$

= $\frac{300 \times 100}{2,500} = 12\%$
1.6 TRADE DISCOUNT and CASH DISCOUNT

When a manufacture or a trade offers a discount to another trader, the discount offered is usually termed as the trade discount. For example, when a wholesaler sells goods in bulk to a retailer, he / she gives a trade discount. It is also sometimes called as the bulk discount. We shall refers to this as the trade discount and write it as T.D.

The rate of Trade Discount is expressed as percentage of the list price. After deducting the trade discount, the remaining amount is what is put down in the invoice or the bill. Hence this is referred to as the invoice price or the reduced list price or the amount of the bill. We will refer to it as the invoice price and denote it as I.P.

Thus \( I.P. = L.P. - T.D. \) Amount.

Even after giving the above trade discount, the seller may offer an extra discount to those buyers who pay immediately by cash and do not ask for credit. This discount is called cash discount and we will denote it as C.D.

The rate of cash discount is expressed as a percentage of the invoice price. Note that the rate of cash discount is not expressed as a percentage of the list price.

After deducting the cash discount from the invoice price, we get the Net Selling Price (N.S.P) = I.P. - C.D.

This can be remembered in the following way:

\[
\text{L.P.} \quad \downarrow \quad \downarrow \quad \text{I.P.}
\]

One discount is given \hspace{1cm} Two discounts given T.D, C.D.

<table>
<thead>
<tr>
<th>L.P. = Rs ( )</th>
<th>L.P. = Rs. ( )</th>
</tr>
</thead>
<tbody>
<tr>
<td>Discount = ( ) % on L.P.</td>
<td>T.D. = ( ) % L.P.</td>
</tr>
<tr>
<td></td>
<td>C.D. = ( ) % ON L.P.</td>
</tr>
<tr>
<td></td>
<td>S.P. = I.P. – C.D.</td>
</tr>
</tbody>
</table>

**Example 20**

A firm allows 20% discount on list price and a further 5% discount for cash payment. What is the net selling price of an article which is marked at Rs. 240

**Solution:** Given TD% = 20, CD% = 5 and L.P. = Rs. 240.

\[
I.P. = L.P. - \text{Trade Discount Amount} = L.P. - 20\% \text{ of L.P.}
\]
\[ 240 - 20\% \text{ of } 240 = 240 - \frac{20}{100} \times 240 = 240 - 48 \]

**Example 21**

A firm allows 25% of trade discount and a further discount for cash payment at 10% rate. Find the list price of an article with a net selling price of Rs. 270.

**Solution:**

Given that T.D. % = 25, CD% = 10 and N.S.P. = Rs 270.

\[ \text{N.S.P.} = \text{I.P.} - \text{Cash Discount Amount} \]

\[ 270 = \text{I.P.} - 10\% \text{ of } \text{I.P} = 90\% \text{ of } \text{I.P.} = \frac{90}{100} \times \text{I.P.} \]

\[ \frac{100}{90} \times 270 = \text{I.P.} \]

\[ \therefore \text{Rs. 300} = \text{I.P.} \]

Now, \[ \text{I.P.} = \text{L.P.} - \text{Trade Discount Amount} \]

\[ 300 = \text{L.P.} - 25\% \text{ of } \text{L.P.} = 75\% \text{ of } \text{L.P.} = \frac{75}{100} \times \text{L.P.} \]

\[ \frac{100}{75} \times 300 = \text{L.P.} \]

\[ \therefore \text{Rs. 400} = \text{L.P.} \]

Hence the list price is Rs. 400.

**Example 22**

An article was listed at Rs. 5,000. It was sold by the shop owner at a 30% discount and thus a profit of 75% on cost was earned. Find the cost price of the shop owner.

**Solution:**

Given that Discount% = 30, Profit% = 75 and L.P. = Rs. 5000

\[ \text{N.S.P.} = \text{L.P.} - \text{Discount Amount} = \text{L.P.} - 30\% \text{ of } \text{L.P.} \]

\[ = 70\% \text{ of } \text{L.P.} = \frac{70}{100} \times 5,000. = \text{Rs. 3,500.} \]

Now, \[ \text{N.S.P} = \text{C.P.} + \text{Profit} \]
\[ 3500 = \text{C.P.} + 75\% \text{ of C.P.} \]
\[ = 175\% \text{ of C.P.} \]
\[ \therefore 3500 = \frac{175}{100} \times \text{C.P.} \]
\[ \therefore \frac{100}{175} \times 3500 = \text{C.P.} \]
\[ \therefore \text{Rs. 2,000} = \text{C.P.} \]
Hence the cost price is Rs. 2,000.

Example 23
An article was sold for a net selling price of Rs. 6,225 after giving a 17\% discount on the list price and thus a 24.5\% profit was gained on cost. Find the list price and the cost price.

Solution:
Given that N.S.P. = Rs 6,225, Discount = 17\% and profit = 24.5\%  
N.S.P. = L.P. – Discount Amount  
6,225 = L.P. – 17\% of L.P  
\[ = 83\% \text{ of } L.P. = \frac{83}{100} \times L.P. \]
\[ \frac{100}{83} \times 6,225 = L.P. \]
Rs. 7,500 = L.P.
Also, N.S.P = C.P. + Profit  
6,225 = C.P. + 24.5\% of C.P.  
6,225 = 124.5\% \text{ of } C.P. = \frac{124.5}{100} \times C.P.  
\[ \frac{100}{124.5} \times 6,225 = C.P. \]
\[ \therefore \text{Rs. 5,000} = \text{C.P.} \]
The list price is Rs. 7,500 and the cost price is Rs.5,000.

Example 24
A mall gave a 10\% discount on a watch and still gained 62\% profit. How much per cent above cost had the mall marked the watch for sale?

Solution:
Given that discount\% = 10 and Profit \% = 62. Since no price is supplied, we assume C.P. = 100
\[ \therefore \text{N.S.P} = \text{C.P.} + \text{Profit} = 100 + 62\% \text{ of } 100 \]
\[ = 100 + 62 = 162 \]
Also, N.S.P = L.P. – Discount Amount
162 = L.P. – 10% of L.P = 90% of L.P.
162 = \frac{90}{100} \times L.P.
\frac{100}{90} \times 162 = L.P.
∴ Rs. 180 = L.P.

Example 25
What percentage of profit is earned by marking goods 60% above cost and then giving a 8% discount?

Solution:
Since no price is given, we assume C.P. = 100
Now L.P. is 60% above C.P.
i.e. L.P = 160% of C.P. = C.P. + \frac{60}{100} \times 100 = 100 + 60 = 160
∴ Discount = 8% on L.P. = \frac{8}{100} \times 160 = 12.80
N.S.P = L.P. – Discount Amount
= 160 – 12.80 = 147.20
Profit = N.S.P. – C.P. = 147.20 – 100
= 47.20
Thus, on C.P = 100, the profit is 47.20. Hence the Profit is 47.20%

Example 26
A trader gives 35% discount and makes 30% profit on cost. Due to recession, the cost price reduces by 20%. By what per cent should he lower the list price, if he wished to keep the profit percentage and the discount percentage same as before?

Solution:
Given the Discount % = 35 and Profit% = 30.
Let the cost price be 100 i.e. C.P. = 100.
N.S.P = C.P. + Profit = C.P. + 30% of C.P.
= 100 + \frac{30}{100} \times 100 = 100 + 30
= 130.
Now, N.S.P = L.P. – Discount Amount
130 = L.P. – 35% of L.P = 65% of L.P = \frac{65}{100} \times L.P.
∴ \frac{100}{65} \times 130 = L.P.
\[ 200 = L.P. \]
Now due to recession, C.P. reduces from 100 by 20% of 100.
i.e. C.P. reduces from 100 by 20
i.e. C.P. reduces from 100 to 100 – 20 = 80.
\[ \text{New C.P.} = 80. \]
The profit percentage is still 30 and the discount percentage is still 35.
\[ \text{New S.P.} = \text{New C.P.} + \text{New Profit} = \text{New C.P.} + 30\% \text{ of New C.P.} \]
\[ \text{New S.P.} = 80 + \frac{30}{100} \times 100 = 80 + 24 = 104. \]
Now, New N.S.P. = New L.P. – New Discount Amount
\[ \text{New N.S.P.} = 104 = \text{New L.P.} – 35\% \text{ of New L.P.} = 65\% \text{ of New L.P.} \]
\[ \frac{65}{100} \times \text{L.P.} = 104 \]
\[ \frac{100}{65} \times 104 = \text{New L.P.} \]
\[ \text{New L.P.} = 160 \]
L.P. has reduced from 200 to 160
L.P. has reduced by 40
We want to find the percentage reduction in L.P.
When L.P. = 200, the reduction = 40.
When L.P. = 100, the reduction = \( \frac{100}{200} \times 40 = 20 \)
\[ \text{He should reduce his list price by 20}\% \]

**Example 27**
An article has a catalogue price of Rs. 400. The seller gave 20% trade discount and a further 5% discount for cash payment and still managed to get a 52% profit on her cost. What was the cost price?

**Solution:**
Given L.P. = Rs 400, T.D.\% = 20, C.D. \% = 5 and profit \% = 52
L.P. = L.P. – T.D. Amount = 400 – 20\% of 400
\[ = 400 - \frac{20}{100} \times 400 = 400 - 80 = \text{Rs. 320} \]
N.S.P. = I.P. – C.D. Amount = 320 – 5\% of 320
\[ = 320 - \frac{5}{100} \times 320 = 320 - 16 = \text{Rs. 304} \]
Now, N.S.P. = C.P. + Profit
\[ 304 = \text{C.P.} - 52\% \text{ of C.P.} = 152\% \text{ of C.P.} \]
\[
\therefore \quad 304 = \frac{152}{100} \text{ C.P.}
\]
\[
\therefore \quad \frac{100}{152} \times 304 = \text{ C.P.}
\]
\[
\therefore \quad \text{Rs. 200} = \text{ C.P.}
\]

The cost of the price is Rs. 200

**Example 28**

A trader allows a trade discount of 8% on the list price of his goods and a further discount of 2% for cash payment and still gains 12.7% profit on this cost. How much percent above cost has he marked his goods for sale?

**Solution:**

Given that T.D% = 8, C.D.% = 2 and profit % = 12.7

Since no price is given, we assume C.P. = 100

\[
\text{N.S.P.} = \text{C.P.} + \text{Profit} = 100 + 12.7\% \text{ on } 100 = 100 + 12.7 = 112.7
\]

Also, \(\text{N.S.P.} = \text{I.P.} - \text{C.D. Amount}\)

\[
112.7 = \text{I.P.} - 2\% \text{ of I.P}
\]

\[
112.7 = 98\% \text{ of I.P.} = \frac{98}{100} \times \text{I.P.}
\]

\[
\frac{100}{98} \times 112.7 = \text{I.P.}
\]

115 = I.P.

Further, \(\text{I.P.} = \text{L.P.} - \text{T.D. Amount}\)

\[
115 = \text{L.P.} - 8\% \text{ of L.P.} = 92\% \text{ of L.P.} = \frac{92}{100} \times \text{L.P}
\]

\[
\frac{100}{92} \times 115 = \text{L.P.}
\]

125 = L.P.

The list price is 125 of goods whose cost is 100.

Thus, the goods are marked 25% above cost.

**Example 29**

An agent is instructed by the manufacture to allow the retailer a discount at the rate of 20% on the list price and received from the manufacture a commission at the rate of 8% on the net selling price. The Agent sold goods worth Rs. 20,000 as per the list price. Calculate the commission and the amount received by the manufacture.
Solution:

Discount Amount = 20% on L.P. = \(\frac{20}{100} \times 20,000\) = Rs. 4,000.

\[\therefore \text{N.S.P.} = \text{L.P.} - \text{Discount Amount} = 20,000 - 4,000 = \text{Rs. 16,000.}\]

Commission = 8% on N.S.P. = \(\frac{8}{100} \times 16,000\) = Rs. 1,280.

\[\therefore \text{Amount received by the manufacturer} = \text{N.S.P.} - \text{Commission} = 16,000 - 1,280 = \text{Rs. 14,720}\]

\[\therefore \text{The commission is Rs. 1,280 and the amount received by the manufacturer is Rs. 14,720.}\]

Example 30

A trader bought a computer for Rs. 50,000 and listed it for Rs. 65,000. He sold it through an agent, giving 4% trade discount and a further 1% cash discount. The agent charged 10% commission on the net selling price. Find the trader’s profit percentage.

Solution:

C.P. = Rs 50,000 and L.P. = Rs. 65,000.
T.D. % = 4, C.D.% = 1 and Commission% = 10

Now, T.D. Amount = 4% on L.P. = \(\frac{4}{100} \times 65,000\) = Rs. 2,600

L.P. = L.P. – T.D. amount = 65,000 – 2,600 = Rs. 62,400

C.P. Amount = 1% of L.P. = \(\frac{1}{100} \times 62,400\) = Rs. 624.

N.S.P. = I.P. – C.D. Amount = 62,400 – 624 = Rs. 61,776

Commission = 10% of N.S.P. = \(\frac{10}{100} \times 61,776\) = Rs. 6,177.60

Amount received by the trader = N.S.P. – Commission
= 61,776 – 6,177.60 = Rs. 55,598.40

Profit = Amount received by the trader – C.P. = 55,598.40 – 50,000
= Rs. 5,598.40

\[\text{Profit\%} = \frac{\text{profit}}{\text{C.P.}} \times 100 = \frac{5,598.40}{50,000} \times 100 = 11.1968 \approx 11.2\]

\[\therefore \text{The profit Percentage is 11.2\%}\]

Example 31

An agent is instructed to give a 10% cash discount for any cash payment. The agent’s commission is 2% on credit sales and 5% on cash sales on the net selling price. In a month, the agent sells goods worth Rs.
Maths_stats_fybcom

40,000 on credit and Rs. 50,000 for cash as per the list price. Find his total commission for the month.

**Solution:**
For credit Sales, there is no discount given, hence N.S.P. = L.P.

- Commission on credit sales = 2% on N.S.P. = 2% on L.P.

\[ = \frac{2}{100} \times 40,000 = \text{Rs} \ 800 \]

Now for the cash sales, Discount Amount = 10% on L.P.

Discount Amount = \[\frac{10}{100} \times 50,000 = \text{Rs} \ 5,000.\]

- N.S.P. = L.P. – Discount Amount = 50,000 - 5,000 = Rs. 45,000.

Commission on cash sales = 5% on N.S.P. = \[\frac{5}{100} \times 45,000 = \text{Rs} \ 2,250.\]

Total Commission = Commission on Credit Sales + Commission on Cash Sales

\[ = 800 + 2,250 = \text{Rs} \ 3,050\]

Thus, the agent’s total commission is Rs. 3,050.

**Check your progress**

1) The printed price of wrist watch is Rs. 880. But the factory supplied to a dealer at Rs 844.80 per watch. Find the rate of discount allowed on each wrist watch.

   **Ans. 4%**

2) The Marked price of radio is Rs. 700. The shopkeeper allows 12 \(\frac{1}{2}\) % discount on it. At what price does he sell the radio?

   **Ans. Rs. 612.5**

3) A manufacture sold utensils for net Price of Rs. 9025 after giving 5% discount on the list price. What was the list price?

   **Ans. Rs. 9500**

4) A firm allows a regular discount of 20% on the listed price and also a further 5% for cash payment. What is the selling price of the goods listed at Rs. 12,500?

   **Ans. Rs. 9,500**

5) A firm allows a trade discount at 25% and a further discount of 10% on cash payment. Find the net price of an article which is listed for Rs. 400.

   **Ans. Rs. 270**

6) An article is marked at Rs. 7,000. Trade discount of 25% is allowed and a further discount of 8% for cash payment is allowed. Find the net selling price of the article.

   **Ans. Rs. 4830**
7) A firm allows a trade discount at 25% and a further discount of 10% for cash payment. Find the selling price of an article which is marked for Rs. 2,000.

   **Ans. Rs. 1350**

8) Find the selling price of T.V. set which is marked for Rs. 10,000 after allowing at 20% trade discount and a 5.5% cash discount.

   **Ans. Rs. 7600**

9) 3% discount is offered on T.V. set whose list price is Rs. 21,000 and a further 1% is offered on cash payment. Find what will be the selling price by cash payment.

   **Ans: Rs. 20166.30**

10) A shopkeeper sold a T.V set for a net amount of Rs. 10,450 after allowing 20% discount on the list price and the further discount of 5% for cash payment. Find list price.

   **Ans. Rs.13750**

11) An article is marked at Rs. 1500. A trader allows a discount of 7% and still gains 20% profit. What was the cost price of the article?

   **Ans. Rs.1162.50**

12) A trader sold an article at 20% discount and still gets 25% profit on the cost. If the cost price was Rs. 1600, find the list price.

   **Ans. Rs.2500**

13) A supermarket gave a 10% discount on an article and thus gained 44% profit. How much percent above cost had the super market marked the article for sale.

   **Ans. 60%**

14) A trader marked an article 40% above his cost and even after giving a discount, gained 19% profit. What was the percentage of the discount given?

   **Ans. 15%**

15) A wholesaler allows 25% trade discount on the list price and earns a profit of 20% on cost. If the cost increases by 10%, by what percent should he raise the list price so as to make the same rate of profit, after allowing the usual trade discount of 25%?

   **Ans. 100%**

16) A trader bought an article for Rs. 80,000 and listed it for Rs. 1,00,000. His agent sold it by giving 14% discount o the list price and charging 2% commission on net selling price. Find traders profit percentage.

   **Ans.5.35%**
1.7 PARTNERSHIP

A form of organization where two or more persons called partners carry on a business together and agree to share profit or losses at a specified proportion is called partnership.

Very often for starting a large business or for expanding an already existing business, a huge capital is required. Many a time a single person does not possess such large amount and in this case, some persons come together and provide necessary capital. Thus the business is run by more than one person by putting capital and sharing the profits and risks at agreed proportion is called *Partnership* and persons who provide capital are called *partners*.

The amount of profit or losses to be shared by partners depends on many factors such as (1) amount invested by partners, (2) the duration for which an amount is invested by a partner, (3) knowledge of a partner regarding the business, (4) effort made by different partner for the business etc.

Goodwill of a business organization is the difference between the market value and the net worth. The value of the goodwill is a quality that keeps changing. The valuation of goodwill is also required to be done in the event of death of a partner for settlement of his dues.

The calculation of goodwill at the time of entry of a new partner or exit of an existing partner is based on average profit of a certain number of previous years.

When a partnership is dissolved, its assets are sold and total liabilities which include the capital are paid. The balance if any is the goodwill; this is shared in the profit sharing ratio. If liabilities are more than the sale proceeds then the losses are also shared in the profit sharing ratio.

**Solved Examples**

**Example 32**

A, B, C start a business by investing Rs. 20,000, Rs. 35,000 and Rs. 45,000 respectively and share the profit of Rs. 10,000 at the of the year. Find the share in profit of each partner.

**Solution:**

Capitals are different but period is the same.

\[\text{The profit of Rs. 10,000 is to be shared in the proportion of capital.}\]

The proportion of capital is \(20,000 : 35,000 : 45,000\) i.e. \(4 : 7 : 9\)

Let A’s share of profit = Rs. 4x,
Example 33

Mr. X, Mr. Y and Mr. Z started a transport business by investing Rs. 1 lakh each. Mr. X left after 5 months from the commencement of business and Mr. Y left 3 months later. At the end of the year the business realized a profit of Rs. 37,500. Find the share of profit of each partner.

Solution:

Capital is the same but the periods are different.

\[ \text{Profit of Rs. 37,500 is shared in proportion of period of investment.} \]

Proportion of period is 5 : 8 : 12

Let Mr X’s share of profit = Rs. 5x

Mr. Y’s share of profit = Rs. 8x

Mr. Z’s share of profit = 12x

\[ 5x + 8x + 12x = 37,500 \]

\[ 25x = 37,500 \]

\[ x = 1,500 \]

Mr X’s share of profit = Rs. 7,500

Mr. Y’s share of profit = Rs. 12,000

Mr. Z’s share of profit = Rs. 18,000.

Example 34

A started a business by investing Rs. 40,000. 4 months after commencement of a business B joined as a partner investing Rs. 60,000 and C joined the partnership one month after B by investing Rs. 60,000. At the end of the year the partnership earned a profit of Rs. 46,000. Find the share in profit of each partner.

Solution:

A’s investment of Rs. 40,000 was for 12 months

B’s investment of Rs. 60,000 was for 8 months

C’s investment of Rs. 60,000 was for 7 months

\[ \text{Total profit of Rs. 46,000 is shared in the proportion} \]

\[ 12 \times 40,000 : 8 \times 60,000 : 7 \times 60,000 \]

i.e. 48 : 48 : 42 i.e. 8 : 8 : 7

Let A’s share of profit be Rs. 8x

B’s share of profit be Rs. 8x

C’s share of profit be Rs. 7x
\[8x + 8x + 7x = 46,000\]
\[23x = 46,000\]
\[x = 2,000\]

\[\text{A’s share of profit} = Rs. 16,000\]
\[\text{B’s share of profit} = Rs. 16,000\]
\[\text{C’s share of profit} = Rs. 14,000\]

**Example 35**

Bala and Shiva are partners in a company and their capitals are in the ratio 5 : 6. Bala and Shiva shared the yearly profit in the ratio 5 : 4 and Shiva had joined the business some months after Bala had started it. Find the period of investment of Y.

**Solution:**

Capital of Bala and Shiva are in the ratio 5 : 6
Let Capital of Bala = Rs. 5a and Capital of Y = Rs. 6a
Period of investment by Bala = 12 months.
Let Period of investment by Shiva = b months.

\[\text{Bala and Shiva share the profit in the ratio} \ 12 \times 5a : b \times 6a \text{ i.e.} \ 10 : b\]

Given : Bala and Shiva share the profit in the ratio 5 : 4

\[10 : b = 5 : 4\]

\[
\begin{align*}
10 & = 5 \\
\frac{b}{4} & = & \\
\therefore \ b & = 8 \\
\therefore \ & \text{Period of Shiva’s investment} = 8 \text{ months.}
\end{align*}
\]

**Example 36**

Rama and Krishna are partners having capital in the ratio 4 : 3. After 4 months Rama withdraws 25% of his capital but Krishna makes up for the lost capital by investing an amount equal to that withdrawn by Rama. At the end of the year Krishna received Rs. 30,000 as his share of yearly profit. Find the total profit.

**Solution:**

Capital of Rama and Krishna are in the ratio 4 : 3
Let Rama’s Capital = Rs. 4x and Krishna’s capital = Rs. 3x
After 4 month Rama withdraws 25% of his capital i.e. Rs. x but Krishna makes up for the lost capital.

\[\therefore \text{Rama’s investment is} \ 4x \text{ for 3 month together with} \ 3x \text{ for 9 months.}\]

\[\text{Krishna’s investment is} \ 3x \text{ for 3 month together with} \ 4x \text{ for 9 months.}\]

\[\therefore \text{The total profit is shared in the ratio} \]

\[(4x \times 3 + 3x \times 9) : (3x \times 3 + 4x \times 9)\]

i.e. \[39x : 45x\]

i.e. \[13 : 15\]
Let Rama’s Share of Profit = Rs. 13y  
& Krishna’s share of yearly profit = 15y
Given : Krishna’s share of yearly profit = Rs. 30,000
∴ 15y = 30,000
∴ y = 2,000
∴ Total Profit = Rs. 28y = Rs. 56,000

**Example 37**
Mr. Wadia and Mr. Jokhi started a business with a capital investment of Rs. 55,000 and Rs. 35,000 respectively. After 5 months, Mr. Wadia put in Rs. 10,000 more as capital while Mr. Jokhi withdrew Rs. 5,000 from his existing capital. At the end of the year, the profit was Rs. 11,150. Determine the proportionate distribution of the profit between the two partners.

**Solution:**
Mr. Wadia invested Rs. 55,000 for the first 5 months and Rs. 65,000 for the next 7 months. Mr. Jokhi invested Rs. 35,000 for the first 5 months and Rs. 30,000 for the next 7 months. Hence the profit distribution will be in the proportion

\[
\frac{55,000 \times 5 + 65,000 \times 7}{35,000 \times 5 + 30,000 \times 7}
\]

i.e., in the proportion \(275 + 455\) : \(175 + 210\)

i.e., in the proportion \(730 : 385\)

Now \(730 + 385 = 1,115\)

∴ Mr. Wadia’s share in the profit \(= \frac{730}{1,115} \times 11,150 = Rs. 7,300\)

Mr. Jokhi’s share in the profit \(= \frac{385}{1,115} \times 11,150 = Rs. 3,850\)

**Example 38**
A and B are partners in a business with their capital as Rs. 2,00,000 and Rs. 3,00,000 respectively. C wishes to join the business with a capital of Rs. 3,00,000 at the beginning of a financial year. They agree that the goodwill will be taken as twice the average annual profit for the last three years. Last three years’ profit are Rs. 90,000, Rs. 70,000 and Rs. 80,000 respectively. Find the goodwill amount that C would be required to pay A and B separately.

**Solution:**
Average of last three years’ annual profits \(= \frac{90,000 + 70,000 + 80,000}{3}\)

\(= Rs. 80,000\)

∴ Goodwill \(= 2 \times 80,000 = Rs. 1,60,000\).

Now, A, B, and C have capital of Rs. 2,00,000, Rs. 3,00,000 and Rs. 3,00,000 respectively.

∴ Their share in the profit as well as their share in the goodwill will be in the proportion \(2,00,000 : 3,00,000 : 3,00,000\),
i.e., in the proportion 2 : 3 : 3.

\[ \text{C’s share of goodwill at the entry point} = \frac{2}{2+3+3} \times 1,60,000 = \frac{2}{8} \times 1,60,000 = \text{Rs.}40,000 \]

\[ \text{C must pay Rs. 40,000 as goodwill to A and B whose capitals are in the ratio 2:3} \]

\[ \text{A’s share of goodwill} = \frac{2}{2+3} \times 40,000 = \frac{80,000}{5} = \text{Rs.}16,000. \]

\[ \text{B’s share of goodwill} = \frac{2}{2+3} \times 40,000 = \frac{120,000}{5} = \text{Rs.}24,000. \]

\[ \text{C should pay Rs. 16,000 and Rs. 24,000 as goodwill to A and B respectively, on being admitted to the partnership.} \]

**Check your progress**

1. Three friends A, B, C hired a small stall in partnership in an exhibition, by investing Rs. 2,100, Rs. 3,300 and Rs. 3,000 respectively. At the end of the day, their profit was Rs. 560. Find the share of each friend.
   
   **Ans.** Rs. 140, Rs.220, Rs.200

2. Three partners opened a bicycle repairing shop with an investment of Rs. 4,000, Rs. 5,000 and Rs. 6,000 respectively. If their profit after one month is Rs. 4,500, find the share of each partner.
   
   **Ans.** Rs. 2,000, Rs. 2,500, Rs. 3,000

3. Three partners floated a company with contribution of Rs. 10,000, Rs. 20,000 and Rs. 18,000 respectively. If the company made a profit of Rs. 4,800, find the share of each partner in the profit.
   
   **Ans.** Rs. 1,000, Rs. 2,000, Rs. 1,800

4. Shikha launched an animation company with certain capital. After 2 months, Miraj joined the company and after 5 more months, Sagar too joined in as a partner. If the capital put in by all the partners in the company is the same and if the profit of the company at the end of the year was Rs. 81,000, find the share of each partner.
   
   **Ans.** Rs. 36,000, Rs. 30,000, Rs. 15,000

5. Priya and Mythili are partners in the company and their capitals are in the ratio 5:6. Mythili has joined the business some months after Priya had started it and they shared the yearly profit in the ratio 5:4. Find the period of investment of Mythili.
   
   **Ans.** 8 months.

6. A and B are partners in a business for some years. Their capital are Rs.3 lakhs and Rs. 2 lakhs respectively. C wants to join the business
with a capital of Rs. 4 lakhs. They agree that the goodwill will be considered as three times the average of the last four years’ profit which are Rs. 60,000, Rs. 75,000, Rs. 65,000 and Rs. 70,000 respectively. What are the amounts to be paid by C to A and B as goodwill?

**Ans.** Rs. 54,000, Rs. 36,000

**Formulae**

1) Commission = \[
\frac{\text{rate of Commission}}{100} \times \text{salevalue}
\]

2) Gross profit = Sale Price – Purchase Price

3) Net Price = Gross profit – Commission

4) Discount \% = \[
\frac{\text{Discount Amount}}{\text{List Price}} \times 100
\]

5) Trade discount (T.D) = \[
\frac{\text{Trade discount \%}}{100} \times \text{List Price}
\]

6) Invoice price (I.P) = List price – Trade discount

7) Cash discount (C.D) = \[
\frac{\text{cash discount \%}}{100} \times \text{Invoice price}
\]

8) Net selling price (N.S.P.) = Invoice price – cash discount

9) Profit \% = \[
\frac{\text{Profit Amount}}{\text{Cost Price}} \times 100
\]

10) Net selling Price = Cost price + profit

11) Profit = N.S.P. – Commission – C.P.

12) Profit = L.P. – Discount – Commission – C.P.

13) Loss = Cost price – selling price

❖❖❖
UNIT II

SHARES AND MUTUAL FUNDS

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).

2.1 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 depositary 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%.

### 2.1.1 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:**

\[
\text{Brokerage per share} = 250 \times \frac{0.5}{100} = 1.25
\]

\[
\therefore \text{cost of purchasing one share} = 250 + 1.25 = 251.25
\]

\[
\therefore \text{Number of shares purchased} = \frac{75375}{251.25} = 300
\]

**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:**

\[
\text{Brokerage per share} = 720 \times \frac{0.4}{100} = 2.88
\]

\[
\text{selling price of a share} = 720 - 2.88 = 717.12
\]

\[
\therefore \text{Number of shares sold} = \frac{430272}{717.12} = 600
\]

**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:**

\[
\text{Face value of 60,000 shares} = 60,000 \times 10 = 6,00,000
\]
Rate of Dividend = \frac{\text{Total Dividend}}{\text{Face value of 60,000 share}} \times 100

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

\therefore \text{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.

**Solution:**

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

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.

\therefore \text{Total Dividend paid to Preference shareholders}

= 2 \times 90000 = 1,80,000/-

Now, dividend to be distributed to the equity shareholders

= 2,40,000 - 1,80,000 = Rs. 60,000/-

\therefore \text{Rate of dividend}

= \frac{60,000}{30,00,000} \times 100 = 2

\therefore \text{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:**

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

= 15 \times 100 = 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**
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**
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
Dividend on 1 share = \( \frac{\text{Rate of dividend}}{100} \times \text{face value of 1 share} \)
= \( \frac{8}{100} \times 100 = 8 \)
Rate of Return on Investment
\( \frac{(\text{Price change}) + (\text{Dividend on 1 share})}{\text{purchase price of 1 share}} = \frac{16 + 8}{240} \times 100 \)
\( = \frac{2400}{240} = 10 \)
\[ \therefore \text{The rate of return on investment was 10\%.} \]

### 2.1.2 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

- Purchase cost of 30 shares = 30 x 36 = 1080/-
- On 3rd April 2007, each Rs. 10/- share became 5 shares of Rs. 2/- each.
- No. of shares = 30 x 5 = 150
- On 4th April 2007, market value of 150 shares was @ Rs. 8 each
  = 150 x 8 = 1200

∴ 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:
- Face value = Rs. 100/-, No. of shares = 500, market price = Rs.700/-
- Dividend = 25%, brokerage = 2%

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

∴ Total purchase = 500 x 714 = Rs.3,57,000/-

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

∴ Total dividend = 500 x 25 = Rs. 12,500/-

Now, bonus shares are 1 for every 4 shares.

∴ No. of bonus shares = \( \frac{1}{4} \times 500 = 125 \)

∴ Total No. of shares = 500 +125 = 625

For sales,
- No. of shares = 625, market price = 610, Brokerage 2%

Sale price of one share = 610 - 2% of 610 = 597.8

∴ Total sale value = sale price of one share \times No. of shares
  = 597.8 x 625 = Rs. 3,73,625/-

Net profit = sale value + Dividend - purchase value
  = 3,73,625 + 12500 - 3,57,000
  = Rs. 29,125/-.

∴ % gain = \( \frac{29,125}{3,57,000} \times 100 = 8.16 \)

= 8.16

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 Rs. 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

### 2.3 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 a given time. This is called Net Asset Value (NAV).

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

or

\[
\text{NAV} = \frac{\text{Total Assets- liabilities}}{\text{Total No. of units outstanding}}
\]
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 |

\[
\text{NAV} = \frac{\text{Rs. 1500 crores} + \text{Rs. 100 crores} - \text{Rs. 200 crores}}{100 \text{ crores}}
\]

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

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 number 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 no 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.

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

(This is for a given period).

Annualised rate of Return = 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% = 0.41

\[ \text{selling price of 1 unit} = 16.4 - 0.41 = 15.99 \]

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

Profit = 39,975 - 25,000 = 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% of 460
= 460 + 10.35 = 470.35

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

Total dividend = 200 x 5 = 1000
Gain = Profit + Dividend

\[ \therefore 1654 = \text{Profit} + 1000 \]

\[ \therefore \text{Profit} = 1654 - 1000 = 654 \]

While selling let NAV of one unit be \( y \)

\[ \therefore \text{sale price of one unit} = \text{NAV} - \text{exit load} \]
= \( y - 0.5\% \text{ of } y \)
= 0.995 \( y \)

\[ \therefore \text{sale price of 200 units} = 200 \times 0.995 \text{ y} = 199 \text{ y} \]
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/-
- Absolute change in NAV = 38 - 28 = Rs. 10/-
- % 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.
- Absolute change in NAV = 12.5 % of x = \( \frac{12.5}{100} \times x \) = 0.125 x
- NAV at the end of the year = x + 0.125 x = 1.125 x
- 1.125 x = 72
- x = \( \frac{72}{1.125} \)
- = 64
- \( \therefore \) 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.

**Solution**:
- Total gain = change in NAV + Dividend
- = (25 - 16) + 5
- = 9 + 5
- = 14
- \( \therefore \) Total percentage gain = \( \frac{\text{Total gain}}{\text{NAV at the beginning}} \times 100 \)
- = \( \frac{14 \times 100}{16} \) = 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/
\[ \therefore \text{annualised change} = \frac{\text{change in NAV}}{\text{NAV at beginning}} \times \frac{12}{\text{No. of months}} \times 100 \]
\[ = \frac{15}{30} \times 12 \times 100 \]
\[ = 33.33 \% \]

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.
\[ \text{Ans} . \ 12,000 , 12440.12 . \]

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)
\[ \text{Ans} . \ 30.2514 \]

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.
\[ \text{Ans} . \ \text{Total gain} = \text{Rs} . \ 1794.46 , \text{Rate of return} = 17.94\% \]

4) Given the following information, calculate NAV of the mutual fund :-
\[ \text{No. of units} = 15000 \]
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
\[ \text{Ans} . \ \text{Rs. 360/-} . \]

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)
\[ \text{Ans} . \ \text{profit} = 6432 , \text{STGT} = 643.2 , \text{Net profit} = 5788.8 \]
(profit- STGT).

2.4 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)

<table>
<thead>
<tr>
<th>Month</th>
<th>Amount (in Rs.)</th>
<th>NAV</th>
<th>No. of units he gets</th>
</tr>
</thead>
<tbody>
<tr>
<td>1</td>
<td>5000</td>
<td>109.48</td>
<td>5000/109.48=45.67</td>
</tr>
<tr>
<td>2</td>
<td>5000</td>
<td>112.36</td>
<td>5000/112.36=44.50</td>
</tr>
<tr>
<td>3</td>
<td>5000</td>
<td>108.14</td>
<td>5000/108.14=46.24</td>
</tr>
<tr>
<td>4</td>
<td>5000</td>
<td>105.62</td>
<td>5000/105.62=47.34</td>
</tr>
<tr>
<td>Total</td>
<td>20,000</td>
<td></td>
<td>183.75</td>
</tr>
</tbody>
</table>

:. 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% throughout 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:

<table>
<thead>
<tr>
<th>Month</th>
<th>NAV</th>
<th>Entry load</th>
<th>Total price</th>
<th>No.of units=1000/Total price</th>
</tr>
</thead>
<tbody>
<tr>
<td>1</td>
<td>34.2600</td>
<td>0.7708</td>
<td>35.0308</td>
<td>285.4627</td>
</tr>
<tr>
<td>2</td>
<td>46.1200</td>
<td>1.0377</td>
<td>47.1577</td>
<td>212.0544</td>
</tr>
<tr>
<td>3</td>
<td>39.3400</td>
<td>0.8851</td>
<td>40.2252</td>
<td>248.6006</td>
</tr>
<tr>
<td>4</td>
<td>41.8500</td>
<td>0.9141</td>
<td>42.7916</td>
<td>233.6906</td>
</tr>
<tr>
<td>TOTAL</td>
<td></td>
<td></td>
<td>165.2053</td>
<td>979.8083</td>
</tr>
</tbody>
</table>

By using Rupee-cost-Averaging method:

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

\[
= \frac{40,000}{979.8083} = 40.8243
\]

A.M. of price = \( \frac{\text{Total price}}{\text{No. of months}} = \frac{165.2053}{4} = 41.3013 \)

\[\therefore\text{Avg. price using Rupee-cost-Averaging method is less than A.M. of prices} .\]

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:

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

Avg. price of units = \( \frac{25000}{613.82} = 40.73 \)
For selling:
selling price of one unit = 51.64 - 2.25% of 51.64 = 50.48
Total sales = 50.48 x 613.82 = 30,991.77
\[ \therefore \text{Net gain} = 30,991.77 - 25,000 = \text{Rs. 5991.77} \]

**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/-, Rs. 45.455 respectively. The entry load was 2.25%. Find his average acquisition cost per unit up to 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 months 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 II

3

LINEAR PROGRAMMING PROBLEMS

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.

<table>
<thead>
<tr>
<th>Machine</th>
<th>Flutes (x)</th>
<th>Drum (y)</th>
<th>Maximum available</th>
</tr>
</thead>
<tbody>
<tr>
<td>A</td>
<td>4</td>
<td>2</td>
<td>12</td>
</tr>
<tr>
<td>B</td>
<td>3</td>
<td>6</td>
<td>18</td>
</tr>
</tbody>
</table>
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
\[
\begin{align*}
4x + 2y & \leq 12 \\
3x + 6y & \leq 18 \\
x & \geq 0, y & \geq 0
\end{align*}
\]

These inequalities are the constrains 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
\[
\text{Profit } z = 20x + 50y, \text{ which is to be maximized }
\]
Thus Maxmise \( z = 20x + 50y \)
Subject to
\[
\begin{align*}
4x + 2y & \leq 12 \\
3x + 6y & \leq 18 \\
x & \geq 0, y & \geq 0
\end{align*}
\]

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.

<table>
<thead>
<tr>
<th>A (x)</th>
<th>B (y)</th>
<th>C (z)</th>
<th>Minimum requirements</th>
</tr>
</thead>
<tbody>
<tr>
<td>Fats</td>
<td>2</td>
<td>6</td>
<td>1</td>
</tr>
<tr>
<td>Carbohydrates</td>
<td>16</td>
<td>4</td>
<td>5</td>
</tr>
<tr>
<td>Proteins</td>
<td>4</td>
<td>3</td>
<td>7</td>
</tr>
</tbody>
</table>

The constraints can be written as
\[
\begin{align*}
2x + 6y + z & \geq 12 \\
16x + 4y + 5z & \geq 30
\end{align*}
\]
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 ≥ 12
16x + 4y + 5z ≥ 30
4x + 3y + 7z ≥ 20,

EXAMPLE 3:

Two types of food packets A and B are available. Each contain vitamins N\textsubscript{1} and N\textsubscript{2}. A person need 4 decigrams of N\textsubscript{1} and 12 decigrams of N\textsubscript{2} per day. Food packet A contain 2 decigram of vitamin N\textsubscript{1} and 4 decigram of vitamin N\textsubscript{2}. Food packet B contain 1 decigrams of vitamin N\textsubscript{1} and 4 decigrams of vitamin N\textsubscript{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.

:. Objective function i.e. cost function

Z = 15x + 10y subject to

Side constraints

<table>
<thead>
<tr>
<th>Vitamin N\textsubscript{1}</th>
<th>Food packed A</th>
<th>Food packet B</th>
<th>Requirement (≥)</th>
</tr>
</thead>
<tbody>
<tr>
<td>2</td>
<td>1</td>
<td>4</td>
<td>4</td>
</tr>
<tr>
<td>4</td>
<td>4</td>
<td>12</td>
<td></td>
</tr>
</tbody>
</table>

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

:. x ≥ 0, y ≥ 0

Mathematical form

Min Z = 15x + 10y

Subject to 2x + y ≥ 4

4x + 4y ≥ 12

x ≥ 0, y ≥ 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.

<table>
<thead>
<tr>
<th></th>
<th>Product A</th>
<th>Product B</th>
<th>Availability</th>
</tr>
</thead>
<tbody>
<tr>
<td>Chemical salt</td>
<td>4</td>
<td>2</td>
<td>1000</td>
</tr>
<tr>
<td>Chemical mixture</td>
<td>2</td>
<td>3</td>
<td>1500</td>
</tr>
</tbody>
</table>

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

Mathematical form

Max \( z = 30x + 20y \)

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 \( c/b \).

![Graph of a linear equation](image)

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 $9x = 27$

\[ \Rightarrow x = 3 \]

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’

\[ \Rightarrow \text{ Graph of equation } 5x + 2y = 0 \text{ is a straight line passing through the origin.} \]
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 \( 2x + y = 6 \)
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. \( y = -2 \).

Answers:

1. \( x + 3y = 6 \)
   - Points: A(6,0), B(0,2)

2. \( 2x - y = 3 \)
   - Points: A(1.5,0), B(0,-3)

3. \( 3x - 5y = 8 \)
   - Points: A(2.7,0), B(0,-1.6)

4. \( x = 5 \)
   - Points: 

   ![Graphs of linear equations](image-url)
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 ≤ 0,
2. ax + by + c ≤ 0,
3. ax + by + c ≥ 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. c.

Case (i): if c < 0, then the half plane P_1 which contain the origin represents the inequality ax + by + cb < 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 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 \( ax + by + c \leq 0 \) 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. 

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

EXAMPLE 4:

Represent the inequality $y \geq 4$ graphically.

$y = 0$ does not satisfy the inequality.

Region represented by $y \geq 4$ is above the line $y = 4$.

EXAMPLE 5:

Represent the inequality $x \leq 3$
EXAMPLE 5: Represent the inequality \( x \geq 2 \) graphically

Solution:

\[
\begin{array}{c|cccccc}
\text{x'} & -4 & -3 & -2 & -1 & 0 & 1 \[2pt]
\hline
\text{y'} & -4 & -3 & -2 & -1 & 0 & 1 \[2pt]
\end{array}
\]

\( (x = 0) -4 \leq x \leq 3 \)

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

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

\( x = 0 \) 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.
**Example 1:** 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 \( 2x + 3y \leq 18 \)

<table>
<thead>
<tr>
<th>x</th>
<th>y</th>
<th>Points</th>
</tr>
</thead>
<tbody>
<tr>
<td>0</td>
<td>6</td>
<td>(0,6)</td>
</tr>
<tr>
<td>9</td>
<td>0</td>
<td>(9,0)</td>
</tr>
</tbody>
</table>

For the line \( 2x + y \leq 10 \)

<table>
<thead>
<tr>
<th>x</th>
<th>y</th>
<th>Points</th>
</tr>
</thead>
<tbody>
<tr>
<td>0</td>
<td>10</td>
<td>(0,10)</td>
</tr>
<tr>
<td>5</td>
<td>0</td>
<td>(5,0)</td>
</tr>
</tbody>
</table>

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 points of this region are O, A, P and D.

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

\[
2x + 3y = 18 \\
2x + y = 10 \\
2y = 8 \\
y = 4
\]

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

So the point of intersection = 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 \( y = 4 \).

**Example 2:** 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:**

<table>
<thead>
<tr>
<th>( x_1 )</th>
<th>( x_2 )</th>
<th>Points</th>
</tr>
</thead>
<tbody>
<tr>
<td>0</td>
<td>3</td>
<td>(3,0)</td>
</tr>
<tr>
<td>5</td>
<td>0</td>
<td>(5,0)</td>
</tr>
</tbody>
</table>

For the line \( 3x_1 + 5x_2 \leq 15 \)

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.

\[
\begin{align*}
3x_1 + 5x_2 & = 15 \quad \text{ ..........(1)} \times 5 \\
5x_1 + 2x_2 & = 10 \quad \text{ ..........(2)} \times 3 \\
15x_1 + 25x_2 & = 75 \\
15x_1 + 6x_2 & = 30 \\
\hline
19x_2 & = 45 \\
x_2 & = \frac{45}{19}
\end{align*}
\]

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

**Example 3:**

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  
$x_1 + 3x_2 = 6$ \hspace{1cm} (I)  
$x_1 + x_2 = 2$ \hspace{1cm} (II)  

For point of intersection, solve the equations simultaneously

\[
\begin{align*}
  x_1 + 3x_2 &= 6 \\
  x_1 + x_2 &= 2
\end{align*}
\]

\[
\begin{array}{c}
  - \\
  + \\
  - \\
  =
\end{array}
\]

\[4x_2 = 4\]

\[x_2 = 1\]

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

\[
\begin{array}{|c|c|c|}
  \hline
  x_1 & x_2 & \text{Points} \\
  \hline
  0 & 2 & (0, 2) \\
  0 & 0 & (2, 0) \\
  \hline
\end{array}
\]

\[
\begin{array}{|c|c|c|}
  \hline
  x_1 & x_2 & \text{Points} \\
  \hline
  0 & -2 & (0, -2) \\
  2 & 0 & (2, 0) \\
  \hline
\end{array}
\]

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 $3x + 4y \geq 12$
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

At \( o \ (0, 0) \), \( Z = 4(0) - 2(0) = 0 \)
At \( (0, 2) \), \( Z = 4x_1 - 2x_2 = 4 \times 0 - 2 \times 2 = -4 \)
At \( P(3, 1) \), \( Z = 4 \times 3 - 2 \times 1 = 10 \)
At \( D (2, 0) \), \( Z = 4 \times 2 - 2 \times 0 = 8 \)
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.

\[ 4x + 6y \leq 120, \quad 10x + 3y \leq 0, \quad y \geq 0 \]

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 from a table for their co-ordinates as follows:

<table>
<thead>
<tr>
<th>( x )</th>
<th>( y )</th>
<th></th>
<th>( x )</th>
<th>( y )</th>
<th></th>
</tr>
</thead>
<tbody>
<tr>
<td>30</td>
<td>0</td>
<td>A(30,0)</td>
<td>18</td>
<td>0</td>
<td>C(18,0)</td>
</tr>
<tr>
<td>0</td>
<td>20</td>
<td>B(0,20)</td>
<td>0</td>
<td>60</td>
<td>D(0,60)</td>
</tr>
</tbody>
</table>
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:

\[4x + 6y = 120 \quad \ldots \ldots (1)\]
\[10x + 3y = 180 \quad \ldots \ldots (2)\]

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

\[20x + 6y = 360 \]
\[4x + 6y = 120 \]
\[-16x = 240 \]
\[x = 15 \quad \ldots \ldots \text{from (2)}\]
\[150 + 3y = 180 \]
\[3y = 30 \]
\[y = 10 \]
\[\therefore \text{P is (15, 10).} \]

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

\[z (O) = 800 \times 0 + 100 \times 0 = 0\]
\[z (P) = 800 \times 15 + 100 \times 10 = 14400\]
\[z (C) = 800 (18) + 100 \times 0 = 14400\]
\[z (B) = 800 \times 0 + 100 \times 20 = 2000\]

\[\therefore z \text{ has the maximum value 14400 at the point C (18, 0).} \]
\[\therefore \text{the manufacturer should produce 18 scooters and 0 bicycles, in order to have maximum profit of Rs. 14400.} \]

**Example 5:**

Minimize \(z = 40x + 37y\)

Subject to constraints:

\[10x + 3y \geq 180\]
\[2x + 3y \geq 60\]
\[x \geq 0, \quad y \geq 0\]

**Solution:**

We first draw the line AB and CD whose equation are
\[10x + 3y = 180 \quad \text{and} \quad 2x + 3y = 60.\]

For the line \(10x + 3y = 180\)

<table>
<thead>
<tr>
<th>x</th>
<th>y</th>
<th>Points</th>
</tr>
</thead>
<tbody>
<tr>
<td>18</td>
<td>0</td>
<td>(18,0)</td>
</tr>
<tr>
<td>0</td>
<td>60</td>
<td>(0,60)</td>
</tr>
</tbody>
</table>

For the line \(2x + 3y = 60\)

<table>
<thead>
<tr>
<th>x</th>
<th>y</th>
<th>Points</th>
</tr>
</thead>
<tbody>
<tr>
<td>30</td>
<td>0</td>
<td>(30,0)</td>
</tr>
<tr>
<td>0</td>
<td>20</td>
<td>(0,20)</td>
</tr>
</tbody>
</table>

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.
\[
\begin{align*}
10x + 3y &= 180 \\
2x + 3y &= 60
\end{align*}
\]
\[
8x = 120 \\
x = 15
\]
\[
\therefore 2x + 3y = 60 \text{ gives } 30 + 3y = 60, \\
i.e., 3y = 30, \text{ i.e. } y = 10
\]
\[
\therefore \text{E is (15, 10)}
\]
The values of the objective function \(z = 40x + 37y\) at these vertices are:

<table>
<thead>
<tr>
<th>Vertex (x, y)</th>
<th>(z = 40x + 37y)</th>
</tr>
</thead>
<tbody>
<tr>
<td>C (30, 0)</td>
<td>(z (C) = 40 \times 30 + 37 \times 0 = 1200)</td>
</tr>
<tr>
<td>B (0, 60)</td>
<td>(z (B) = 40 \times 0 + 37 \times 60 = 220)</td>
</tr>
<tr>
<td>E (15, 10)</td>
<td>(z (E) = 40 \times 15 + 37 \times 10 = 600 + 370 = 970).</td>
</tr>
</tbody>
</table>

\[
\therefore z \text{ has minimum value 970 at the point E (15, 10), where } x = 15 \text{ and } y = 10.
\]

**Example 6:** 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 \(x + 3y = 3\)  

\[
\begin{array}{ccc}
\text{x} & \text{y} & \text{Point} \\
3 & 0 & A(3,0) \\
0 & 1 & B(0,1)
\end{array}
\]

For the line \(2x + y = 2\)

\[
\begin{array}{ccc}
\text{x} & \text{y} & \text{Point} \\
1 & 0 & C(1,0) \\
0 & 2 & D(0,2)
\end{array}
\]

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.

\[
\begin{align*}
2x + 6y &= 6 \\
2x + y &= 2
\end{align*}
\]

\[
\begin{array}{c}
2x + y = 2 \\
x + 3y = 3
\end{array}
\]

\[
\begin{array}{r}
-5y = 4 \\
x = \frac{4}{5} = 0.8
\end{array}
\]

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

\[
\begin{align*}
\bullet & \quad x + 3 \times 0.8 = 3 \\
\bullet & \quad x + 2.4 = 3 \\
\bullet & \quad x = 3 - 2.4 = 0.6
\end{align*}
\]

Thus E is (0.6, 0.8)

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

<table>
<thead>
<tr>
<th>Vertex (x, y)</th>
<th>( z = 4x + 2y )</th>
</tr>
</thead>
<tbody>
<tr>
<td>C (3, 0)</td>
<td>( z (A) = 4 \times 3 + 2 \times 0 = 12 )</td>
</tr>
<tr>
<td>E (0.6, 0.8)</td>
<td>( z (E) = 4 \times 0.6 + 2 \times 0.8 = 2.4 + 1.6 = 4 )</td>
</tr>
<tr>
<td>D(0, 2)</td>
<td>( z (D) = 4 \times 0 + 2 \times 2 = 4 )</td>
</tr>
</tbody>
</table>

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)

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 at 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)
3x + 4y = 360
2x + y = 100

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

INTRODUCTION TO STATISTICS AND DATA COLLECTION

OBJECTIVES:

The objectives of this chapter are to give an overview of:
1. What is statistics and why one should learn it?
2. How did it originate?
3. Scope and limitations.
4. Definition of some basic terms that are used in the subject.
5. Types of data and how they are collected.
6. How should the data be arranged?

4.1 INTRODUCTION

In almost all areas of our daily living in which we make simple statements we are actually doing what one may call statistical thinking. What may seem like a simple “I walk on an average 4 kms per day” is a statement which involves statistics. “There are 60 percent chances that a particular political party will get reelected at the next election”, “This year it was much hotter than the same month previous year”, are all examples of the common man doing statistics without realizing it. Due to the simple capacity to observe people, things and event in our environment we notice similarities and differences and look for patterns and regularities as part of a survival mechanism to sense danger or opportunity in an effort to grow wiser and understand our environment better. The simple act of observation results in us getting data. The word data means fact, and is plural of the word datum. This in turn makes us look for connections among the data that we have noticed. While most of us make connections and make conclusions which affects us and others around us the question that we must ask is “How mature or informed our conclusions are?”, “How can we be sure that our conclusions are factually correct?” These questions also lead us to ask “Is there a scientific, rational way of determining whether our conclusions are valid? Can we use structured, methodical, time tested ways to make these conclusions?” The remaining part of this chapter is devoted to answer these questions.

4.1.1 What is statistics?

In its simplest sense statistics refers to the science of collecting data, assortment of data and analyzing them to reach a definite conclusion. The methods by which the data are analysed are called as Statistical
Methods.

The word statistics seems to have originated from the German word “statistik”, the Italian word ‘statistica’ or the Latin word ‘status’ all meaning ‘state’ in the political sense of the word. This in turn seems to have come because statistics was widely used initially to collect data for the state (Government) so that officials could use the data for better planning in the future. Statistical thinking however has a long history as from very early times kings and rulers have been collecting data about their populations and resources.

4.1.2 Functions of Statistics

Following are the important functions of statistics

1. *It represents data in a definite form*
   One of the important function of statistics is that it enables one to make statements which are precise and in quantitative terms. To say that India is a overpopulated country, or that a country has high level of poverty are all general statements but do not convey any precise meaning. Words such as “high”, “low”, “good”, “bad”, are very vague and subject to interpretation by different people in different manner. Statements of facts made in exact quantitative terms are more convincing.

2. *It simplifies complex data or a mass of figures.*
   Statistics not only helps one to express data in a concrete definite form but also reduces the data to a few significant figures which give the essence of the issue under consideration.

3. *It facilitates comparison.*
   Generally plain facts by themselves have little value unless they are seen in the correct context in which they occur. This requires that we place the facts in relation to facts of similar type but in other situations. For example the fact that the population of a country is 12 million by itself would have a restricted meaning but in the context of how big the geographical area is and what are the populations of other similarly placed countries would throw light on some very important aspects of population density and other issues such as crowding and pollution.

4. *It helps in forecasting.*
   Statistics helps in predicting what may happen in future on the basis of past data and analysis. It is useful for anybody in business to have an idea of what might be the possibility of selling goods or services so that the person can plan for any future changes in demands. Production Planning, inventories are all based on statistical modes which help by predicting excess or shortfalls in demands.

5. *It helps in formulating policies*
   Due to its potential to do forecasting statistics helps policy makers in
drafting policies which help in better governance. The annual finance budget relies heavily on statistical methods to decide issues of direct and indirect taxation and budgetary allocations based on past data and its analysis.

4.1.3 Scope of Statistics
There has been hardly any area whether from the exact sciences or social sciences where statistics has not been applied, whether it be trade, industry, commerce, economics, life sciences, education where statistical methods have not been applied. However certain fields have used statistics very frequently and effectively. We list some of the important fields.

1. **State**: As we have stated earlier from very early times statistics have been used by governments in framing policies on the basis of data about population, military, crimes, education etc. The present day governments have special departments which maintain a variety of data of significance to the state. The extensive proliferation of technologies such as the internet has made huge amount of data available making statistical tools absolutely indispensible.

2. **Business**: With the advent of globalization and the extensive range of operations Business has grown almost exponentially in many areas. Business now extends across geographical and political boundaries. The large amount of data that it generates is used for analysis and forecasting of consumption patterns and has given rise to advance techniques such as Data Mining.

4.1.4 Limitations of statistics
Although there are many areas where statistics is useful there are certain limitations also as there are areas where it is not applicable and there are possibilities of its misuse. We indicate some of them as under.

1. **Statistics does not deal with individual measurement.**
Statistical measurements are generally mass measurements and it deals with a collection of data and not any particular measurement. In fact a particular measurement can be very different from an average.

2. **Statistics deals with quantitative characteristics.**
Statistics are numerical statement of facts. There are many aspects of life which cannot be quantified and are therefore not within the purview of statistics. For e.g. The honesty or integrity of person, whether a person is affectionate or not, whether or not a person is intelligent are some qualities which are not easily quantifiable and not available for statistical analysis.

3. **Statistical results are true on an average.**
Unlike exact sciences in which there is a clear cause effect relationship in the phenomena which is being studied, in statistics the results are
true only on an average. Any individual element in the data collected may not have any resemblance with the average.

4. **Statistics is one of the many methods of studying phenomena.**
   A statistical study is not complete in itself. There are many ways in which phenomena can be studied and one of it is statistics. Generally statistics complements other methods to substantiate the understanding of phenomena.

5. **Statistics can be misused.**
   Perhaps the greatest limitations of statistics are that it can be misused. Many people use statistics to prove a point to which they have already subscribed. For e.g if a person wants to prove that malnutrition deaths have been high in an area then the researcher can deliberately collect data that will favour that conclusion and ignore the data that will not be favorable. Although there are ways in which this bias can be detected an intelligent researcher can easily mislead people. One of the reasons why one should learn statistics to understand how one can be mislead by such people and be cautious.

### 4.2 BASIC STATISTICAL CONCEPTS.

In using statistics it is necessary that we have a clear understanding of the words that are used. We give some definitions of some of the very basic concepts in statistics.

4.2.1 **Data**

The word *data* is plural of the word *datum* which in Greek means fact. It is a collection of observations expressed in numerical quantities. Data is always used in the collective sense and not in singular.

4.2.2 **Population**

The word population in statistics means the *totality of the set of objects under study*. It should not be understood in the limited sense in which it is generally used to mean people in a certain city or country.

4.2.3 **Sample**

A sample is a selected number of entities or individuals which form a part of the population under study. The study of a sample is more practical and economical in most situations where the population is large and is used to make conclusions about the entire population.

4.2.4 **Characteristic**

The word characteristic means an aspect possessed by an individual entity. We may study the rainfall of a certain region, or the marks scored by
students in a certain school. These are referred to as characteristics.

### 4.2.5 Variables and attributes.

In statistics characteristics are of two types. Measurable and non-measurable. Measurable characteristics are those that can be quantified as expressed in numerical terms. The measurable characteristics are known as \textit{variables}. A non-measurable characteristic is qualitative in nature and cannot be quantified. Such a characteristic e.g nationality, religion, etc are called as \textit{attributes}.

### 4.3.1 Collection of Data

To apply statistical methods or to study a problem we must collect data. Collection of data is very important as it forms the basis of the analysis. It is important to understand the techniques of collection of data because if the data is not collected properly its reliability itself becomes questionable and our entire analysis will be on weak foundations.

Based on whether the investigator collects data himself/herself or uses data collected by some other person or agency data is classified into two types. Primary data and Secondary data.

1. **Primary Data**

   Primary data are those which are collected directly from the field of enquiry for a specific purpose. This is raw data original in nature and directly collected from the population. The collection of the data can be made through two methods. a) \textbf{Complete enumeration} or \textit{census method} or b) \textbf{Sampling survey} methods.

   The complete enumeration or census method is a study of the entire population and data are collected about each individual of the population. Generally this is a laborious, time consuming and expensive method. Large organizations and Government semi-government organizations, Research institutions and Public sector bodies such as the RBI ( Reserve Bank Of India ) can afford such methods of collecting data.

2. **Secondary Data**

   Secondary data are such information which has already been collected by some agency for a specific purpose and is subsequently compiled by the investigator from that source for application in a different area. Data used by any other person or agency other than the one which collected it constitutes secondary data. The same data is primary when collected by the source agency and becomes secondary when used by any other agency. Data after analysis are also termed as secondary data.

### 4.3.2 Collection of primary data

Primary data can be collected by the investigator in following ways.
i) By direct personal observation
The investigator may collect data by direct observation. This can be done by meeting and interrogating people who may supply the desired information. Such data is directly obtained and can be very reliable but it is a time consuming and costly process.

ii) By indirect oral investigation
Here the information is collected not by questioning the concerned people but by asking people connected with the concerned people. These people can be called as witnesses who have knowledge about the persons concerned or situation involved. Here the investigator has the added responsibility of ensuring that the witnesses are not biased and therefore the reliability of data cannot be questioned. Sometimes this can be the only way to get the desired information as the people themselves are either not accessible or not willing to give the information.

iii) By sending questionnaires by mail or email
A questionnaire is a proforma containing a set of questions. A collection of questions relevant to the area of study is created and sent by post or email to selected people with a request to fill up and returned by post or email. This method can cover a large population economically but does not necessarily result in good response. The questionnaire has also to be carefully drafted so as to not have leading questions which direct the person to an answer desired by the investigator.

iv) By sending schedules through paid investigation
This method is used quite widely particularly by market researchers. Here schedules are prepared and the investigators are trained to meet people concerned with the schedules. A schedule is a form where information is to be noted by an enumerator who questions people. The investigators are to fill up the schedules on the basis of answers given by the respondents. The success of this method largely depends on how efficient the investigators are and how tactfully they collect the needed data.

4.3.3 Collection of Secondary Data
Secondary data are those which are collected by some other agency and are used for further investigation. The sources of secondary data can be classified into two:

a) Published sources  b) Unpublished sources

Published Sources
Some of the published sources providing secondary data are:
a. Government Publications: Government, semi-government and private organizations collect data related to business, trade, prices, consumption; production, industries, income, health, etc. These publications are very powerful source of secondary data. Central Statistical Organization (C.S.O.), National Sample Survey Organization (N.S.S.O.), office of the
b. International Publications: Various governments in the world and international agencies regularly publish reports on data collected by them on various aspects. For example, U.N.O.’s Statistical Year book, Demography Year Book.

c. Semi-official Publications: Local bodies like District Boards, Municipal Corporations publish periodicals providing information about vital factors like health, births, deaths etc.

d. Reports of Committees and Commissions: At times state and central governments appoint committees and commissions with a specific reference to study a phenomenon. The reports of these committees and commissions provide important secondary data. For example, Kothari commission report on education reforms, Report of National Agricultural Commission.

e. Private Publications: The following private publications may also be enlisted as the source of secondary data:

Journals and Newspapers such as the Financial Express and Economic times Research Publications of well known agencies such as Indian Statistical Institute (I.S.I.) Calcutta and Delhi, I.C.A.R., N.C.E.R.T., I.C.M.R., Publications of Business and Financial Institutions such as the Stock Exchanges, Chambers of Commerce, Newspaper Articles are also sources of secondary data.

Although secondary data are easier and cheaper to procure they have certain limitations which one must be aware of. The limitations include:

1. The possibility that proper procedure might not have been followed in their collection.
2. These may not be relevant in the present contest.
3. These may not be free from personal bias and prejudices.
4. These may not have the needed accuracy or reliability.
5. These may not be adequate.

### 4.4 CLASSIFICATION AND TABULATION OF DATA

In the previous unit we learned how to collect data. After data collection comes the important task how to present that data so that it is meaningful and can be used effectively. The collected data is known as raw data and these are in an unorganized form and need to be organized and presented in a meaningful and understandable form in order to carry out subsequent statistical analysis. The collected data in following ways:

1. Classification and Tabulation
2. Diagrammatic Presentation
3. Graphical Presentation

Our emphasis will be mainly on Classification and Tabulation.

4.4.1 What is Classification of Data?
Classification is the method by which things are arranged in groups according to their similarity. For example, students in a class may be grouped according to their sex, age, etc. One may classify documents in an organization such as purchase, sales, tax related, etc.

4.4.2 Why Do We Need Classification of Data?

The following are the main objectives of classifying the data:
1. It condenses the mass of data in an easily understandable form.
2. It removes unnecessary details.
3. It helps comparison and emphasizes the important elements of data.
4. It enables one to get a good idea of the information and helps in making conclusions.
5. It helps in the statistical treatment of the information collected.

4.4.3 What is Tabulation?

Tabulation is the process of summarizing classified or grouped data in the form of a table so that it is easily understood and an investigator is able to locate the desired information promptly. A Table is a systematic arrangement of classified data in columns and rows. A statistical table makes it possible for the investigator to present voluminous data in a detailed and orderly form. It helps comparison and reveals certain patterns in data, which are otherwise not obvious. Classification and Tabulation, as a matter of fact; are not two distinct processes. Actually they go together. Before tabulation data are classified and then displayed under different columns and rows of a table.

4.4.4 Advantages of Tabulation
Statistical data arranged in a tabular form serve following objectives:
1. It simplifies complex data and the data presented are easily understood.
2. It helps in comparison of related facts.
3. It helps in computation of various statistical measures like averages, dispersion, correlation etc.
4. It presents facts in least amount of space and avoids repetitions and explanations. The needed information can be easily located.
5. Tabulated data are good for references and they make it easier to present the information in the form of graphs and diagrams.

4.4.5 Preparing a Table
The making of a compact table is a desirable skill. It should contain all the information needed within the minimum space. What the purpose of tabulation is and how the tabulated information is to be used are the main points to be kept in mind while preparing for a statistical table. An ideal
Maths_stats_fybcom

Table should consist of the following main parts:
1. Table number.
2. Title of the table
3. Headnote
4. Captions or column headings.
5. Stubs or row designations
7. Footnotes.
8. Sources of data.

**Table Number**
A table should be numbered for easy reference and identification. This number, if possible, should be written in the center at the top of the table. Sometimes it is also written just before the title of the table.

**Title**
A good table must have a clearly worded, brief but clear title explaining the nature of data contained in the table. It should also state arrangement of data and the period covered. The title should be placed centrally on the top of a table just below the table number (or just after table number in the same line).

**Headnote**
If figure are expressed in percentage or any other unit it should be written in the head-note.

**Captions or Column Headings**
Captions in a table stand for brief and self-explanatory headings of vertical columns. Captions may involve headings and subheadings as well. The unit of data contained should also be given for each column. Usually, a relatively less important and shorter classification should be tabulated in the columns.

**Stubs or Row Designations**
Stubs stands for brief and self explanatory headings of horizontal rows. Normally, a relatively more important classification is given in rows. Also a variable with a large number of classes is usually represented in rows. For example, rows may stand for score classes and columns for data related to sex of students. In the process, there will be many rows for score classes but only two columns for male and female students.

**Body**
The body of the table contains the numerical information of frequency of observations in the different cells. This arrangement of data is according to the description of captions and stubs.

**Footnotes**
Footnotes are given at the foot of the table for explanation of any fact or information included in the table, which needs some explanation. Thus,
they are meant for explaining or providing further details about the data that have not been covered in title, captions and stubs. e.g. Some figures are projected and not collected.

**Sources of Data**
Lastly one should also mention the source of information from which data are taken. This may preferably include the name of the author, volume, page and the year of publication. This should also state whether the data contained in the table is of primary or secondary nature.

A model structure of a table is given below

**Model Structure of a Table**

<table>
<thead>
<tr>
<th>Table Number</th>
<th>Title of the Table</th>
<th>Headnote</th>
</tr>
</thead>
<tbody>
<tr>
<td></td>
<td></td>
<td></td>
</tr>
</tbody>
</table>

<table>
<thead>
<tr>
<th>Caption</th>
<th>Caption sub headings</th>
<th>Total</th>
</tr>
</thead>
<tbody>
<tr>
<td>Caption</td>
<td></td>
<td></td>
</tr>
<tr>
<td>Caption sub headings</td>
<td></td>
<td></td>
</tr>
<tr>
<td>Total</td>
<td></td>
<td></td>
</tr>
</tbody>
</table>

<table>
<thead>
<tr>
<th>Sub Stub Heading</th>
<th>Caption sub headings</th>
<th>Total</th>
</tr>
</thead>
<tbody>
<tr>
<td>Sub Stub</td>
<td></td>
<td></td>
</tr>
<tr>
<td>Heading</td>
<td></td>
<td></td>
</tr>
<tr>
<td>Caption sub headings</td>
<td></td>
<td></td>
</tr>
<tr>
<td>BODY</td>
<td></td>
<td></td>
</tr>
<tr>
<td>Total</td>
<td></td>
<td></td>
</tr>
</tbody>
</table>

Foot notes: 1- 2-
Sources Note: 1- 2-

### 4.4.6 Requirements of a Good Table
A good statistical table is not just a careless grouping of columns and rows but should be such that it summarizes the total information in an easily accessible and understandable form in least possible space. While preparing a table, one must have a clear idea of the information to be presented, the facts to be compared and the points to be stressed. The following points should be kept in mind:

1. A table should be formed in keeping with the objects of a statistical enquiry.
2. A table should be prepared meticulously so that it is easily understandable.
3. A table should be formed so as to suit the size of the paper. But such an adjustment should not be at the cost of legibility.
4. If the figures in the table are large, they should be rounded or approximated. The method of approximation and units of measurements too should be specified.
5. Rows and columns in a table should be numbered and certain figures
to be stressed may be put in ‘box’ or ‘circle’ or in bold letters.

6. The arrangement of rows and columns should be in a logical and systematic order. This arrangement may be an alphabetical, chronological or according to magnitude.

7. The rows and columns are separated by single, double or thick lines to represent various classes and sub-classes used. The corresponding proportions or percentages should be given in adjoining rows and columns to allow easy comparison. A vertical expansion of the table is generally more desirable than the horizontal one.

8. The averages or totals of different rows should be given at the right of the table and that of columns at the bottom of the table. Totals for every sub-class too should be mentioned.

9. In case it is not possible to accommodate all the information in a single table, it is better to have two or more related tables.

**Type of Tables**

Tables can be classified according to their purpose, number of characteristics used, stage of enquiry and other parameters. On the basis of the number of characteristics, tables may be classified as follows:

1. Simple or one-way table.
2. Two-way table.

**Simple or One-Way Table**

A simple or one-way table is the simplest table, which contains data of one characteristic only. A simple table is easy-to construct and simple to follow. For example the table below may be used to show the number of students in different faculties of a college.

<table>
<thead>
<tr>
<th>Faculties</th>
<th>No of students</th>
</tr>
</thead>
<tbody>
<tr>
<td>Commerce</td>
<td></td>
</tr>
<tr>
<td>Arts</td>
<td></td>
</tr>
<tr>
<td>Science</td>
<td></td>
</tr>
</tbody>
</table>

If we consider the strength of the students in various academic years for the different faculties we get a **two-way table** as we are now considering two characteristics year and faculties.

<table>
<thead>
<tr>
<th>Year</th>
<th>Commerce</th>
<th>Arts</th>
<th>Science</th>
</tr>
</thead>
<tbody>
<tr>
<td>2001-02</td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>2002-03</td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>2003-04</td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>2004-05</td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>2005-06</td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td><strong>Totals</strong></td>
<td></td>
<td></td>
<td></td>
</tr>
</tbody>
</table>

If we consider finding out how many males /female students exist in each faculty for each of the academic year and their totals we have a **three-way**
Example 1:
The number of students in AVS College in the year 1990 was 500, of which 200 were rural students. In 1991, the number of students increased by 150 and the number of urban students increased by 75. In 1992, the number rural students increased by 20%, while the total number of students increased by 50%. Tabulate the above data.

Solution: As we can observe that there are two characteristics in the given data i.e. (i) Year and (ii) residence of students.

Thus, the data can be tabulated in the table with the given information as follows:

<table>
<thead>
<tr>
<th>Year</th>
<th>Students</th>
<th>Total</th>
</tr>
</thead>
<tbody>
<tr>
<td></td>
<td>Rural</td>
<td>Urban</td>
</tr>
<tr>
<td>1990</td>
<td>200</td>
<td>?</td>
</tr>
<tr>
<td>1991</td>
<td>?</td>
<td>?</td>
</tr>
<tr>
<td>1992</td>
<td>?</td>
<td>?</td>
</tr>
<tr>
<td>Total</td>
<td>?</td>
<td>?</td>
</tr>
</tbody>
</table>

The data for blank cells are calculated as follows:

In the year 1990, total number of students = 500, rural students = 200.
Thus, urban students = 500 – 200 = 300.
In the year 1991, number of students = 500 + 150 = 650.
Number of urban students = 300 + 75 = 375.
Hence, rural students = 650 – 375 = 275.

In the year 1992, rural students = 275 + \( \frac{20}{100} \times 275 = 275 + 55 = 330 \)
Total number of students = 650 + \( \frac{50}{100} \times 650 = 650 + 325 = 975 \).
Hence, urban students = 975 – 330 = 645.
Thus, the completed table is given by

<table>
<thead>
<tr>
<th>Year</th>
<th>Students</th>
<th>Total</th>
</tr>
</thead>
<tbody>
<tr>
<td></td>
<td>Rural</td>
<td>Urban</td>
</tr>
<tr>
<td>1990</td>
<td>200</td>
<td>300</td>
</tr>
<tr>
<td></td>
<td></td>
<td>500</td>
</tr>
</tbody>
</table>
Example 2:
In a Factory, the total numbers of employees were 1600. In its Production Dept. there were 550 workers, of which 150 were female. The R&D Dept. had 350 male workers. The total number of workers in the Administrative Dept. was 500. The ratio of male to female workers in the factory is 3:2. Tabulate the above information.

Solution:
From the given information we can see that there are two characteristics i.e. (i) Sex of the workers and (ii) Departments of the Factory.
Thus, the data can be tabulated in the table with the given information as follows:

<table>
<thead>
<tr>
<th>Workers Dept.</th>
<th>Male</th>
<th>Female</th>
<th>Total</th>
</tr>
</thead>
<tbody>
<tr>
<td>Production</td>
<td>?</td>
<td>150</td>
<td>550</td>
</tr>
<tr>
<td>R &amp; D</td>
<td>350</td>
<td>?</td>
<td>?</td>
</tr>
<tr>
<td>Administrative</td>
<td>?</td>
<td>?</td>
<td>500</td>
</tr>
<tr>
<td>Total</td>
<td>?</td>
<td>?</td>
<td>1600</td>
</tr>
</tbody>
</table>

The data for blank cells are calculated as follows:
The number of male workers in the Production dept. = 550 – 150 = 400.
The total number of workers in R & D Dept. = 1600 – 550 – 500 = 550.
Hence, the number of female workers in this dept. = 550 – 350 = 200.
The ratio of male to female workers is 3:2,
\[
\therefore \text{number of male workers} = \frac{3}{5} \times 1600 = 960
\]
and number of female workers = \[
\frac{2}{5} \times 1600 = 640
\]
Now, the number of female workers in the administrative dept. = 640 – 150 – 200 = 290.
Hence, the number of male workers in this dept. = 500 – 290 = 210.
Thus, the completed table is given by
Example 3:
The ratio of male to female workers in ABC Company was 5:3. Out of the total of 160 workers, 25 men were Post Graduates and 20 women were Graduates. The number of men which were HSC was 15 more than that of Post Graduates. Number of Post Graduate women is one-third of HSC women. Tabulate the above data.

Solution:
From the given information we can see that there are two characteristics i.e. (i) Qualification and (ii) Sex of workers.
Thus, the data can be tabulated in the table with the given information as follows:

<table>
<thead>
<tr>
<th></th>
<th>Workers</th>
<th>Men</th>
<th>Women</th>
<th>Total</th>
</tr>
</thead>
<tbody>
<tr>
<td>Post Graduate</td>
<td></td>
<td>25</td>
<td>?</td>
<td>?</td>
</tr>
<tr>
<td>Graduates</td>
<td></td>
<td>?</td>
<td>20</td>
<td>?</td>
</tr>
<tr>
<td>HSC</td>
<td></td>
<td>?</td>
<td>?</td>
<td>?</td>
</tr>
<tr>
<td>Total</td>
<td></td>
<td>?</td>
<td>?</td>
<td>160</td>
</tr>
</tbody>
</table>

Now, the ratio of men to women workers is 5:3 and total number of workers = 160

\[ \text{number of men} = \frac{5}{8} \times 160 = 100 \] 
\[ \text{number of women} = \frac{3}{8} \times 160 = 60 \]

The number of HSC men = 15 + 25 = 40.
Hence, number of Graduate men = 100 - 25 - 40 = 35
\[ \text{total number of Graduates} = 35 + 20 = 55. \]

Now, let number of women with HSC = \( x \), then number of Post Graduate women = \( \frac{x}{3} \)

\[ \therefore x + \frac{x}{3} + 20 = 60 \]
\[ \Rightarrow x + \frac{x}{3} = 40 \]
\[ \Rightarrow \frac{4x}{3} = 40 \]
\[ \therefore x = 30 \]

Thus, women with HSC = 30 and Post Graduate women = 10.
Total number of Post Graduates = 25 + 10 = 35
Total number of HSC workers = 40 + 30 = 70.
Thus, the completed table is given by

<table>
<thead>
<tr>
<th></th>
<th>Workers</th>
<th>Men</th>
<th>Women</th>
<th>Total</th>
</tr>
</thead>
<tbody>
<tr>
<td>Post Graduate</td>
<td></td>
<td>25</td>
<td>?</td>
<td>?</td>
</tr>
<tr>
<td>Graduates</td>
<td></td>
<td>?</td>
<td>20</td>
<td>?</td>
</tr>
<tr>
<td>HSC</td>
<td></td>
<td>?</td>
<td>?</td>
<td>?</td>
</tr>
<tr>
<td>Total</td>
<td></td>
<td>?</td>
<td>?</td>
<td>160</td>
</tr>
</tbody>
</table>
Example 4:

A survey of the number of students of Almighty Residential School from 2002 to 2005 revealed the following details: In 2002-03 the total number of students was 360, of which 100 were girls. The number of Indian boys outnumbered the NRI boys by 20. The ratio of Indian and NRI girls was 3:2. In 2003-04, the number of Indian boys increased by 60, while the total number of Indians increased by 50%. The number of NRI girls increased by 20 and the total number of NRI’s increased by 25%. In 2004-05, in comparison with the previous year, the total number of Indians increased by 200 and the Indian girls increased by 50%. The total number of students were 750, of which the boys were 550. Tabulate the above information.

Solution:

From the given information we can see that there are three characteristics i.e. (i) Gender, (ii) Nationality and (iii) year of enrollment. To tabulate this we require a three-way table.

Thus, the data can be tabulated in the 3-way table with the given information as follows:

<table>
<thead>
<tr>
<th>Year</th>
<th>Students</th>
<th>Total</th>
</tr>
</thead>
<tbody>
<tr>
<td></td>
<td>Indian</td>
<td>NRI</td>
</tr>
<tr>
<td></td>
<td>Boys</td>
<td>Girls</td>
</tr>
<tr>
<td></td>
<td>a</td>
<td>b</td>
</tr>
</tbody>
</table>

For the year 2002-03:

No. of Indian students = 360 – 100 = 260

Since the Indian boys outnumbered the NRI boys by 20, if the no. of NRI boys is \( x \) then no. of Indian boys is \( x + 20 \).

Hence, \((x + 20) + x = 260 \Rightarrow 2x + 20 = 260 \Rightarrow 2x = 240 \Rightarrow x = 120\)
Thus, the no. NRI boys = 120 and no. of Indian boys = 140
The ratio of Indian and NRI girls is 3:2

<table>
<thead>
<tr>
<th>Year</th>
<th>Students</th>
<th>Total</th>
</tr>
</thead>
<tbody>
<tr>
<td></td>
<td>Indian</td>
<td></td>
</tr>
<tr>
<td></td>
<td>Boys a</td>
<td>Girls b</td>
</tr>
<tr>
<td></td>
<td>NRI</td>
<td></td>
</tr>
<tr>
<td></td>
<td>Boys c</td>
<td>Girls d</td>
</tr>
<tr>
<td></td>
<td>Boys a+c</td>
<td>Girls b+d</td>
</tr>
<tr>
<td>2002-03</td>
<td>140</td>
<td>60</td>
</tr>
<tr>
<td></td>
<td>120</td>
<td>40</td>
</tr>
<tr>
<td></td>
<td>260</td>
<td>100</td>
</tr>
<tr>
<td>2003-04</td>
<td>200</td>
<td>100</td>
</tr>
<tr>
<td></td>
<td>140</td>
<td>60</td>
</tr>
<tr>
<td></td>
<td>340</td>
<td>160</td>
</tr>
<tr>
<td>2004-05</td>
<td>350</td>
<td>150</td>
</tr>
<tr>
<td></td>
<td>200</td>
<td>50</td>
</tr>
<tr>
<td></td>
<td>550</td>
<td>200</td>
</tr>
</tbody>
</table>

UNIT IV

5

DIAGRAMS AND GRAPHS

OBJECTIVES

After learning this topic you will be able to know
- What is the importance of diagrams?
- How to interpret the given diagrams i.e. read the given diagram/s and answer the questions asked on the diagram.

5.2 Introduction

5.1.1 Importance of diagram

To understand various trends of the data (which is already classified and tabulated) at a glance and to facilitate the comparison of various situations, the data are presented in the form of diagrams & graphs. Diagrams are useful in many ways, some of the uses are given below:-

i) Diagrams are attractive and impressive.
ii) Diagrams are useful in simplifying the data.
iii) Diagrams save time and labour.
iv) Diagrams are useful in making comparisons.
v) Diagrams have universal applicability.
(i.e. They are used in almost in every field of study like business,
5.1.2 Limitations of Diagrams
Some of the limitations of diagrams are listed below:
i) Diagrams give only a vague idea of the problem which may be useful for a common man but not for an expert who wishes to have an exact idea of the problem.

ii) The level of precision of values indicated by diagrams is very low.

iii) Diagrams are helpful only when comparisons are desired. They don't lead to any further analysis of data.

iv) They can be a supplement to the tabular presentation but not an alternative to it.

* Note that we will discuss here only interpretation of diagrams and no construction of diagrams, since construction of diagrams is not in the F.Y.BCom syllabus.

5.2 TYPES OF DIAGRAM

1) **One dimensional diagrams:**
   a) Simple Bar diagrams
   b) Multiple Bar diagrams (joint Bar Diagrams)
   c) Sub-divided Bar Diagram
   d) Percent Bar Diagrams.

2) **Two dimensional diagrams:** Pie or Circular diagram.

Let us solve some examples based on the above diagrams.

a) **Simple Bar Diagram:**

Example 1
Study the following diagram giving enrollments for C.F.A. and answer the questions given below:

![Simple Bar Diagram]

**Fig. 5.1**

i) Name the above diagram – *Simple Bar Diagram*.

ii) Which year has least enrollment? – *The year 1990.*
iii) What is the percentage increase in enrollments from the year 1991 to 1992? \[ \frac{12-11}{11} \times 100 = 9.09\% \]

iv) What is the enrollment for the year 1993? – 15,000.

b) Multiple Bar Diagrams:
Example 2:
The following diagram gives the figures of Indo-US trade during 1987 to 1990. The figures of Indian Exports & Imports are in billion. Answer the following questions from the diagram given below:

![Fig. 5.2](image)

i) Name the above diagram.
ii) Which year shows the maximum exports?
iii) The difference in exports and imports is least in the year ---- and maximum in the year -----
iv) The percentage decrease in exports from year 1989 to year 1990 is -----
v) The import for the year 1987 is -----

c) Sub-divided Bar Diagrams:
Example 3:
The following diagram shows the number of Govt. and Non-Govt. Companies in various years. Study the diagram and answer the questions given below:

![Fig. 5.3](image)
i) Name the above diagram.
ii) The number of Govt. companies in year 1976 is -----.
iii) The increase in Non-Govt. companies from the year 1986 to year 1991 is -----.
iv) In which two years the number of non-Govt. companies is equal?

d) Percentage Bar Diagrams:
Example 4:
The following diagram gives monthly income and expenditure of two families A and B. Study the diagram and answers the questions given below:

![Percentage Bar Diagram](image)

i) Name the above diagram
ii) Which family spends more on clothing?
iii) If the total income of family A is Rs. 20,000/- then family A spends Rs. ----- on Rent.
iv) Family A spends equal amount on which two items?
v) The expenses on item ------ is double than the expenses on item ---- -- in Family B.
e) pie or circular Diagram

Example 5:
The following diagram shows the expenditure of an average working class family. Study the diagram & answer the questions given below.

![Pie Chart](image)

i) Name the above diagram.
ii) Expenses on ------- are maximum and ------ are minimum.
iii) If the total income is Rs.30,000/- then expenses on Housing & fuel together is Rs.-----.
(Hint: Housing + fuel = (12+5)% i.e. 17%)
Required Expenses = 17 \times 30,000 = 5100/- \)

Limitations of pie-diagrams:
Pie-diagrams are less effective than bar diagrams for accurate reading and interpretation, particularly when the data is divided into a large number of components is very small.

5.3 GRAPHS

A large variety of graphs are used in practice. Here we will be discussing the graphs of frequency distributions only.

A frequency distribution can be presented graphically in any of the following ways:
1) Histogram
2) Frequency Polygon
3) Frequency Curve
4) 'Ogives' or cumulative frequency curves.

Remark: We will deal with equal class intervals’ Histograms only.

1) Histogram: It is a graph of a frequency distribution in which the class intervals are plotted on x-axis and their respective frequencies on y-axis. On each class a rectangle is drawn, the height of each rectangle is taken to be equal to the frequency of the corresponding class. The construction of such a Histogram is shown in the following example.

Example 6:
Draw a Histogram for the following distribution giving the marks obtained by 60 students of a class in a college.

Marks: 20-24 25-29 30-34 35-39 40-44 45-49 50-54
No. of students: 3 5 12 18 14 6 2

Solution: Here class intervals given are of inclusive type. State the upper limit of a class is not equal to the lower limit of its following class, the class boundaries will have to be determined. After the adjustment, the distribution will be as below.

Marks: 19.5-24.5 24.5-29.5 29.5-34.5 34.5-39.5 39.5-44.5 44.5-49.5 49.5-54.5
No. of students: 3 5 12 18 14 6 2

scale: x-axis 1 cm. = 10 marks
y-axis 1 cm. = 5 students.
5.4 FREQUENCY POLYGON

It is another method of representing a frequency distribution of a graph. Frequency polygons are more suitable than histograms whenever two or more frequency distributions are to be compared.

Frequency polygon of a grouped or continuous freq. distribution is a straight line graph. The frequencies of the classes are plotted against the mid-values of the corresponding classes. The points so obtained are joined by straight lines (segments) to obtain the frequency polygon.

Example 7:
The following data show the number of accidents sustained by 313 drivers of a public utility company over a period of 5 years. Draw the frequency polygon

<table>
<thead>
<tr>
<th>No. of accidents</th>
<th>0</th>
<th>1</th>
<th>2</th>
<th>3</th>
<th>4</th>
<th>5</th>
<th>6</th>
<th>7</th>
<th>8</th>
<th>9</th>
<th>10</th>
<th>11</th>
</tr>
</thead>
<tbody>
<tr>
<td>No. of drivers</td>
<td>80</td>
<td>44</td>
<td>68</td>
<td>41</td>
<td>25</td>
<td>20</td>
<td>13</td>
<td>7</td>
<td>5</td>
<td>4</td>
<td>3</td>
<td>2</td>
</tr>
</tbody>
</table>

Frequency polygon of a number of accidents
Example 8:
Draw a Histogram & Frequency polygon from the following distribution giving marks of 50 students in statistics.

Marks in Statistics: 0-10 10-20 20-30 30-40 40-50 50-60 60-70 70-80 80-90
No. of students: 0 2 3 7 13 13 9 2 1

Note that, here we will first draw Histogram and then the mid-points of the top of bars are joined by line segments to get the frequency polygon.

Remark: Note that frequency polygon can be drawn even without converting the given distribution into classes. The frequencies are plotted against the corresponding mid-points (given) and joined by line segments.
Example 9:
Draw a frequency polygon to represent the following distribution.

<table>
<thead>
<tr>
<th>Monthly Income (in Rs.)</th>
<th>Number of families</th>
</tr>
</thead>
<tbody>
<tr>
<td>0-500</td>
<td>10</td>
</tr>
<tr>
<td>500-1000</td>
<td>15</td>
</tr>
<tr>
<td>1000-1500</td>
<td>18</td>
</tr>
<tr>
<td>1500-2000</td>
<td>12</td>
</tr>
<tr>
<td>2000-2500</td>
<td>8</td>
</tr>
<tr>
<td>2500-3000</td>
<td>4</td>
</tr>
</tbody>
</table>

5.5 FREQUENCY CURVE

Frequency curve is similar like frequency polygon, only the difference is that the points are joined by a free hand curve instead of line segments as we join in frequency polygon.

Let us study the following examples to study the concept of frequency curve.

Example 10: Draw a frequency curve for the following data:
Age (yrs) : 17-19 19-21 21-23 23-25 25-27 27-29 29-31
No. of students : 7 13 24 30 22 15 6

Solution: Here we will take ages on horizontal axis (i.e. x-axis) and no. of students on vertical axis (i.e. say y-axis)

We will plot the given frequencies (No. of students) against mid-points of the given class interval & then join these points by free hand curve. Extremities (first and the last points plotted) are joined to the mid-points of the neighboring class intervals.
Example 11: Draw a Histogram & frequency curve for the following data:

<table>
<thead>
<tr>
<th>Age (yrs)</th>
<th>No. of students</th>
</tr>
</thead>
<tbody>
<tr>
<td>17-19</td>
<td>7</td>
</tr>
<tr>
<td>19-21</td>
<td>13</td>
</tr>
<tr>
<td>21-23</td>
<td>24</td>
</tr>
<tr>
<td>23-25</td>
<td>30</td>
</tr>
<tr>
<td>25-27</td>
<td>22</td>
</tr>
<tr>
<td>27-29</td>
<td>15</td>
</tr>
<tr>
<td>29-31</td>
<td>6</td>
</tr>
</tbody>
</table>

Solution: First we will draw histogram & then the mid-points of the top of the bars are joined by free hand curve to get the frequency curve.

5.6 Ogive or Cumulative Frequency Curve

Ogive (pronounced as ogive) is a graphic presentation of the cumulative frequency (c.f) distribution of continuous variable. It consists in plotting the c.f. (along the y-axis) against the class boundaries (along x-axis). Since there are two types of c.f. distributions namely 'less than' c.f. and 'more than' c.f., we have accordingly two types of ogives, namely, i) 'less than and equal to' ogive, ii) 'more than and equal to' ogive.

'Less than and equal to ogive': This consists of plotting the 'less than' cumulative frequencies against the upper class boundaries of the respective classes. The points so obtained are joined by a smooth free
hand curve to give 'less than ' ogive. Obviously, 'less than ' ogive is an increasing curve, sloping upwards from left to right.

**Note:** Since the frequency below the lower limit of the first class (i.e. upper limit of the class preceding the first class) is zero, the ogive curve should start on the left with a cum. freq. zero of the lower boundary of the first class.

**‘More than and equal to’ ogive:** This consists of plotting the 'more than ' cum. frequencies against the lower class boundaries of the respective classes. The points so obtained are joined by a smooth free hand curve to give 'equal give. Obviously, 'more than ' ogive is a decreasing curve, slopes downwards from left to right.

**Note:** We may draw both the 'less than' and 'more than and equal to' ogives on the same graph. If done so, they intersect at a point. The foot of the perpendicular from their point of intersection on the x-axis gives the value of the median.

**Remark:** Ogives are particularly useful for graphic computation of partition values namely median, quartiles, deciles, percentiles etc. They can also be used to determine graphically the number of proportion of observations below or above a given value of the variable or lying between certain interval of the values of the variable.

We will now study the following examples to understand the concepts discussed above:-

**Example 12:** The data below give the marks secured by 70 students to a certain examination:
i) Draw the 'less than ' type ogive curve.
ii) Use the ogive curve to estimate the percentage of students getting marks less than 45.

<table>
<thead>
<tr>
<th>Marks</th>
<th>No. of students</th>
</tr>
</thead>
<tbody>
<tr>
<td>0-10</td>
<td>2</td>
</tr>
<tr>
<td>10-20</td>
<td>3</td>
</tr>
<tr>
<td>20-30</td>
<td>6</td>
</tr>
<tr>
<td>30-40</td>
<td>11</td>
</tr>
<tr>
<td>40-50</td>
<td>12</td>
</tr>
<tr>
<td>50-60</td>
<td>15</td>
</tr>
<tr>
<td>60-70</td>
<td>10</td>
</tr>
<tr>
<td>70-80</td>
<td>7</td>
</tr>
<tr>
<td>80-90</td>
<td>4</td>
</tr>
</tbody>
</table>

**Solution:** We will take marks on x-axis & No. of students on y-axis. Also we will calculate 'less than ' cumulative frequencies & then we will plot these calculated frequencies against the upper limit of the corresponding class & joining the points so obtained by a smooth free hand curve to get a 'less than' ogive

<table>
<thead>
<tr>
<th>Marks</th>
<th>No.of students</th>
<th>Less than cf</th>
</tr>
</thead>
<tbody>
<tr>
<td>0-10</td>
<td>2</td>
<td>→ 2</td>
</tr>
<tr>
<td>10-20</td>
<td>3</td>
<td>→ 2+3=5</td>
</tr>
</tbody>
</table>
ii) To estimate the number of students getting marks less than 45, draw perpendicular to the x-axis (representing marks) at x=45, meeting the 'less than' ogive at point P. From P draw a perpendicular PM on the y-axis (representing number of students).

Then, from the graph OM = 26.8 (approximate) ≈ 27 is the number of candidates getting score 45 or less.

Hence, the percentage of students getting less than 45 marks is given by \( \frac{27}{70} \times 100 \) i.e. 38.57\%

(Here total students =70).

Example 13: The following table gives the distribution of monthly income of 600 families in a certain city.

<table>
<thead>
<tr>
<th>Monthly Income Below 75</th>
<th>75-150</th>
<th>150-225</th>
<th>225-300</th>
<th>300-375</th>
<th>375-450</th>
<th>Above 450</th>
</tr>
</thead>
<tbody>
<tr>
<td>No. of families</td>
<td>60</td>
<td>170</td>
<td>200</td>
<td>60</td>
<td>50</td>
<td>40</td>
</tr>
</tbody>
</table>

Draw a 'less than' and a 'more than equal to' ogive curve for the above data on the same graph and from there read the median income.

Solution: For drawing the 'less than' and 'more than' equal to ogive we convert the given distribution into 'less than' and 'more than equal to' cumulative frequencies (c.f.) as given in the following table.
<table>
<thead>
<tr>
<th>Monthly Income ('00 Rs.)</th>
<th>No. of families (f)</th>
<th>Less than c.f.</th>
<th>More than c.f</th>
</tr>
</thead>
<tbody>
<tr>
<td>Below 75</td>
<td>60</td>
<td>60</td>
<td>600</td>
</tr>
<tr>
<td>75-150</td>
<td>170</td>
<td>230</td>
<td>540</td>
</tr>
<tr>
<td>150-225</td>
<td>200</td>
<td>430</td>
<td>370</td>
</tr>
<tr>
<td>225-300</td>
<td>60</td>
<td>490</td>
<td>170</td>
</tr>
<tr>
<td>300-375</td>
<td>50</td>
<td>540</td>
<td>110</td>
</tr>
<tr>
<td>375-450</td>
<td>40</td>
<td>580</td>
<td>60</td>
</tr>
<tr>
<td>450 and above</td>
<td>20</td>
<td>600</td>
<td>20</td>
</tr>
<tr>
<td><strong>TOTAL</strong></td>
<td><strong>600</strong></td>
<td></td>
<td></td>
</tr>
</tbody>
</table>

'Less than' and 'More than & equal to' ogive

Example 14:
Convert the following distribution into ' more than and equal to ' frequency distribution.

Weekly wages less than ('00 Rs.) : 20 40 60 80 100
No. of workers : 41 92 156 194 201

For the data given above, draw ' less than ' and equal to more than ' ogives and hence find the value of median.
Solution:

<table>
<thead>
<tr>
<th>Weekly wages (in '00 Rs.)</th>
<th>No. of Workers (f)</th>
<th>Cumulative Freq. (c.f) Less than</th>
<th>Cumulative Freq. (c.f) More than</th>
</tr>
</thead>
<tbody>
<tr>
<td>0-20</td>
<td>41</td>
<td>41</td>
<td>160+41=201</td>
</tr>
<tr>
<td>20-40</td>
<td>92 -41=51</td>
<td>92</td>
<td>51+109=160</td>
</tr>
<tr>
<td>40-60</td>
<td>156-92=64</td>
<td>156</td>
<td>64+45=109</td>
</tr>
<tr>
<td>60-80</td>
<td>194-156=38</td>
<td>194</td>
<td>38+7=45</td>
</tr>
<tr>
<td>80-100</td>
<td>201-194=7</td>
<td>201</td>
<td>7</td>
</tr>
</tbody>
</table>

'Less than' and 'More than & equal to' ogives are shown in the following figure:

Example 15: Draw 'more than and equal to' ogive for the following data

Marks : 0-10 10-20 20-30 30-40 40-50 50-60 60-70 70-80
No.of students : 4 6 10 15 25 22 11 7

Here find i) median marks ii) Number of students getting 60 or more marks.

Solution:

First we prepare 'more than and equal to' c.f. distribution table for the given data.
<table>
<thead>
<tr>
<th>Marks more than (lower limit of class interval)</th>
<th>No. of students</th>
<th>More than c.f.</th>
</tr>
</thead>
<tbody>
<tr>
<td>0</td>
<td>4</td>
<td>100</td>
</tr>
<tr>
<td>10</td>
<td>6</td>
<td>96</td>
</tr>
<tr>
<td>20</td>
<td>10</td>
<td>90</td>
</tr>
<tr>
<td>30</td>
<td>15</td>
<td>80</td>
</tr>
<tr>
<td>40</td>
<td>25</td>
<td>65</td>
</tr>
<tr>
<td>50</td>
<td>22</td>
<td>40</td>
</tr>
<tr>
<td>60</td>
<td>11</td>
<td>18</td>
</tr>
<tr>
<td>70</td>
<td>7</td>
<td>7</td>
</tr>
<tr>
<td>TOTAL</td>
<td>100</td>
<td></td>
</tr>
</tbody>
</table>

We draw more than c.f. curve by taking lower limit of class interval on x-axis and more than c.f. on y-axis.

![Graph of more than c.f. curve](image)

i) To locate median, we find \( \frac{N}{2} = \frac{100}{2} = 50 \) (N is frequency total)

Now draw a horizontal line parallel to x-axis passing through a point on y-axis (0,50). The line will meet the curve at some point. From that point draw a perpendicular line meeting x-axis. The foot of the perpendicular will be the value of the median.

Therefore, Graphical value of median = 45 (M = median)

ii) To find the number of students getting 60 or more marks, we draw a vertical line passing through (60,0) on x-axis. The line will meet the curve
at some point. From that point draw a horizontal line which will meet y-axis at (0,18).

Therefore, the number of students getting 60 or more marks = 18.

Note: We can find no. of students getting less than 60 marks by simply calculating:- Total no. of students – 18 = 100 – 18 = 82.

EXERCISE

1) What is the importance of diagrams & graphs in statistical analysis?

2) What is an ogive? Explain its uses.

3) Name different types of diagrams.

4) Draw histogram and locate mode graphically.
   Marks : 0-10 10-20 20-30 30-40 40-50
   No. of students : 6 11 15 8 3

5) Draw less than ogive curve for the following data.
   Wages (in Rs.) : 70-80 80-90 90-100 100-110 110-120 120-130 130-140
   No. of workers : 85 109 126 134 115 83 68

Hence find the number of workers whose wages are less than Rs. 117.
   Also find no. of workers whose wages are Rs. 117 or above.

6) Study the following diagram giving the occupation of males and females in India and answer the following questions:
i) Name the above diagram.

ii) In which occupation the proportion of males was the highest in India?

iii) In which industry the absolute difference between the proportion of males and females was the highest?

7) Study the following diagram giving percentage of investment pattern in the fifth plan & answer the questions given below:

i) Name the above diagram.

ii) What is the highest percentage of investment pattern in the fifth plan?

iii) Which is the smallest sector in the above diagram?

iv) What is the contribution of irrigation and power sector?

8) Draw a Histogram & freq. polygon for the following data:

Marks : 0-10  10-20  20-30  30-40  40-50  50-60
No. of students : 3  8  12  18  10  4
REFERENCES:

1) Business statistics by R.S. Bharadwaj - Excel Books - New Delhi
3) Statistical Methods by Dr. S. P. Gupta - Sultan Chand & sons - New Delhi

UNIT IV

MEASURES OF CENTRAL TENDENCY

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.

6.1 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.

### 6.2 REQUISITES FOR AN IDEAL AVERAGE

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

1. It should be rigidly defined.
2. It should be easy to calculate and easy to understand.
3. It should be based on all observations.
4. It should be suitable for further mathematical treatment.
5. It should be least affected by the fluctuations of sampling.
6. It should not be affected by extreme values.

### 6.3 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.

### 6.4 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 \( x_1, x_2, \ldots, x_n \) are the \( n \) observations of variable \( x \), then the arithmetic mean is given by

\[
\bar{x} = \frac{x_1 + x_2 + \cdots + x_n}{n}
\]

If we denote the sum \( x_1 + x_2 + \cdots + x_n \) as \( \sum x \), then

\[
\bar{x} = \frac{\sum x}{n}
\]

**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:**
\( n = 6 \)

\[
\bar{x} = \frac{\sum x}{n} \Rightarrow \bar{x} = \frac{74 + 89 + 93 + 68 + 85 + 76}{6} = \frac{485}{6} = 80.83
\]

**Example 2:** Find arithmetic mean for the following data.

\( 425, 408, 441, 435, 418 \)

**Solution:**
\( n = 5 \) and \( \bar{x} = \frac{\sum x}{n} \)

\[
\therefore \bar{x} = \frac{425 + 408 + 441 + 435 + 418}{5} = \frac{2127}{5} = 425.4
\]

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

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

\[
\bar{x} = \frac{f_1 x_1 + f_2 x_2 + \cdots + f_n x_n}{f_1 + f_2 + \cdots + f_n}
\]

\[
\bar{x} = \frac{\sum fx}{\sum f}
\]
Or

Where \( \sum f = f_1 + f_2 + \ldots + f_n = \) sum of the frequencies

If 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.

<table>
<thead>
<tr>
<th>Age in years</th>
<th>No. of Students</th>
</tr>
</thead>
<tbody>
<tr>
<td>11</td>
<td>7</td>
</tr>
<tr>
<td>12</td>
<td>10</td>
</tr>
<tr>
<td>13</td>
<td>16</td>
</tr>
<tr>
<td>14</td>
<td>12</td>
</tr>
<tr>
<td>15</td>
<td>8</td>
</tr>
<tr>
<td>16</td>
<td>11</td>
</tr>
<tr>
<td>17</td>
<td>5</td>
</tr>
</tbody>
</table>

\[
\sum f = N = 69, \quad \sum fx = 954
\]

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

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

<table>
<thead>
<tr>
<th>Bonus (in Rs.)</th>
<th>No. of persons</th>
</tr>
</thead>
<tbody>
<tr>
<td>40</td>
<td>2</td>
</tr>
<tr>
<td>50</td>
<td>5</td>
</tr>
<tr>
<td>60</td>
<td>7</td>
</tr>
<tr>
<td>70</td>
<td>6</td>
</tr>
<tr>
<td>80</td>
<td>4</td>
</tr>
<tr>
<td>90</td>
<td>8</td>
</tr>
<tr>
<td>100</td>
<td>3</td>
</tr>
</tbody>
</table>

\[
\sum f = N = 69, \quad \sum fx = 954
\]

\[
\bar{x} = \frac{\sum fx}{N} = 13.83 \text{ years.}
\]
\[ \bar{x} = \frac{\sum fx}{N} = \frac{2510}{35} = \text{Rs. 71.71} \]

(iii) \textbf{Grouped data} (class intervals and frequencies are given)

Formula: \[ \bar{x} = \frac{\sum fx}{N} \] where \( N = \Sigma f \) and \( x = \text{Mid-point of the class intervals} \)

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

\textbf{Example 5:}
Find the arithmetic mean for the following data representing marks of 60 students.

<table>
<thead>
<tr>
<th>Marks</th>
<th>10-20</th>
<th>20-30</th>
<th>30-40</th>
<th>40-50</th>
<th>50-60</th>
<th>60-70</th>
<th>70-80</th>
</tr>
</thead>
<tbody>
<tr>
<td>No. of Students</td>
<td>8</td>
<td>15</td>
<td>13</td>
<td>10</td>
<td>7</td>
<td>4</td>
<td>3</td>
</tr>
</tbody>
</table>

**Solution:**

\[
\begin{array}{c|c|c|c|c|c|c|c}
\hline
\text{Marks} & \text{No. of students} & \text{Mid-point (x)} & \text{fx} \\
\hline
10-20 & 8 & 15 & 120 \\
20-30 & 15 & 25 & 375 \\
30-40 & 13 & 35 & 455 \\
40-50 & 10 & 45 & 450 \\
50-60 & 7 & 55 & 385 \\
60-70 & 4 & 65 & 260 \\
70-80 & 3 & 75 & 225 \\
\hline
\text{Total} & \text{N} = 60 & \sum fx = 2270 \\
\hline
\end{array}
\]

\[ \bar{x} = \frac{\sum fx}{N} = \frac{2270}{60} = 37.83 \]
\[ \therefore \text{The average marks are 37.83} \]

\textbf{Example 6:}
Calculate the arithmetic mean of heights of 80 Students for the following data.

<table>
<thead>
<tr>
<th>Heights in cms</th>
<th>130-134</th>
<th>135-139</th>
<th>140-144</th>
<th>145-149</th>
<th>150-154</th>
<th>155-159</th>
<th>160-164</th>
</tr>
</thead>
<tbody>
<tr>
<td>No. of</td>
<td>7</td>
<td>11</td>
<td>15</td>
<td>21</td>
<td>16</td>
<td>6</td>
<td>4</td>
</tr>
</tbody>
</table>
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
130.5 = 129.5 and 134.5 = 134.5
135.5 = 134.5 and 139.5 = 139.5 and so on.

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

\[ N = \sum f = 80 \text{ and } \sum fx = 11270 \]

\[ \therefore \bar{x} = \frac{\sum fx}{N} = \frac{11270}{80} = 140.88 \]

\[ \therefore \text{The average height of students} = 140.88 \text{ cm}. \]

6.5 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 \( x_1, x_2, \ldots, x_n \) are the \( n \) values of the variable \( x \) with the corresponding weights \( w_1, w_2, \ldots, w_n \), then the weighted mean is given by

\[ x_w = \frac{w_1x_1 + w_2x_2 + \ldots + w_nx_n}{w_1 + w_2 + \ldots + w_n} = \frac{\sum wx}{\sum w} \]
Where \( x_w = \text{Weighted Arithmetic Mean} \) and \( \sum w = \text{Sum of the weights} \).

**Steps:**
(1) Multiply the weights \( w \) by the variable \( x \) and obtain the total \( \sum x_w \).
(2) Divide the total \( \sum x_w \) by the sum of the weights \( \sum w \).

**Example 7:**
Find the weighted mean for the following data.

<table>
<thead>
<tr>
<th>( x )</th>
<th>28</th>
<th>25</th>
<th>20</th>
<th>32</th>
<th>40</th>
</tr>
</thead>
<tbody>
<tr>
<td>( W )</td>
<td>3</td>
<td>6</td>
<td>4</td>
<td>5</td>
<td>8</td>
</tr>
</tbody>
</table>

<table>
<thead>
<tr>
<th>( x )</th>
<th>( W )</th>
<th>( xw )</th>
</tr>
</thead>
<tbody>
<tr>
<td>28</td>
<td>3</td>
<td>84</td>
</tr>
<tr>
<td>25</td>
<td>6</td>
<td>150</td>
</tr>
<tr>
<td>20</td>
<td>4</td>
<td>80</td>
</tr>
<tr>
<td>32</td>
<td>5</td>
<td>160</td>
</tr>
<tr>
<td>40</td>
<td>8</td>
<td>320</td>
</tr>
<tr>
<td><strong>Total</strong></td>
<td><strong>26</strong></td>
<td><strong>794</strong></td>
</tr>
</tbody>
</table>

\[ x_w = \frac{\sum xw}{\sum w} = \frac{794}{26} = 30.54 \text{ units} \]

**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:**

<table>
<thead>
<tr>
<th>Subject</th>
<th>Percentage of Marks (( x ))</th>
<th>Weights (( w ))</th>
<th>( wx )</th>
</tr>
</thead>
<tbody>
<tr>
<td>English</td>
<td>64</td>
<td>2</td>
<td>128</td>
</tr>
<tr>
<td>Mathematics</td>
<td>93</td>
<td>3</td>
<td>279</td>
</tr>
<tr>
<td>Economics</td>
<td>72</td>
<td>3</td>
<td>216</td>
</tr>
<tr>
<td>Accountancy</td>
<td>85</td>
<td>4</td>
<td>340</td>
</tr>
<tr>
<td>Statistics</td>
<td>79</td>
<td>1</td>
<td>79</td>
</tr>
<tr>
<td><strong>Total</strong></td>
<td></td>
<td><strong>13</strong></td>
<td><strong>1042</strong></td>
</tr>
</tbody>
</table>

\[ \sum w = 13, \quad \sum wx = 1042 \]

\[ x_w = \frac{\sum wx}{\sum w} = \frac{1042}{13} = 80.15 \quad \therefore \text{Weighted Mean Marks} = 80.15 \]

**6.6 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

\[
\bar{x} = \frac{n_1\bar{x}_1 + n_2\bar{x}_2}{n_1 + n_2}
\]

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

\[
\bar{x} = \frac{n_1x_1 + n_2x_2 + \ldots + n_kx_k}{n_1 + n_2 + \ldots + n_k}
\]

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
\( n_1 = 80 \quad n_2 = 70 \)
\( \bar{x}_1 = 1500 \quad \bar{x}_2 = 2200 \)

\[\therefore \bar{x} = \frac{n_1\bar{x}_1 + n_2\bar{x}_2}{n_1 + n_2} \]
\[= \frac{80 \times 1500 + 70 \times 2200}{80 + 70} \]
\[= \frac{120000 + 154000}{150} \]
\[= \frac{274000}{150} \]
\[= 1826.67 \]

\[\therefore \text{The average monthly salary of the combined group of 150 employees is Rs.} \text{ 1826.67} \]

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 \( n_1 = 50 \quad n_2 = 60 \quad n_3 = 90 \)
Mean weight \( x_1 = 58 \quad x_2 = 62 \quad x_3 = 56 \)

\[ \bar{x} = \frac{n_1 x_1 + n_2 x_2 + n_3 x_3}{n_1 + n_2 + n_3} \]
\[ = \frac{50 (58) + 60 (62) + 90 (56)}{50 + 60 + 90} \]
\[ = \frac{11660}{200} = 58.3 \]

\[ \therefore \text{The average weight of the combined group of 200 workers is 58.3 kgs} \]

**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 6.1**

1. Find the arithmetic mean for the following sets of observations.
   (i) 48, 55, 83, 65, 38, 74, 58
   (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:

<table>
<thead>
<tr>
<th>( x )</th>
<th>15</th>
<th>17</th>
<th>19</th>
<th>21</th>
<th>23</th>
<th>25</th>
</tr>
</thead>
<tbody>
<tr>
<td>( f )</td>
<td>6</td>
<td>11</td>
<td>8</td>
<td>15</td>
<td>5</td>
<td>4</td>
</tr>
</tbody>
</table>

   **Ans:** \( \bar{x} = 19.57 \)

3. Find the mean for the following data.

<table>
<thead>
<tr>
<th>Size of Shoe</th>
<th>5</th>
<th>6</th>
<th>7</th>
<th>8</th>
<th>9</th>
<th>10</th>
</tr>
</thead>
<tbody>
<tr>
<td>No. of pairs</td>
<td>22</td>
<td>35</td>
<td>28</td>
<td>42</td>
<td>15</td>
<td>12</td>
</tr>
</tbody>
</table>

   **Ans:** \( \bar{x} = 7.19 \)

4. Calculate the arithmetic mean for the following data

<table>
<thead>
<tr>
<th>Daily Wages in Rs</th>
<th>40-80</th>
<th>80-20</th>
<th>120-60</th>
<th>160-00</th>
<th>200-40</th>
<th>240-80</th>
<th>280-20</th>
</tr>
</thead>
</table>
| \( f \) | }
5. Calculate the mean for the following data representing monthly salary of a group of employees.

<table>
<thead>
<tr>
<th>Salary in Rs:</th>
<th>500-1000</th>
<th>1000-1500</th>
<th>1500-2000</th>
<th>2000-2500</th>
<th>2500-3000</th>
<th>3000-3500</th>
</tr>
</thead>
<tbody>
<tr>
<td>No. of persons:</td>
<td>7</td>
<td>11</td>
<td>13</td>
<td>8</td>
<td>5</td>
<td>4</td>
</tr>
</tbody>
</table>

\[
\bar{x} = \frac{\sum (\text{Salary} \times \text{No. of persons})}{\sum \text{No. of persons}}
\]

\[
\bar{x} = \frac{(7 \times 500) + (11 \times 1000) + (13 \times 1500) + (8 \times 2000) + (5 \times 2500) + (4 \times 3000)}{7 + 11 + 13 + 8 + 5 + 4}
\]

\[
\bar{x} = \frac{3500 + 11000 + 19500 + 16000 + 12500 + 12000}{46}
\]

\[
\bar{x} = 181.32
\]

6. Find the mean for the following data.

<table>
<thead>
<tr>
<th>Weekly wages in Rs:</th>
<th>50-99</th>
<th>100-149</th>
<th>150-199</th>
<th>200-249</th>
<th>250-299</th>
<th>300-349</th>
</tr>
</thead>
<tbody>
<tr>
<td>No. of workers:</td>
<td>12</td>
<td>17</td>
<td>15</td>
<td>20</td>
<td>8</td>
<td>7</td>
</tr>
</tbody>
</table>

\[
\bar{x} = \frac{\sum (\text{Weekly wages} \times \text{No. of workers})}{\sum \text{No. of workers}}
\]

\[
\bar{x} = \frac{(12 \times 55) + (17 \times 124) + (15 \times 174) + (20 \times 224) + (8 \times 254) + (7 \times 304)}{12 + 17 + 15 + 20 + 8 + 7}
\]

\[
\bar{x} = \frac{660 + 2082 + 2610 + 4480 + 2032 + 2128}{70}
\]

\[
\bar{x} = 1802.08
\]

7. Find the arithmetic mean for the following data

<table>
<thead>
<tr>
<th>Age in yrs:</th>
<th>10-19</th>
<th>20-29</th>
<th>30-39</th>
<th>40-49</th>
<th>50-59</th>
<th>60-69</th>
<th>70-79</th>
</tr>
</thead>
<tbody>
<tr>
<td>No. of persons:</td>
<td>6</td>
<td>12</td>
<td>19</td>
<td>23</td>
<td>15</td>
<td>8</td>
<td>6</td>
</tr>
</tbody>
</table>

\[
\bar{x} = \frac{\sum (\text{Age} \times \text{No. of persons})}{\sum \text{No. of persons}}
\]

\[
\bar{x} = \frac{(6 \times 15) + (12 \times 29) + (19 \times 39) + (23 \times 49) + (15 \times 59) + (8 \times 69) + (6 \times 79)}{70}
\]

\[
\bar{x} = \frac{90 + 348 + 741 + 1127 + 885 + 552 + 474}{70}
\]

\[
\bar{x} = 43.15
\]

8. Find the weighted mean for the following data.

<table>
<thead>
<tr>
<th>(x)</th>
<th>25</th>
<th>22</th>
<th>15</th>
<th>30</th>
<th>18</th>
</tr>
</thead>
<tbody>
<tr>
<td>(w)</td>
<td>6</td>
<td>4</td>
<td>3</td>
<td>8</td>
<td>2</td>
</tr>
</tbody>
</table>

\[
\bar{xw} = \frac{\sum (\text{\(x\)} \times \text{\(w\)})}{\sum \text{\(w\)}}
\]

\[
\bar{xw} = \frac{(25 \times 6) + (22 \times 4) + (15 \times 3) + (30 \times 8) + (18 \times 2)}{6 + 4 + 3 + 8 + 2}
\]

\[
\bar{xw} = \frac{150 + 88 + 45 + 240 + 36}{23}
\]

\[
\bar{xw} = 24.30
\]

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.

\[
\bar{x} = \frac{\sum (\text{Score} \times \text{Weight})}{\sum \text{Weight}}
\]

\[
\bar{x} = \frac{(44 \times 4) + (32 \times 3) + (27 \times 3) + (38 \times 4)}{4 + 3 + 3 + 4}
\]

\[
\bar{x} = \frac{176 + 96 + 81 + 152}{14}
\]

\[
\bar{x} = 36.07
\]

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.

\[
\bar{x} = \frac{\sum \text{Wages}}{\sum \text{Workers}}
\]

\[
\bar{x} = \frac{120 \times 78}{120} = 78
\]

\[
\bar{x} = \frac{80 \times 92}{80} = 92
\]

\[
\bar{x} = \frac{\sum \text{Wages}}{\sum \text{Workers}} = \frac{(120 \times 78) - (80 \times 92) + \bar{x}_f \times 40}{120 - 80 + 40}
\]

\[
\bar{x}_f = \frac{15600 - 7360 + \bar{x}_f \times 40}{40}
\]

\[
\bar{x}_f = \frac{8240 + \bar{x}_f \times 40}{40}
\]

\[
\bar{x}_f = \frac{8240}{40} + \bar{x}_f
\]

\[
\bar{x}_f = 206 + \bar{x}_f
\]

\[
\bar{x}_f = 206
\]

11. There are three groups in a class of 200 Students. The first groups 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.

\[
\bar{x} = \frac{\sum \text{Marks}}{\sum \text{Students}}
\]

\[
\bar{x} = \frac{(80 \times 65) + (70 \times 74)}{150}
\]

\[
\bar{x} = \frac{5200 + 5180}{150}
\]

\[
\bar{x} = 44.4
\]
12. Find the combined arithmetic mean for the following data:

<table>
<thead>
<tr>
<th></th>
<th>Group I</th>
<th>Group II</th>
<th>Group III</th>
</tr>
</thead>
<tbody>
<tr>
<td>Mean</td>
<td>125</td>
<td>141</td>
<td>131</td>
</tr>
<tr>
<td>No.of observations</td>
<td>150</td>
<td>100</td>
<td>50</td>
</tr>
</tbody>
</table>

Ans: 131.67

6.7 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:
   \[
   \text{Median} = \text{value of } \left( \frac{n+1}{2} \right)^{th} \text{term.}
   \]

3) When n is even:
   \[
   \text{Median} = \text{Arithmetic Mean of value of } \left( \frac{n}{2} \right) \text{ & } \left( \frac{n}{2} + 1 \right)^{th} \text{ terms.}
   \]

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

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

= 4 th term

Median = 65 units.

**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.

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

Median = 56

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

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

= A. M. of (4, 4+1) th terms

= A. M. of (4, 5) th terms

= A. M. of 45 and 67

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

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.

<table>
<thead>
<tr>
<th>Income in Rs:</th>
<th>1500</th>
<th>2500</th>
<th>3000</th>
<th>3500</th>
<th>4500</th>
</tr>
</thead>
<tbody>
<tr>
<td>No. of persons</td>
<td>12</td>
<td>8</td>
<td>15</td>
<td>6</td>
<td>5</td>
</tr>
</tbody>
</table>

**Solution:**
<table>
<thead>
<tr>
<th>Income (in Rs.) (x)</th>
<th>No. of persons (f)</th>
<th>c.ƒ less than type</th>
</tr>
</thead>
<tbody>
<tr>
<td>1500</td>
<td>12</td>
<td>12</td>
</tr>
<tr>
<td>2500</td>
<td>8</td>
<td>20</td>
</tr>
<tr>
<td>3000</td>
<td>15</td>
<td>35</td>
</tr>
<tr>
<td>3500</td>
<td>6</td>
<td>41</td>
</tr>
<tr>
<td>4500</td>
<td>5</td>
<td>46</td>
</tr>
</tbody>
</table>

\[ \text{N} = 46 \]

\[ \text{N} = \frac{46}{2} = 23 \]

In c.ƒ column, we get 35 as the first cumulative frequency greater than 23.

The value of \( \hat{x} \) corresponding to the c.ƒ 35 is 3000.

\[ \therefore \text{Median} = \text{Rs. 3000} \]

**Example : 14**

Find the median for the following data.

<table>
<thead>
<tr>
<th>( \hat{x} )</th>
<th>25</th>
<th>20</th>
<th>30</th>
<th>40</th>
<th>35</th>
<th>15</th>
</tr>
</thead>
<tbody>
<tr>
<td>( f )</td>
<td>19</td>
<td>14</td>
<td>23</td>
<td>12</td>
<td>20</td>
<td>8</td>
</tr>
</tbody>
</table>

**Solution :** Arrange the values in ascending order.

<table>
<thead>
<tr>
<th>( x )</th>
<th>( \hat{f} )</th>
<th>c.( \hat{f} ) less than type</th>
</tr>
</thead>
<tbody>
<tr>
<td>15</td>
<td>8</td>
<td>8</td>
</tr>
<tr>
<td>20</td>
<td>14</td>
<td>22</td>
</tr>
<tr>
<td>25</td>
<td>19</td>
<td>41</td>
</tr>
<tr>
<td>30</td>
<td>23</td>
<td>64</td>
</tr>
<tr>
<td>35</td>
<td>20</td>
<td>84</td>
</tr>
<tr>
<td>40</td>
<td>12</td>
<td>96</td>
</tr>
</tbody>
</table>

\[ \text{N} = \frac{96}{2} = 48 \]

In c.\( \hat{f} \) column, greater than 48 is 64.

\[ \therefore \text{The corresponding value of } x \text{ to the c.}\( \hat{f} \) is 30. \]

\[ \text{Median} = 30 \text{ units.} \]

(iii) For grouped data.

**Steps:**
1) Find the c.\( \hat{f} \) less than type
2) Find \( \text{N}/2 \), \( \text{N} \)= total frequency.
3) See the c.\( \hat{f} \) column just greater than \( \text{N}/2 \).
4) The Corresponding class interval is called the Median class.
5) To find Median, use the following formula.
M = \text{l}_1 + \left( \frac{N/2 - \text{c.f}}{f} \right) (\text{l}_2 - \text{l}_1)

Where
\text{l}_1 = \text{lower class-bounding of the median class.}
\text{l}_2 = \text{upper class-bounding of the median class.}
f = \text{frequency of the median class.}
c.\text{f} = \text{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:

<table>
<thead>
<tr>
<th>No. of patients :</th>
<th>1000-1200</th>
<th>1200-1400</th>
<th>1400-1600</th>
<th>1600-1800</th>
<th>1800-2000</th>
<th>2000-2200</th>
</tr>
</thead>
<tbody>
<tr>
<td>No. of days :</td>
<td>15</td>
<td>21</td>
<td>24</td>
<td>18</td>
<td>12</td>
<td>10</td>
</tr>
</tbody>
</table>

**Solution :**

<table>
<thead>
<tr>
<th>No. of patients</th>
<th>No. of days</th>
<th>c.f. less than type</th>
</tr>
</thead>
<tbody>
<tr>
<td>1000-1200</td>
<td>15</td>
<td>15</td>
</tr>
<tr>
<td>1200-1400</td>
<td>21</td>
<td>36</td>
</tr>
<tr>
<td>1400-1600</td>
<td>24</td>
<td>60</td>
</tr>
<tr>
<td>1600-1800</td>
<td>18</td>
<td>78</td>
</tr>
<tr>
<td>1800-2000</td>
<td>12</td>
<td>90</td>
</tr>
<tr>
<td>2000-2200</td>
<td>10</td>
<td>100</td>
</tr>
</tbody>
</table>

N/2 = 100/2 = 50

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

\therefore \quad \text{l}_1 = 1400; \text{l}_2 = 1600; f = 24; c.\text{f} = 36

\therefore \quad M = \text{l}_1 + \left( \frac{N/2 - \text{c.f}}{f} \right) (\text{l}_2 - \text{l}_1) = 1400 + \left[ \frac{50 - 36}{24} \right] (1600-1400)

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

\text{M} = 1516.67 \text{ patients}

**Example 16:** Find the median for the following distribution:

<table>
<thead>
<tr>
<th>Weights (in kgs)</th>
<th>30-34</th>
<th>35-39</th>
<th>40-44</th>
<th>45-49</th>
<th>50-54</th>
<th>55-59</th>
<th>60-64</th>
</tr>
</thead>
<tbody>
<tr>
<td>No. of students</td>
<td>4</td>
<td>7</td>
<td>15</td>
<td>21</td>
<td>18</td>
<td>10</td>
<td>5</td>
</tr>
</tbody>
</table>

**Solution :** Convert the given class intervals into exclusive type
<table>
<thead>
<tr>
<th>Weights</th>
<th>No. of students</th>
<th>c.f. less than type</th>
</tr>
</thead>
<tbody>
<tr>
<td>29.5-34.5</td>
<td>4</td>
<td>4</td>
</tr>
<tr>
<td>34.5-39.5</td>
<td>7</td>
<td>11</td>
</tr>
<tr>
<td>39.5-44.5</td>
<td>15</td>
<td>26</td>
</tr>
<tr>
<td>44.5-49.5</td>
<td>21</td>
<td>47</td>
</tr>
<tr>
<td>49.5-54.5</td>
<td>18</td>
<td>65</td>
</tr>
<tr>
<td>54.5-59.5</td>
<td>10</td>
<td>75</td>
</tr>
<tr>
<td>59.5-64.5</td>
<td>5</td>
<td>80</td>
</tr>
<tr>
<td>N=80</td>
<td></td>
<td>N/2 =80/2 = 40</td>
</tr>
</tbody>
</table>

Median class is 44.5-49.5

\[ l_1= 44.5 ; l_2 = 49.5 , f =21, c.f. =26 \]

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

\[ = 44.5 + \left[ \frac{40-26}{21} \right] (49.5-44.5) \]

\[ = 44.5 + [14/21] (5) \]

\[ = 44.5 + 3.33 \]

\[ M = 47.83 \] units.

**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.

### 6.8 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.
Q₃ is called third quartile or upper quartile.  
Q₂ is called second quartile which coincides with median. Therefore Q₂ is nothing but Median.

\[
\begin{align*}
Q_1 & , Q_2 & , Q_3 \\
N/4 & , N/4 & , N/4 & , N/4
\end{align*}
\]

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

**To find Q₁**: (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 Q₁, use the following formula:

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

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 Q₃**: (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 Q₃, use the following formula:

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

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 Q₂ (Second Quartile)**

\(Q_2 = \text{Median (Discussed above)}\)

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

<table>
<thead>
<tr>
<th>Commission (in Rs):</th>
<th>100-120</th>
<th>120-140</th>
<th>140-160</th>
<th>160-180</th>
<th>180-200</th>
<th>200-220</th>
<th>220-240</th>
</tr>
</thead>
<tbody>
<tr>
<td>No. of Salesmen:</td>
<td>13</td>
<td>38</td>
<td>45</td>
<td>56</td>
<td>27</td>
<td>17</td>
<td>12</td>
</tr>
</tbody>
</table>
Solution:

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

To find $Q_1$:

$N = 208 = 52$

4  4

See the c.f. column just greater than 52.
Quartile class: 140-160

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

\[
Q_1 = l_1 + \left( \frac{n/4 - c.f.}{f} \right) (l_2 - l_1)
\]

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

\[
= 140 + \left( \frac{1}{45} \right) (20)
\]

\[
= 140 + 0.44
\]

$Q_1 = 140.44$

To find $Q_2$ = Median:

$N/2 = 208/2 = 104$

See the c.f. column just greater than 104.
Median class = 160-180

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

\[
M = l_1 + \left( \frac{n/2 - c.f.}{f} \right) (l_2 - l_1)
\]

\[
= 160 + \left( \frac{104 - 96}{56} \right) (180 - 160)
\]

\[
= 160 + [8/56] (20)
\]

\[
= 160 + 2.86
\]

$Q_2 = 162.86$

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

Q_3=l_1+\frac{[3N/4-c.f]}{f} (l_2-l_1)

= 180+\frac{[156-152]}{27} (200-180)
= 180+\frac{4}{27} \times 20
= 180+2.96

Q_3= 182.96

\therefore \quad Q_1 = 140.44
\quad Q_2 = M = 162.86
\quad Q_3 = 182.96

**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 = 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.

<table>
<thead>
<tr>
<th>Age in yrs.</th>
<th>0-10</th>
<th>10-20</th>
<th>20-30</th>
<th>30-40</th>
<th>40-50</th>
<th>50-60</th>
<th>60-70</th>
<th>70-80</th>
</tr>
</thead>
<tbody>
<tr>
<td>No. of persons</td>
<td>15</td>
<td>13</td>
<td>25</td>
<td>22</td>
<td>25</td>
<td>10</td>
<td>5</td>
<td>5</td>
</tr>
</tbody>
</table>

**Solution :**

<table>
<thead>
<tr>
<th>Age (in yrs.)</th>
<th>No. of persons.</th>
<th>c.f. less than type</th>
</tr>
</thead>
<tbody>
<tr>
<td>0-10</td>
<td>15</td>
<td>15</td>
</tr>
<tr>
<td>10-20</td>
<td>13</td>
<td>28</td>
</tr>
<tr>
<td>X-axis (upper limit)</td>
<td>Y-axis c.f. &lt; type</td>
<td></td>
</tr>
<tr>
<td>---------------------</td>
<td>------------------</td>
<td></td>
</tr>
<tr>
<td>10</td>
<td>15</td>
<td></td>
</tr>
<tr>
<td>20</td>
<td>28</td>
<td></td>
</tr>
<tr>
<td>30</td>
<td>53</td>
<td></td>
</tr>
<tr>
<td>40</td>
<td>75</td>
<td></td>
</tr>
<tr>
<td>50</td>
<td>100</td>
<td></td>
</tr>
<tr>
<td>60</td>
<td>110</td>
<td></td>
</tr>
<tr>
<td>70</td>
<td>115</td>
<td></td>
</tr>
<tr>
<td>80</td>
<td>120</td>
<td></td>
</tr>
</tbody>
</table>

\[
n/4 = 120/4 = 30
\]

\[
n/2 = 120/2 = 60
\]

\[
3n/4 = 3 (120/4)
3n/4 = 90
\]
From the graph:
- $Q_1 = 21$ approximately
- $M = 33$ approximately
- $Q_3 = 46$ approximately

**Exercise 6.2**

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:

<table>
<thead>
<tr>
<th>$x$</th>
<th>25</th>
<th>30</th>
<th>35</th>
<th>40</th>
<th>45</th>
<th>50</th>
<th>55</th>
</tr>
</thead>
<tbody>
<tr>
<td>$f$</td>
<td>8</td>
<td>14</td>
<td>23</td>
<td>28</td>
<td>10</td>
<td>6</td>
<td>4</td>
</tr>
</tbody>
</table>

   Ans: $M = 40$

3. Find the median for the following data.

<table>
<thead>
<tr>
<th>Weekly Wages (in Rs)</th>
<th>50-100</th>
<th>100-150</th>
<th>150-200</th>
<th>200-250</th>
<th>250-300</th>
<th>300-350</th>
</tr>
</thead>
<tbody>
<tr>
<td>No. of Workers:</td>
<td>20</td>
<td>18</td>
<td>32</td>
<td>18</td>
<td>12</td>
<td>15</td>
</tr>
</tbody>
</table>
4. Find the median height for the following distribution:

<table>
<thead>
<tr>
<th>Height in cms:</th>
<th>150-154</th>
<th>154-158</th>
<th>158-162</th>
<th>162-166</th>
<th>166-170</th>
<th>170-174</th>
</tr>
</thead>
<tbody>
<tr>
<td>No. of Students:</td>
<td>6</td>
<td>15</td>
<td>23</td>
<td>20</td>
<td>12</td>
<td>10</td>
</tr>
</tbody>
</table>

Ans: M = 180.47

5. Find the three quartiles for the following distribution:

<table>
<thead>
<tr>
<th>Income in Rs:</th>
<th>0-500</th>
<th>500-1000</th>
<th>1000-1500</th>
<th>1500-2000</th>
<th>2000-2500</th>
<th>2500-3000</th>
<th>3000-3500</th>
</tr>
</thead>
<tbody>
<tr>
<td>No. of families</td>
<td>3</td>
<td>5</td>
<td>10</td>
<td>18</td>
<td>15</td>
<td>9</td>
<td>6</td>
</tr>
</tbody>
</table>

Ans: Q₁ = 1425, M = 1916.67, Q₃ = 2450

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

<table>
<thead>
<tr>
<th></th>
<th></th>
<th></th>
<th></th>
<th></th>
<th></th>
<th></th>
<th></th>
</tr>
</thead>
<tbody>
<tr>
<td>No. of years:</td>
<td>3</td>
<td>7</td>
<td>10</td>
<td>16</td>
<td>12</td>
<td>8</td>
<td>4</td>
</tr>
</tbody>
</table>

Ans: Q₁ = 27.5, M = 33.13, Q₃ = 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.

<table>
<thead>
<tr>
<th>Profit:</th>
<th>70-80</th>
<th>80-90</th>
<th>90-100</th>
<th>100-110</th>
<th>110-120</th>
<th>120-130</th>
</tr>
</thead>
<tbody>
<tr>
<td>No. of Companies</td>
<td>11</td>
<td>15</td>
<td>19</td>
<td>23</td>
<td>22</td>
<td>10</td>
</tr>
</tbody>
</table>

Ans: Q₁ = 89.33, M = 102.17, Q₃ = 113.18

6.9 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.

∴ 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.

<table>
<thead>
<tr>
<th>Size of Shoe</th>
<th>5</th>
<th>6</th>
<th>7</th>
<th>8</th>
<th>9</th>
<th>10</th>
</tr>
</thead>
<tbody>
<tr>
<td>No. of Pairs</td>
<td>38</td>
<td>43</td>
<td>48</td>
<td>56</td>
<td>25</td>
<td>22</td>
</tr>
</tbody>
</table>

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

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

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.

<table>
<thead>
<tr>
<th>Daily wages (in Rs.)</th>
<th>20-40</th>
<th>40-60</th>
<th>60-80</th>
<th>80-100</th>
<th>100-120</th>
<th>120-140</th>
<th>140-160</th>
</tr>
</thead>
<tbody>
<tr>
<td>No. of employees</td>
<td>21</td>
<td>28</td>
<td>35</td>
<td>40</td>
<td>24</td>
<td>18</td>
<td>10</td>
</tr>
</tbody>
</table>

**Solution:**

<table>
<thead>
<tr>
<th>Daily wages (in Rs.)</th>
<th>No. of employees</th>
</tr>
</thead>
<tbody>
<tr>
<td>20-40</td>
<td>21</td>
</tr>
<tr>
<td>40-60</td>
<td>28</td>
</tr>
<tr>
<td>60-80</td>
<td>35 ( f_0 )</td>
</tr>
<tr>
<td>80-100</td>
<td>40 ( f_1 )</td>
</tr>
<tr>
<td>100-120</td>
<td>24 ( f_2 )</td>
</tr>
<tr>
<td>120-140</td>
<td>18</td>
</tr>
<tr>
<td>140-160</td>
<td>10</td>
</tr>
</tbody>
</table>

The maximum frequency is 40.

∴ Modal class is 80-100.

∴ \( l_1 = 80; l_2 = 100; f_0 = 35; f_1 = 40; f_2 = 24 \)
\[ Z = l_1 + \left[ \frac{f_i - f_0}{(f_i - f_0) + (f_i - f_2)} \right] (l_2 - l_i) \]

\[ = 80 + \left[ \frac{40 - 35}{(40 - 35) + (40 - 24)} \right] (100 - 80) \]

\[ = 80 + \left[ \frac{5}{5 + 16} \right] \times 20 = 80 + 4.76 \]

\[ Z = 84.76 \]

Modal daily wages = Rs. 84.76

**Example: 21**

Find the mode for the following data.

<table>
<thead>
<tr>
<th>Marks</th>
<th>10-19</th>
<th>20-29</th>
<th>30-39</th>
<th>40-49</th>
<th>50-59</th>
<th>60-69</th>
<th>70-79</th>
</tr>
</thead>
<tbody>
<tr>
<td>No of Students</td>
<td>8</td>
<td>22</td>
<td>31</td>
<td>44</td>
<td>15</td>
<td>13</td>
<td>10</td>
</tr>
</tbody>
</table>

**Solution:**

<table>
<thead>
<tr>
<th>Marks</th>
<th>Class boundaries</th>
<th>No. of Students</th>
</tr>
</thead>
<tbody>
<tr>
<td>10-19</td>
<td>9.5-19.5</td>
<td>8</td>
</tr>
<tr>
<td>20-29</td>
<td>19.5-29.5</td>
<td>22</td>
</tr>
<tr>
<td>30-39</td>
<td>29.5-39.5</td>
<td>31</td>
</tr>
<tr>
<td>40-49</td>
<td>39.5-49.5</td>
<td>44</td>
</tr>
<tr>
<td>50-59</td>
<td>49.5-59.5</td>
<td>15</td>
</tr>
<tr>
<td>60-69</td>
<td>59.5-69.5</td>
<td>13</td>
</tr>
<tr>
<td>70-79</td>
<td>69.5-79.5</td>
<td>10</td>
</tr>
</tbody>
</table>

The maximum frequency is 44.

\[ l_1 = 39.5 \quad f_0 = 31 \]
\[ l_2 = 49.5 \quad f_1 = 44, \quad f_2 = 15 \]

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

\[ = 39.5 + \left[ \frac{44 - 31}{(44 - 31) + (44 - 15)} \right] (49.5 - 39.5) = 39.5 + \left[ \frac{13}{13 + 29} \right] \times 10 \]

\[ = 39.5 + 3.095 \]

\[ = 42.595 \]

\[ Z = 42.60 \]

Modal Marks = 42.60

---

### 6.10 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 the 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.

<table>
<thead>
<tr>
<th>Profit (in lakhs):</th>
<th>10-20</th>
<th>20-30</th>
<th>30-40</th>
<th>40-50</th>
<th>50-60</th>
<th>60-70</th>
</tr>
</thead>
<tbody>
<tr>
<td>No. of Companies</td>
<td>12</td>
<td>25</td>
<td>31</td>
<td>45</td>
<td>20</td>
<td>15</td>
</tr>
</tbody>
</table>

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 6.3

(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.

<table>
<thead>
<tr>
<th>x</th>
<th>15</th>
<th>18</th>
<th>20</th>
<th>22</th>
<th>24</th>
</tr>
</thead>
<tbody>
<tr>
<td>f</td>
<td>8</td>
<td>6</td>
<td>13</td>
<td>18</td>
<td>10</td>
</tr>
</tbody>
</table>

Ans: Z = 22

(3) Calculate the modal wages for the following distribution.

<table>
<thead>
<tr>
<th>Wages in Rs:</th>
<th>50-65</th>
<th>65-80</th>
<th>80-95</th>
<th>95-110</th>
<th>110-125</th>
<th>125-140</th>
</tr>
</thead>
<tbody>
<tr>
<td>No. of workers</td>
<td>4</td>
<td>7</td>
<td>15</td>
<td>28</td>
<td>11</td>
<td>5</td>
</tr>
</tbody>
</table>

Ans: Z = 101.5

(4) Find mode for the following data.

<table>
<thead>
<tr>
<th>Life in hrs:</th>
<th>200-400</th>
<th>400-600</th>
<th>600-800</th>
<th>800-1000</th>
<th>1000-1200</th>
<th>1200-1400</th>
<th>1400-1600</th>
</tr>
</thead>
<tbody>
<tr>
<td>No. of bulbs:</td>
<td>15</td>
<td>22</td>
<td>40</td>
<td>65</td>
<td>38</td>
<td>26</td>
<td>10</td>
</tr>
</tbody>
</table>

Ans: Z = 896.15

(5) Find the mode for the following data and hence locate mode graphically.

<table>
<thead>
<tr>
<th>Salary in</th>
<th>1000-</th>
<th>1500-</th>
<th>2000-</th>
<th>2500-</th>
<th>3000-</th>
<th>3500-</th>
</tr>
</thead>
</table>
(6) Find the mode for the following data, also locate mode graphically.

<table>
<thead>
<tr>
<th>Class Interval</th>
<th>100-150</th>
<th>150-200</th>
<th>200-250</th>
<th>250-300</th>
<th>300-350</th>
<th>350-400</th>
<th>400-450</th>
</tr>
</thead>
<tbody>
<tr>
<td>Frequency</td>
<td>15</td>
<td>29</td>
<td>31</td>
<td>42</td>
<td>56</td>
<td>35</td>
<td>17</td>
</tr>
</tbody>
</table>

**Ans:** $z = 320$

(7) Find Mean Median and mode for the following data.

<table>
<thead>
<tr>
<th>Income (in Rs.):</th>
<th>1000-2000</th>
<th>2000-3000</th>
<th>3000-4000</th>
<th>4000-5000</th>
<th>5000-6000</th>
<th>6000-7000</th>
</tr>
</thead>
<tbody>
<tr>
<td>No.of employees</td>
<td>15</td>
<td>26</td>
<td>45</td>
<td>32</td>
<td>12</td>
<td>10</td>
</tr>
</tbody>
</table>

**Ans:** $x = 3714.29; M = 3644.44; Z = 3593.75$

(8) Prove that median lies between mean and mode from the following data:

<table>
<thead>
<tr>
<th>Weekly wages (in Rs.)</th>
<th>100-200</th>
<th>200-300</th>
<th>300-400</th>
<th>400-500</th>
<th>500-600</th>
<th>600-700</th>
<th>700-800</th>
</tr>
</thead>
<tbody>
<tr>
<td>No. of workers</td>
<td>18</td>
<td>23</td>
<td>32</td>
<td>41</td>
<td>35</td>
<td>20</td>
<td>15</td>
</tr>
</tbody>
</table>

**Ans:** $x = 443.48; M = 446.34; Z = 460$

(9) Calculate mean median and mode for the following data.

<table>
<thead>
<tr>
<th>Height (in cms)</th>
<th>125-130</th>
<th>130-135</th>
<th>135-140</th>
<th>140-145</th>
<th>145-150</th>
<th>150-155</th>
</tr>
</thead>
<tbody>
<tr>
<td>No. of children</td>
<td>6</td>
<td>13</td>
<td>19</td>
<td>26</td>
<td>15</td>
<td>7</td>
</tr>
</tbody>
</table>

**Ans:** $x = 140.52; M = 140.96; Z = 141.94$

(10) Find arithmetic mean median and mode for the following data.

<table>
<thead>
<tr>
<th>Age in yrs:</th>
<th>10-20</th>
<th>20-30</th>
<th>30-40</th>
<th>40-50</th>
<th>50-60</th>
<th>60-70</th>
<th>70-80</th>
</tr>
</thead>
<tbody>
<tr>
<td>No. of persons</td>
<td>12</td>
<td>18</td>
<td>23</td>
<td>28</td>
<td>20</td>
<td>16</td>
<td>7</td>
</tr>
</tbody>
</table>

**Ans:** $x = 43.23; M = 43.21; Z = 43.85$
UNIT IV

MEASURES OF DISPERSION

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.

7.1 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.

<table>
<thead>
<tr>
<th>Region</th>
<th>Rainfall (cms)</th>
<th>Rainfall (cms)</th>
<th>Rainfall (cms)</th>
<th>Rainfall (cms)</th>
</tr>
</thead>
<tbody>
<tr>
<td>East</td>
<td>106</td>
<td>106</td>
<td>106</td>
<td>106</td>
</tr>
<tr>
<td>West</td>
<td>90</td>
<td>98</td>
<td>106</td>
<td>114</td>
</tr>
</tbody>
</table>
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 relatives measures can be used.

### 7.2 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]

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.

### 7.3 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.

\[
Q.D. = \frac{Q_3 - Q_1}{2}
\]

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.
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:

\( Q_1 = (11 + 1)/4 = 3^{\text{rd}} \) observation = 50.

\( \ldots (1) \)

\( Q_3 = 3(11 + 1)/4 = 9^{\text{th}} \) observation = 90.

\( \ldots (2) \)
Q.D. = \( \frac{Q_3 - Q_1}{2} = (90 - 50)/2 = 20 \)  
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.

<table>
<thead>
<tr>
<th>Weight in kg</th>
<th>30 – 35</th>
<th>35 – 40</th>
<th>40 – 45</th>
<th>45 – 50</th>
<th>50 – 55</th>
<th>55 – 60</th>
</tr>
</thead>
<tbody>
<tr>
<td>No. of students</td>
<td>4</td>
<td>16</td>
<td>12</td>
<td>8</td>
<td>10</td>
<td>5</td>
</tr>
</tbody>
</table>

**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:

<table>
<thead>
<tr>
<th>Weight in kg</th>
<th>30 – 35</th>
<th>35 – 40</th>
<th>40 – 45</th>
<th>45 – 50</th>
<th>50 – 55</th>
<th>55 – 60</th>
</tr>
</thead>
<tbody>
<tr>
<td>No. of students</td>
<td>4</td>
<td>16</td>
<td>12</td>
<td>8</td>
<td>10</td>
<td>6</td>
</tr>
<tr>
<td>( cf )</td>
<td>( 4 )</td>
<td>20</td>
<td>32</td>
<td>40</td>
<td>50</td>
<td>56</td>
</tr>
</tbody>
</table>

\[ Q_1 = l_1 + \left( \frac{m - pcf}{f} \times i \right) = 35 + \left( \frac{14 - 4}{16} \times 5 \right) = 35 + 3.125 = 38.125 \]

To find \( Q_3 \): \( N = 56 \). Thus \( m = N/4 = 14 \).
The \( cf \) just greater than 14 is 20, so 35 – 40 is the 1\(^{st} \) quartile class and
\[ l_1 = 35, l_2 = 40, i = 40 - 35 = 5, f = 16 \text{ and } pcf = 4. \]
\[ \therefore Q_3 = l_1 + \left[ \frac{m - pcf}{f} \times i \right] = 50 + \left[ \frac{42 - 40}{10} \times 5 \right] = 50 + 1 = 51 \]

Thus, the range of weight for the central 50% students = 12.875 kg

**Example**
Find the semi – inter quartile range and its coefficient for the following data:

<table>
<thead>
<tr>
<th>Size of shoe</th>
<th>0</th>
<th>1</th>
<th>2</th>
<th>3</th>
<th>4</th>
<th>5</th>
<th>6</th>
<th>7</th>
<th>8</th>
<th>9</th>
<th>10</th>
</tr>
</thead>
<tbody>
<tr>
<td>No. of boys</td>
<td>7</td>
<td>10</td>
<td>15</td>
<td>16</td>
<td>18</td>
<td>10</td>
<td>16</td>
<td>5</td>
<td>12</td>
<td>6</td>
<td>2</td>
</tr>
</tbody>
</table>

**Solution:**
The less than \( cf \) are computed and the table is completed as follows:

<table>
<thead>
<tr>
<th>Size of shoe</th>
<th>0</th>
<th>1</th>
<th>2</th>
<th>3</th>
<th>4</th>
<th>5</th>
<th>6</th>
<th>7</th>
<th>8</th>
<th>9</th>
<th>10</th>
</tr>
</thead>
<tbody>
<tr>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
</tbody>
</table>
Here \( N = 112 \).

To find \( Q_1 \):  
\[
m = \frac{N}{4} = 28.
\]
The first \( c.f \) just greater than 28 is 32, so the 1\textsuperscript{st} quartile is \( Q_1 = 2 \).

To find \( Q_3 \):  
\[
m = \frac{3N}{4} = 84.
\]
The first \( c.f \) just greater than 84 is 87, so the 3\textsuperscript{rd} quartile is \( Q_3 = 6 \).

\[\therefore \text{the semi inter quartile range i.e. Q.D.} = \frac{Q_3 - Q_1}{2} = \frac{6 - 2}{2} = 2\]

Coefficient of Q.D. = \[\frac{Q_3 - Q_1}{Q_3 + Q_1} = \frac{6 - 2}{6 + 2} = 0.5\]

### 7.4 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.
So in the above case the M.D. = \[\frac{\sum|x - \bar{x}|}{n} = \frac{3 + 1 + 0 + 1 + 1 + 2}{6} = \frac{8}{6}\]

For a frequency distribution the calculation of M.D. is done as follows

\[\text{M.D.} = \frac{\sum f|d|}{\sum f}\]
where \(f\) is the frequency of the observation and \(d = |x_i - \bar{x}|\)
where \(\bar{x}\) is the mean.

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:

<table>
<thead>
<tr>
<th>(x)</th>
<th>10</th>
<th>15</th>
<th>20</th>
<th>25</th>
<th>30</th>
<th>25</th>
<th>40</th>
<th>90</th>
</tr>
</thead>
<tbody>
<tr>
<td>(f)</td>
<td>6</td>
<td>17</td>
<td>29</td>
<td>38</td>
<td>25</td>
<td>14</td>
<td>9</td>
<td>1</td>
</tr>
</tbody>
</table>

**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 | | | | \(\sum f \cdot dx = 887.63\) |

In the above example the mean is = \(\bar{x} = \frac{\sum f.x}{n} = \frac{3535}{139} = 25.43\)
\(n = \Sigma f\)

And mean Deviation from Mean = \(\frac{\sum (f . |dx|)}{n} = \frac{887.63}{139} = 6.386\)
\(n = \Sigma f\)

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.

\[
\overline{x} = \frac{\sum x}{n} = \frac{715}{10} = 71.5
\]

| x   | d = |x - \overline{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|                 |

\[
\sum x = 715 \quad \sum d = 182
\]

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.

Since \(N = 10\),

\[\text{Median} = \text{A.M. of } 5^{th} \text{ and } 6^{th} \text{ observation}\]

\[\therefore M = \frac{6 + 10}{2} = 8\]

M.D. from median = \(\frac{\sum d}{n} = \frac{42}{10}\)

\[\therefore \delta M = 4.2\]

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.

<p>| No. of | No. of |</p>
<table>
<thead>
<tr>
<th>accidents</th>
<th>days</th>
</tr>
</thead>
<tbody>
<tr>
<td>0</td>
<td>26</td>
</tr>
<tr>
<td>1</td>
<td>32</td>
</tr>
<tr>
<td>2</td>
<td>41</td>
</tr>
<tr>
<td>3</td>
<td>12</td>
</tr>
<tr>
<td>4</td>
<td>22</td>
</tr>
<tr>
<td>5</td>
<td>10</td>
</tr>
<tr>
<td>6</td>
<td>05</td>
</tr>
<tr>
<td>7</td>
<td>01</td>
</tr>
<tr>
<td>8</td>
<td>06</td>
</tr>
<tr>
<td>9</td>
<td>15</td>
</tr>
<tr>
<td>10</td>
<td>10</td>
</tr>
<tr>
<td>Total</td>
<td>Σd = 42</td>
</tr>
</tbody>
</table>

Solution:

<p>| No. of | No. of days | cf  | (d = |x - M|) | (f \cdot d) |</p>
<table>
<thead>
<tr>
<th>accidents</th>
<th>(f)</th>
<th></th>
<th></th>
<th></th>
</tr>
</thead>
<tbody>
<tr>
<td>0</td>
<td>26</td>
<td>26</td>
<td>2</td>
<td>52</td>
</tr>
<tr>
<td>1</td>
<td>32</td>
<td>58</td>
<td>1</td>
<td>32</td>
</tr>
<tr>
<td>2</td>
<td>41</td>
<td>99</td>
<td>0</td>
<td>0</td>
</tr>
<tr>
<td>3</td>
<td>12</td>
<td>111</td>
<td>1</td>
<td>12</td>
</tr>
<tr>
<td>4</td>
<td>22</td>
<td>133</td>
<td>2</td>
<td>44</td>
</tr>
<tr>
<td>5</td>
<td>10</td>
<td>143</td>
<td>3</td>
<td>30</td>
</tr>
<tr>
<td>6</td>
<td>05</td>
<td>148</td>
<td>4</td>
<td>20</td>
</tr>
<tr>
<td>7</td>
<td>01</td>
<td>149</td>
<td>5</td>
<td>05</td>
</tr>
<tr>
<td>8</td>
<td>06</td>
<td>155</td>
<td>6</td>
<td>36</td>
</tr>
<tr>
<td>9</td>
<td>15</td>
<td>170</td>
<td>7</td>
<td>105</td>
</tr>
<tr>
<td>10</td>
<td>10</td>
<td>180</td>
<td>8</td>
<td>80</td>
</tr>
<tr>
<td>Total</td>
<td>N = 180</td>
<td>-</td>
<td>Σf \cdot d = 416</td>
<td></td>
</tr>
</tbody>
</table>

Now, \(N = 180\) \(\Rightarrow m = N/2 = 180/2 = 90\).

The \(cf\) just greater than 90 is 99. The corresponding observation is 9.

2. \(\therefore M = 2\)

M.D. from median = \(\delta M = \frac{\Sigma f \cdot d}{N} = \frac{416}{180} = 2.31\)

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.

<table>
<thead>
<tr>
<th>Wages less than</th>
<th>80</th>
<th>100</th>
<th>120</th>
<th>140</th>
<th>160</th>
<th>180</th>
<th>200</th>
</tr>
</thead>
<tbody>
<tr>
<td>No. of workers</td>
<td>30</td>
<td>45</td>
<td>77</td>
<td>98</td>
<td>128</td>
<td>172</td>
<td>200</td>
</tr>
</tbody>
</table>
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.

\[
\begin{array}{|c|c|c|c|c|c|}
\hline
\text{Wages in Rs} & \text{No. of workers (f)} & \text{cf} & x & d = |x - 141.33| & fd \\
\hline
60 - 80 & 30 & 30 & 70 & 71.33 & 2139.9 \\
80 - 100 & 15 & 45 & 90 & 51.33 & 769.95 \\
100 - 120 & 32 & 77 & 110 & 31.33 & 1002.56 \\
120 - 140 & 21 & 98 & 130 & 11.33 & 237.93 \\
140 - 160 & 30 & 128 & 150 & 8.67 & 260.1 \\
160 - 180 & 44 & 172 & 170 & 28.67 & 1261.48 \\
180 - 200 & 28 & 200 & 190 & 48.67 & 1362.76 \\
\hline
\text{Total} & N = 200 & - & - & - & \Sigma fd = 7034.68 \\
\hline
\end{array}
\]

The median class is 140 – 160. ∴ \( l_1 = 140, l_2 = 160, i = 160-140 = 20, f = 30 \) and \( pcf = 98 \)

\[
\therefore M = l_1 + \left[ \frac{m - pcf}{f} \times i \right] = 140 + \left[ \frac{100 - 98}{30} \times 20 \right] = 140 + 1.33 = 141.33
\]

M.D. from median = \( \delta M = \frac{\Sigma fd}{N} = \frac{7034.68}{200} = 35.1734 \)

Coefficient of M.D. = \( \frac{\delta M}{M} = \frac{35.1734}{141.33} = 0.25 \)

7.5 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 \( \sigma \) (read as ‘sigma’); \( \sigma^2 \) 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.

\[ s.d = \sigma = \sqrt{\frac{\sum(x - \bar{x})^2}{n}} \]

For a frequency distribution standard deviation is

\[ \sqrt{\frac{\sum f(x - \bar{x})^2}{\sum f}} = \sqrt{\frac{fx^2}{\sum f} - \left(\frac{\sum fx}{\sum f}\right)^2} \]

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:

<table>
<thead>
<tr>
<th>Class interval</th>
<th>Frequency</th>
</tr>
</thead>
<tbody>
<tr>
<td>10 – 20</td>
<td>9</td>
</tr>
<tr>
<td>20 – 30</td>
<td>18</td>
</tr>
<tr>
<td>30 – 40</td>
<td>31</td>
</tr>
<tr>
<td>40 – 50</td>
<td>17</td>
</tr>
<tr>
<td>50 – 60</td>
<td>16</td>
</tr>
<tr>
<td>60 - 70</td>
<td>9</td>
</tr>
<tr>
<td><strong>Total</strong></td>
<td><strong>100</strong></td>
</tr>
</tbody>
</table>

<table>
<thead>
<tr>
<th>Interval</th>
<th>Mid-point X</th>
<th>Frequency</th>
<th>fx</th>
<th>dx</th>
<th>dx²</th>
<th>f dx²</th>
</tr>
</thead>
<tbody>
<tr>
<td>10 – 20</td>
<td>15</td>
<td>9</td>
<td>135</td>
<td>24</td>
<td>576</td>
<td>5184</td>
</tr>
<tr>
<td>20 – 30</td>
<td>25</td>
<td>18</td>
<td>450</td>
<td>14</td>
<td>196</td>
<td>3528</td>
</tr>
<tr>
<td>30 – 40</td>
<td>35</td>
<td>31</td>
<td>1085</td>
<td>4</td>
<td>16</td>
<td>496</td>
</tr>
<tr>
<td>40 – 50</td>
<td>45</td>
<td>17</td>
<td>765</td>
<td>6</td>
<td>36</td>
<td>612</td>
</tr>
<tr>
<td>50 – 60</td>
<td>55</td>
<td>16</td>
<td>880</td>
<td>16</td>
<td>256</td>
<td>4096</td>
</tr>
<tr>
<td>60 - 70</td>
<td>65</td>
<td>9</td>
<td>585</td>
<td>26</td>
<td>676</td>
<td>6084</td>
</tr>
<tr>
<td><strong>Total</strong></td>
<td><strong>100</strong></td>
<td><strong>3900</strong></td>
<td></td>
<td></td>
<td></td>
<td><strong>20000</strong></td>
</tr>
</tbody>
</table>
Standard Deviation $\sigma = \sqrt{\frac{\sum f d^2}{\sum f}} = \sqrt{\frac{20000}{100}} = \sqrt{200} = 14.1$

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 $c = 10$

<table>
<thead>
<tr>
<th>Interval</th>
<th>Mid-point $X$</th>
<th>Frequency $f$</th>
<th>$dx = \frac{x - 35}{c}$</th>
<th>$dx^2$</th>
<th>$f dx$</th>
<th>$f dx^2$</th>
</tr>
</thead>
<tbody>
<tr>
<td>10 – 20</td>
<td>15</td>
<td>9</td>
<td>-2</td>
<td>4</td>
<td>-18</td>
<td>36</td>
</tr>
<tr>
<td>20 – 30</td>
<td>25</td>
<td>18</td>
<td>-1</td>
<td>1</td>
<td>-18</td>
<td>18</td>
</tr>
<tr>
<td>30 – 40</td>
<td>35</td>
<td>31</td>
<td>0</td>
<td>0</td>
<td>0</td>
<td>0</td>
</tr>
<tr>
<td>40 – 50</td>
<td>45</td>
<td>17</td>
<td>1</td>
<td>1</td>
<td>17</td>
<td>17</td>
</tr>
<tr>
<td>50 – 60</td>
<td>55</td>
<td>16</td>
<td>2</td>
<td>4</td>
<td>32</td>
<td>64</td>
</tr>
<tr>
<td>60 - 70</td>
<td>65</td>
<td>9</td>
<td>3</td>
<td>9</td>
<td>27</td>
<td>81</td>
</tr>
<tr>
<td>Total</td>
<td></td>
<td>100</td>
<td></td>
<td>40</td>
<td>216</td>
<td></td>
</tr>
</tbody>
</table>

S.D. $= \sigma = \sqrt{\frac{\sum f d^2}{\sum f} - \left( \frac{\sum f dx}{\sum f} \right)^2} x c = \sqrt{\frac{216}{100} - \left( \frac{40}{100} \right)^2 x 10}$

$= \sqrt{2 x 10} = 14.1421$

Relative measure of standard deviation or coefficient of variation

$CV = \frac{S.D.}{\text{mean}} x 100 = \frac{14.14}{39} x 100 = 36.26$

**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.

<table>
<thead>
<tr>
<th>$x$</th>
<th>$d = x - \bar{x}$</th>
<th>$d^2$</th>
</tr>
</thead>
<tbody>
<tr>
<td>22</td>
<td>-1</td>
<td>1</td>
</tr>
<tr>
<td>30</td>
<td>7</td>
<td>49</td>
</tr>
<tr>
<td>36</td>
<td>13</td>
<td>169</td>
</tr>
<tr>
<td>12</td>
<td>-11</td>
<td>121</td>
</tr>
<tr>
<td>15</td>
<td>-8</td>
<td>64</td>
</tr>
<tr>
<td>25</td>
<td>2</td>
<td>4</td>
</tr>
<tr>
<td>18</td>
<td>-5</td>
<td>25</td>
</tr>
<tr>
<td>10</td>
<td>-13</td>
<td>169</td>
</tr>
<tr>
<td>33</td>
<td>10</td>
<td>100</td>
</tr>
<tr>
<td>29</td>
<td>6</td>
<td>36</td>
</tr>
</tbody>
</table>
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:

<table>
<thead>
<tr>
<th>x</th>
<th>x²</th>
</tr>
</thead>
<tbody>
<tr>
<td>3</td>
<td>9</td>
</tr>
<tr>
<td>12</td>
<td>144</td>
</tr>
<tr>
<td>17</td>
<td>289</td>
</tr>
<tr>
<td>29</td>
<td>841</td>
</tr>
<tr>
<td>10</td>
<td>100</td>
</tr>
<tr>
<td>05</td>
<td>25</td>
</tr>
<tr>
<td>18</td>
<td>324</td>
</tr>
<tr>
<td>14</td>
<td>196</td>
</tr>
<tr>
<td>12</td>
<td>144</td>
</tr>
<tr>
<td>20</td>
<td>400</td>
</tr>
</tbody>
</table>

\[ \Sigma x = 230 \quad \Sigma x^2 = 738 \]

I. \[ \bar{x} = \frac{\Sigma x}{n} = \frac{230}{10} = 23 \]

II. \[ \sigma = \sqrt{\frac{\Sigma d^2}{n}} = \sqrt{\frac{738}{10}} \]
\[ = \sqrt{73.8} = 8.59 \]
\[ \therefore \sigma = 8.59 \]

Example

Compute the standard deviation for the following data:

<table>
<thead>
<tr>
<th>x</th>
<th>100</th>
<th>102</th>
<th>104</th>
<th>106</th>
<th>108</th>
<th>110</th>
<th>112</th>
</tr>
</thead>
<tbody>
<tr>
<td>f</td>
<td>5</td>
<td>11</td>
<td>7</td>
<td>9</td>
<td>13</td>
<td>10</td>
<td>12</td>
</tr>
</tbody>
</table>

Solution:  

*Short-cut method:*

In problems were 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:

<table>
<thead>
<tr>
<th>( x )</th>
<th>( u = x - 106 )</th>
<th>( f )</th>
<th>( fu )</th>
<th>( fu^2 )</th>
</tr>
</thead>
<tbody>
<tr>
<td>100</td>
<td>-6</td>
<td>5</td>
<td>-30</td>
<td>180</td>
</tr>
<tr>
<td>102</td>
<td>-4</td>
<td>11</td>
<td>-44</td>
<td>176</td>
</tr>
<tr>
<td>104</td>
<td>-2</td>
<td>7</td>
<td>-14</td>
<td>28</td>
</tr>
<tr>
<td>106</td>
<td>0</td>
<td>9</td>
<td>0</td>
<td>0</td>
</tr>
<tr>
<td>108</td>
<td>2</td>
<td>13</td>
<td>26</td>
<td>52</td>
</tr>
<tr>
<td>110</td>
<td>4</td>
<td>10</td>
<td>40</td>
<td>160</td>
</tr>
<tr>
<td>112</td>
<td>6</td>
<td>12</td>
<td>72</td>
<td>432</td>
</tr>
<tr>
<td>Total</td>
<td></td>
<td>( N = 67 )</td>
<td>( \Sigma fu = 50 )</td>
<td>( \Sigma fu^2 = 1028 )</td>
</tr>
</tbody>
</table>

From the table we have: \( \Sigma fu^2 = 1028 \), \( \Sigma fu = 50 \) and \( N = 67 \).

\[ \therefore \sigma = \sqrt{\frac{\Sigma fu^2}{N} - \left(\frac{\Sigma fu}{N}\right)^2} = \sqrt{\frac{1028}{67} - \left(\frac{50}{67}\right)^2} \]

\[ \therefore \sigma = \sqrt{15.34 - 0.56} = \sqrt{14.78} \]

\[ \therefore \sigma = 3.84 \]

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 V

8

ELEMENTARY PROBABILITY THEORY

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.

8.1 INTRODUCTION

In our day-to-day life conversation we here 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, or 6.

**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

\[ S = \{HH, HT, TH, TT\} \quad n(S) = 4 \]

Or when a cubic die is thrown the sample space is

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

**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.

\[ A = \{3,6\} \quad n(A) = 2 \]

**Event B:** The no of dots appeared is a prime number.

\[ B = \{2,3,5\} \quad n(B) = 3 \]
<table>
<thead>
<tr>
<th>Random Experiment</th>
<th>Event</th>
</tr>
</thead>
</table>
| **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}  
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) … (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)}  
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)}  
n(B) = 9 |
| **Selecting a two digit number**  
S = {10, 11, 12, … , 99}  
n(S) = 100 | A: The number is a perfect square  
A: {16, 25, 36, 49, 64, 81}  
n(A) = 6  
B: The number is > 99  
B: { }  
n(B) = 0 |

### 8.2 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 = Φ and n(A) = 0.

### 8.3 ALGEBRA OF EVENTS

**Union of two events** A and B (AUB): union AUB is the event containing all the sample points in A and B together.

AUB = {Elements in either A or B or in both A and B together}  
= {x such that x ∈ A or x ∈ B} or ∈ A ∪ B both

* e.g. When A = {1, 2, 3} and B = {2, 3, 5}
AUB= \{1,2,3,5\}

**Intersection of two events** \( A \cap B \): 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.
\[
A \cap B = \{\text{Elements in } A \text{ and } B \text{ both}\}
\]
\[
= \{x \text{ such that } x \in A \text{ and } x \in B\}
\]
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
\[
A=\{1,2,3\} \text{ and } B=\{4,5,6\}
\]

then \( A \cap B = \emptyset \) and \( AUB=S \)

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 it’s complement \( A' \) is the event containing all; the sample points in \( S \) which are not in \( A \).

\[
A' = \{\text{elements in } S \text{ which are not in } A\}
\]
\[
A = \{x \text{ s.t. } x \in S \text{ but } x \notin A\}
\]
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 it’s 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(Sample space) = P(S) = 1 and P(\emptyset) = 0.
Also, when events A and B are such that A \subseteq B then, P(A) \leq P(B).

8.4 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 in 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\} 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

\[
S = \{(1,1), (1,2), (1,3), (1,4), (1,5), (1,6),
(2,1), \ldots, (5,6), (6,6)\}
\]

\[n(S) = 36.\]

To find the probability we define the events,

a) event A: the sum of dots on the uppermost face is 6 or 10.
\[A = \{(1,5), (5,1), (2,4), (4,2), (3,3), (4,6), (6,4), (5,5)\} \quad n(A) = 8.\]

\[P(A) = \frac{n(A)}{n(S)} = \frac{8}{36} = \frac{2}{9}.\]

b) event B: The sum of the dots on the uppermost faces is divisible by 4.
\[B: \{(1,3),(3,1),(2,2),(2,6),(6,2),(3,5),(5,3),(4,4),(6,6)\} \quad n(B) = 9.\]

\[P(B) = \frac{n(B)}{n(S)} = \frac{9}{36} = 0.25.\]

c) event C: the sum of the dots is < 6.
\[C: \{(1,1),(1,2),(1,3),(1,4),(2,1),(2,2),(2,3),(3,1)\} \quad n(C) = 8.\]

\[P(C) = \frac{n(C)}{n(S)} = \frac{8}{36} = \frac{2}{9}.\]

**8.5 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,

\[\binom{n}{r} = \frac{n!}{r!(n-r)!}\]

where \(n! = n(n-1)(n-2)\ldots3x2x1.\)

(e.g. \(5! = 5x4x3x2x1 = 120\))

**Illustrations**-

(i) We can select a group of 3 students from 5 students in \(\binom{5}{3}\) ways and \(\binom{5}{3} = \frac{5!}{2!(5-2)!} = \frac{5!}{2!3!} = \frac{5x4x3x2x1}{2x1x3x2x1} = 10.\)
(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

\[
\text{n(S)} = \frac{52!}{3! 52 - 3!} = \frac{51 \times 51 \times 50}{3 \times 2 \times 1} = 22100.
\]

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 \( _{13}C_3 \)

\[
n(A) = \frac{13!}{3! 10!} = \frac{13 \times 12 \times 11}{3 \times 2 \times 1} = 286
\]

\[
\therefore P(A) = \frac{286}{22100} = 0.0129.
\]

b) 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,

\[
n(B) = _{13}C_3 + _{13}C_3 + _{13}C_3 + _{13}C_3 = 4 \times 286 = 1144. \quad \text{(refer to illustrations above)}
\]

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,

\[
n(C) = _4C_2 \times _4C_1 = 6 \times 4 = 24.
\]

Where, \( _4C_2 = \frac{4!}{2!2!} = \frac{4 \times 3 \times 2 \times 1}{2 \times 1 \times 2 \times 1} = 6 \) & \( _4C_1 = \frac{4!}{1!3!} = \frac{4 \times 3 \times 2 \times 1}{3 \times 2 \times 1} = 4 \).

\[
\therefore P(C) = \frac{n(C)}{n(S)} = \frac{24}{22100}
\]
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 \(^9\text{C}_2\) ways.

\[
n(S) = \frac{9!}{2!7!} = \frac{9 \times 8}{2 \times 1} = 36.
\]

(i) We define the event A: Both are of same color

\[
n(A) = \binom{4}{2} + \binom{5}{2} = \frac{4!}{2!2!} + \frac{5!}{2!3!} = 6 + 10 = 16.
\]

\[
\therefore P(A) = \frac{n(A)}{n(S)} = \frac{16}{36} = 0.44
\]

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

\[
n(B) = \binom{4}{2} = 6
\]

\[
\therefore P(B) = \frac{n(B)}{n(S)} = \frac{6}{36} = 0.16
\]

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 bares 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 ∩ B)= 0.4. Find P(AUB); P(A/B); P(B/A).
20. For two events A and B; P(A)=\( \frac{2}{5} \), P(B’) = \( \frac{1}{3} \), P(AUB)= \( \frac{5}{6} \). Find P(A \( \cap \) B); P(A/B); P(only A); P(Only one).
21. For two mutually exclusive events A and B, P(A)=0.7 and P(B)=0.5, find P(A ∩ 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(AUB); P(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.
8.6 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

\[ \{X=0\} \leftrightarrow \{TTT\}; \quad \{X=1\} \leftrightarrow \{HTT, THT, TTH\}; \]
\[ \{X=2\} \leftrightarrow \{HHT, HTH, THH\}. \]

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

\[ P(\{X=0\}) = P(\{TTT\}) = 1/8, \quad P(\{X=1\}) = P(\{HTT, THT, TTH\}) = 3/8, \]
\[ P(\{X=3\}) = P(\{HHH\}) = 1/8, \quad P(\{X=2\}) = P(\{HHT, HTH, THH\}) = 3/8. \]

Now we can express these probabilities in the form of a table,

<table>
<thead>
<tr>
<th>( X )</th>
<th>0</th>
<th>1</th>
<th>2</th>
<th>3</th>
</tr>
</thead>
<tbody>
<tr>
<td>( P(x) )</td>
<td>1/8</td>
<td>3/8</td>
<td>3/8</td>
<td>1/8</td>
</tr>
</tbody>
</table>

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) \( \Sigma p(x) = 1 \) 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 \( x_1, x_2, x_3, \ldots, x_n \) with respective probabilities \( p_1, p_2, p_3, \ldots, p_n \); \( P(x = x_1) = p_1 \), it’s expected value is defined as,

\[ E(X) = \Sigma x_i \cdot p(x) \]

Expected value is also called as the mean of \( X \).

Variance of \( X \); \( V(X) \): For the random variable \( X \), variance is defined as,
\[ V(X) = E(X - E(X))^2 \]
\[ = E(X^2) - [E(X)]^2. \]
\[ = \sum x^2.p(x) - (\sum x.p(x))^2 \]

Root of variance is called as standard deviation S.D.

**Solved examples:**

**Example 6:**
A discrete random variable X has the following probability distribution.

\[
\begin{array}{cccc}
\text{x:} & -2 & -1 & 0 & 1 & 2 \\
p(x) & k & 0.2 & 2k & 2k & 0.1 \\
\end{array}
\]

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.

\[ \sum p(x) = 1 \implies 5k + 0.3 = 1 \implies k = 0.7/5 = 0.14 \]

p(x) values are 0.14; 0.2; 0.28; 0.28; 0.1

Now we calculate the expected value by the formula,

\[ E(X) = \sum x.p(x) = 0 \text{ ------ from the table.} \]

<table>
<thead>
<tr>
<th>x</th>
<th>p(x)</th>
<th>xp(x)</th>
</tr>
</thead>
<tbody>
<tr>
<td>-2</td>
<td>0.14</td>
<td>-0.28</td>
</tr>
<tr>
<td>-1</td>
<td>0.2</td>
<td>0.2</td>
</tr>
<tr>
<td>0</td>
<td>0.28</td>
<td>0.28</td>
</tr>
<tr>
<td>1</td>
<td>0.28</td>
<td>0.28</td>
</tr>
<tr>
<td>2</td>
<td>0.1</td>
<td>0.2</td>
</tr>
<tr>
<td>Total</td>
<td>1.00</td>
<td>1=\Sigma x.p(x)</td>
</tr>
</tbody>
</table>

**Example 7:**
A random variable follows the probability distribution given below,

\[
\begin{array}{cccc}
\text{X:} & 0 & 1 & 2 & 3 & 4 \\
p(x) & 0.12 & 0.23 & 0.35 & 0.20 & 0.10. \\
\end{array}
\]

Obtain the expected value and variance of X.

**Solution:** The expected value and variance are given by the formula,

\[ E(X) = \sum x.p(x) \text{ and } V(X) = \sum x^2.p(x) - (\sum x.p(x))^2 \]

<table>
<thead>
<tr>
<th>x</th>
<th>p(x)</th>
<th>xp(x)</th>
<th>x^2.p(x)</th>
</tr>
</thead>
<tbody>
<tr>
<td>0</td>
<td>0.12</td>
<td>0x0.12=0</td>
<td>0x0=0</td>
</tr>
<tr>
<td>1</td>
<td>0.23</td>
<td>0.23</td>
<td>0.23</td>
</tr>
<tr>
<td>2</td>
<td>0.35</td>
<td>0.70</td>
<td>1.40</td>
</tr>
<tr>
<td>3</td>
<td>0.20</td>
<td>0.60</td>
<td>1.80</td>
</tr>
<tr>
<td>4</td>
<td>0.10</td>
<td>0.40</td>
<td>1.60</td>
</tr>
<tr>
<td>Total</td>
<td>1.00</td>
<td>1.93=\Sigma x.p(x)</td>
<td>4.03 = \Sigma x^2.p(x)</td>
</tr>
</tbody>
</table>

Now from the table,

\[ E(X) = \sum x.p(x) = 1.93. \]

\[ V(X) = \sum x^2.p(x) - (\sum x.p(x))^2 \]
\[ = 4.03 - (1.93)^2 \]

\[ V(X) = 0.35. \]

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
Maths_stats_fybcom

Solution: For the random variable X, we have
From the table we get,
\[ E(X) = \sum x \cdot p(x) = 0. \]
And \[ V(X) = \sum x^2 \cdot p(x) - (\sum x \cdot p(x))^2 \]
\[ = \frac{12}{16} - 0 = \frac{12}{16} \]

<table>
<thead>
<tr>
<th>x</th>
<th>p(x)</th>
<th>xp(x)</th>
<th>(x^2\cdot p(x))</th>
</tr>
</thead>
<tbody>
<tr>
<td>-2</td>
<td>1/16</td>
<td>-2/16</td>
<td>-2\cdot(-2/16) = 4/16</td>
</tr>
<tr>
<td>-1</td>
<td>1/8</td>
<td>-1/8</td>
<td>1/8</td>
</tr>
<tr>
<td>0</td>
<td>5/8</td>
<td>0</td>
<td>0</td>
</tr>
<tr>
<td>1</td>
<td>1/8</td>
<td>1/8</td>
<td>1/8</td>
</tr>
<tr>
<td>2</td>
<td>1/16</td>
<td>2/16</td>
<td>4/16</td>
</tr>
<tr>
<td>Total</td>
<td>1.00</td>
<td>0 = (\sum x \cdot p(x))</td>
<td>12/16 = (\sum x^2 \cdot p(x))</td>
</tr>
</tbody>
</table>

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,

\[ E(X) = \sum x \cdot p(x) \]
\[ = \frac{21}{6} = 3.5 \text{ ------ from the table.} \]

Hence the mean of X is 3.5.

EXERCISE II
1. A random variable X has the following probability distribution:
   \[ X: \quad -2 \quad -1 \quad 0 \quad 1 \quad 2 \quad 3 \]
   \[ P(x): \quad 0.1 \quad k \quad 0.2 \quad 2k \quad 0.3 \quad k \]
Find the value of k. Find the expected value and variance of x.

2. A random variable X has the following probability distribution:
   \[ X: \quad 0 \quad 1 \quad 2 \quad 3 \quad 4 \quad 5 \]
   \[ P(x): \quad 0.1 \quad 0.1 \quad 0.2 \quad 0.3 \quad 0.2 \quad 0.1 \]
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.

### 8.7 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 $\mu$ and $\sigma^2$, written as $X \sim N(\mu, \sigma^2)$ if its probability function is given by

$$p(x) = \frac{1}{\sigma \sqrt{2\pi}} e^{-\frac{1}{2} \left( \frac{x-\mu}{\sigma} \right)^2}$$

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 its mean $\mu$. 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 $\mu$ divides the curve into two equal parts so it is also the median.

The curve has its maximum height at $x = \mu$, therefore it the mode of the distribution.

v) Hence for normal distribution $\text{Mean} = \text{Median} = \text{Mode} = \mu$

For the probability calculations, we define the variable

$$Z = \frac{x - \text{Mean}}{S.D.} = \frac{x - \mu}{\sigma}$$
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 (Q₁) and upper quartiles (Q₃) are equidistant from the mean
\[ \mu , \text{i.e.} \quad \mu - Q₁ = Q₃ - \mu \Rightarrow \mu = \frac{Q₁ + Q₃}{2} \]

vii) The mean deviation (M.D.) of normal distribution is \(\frac{4}{5} \sigma\)

viii) The quartile deviation (Q.D.) of normal distribution is 0.67 \(\sigma\).

Area under the normal curve between

(i) \(\mu ± \sigma\) is 68.27%
(ii) \(\mu ± 2\sigma\) is 95.45%
(iii) \(\mu ± 3\sigma\) is 99.73%.
(iv) \(\mu ± 1.645\sigma\) is 90%
(v) \(\mu ± 1.96\sigma\) is 95%
(vi) \(\mu ± 0.67\sigma\) 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) \(P(X≥55)\)  
b) \(P(45≤X≤60)\)  
c) \(P(X≤45)\).

Solution: For the normal variable \(X\), we have

Mean \((X) = \mu = 50\) and standard deviation = \(\sigma = 5\).

\(\therefore \ X \sim N(\mu, \sigma^2) = N(50, 10^2).\)

We define the variable \(Z = \frac{x - \mu}{\sigma} = \frac{x - 50}{5}\)

Mean \((Z) = 0\) and S.D.(Z) = 1. \(Z\) is called a standard normal variable (s.n.v.)

Also \(P(X) = P(Z) = \text{Area}(Z)\).

a) \(P(X≥55) = P(\frac{x - 50}{5} ≥ 55 - 50)\)

\(= P(Z≥1)\)
\(= \text{Area r.h.s. of +1}\)
\(= 0.5 - \text{Area from 0 to 1}\)
\(= 0.5 - 0.3413 = 0.1587.\)

b) \(P(45≤X≤60) = P(\frac{45 - 50}{5} ≤ \frac{x - 50}{5} ≤ \frac{60 - 50}{5})\)

\(= P(-1 ≤ Z ≤ 2)\)
\(= \text{Area between -1 & +2}\)
\(= \text{Area from -1 to 0} + \text{Area from 0 to 2}\).
\(= 0.3413 + 0.4772 = 0.8185.\)

c) \(P(X≤45) = P(\frac{x - 50}{5} ≤ \frac{45 - 50}{5})\)

\(= P(Z≤-1)\)
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) = 6 = 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 \( P(\text{marks less than 45}) \)
\[ = P(X < 45) = P\left(\frac{x - 60}{10} < \frac{45 - 60}{10}\right) \]
\[ = P(Z < -1.5) \]
\[ = \text{Area on l.h.s. of } -1.5 \]
\[ = 0.5 - \text{Area from } -1.5 \text{ to } 0 \]
\[ = 0.5 - 0.3413 = 0.1587 = 15.87\% \]

Expected no of students = 15.87\% (150) = 10.
Similarly, to find percentage of students scoring marks between 55 and 70.
Consider, \( P(\text{marks between 55 and 70}) \)
\[ = P(55 < X < 70) \]
\[ = P\left(\frac{55 - 60}{10} < \frac{x - 60}{10} < \frac{70 - 60}{10}\right) \]
\[ = P(-0.5 < Z < 1) \]
\[ = \text{Area between } -0.5 \text{ & } +1 \]
\[ = \text{Area from } -0.5 \text{ to } 0 \text{+ Area from } 0 \text{ to } 1 \]
\[ = 0.1915 + 0.3413 = 0.5328 = 53.28\% \]
\( \therefore \) 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 155 cms. and S.D. of 20 cms. 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 \( Z = \frac{x - \mu}{\sigma} = \frac{x - 155}{20} \).

\[ \therefore \text{Mean (Z) = 0 and S.D.(Z)= 1. Z is called a standard normal variable (s.n.v.)} \]

Also \( P(X) = P(Z) = \text{Area (Z)} \).

Now, to find the proportion of soldiers with height above 170 cms, we find,

\[ P(\text{soldier's height is above 170 cms}) \]

\[ = P(x > 170) = P\left(20 \left(\frac{x - 155}{20}\right) \geq 20 \left(\frac{170 - 155}{20}\right)\right) \]

\[ = P(Z \geq 0.75) \]

\[ = \text{Area r.h.s. of 0.75} \]

\[ = 0.5 - \text{Area from 0 to 0.75.} \]

\[ = 0.5 - 0.2734 = 0.2766. \]

\[ \therefore \text{Proportion of soldiers with height above 170 cms is } (0.2766 \times 250) : 250 \]

\[ \text{i.e. } 69 : 250. \]

Now, let the height of the shortest soldier in the group of tallest 20% soldiers be \( h \).

\( P(\text{height less than } h) = 20\% = 0.2 \quad \text{i.e. } P(X \leq h) = 0.2 \)

Consider, \( P(X \leq h) = P\left(20 \left(\frac{x - 155}{20}\right) \leq 20 \left(\frac{h - 155}{20}\right) = t \text{ say}\right) \quad \text{i.e. } P(Z \leq t) = 0.2. \)

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.(\( X \)) = \( \sigma \) = 500.

\[ \therefore \text{X has normal distribution with } \mu = 2500 \text{ and } \sigma = 500. \]

i.e. \( X \sim N(\mu, \sigma^2) = N(2500, 500^2) \)

We define \( Z = \frac{x - \mu}{\sigma} = \frac{x - 2500}{500} \)
Mean (Z) = 0 and S.D.(Z)= 1. Z is called a standard normal variable (s.n.v.)
Also P (X) = P (Z) = Area (Z).

Now, to find the percentage of workers earning wages between Rs 3000 and Rs 4000.
Consider, P(wages between Rs.3000 and Rs 4000)
\[= P(3000 \leq X \leq 4000)\]
\[= P\left(\frac{3000 - 2500}{500} \leq \frac{x - 2500}{500} \leq \frac{4000 - 2500}{500}\right)\]
\[= P(1 \leq Z \leq 1.5 )\]
= Area between +1 & +1.5
= Area from 0 to 1.5 - Area from 0 to 1.
= 0.4332 - 0.3413= 0.0919 = 9.19%.

∴ 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(wages greater than \(h\)) =30%=0.3 i.e. \(P(X \geq h) = 0.3\).

Consider, \(P(X \geq h) = P(\frac{x - 2500}{500} \geq \frac{h - 2500}{500} = t \text{ say})\)
i.e. \(P(Z \geq t) = 0.3\).

∴ 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.
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 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,
   a) \(P(X \geq 55)\)  
   b) \(P(45 \leq X \leq 60)\)  
   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.
5. The height of 250 soldiers in a military camp confirms a normal distribution with mean height of 155 cms and S.D. of 20 cms. 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.

$P(X \geq 20)$  $P(10 \leq X \leq 17.5)$  $P(X \leq 12)$.

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 VI

FUNCTIONS, DERIVATIVES AND THEIR APPLICATIONS

OBJECTIVES

After reading this chapter you will be able to recognize.
1) Definition of function.
2) Standard Mathematical function.
3) Definition of derivative.
4) Derivatives of standard functions.
5) Second order derivatives.
6) Application of derivatives.
7) Maxima and Minima.

9.1 FUNCTIONS

If \( y = f(x) \) is a function then the set of all values of \( x \) for which this function is defined is called the domain of the function \( f \). Here \( x \) is called an independent variable and \( y \) is called the dependent variable. The set of all corresponding values of \( y \) for \( x \) in the domain is called the range of the function \( f \).

The function \( f \) is defined from the domain to the range.

We shall discuss only those functions where the domain and the range are subsets of real numbers. Such functions are called 'real valued functions'.

9.1.1 Standard Mathematical Functions:

(1) Constant function:
   The constant function is defined by
   \[ y = f(x) = C \]
   where \( C \) is a constant.
   The constants are denoted by real numbers or alphabets. The graph of a constant function is a straight line parallel to \( x \)-axis.

Examples:
\( y = f(x) = 5 \)
\( y = f(x) = -10 \)
\( y = f(x) = K \)
\( y = f(x) = a \)

(2) Linear function:
The linear function is defined by $y = f(x) = ax + b$ where $a$ and $b$ are constants.

Examples: $y = f(x) = 2x + 5$
$y = f(x) = -3x + 10$
$y = f(x) = 5x - 7$

(3) Functions with power of $x$: A function $f(x) = x^n$ is called power function or function with power of $x$. Here $x$ is called base and $n$ is called power.

Examples: $f(x) = x^2$
$f(x) = x^{-5}$
$f(x) = x^{-4/3}$
$f(x) = x^{3/2}$

(4) Exponential functions: These functions are of the type.
$f(x) = e^x$ and
$f(x) = a^x$, $a > 0$

(5) Logarithmic function: The logarithmic function is defined by
$y = f(x) = \log_e x$, $x > 0$

9.1.2. Standard functions from Economics:

(1) Demand: It refers to the quantity of a product desired by the buyers. The demand depends on the price. Therefore, there is a relationship between the price and the quantity demanded. Such relationship is called a demand function.

Hence the demand function is defined as $D = g(p)$ where $D =$ demand and $p =$ price.

Here demand is a dependent variable and the price is an independent variable.

For example, $D = 50 + 4p - 3p^2$

(2) Supply: It refers to the quantity of a product the market can offer. The supply depends on the price. Therefore, there is a relationship between the price and the quantity supplied. Such relationship is called a supply function.

Hence the supply function is defined as $S = f(p)$ where $S =$ supply and $p =$ price.

Here supply is a dependent variable and price is an independent variable.

For example, $S = 2p^2 - 6p + 25$

(3) Break-even Point: Equilibrium point.

(i) By the law of demand, the demand decreases when the price increases, the demand curve is a decreasing curve as shown in the figure:
(ii) By the law of supply, the supply increases when the price increases, the supply curve is an increasing curve as shown in the figure.

(iii) The demand and supply curves $D = f(p)$ and $S = g(p)$ are intersecting at a point. The point of intersection of the demand and supply curves represents a specific price at which the demand and supply are equal. This point is called the Break-even point or equilibrium point. The corresponding price at which this point occurs is called an equilibrium price and is denoted by $p_e$.

At equilibrium price, the amount of goods supplied is equal to the amount of goods demanded.

(4) The total cost function:

The total cost function or cost function is denoted by $C$ and it is expressed in terms of $x$. If $C$ is the cost of producing $x$ units of a product, then $C$ is generally a function of $x$ and is called the total cost function.

\[ C = f(x) \]

For example, $C = 2x^2 - 5x + 10$

(5) Average cost function:
The ratio between the cost function and the number of units produced is called average cost function. i.e. \( AC = \frac{C}{x} \)

For example, \( AC = \frac{x^2 + 2x + 5}{x} \)

(6) Total Revenue function:

The total revenue function is defined as in terms of the demand and the price per item. If \( D \) units are demanded with the selling price of \( p \) per unit, then the total revenue function \( R \) is given by

\[
R = p \times D \quad \text{where} \quad p = \text{price and} \quad D = \text{demand}
\]

For example,

If \( D = p^2 + 2p + 3 \) then \( R = p \times (p^2 + 2p + 3) = p^3 + 2p^2 + 3p \)

(7) Average revenue:

Average revenue is defined as the ratio between the revenue and the demand and is denoted by \( AR \).

\[ i.e., \ AR = \frac{R}{D} \quad AR = \frac{p \times D}{D} \quad (\text{as} \quad R = p \times D) \]

\[ \therefore \ AR = p \]

\[ \therefore \text{Average revenue is nothing but the selling price per unit.} \]

(8) The Profit function:

The profit function or the total profit function is denoted by \( P \) and is defined by the difference between the total revenue and the total cost.

\[ \therefore \text{Total Profit} = \text{Total Revenue} - \text{Total cost} \]

\[ i.e. \ P = R - C \]

Example 1:

Find the total profit function if the cost function \( C = 40 + 15x - x^2 \), \( x \) = number of items produced and the demand function is \( p = 200 - x^2 \)

Solution:

Given \( C = 40 + 15x - x^2 \)
\[
p = 200 - x^2
\]
\[
R = p \times D \quad (D = x)
= (200 - x^2) \times x
\]
\[ \therefore R = 200x - x^3 \]

Profit = Revenue - Cost
\[
P = R - C
= (200x - x^3) - (40 + 15x - x^2)
= 200x - x^3 - 40 - 15x + x^2
P = 185x + x^2 - x^3 - 40
\]

Example 2: The total cost function is \( C = 20 - 3x^2 \) and the demand function is \( p = 5 + 6x \). Find the profit when \( x = 100 \).

Solution:

Given: \( C = 20 - 3x^2 \)
\[
R = p \times D \quad (D = x)
= p \times x
\]
\[
\text{Profit} = \text{Revenue} - \text{Cost} \\
= (5x + 6x^2) - (20 - 3x^2) \\
= 5x + 6x^2 - 20 + 3x^2 \\
= 5x + 9x^2 - 20
\]

When \( x = 100 \), \( P = 5(100) + 9(100)^2 - 20 \\
= 500 + 90000 - 20 \\
= 90480.

### 9.2 DERIVATIVES

#### 9.2.1 Derivative as rate measure:

Definition: Let \( y = f(x) \) be the given function.

If \( \lim_{h \to 0} \frac{f(x+h) - f(x)}{h} \) exists,

then we say that the function \( f(x) \) has derivative at \( x \) and is denoted by \( f'(x) \).

\[
\text{i.e.}, \ f'(x) = \lim_{h \to 0} \frac{f(x+h) - f(x)}{h}
\]

The rate of change is called the "derivative" of \( y = f(x) \) with respect to \( x \) and is denoted by \( \frac{dy}{dx} \) or \( f'(x) \).

\[
\frac{dy}{dx} = \text{the rate of change of } y \text{ with respect to } x \text{ or the derivative of } y \text{ with respect to } x.
\]

**Note**: (1) Derivative means "rate of change"

(2) The process of finding the derivative of a function is called "differentiation".

\[
\frac{dc}{dx} = \text{the rate of change cost with respect to } x.
\]

For example \( \frac{dD}{dp} = \text{the rate of change of demand with respect to } p. \)

#### 9.2.2 Derivatives of Standard functions:

(1) If \( y = x^n \), where \( n \) is a real number, then

\[
\frac{dy}{dx} = nx^{n-1}
\]

\[
\text{i.e.}, \ \frac{dy}{dx} = \frac{d(x^n)}{dx} = nx^{n-1}
\]

(2) If \( y = C \), where \( C \) is a constant, then \( \frac{dy}{dx} = 0 \)

\[
\text{i.e.}, \ \frac{dy}{dx} = \frac{d(C)}{dx} = 0
\]
(3) If \( y = e^x \), then \( \frac{dy}{dx} = e^x \)

i.e., \( \frac{dy}{dx} = d(e^x) = e^x \frac{dx}{dx} \)

(4) If \( y = a^x \), where \( a \) is a positive real number, then

\[
\frac{dy}{dx} = a^x \log a
\]

i.e., \( \frac{dy}{dx} = d(a^x) = a^x \log a \frac{dx}{dx} \)

(5) If \( y = \log x \), then \( \frac{dy}{dx} = \frac{1}{x} \)

where \( x > 0 \)

i.e., \( \frac{dy}{dx} = d(\log x) = \frac{1}{x} \frac{dx}{dx} \)

Examples:

(1) \( y = x = x^1 \) \( \therefore \) \( \frac{dy}{dx} = 1 \cdot x^{1-1} = 1 \cdot x^0 = 1 \)

(2) \( y = x^4 \) \( \therefore \) \( \frac{dy}{dx} = 4 \cdot x^{4-1} = 4 \cdot x^3 \)

(3) \( y = x^{10} \) \( \therefore \) \( \frac{dy}{dx} = 10 \cdot x^{10-1} = 10 \cdot x^9 \)

(4) \( y = \frac{1}{x} = x^{-1} \) \( \therefore \) \( \frac{dy}{dx} = -1 \cdot x^{-1-1} = -1 \cdot x^{-2} = - \frac{1}{x^2} \)

(5) \( y = \frac{1}{x^3} \) \( \therefore \) \( \frac{dy}{dx} = -3 \cdot x^{-3-1} = -3 \cdot x^{-4} = - \frac{3}{x^4} \)

(6) \( y = \sqrt{x} = x^{1/2} \) \( \therefore \) \( \frac{dy}{dx} = \frac{1}{2} \cdot x^{1/2-1} = \frac{1}{2} \cdot x^{-1/2} = \frac{1}{2 \sqrt{x}} \)

(7) \( y = x^{5/2} \) \( \therefore \) \( \frac{dy}{dx} = \frac{5}{2} \cdot x^{5/2-1} = \frac{5}{2} \cdot x^{3/2} \)

(8) \( y = x^{3/2} \) \( \therefore \) \( \frac{dy}{dx} = - \frac{3}{2} \cdot x^{3/2-1} = - \frac{3}{2} \cdot x^{-5/2} \)

(9) \( y = x^{-7/2} \)

\[
\frac{dy}{dx} = -\frac{7}{2} \cdot x^{-7/2-1} = - \frac{7}{2} \cdot x^{-9/2}
\]

(10) \( y = 5 \), 5 is a constant
\[ \frac{dy}{dx} = 0 \]

(11) \( y = K \), \( \frac{dy}{dx} = 0 \)

(12) \( y = \log_2 \), \( \frac{dy}{dx} = 0 \)

(13) \( y = -10 \); \( \frac{dy}{dx} = 0 \)

(14) \( y = e^x \); \( \frac{dy}{dx} = e^x \)

(15) \( y = 2^x \); \( \frac{dy}{dx} = 2^x \log 2 \)

(16) \( y = 10^x \); \( \frac{dy}{dx} = 10^x \log 10 \)

(17) \( y = \log x \); \( \frac{dy}{dx} = \frac{1}{x} \)

**Exercise: 9.1**

Find \( \frac{dy}{dx} \) for the following:

\[ \frac{dx}{dx} \]

(1) \( y = x^6 \)

(2) \( y = \frac{1}{x^2} \)

(3) \( y = x^{7/2} \)

(4) \( y = x^{-5/2} \)

(5) \( y = 3 \)

(6) \( y = \log 10 \)

(7) \( y = -8 \)

(8) \( y = 4^x \)

(9) \( y = 9^x \)

(10) \( y = 15^x \)

**Answers:**

(1) \( 6x^5 \)

(2) \( -\frac{2}{x^3} \)

(3) \( \frac{7x^{5/2}}{2} \)

(4) \( -\frac{5}{x^{7/2}} \)

(5) \( 0 \)

(6) \( 0 \)

(7) \( 0 \)

(8) \( 4^x \log 4 \)

(9) \( 9^x \log 9 \)

(10) \( 15^x \log 15 \)

**9.2.3 Rules of derivatives:**

Rule: 1 Addition Rule (or) Sum rule:

If \( y = u + v \) where \( u \) and \( v \) are differentiable functions of \( x \) then

\[ \frac{dy}{dx} = \frac{du}{dx} + \frac{dv}{dx} \]
\[
\text{i.e., } \frac{dy}{dx} = \frac{d(u+v)}{dx} = \frac{du}{dx} + \frac{dv}{dx}
\]

Examples:

(1) If \( y = x^2 + e^x \), find \( \frac{dy}{dx} \)

Solution: Given: \( y = x^2 + e^x \)
\[
\frac{dy}{dx} = \frac{d(x^2 + e^x)}{dx}
\]
\[
= \frac{d(x^2)}{dx} + \frac{d(e^x)}{dx}
\]
\[
\therefore \quad \frac{dy}{dx} = 2x + e^x
\]

(2) If \( y = x^{10} + \log x \), find \( \frac{dy}{dx} \)

Solution: Given: \( y = x^{10} + \log x \)
\[
\frac{dy}{dx} = \frac{d(x^{10} + \log x)}{dx}
\]
\[
= \frac{d(x^{10})}{dx} + \frac{d(\log x)}{dx}
\]
\[
\therefore \quad \frac{dy}{dx} = 10x^9 + \frac{1}{x}
\]

**Rule 2**  **Subtraction Rule (or) Difference rule:**

If \( y = u - v \) where \( u \) and \( v \) are differentiable functions of \( x \) then
\[
\frac{dy}{dx} = \frac{du}{dx} - \frac{dv}{dx}
\]
\[
\text{i.e., } \frac{dy}{dx} = \frac{d(u-v)}{dx} = \frac{du}{dx} - \frac{dv}{dx}
\]

Examples:

(1) If \( y = x^5 - 2x \), find \( \frac{dy}{dx} \)

Solution: Given: \( y = x^5 - 2x \)
\[
\therefore \quad \frac{dy}{dx} = \frac{d(x^5 - 2x)}{dx}
\]
\[
= \frac{d(x^5)}{dx} - \frac{d(2x)}{dx}
\]
\[
\therefore \quad \frac{dy}{dx} = 5x^4 - 2x \log 2
\]
(2) If \( y = 100 - \log x \), find \( \frac{dy}{dx} \)

**Solution**:

Given: \( y = 100 - \log x \)

\[
\frac{dy}{dx} = \frac{d}{dx}(100 - \log x)
\]

\[
= \frac{d}{dx}(100) - \frac{d}{dx}(\log x)
\]

\[
= 0 - \frac{1}{x}
\]

\[
\therefore \frac{dy}{dx} = -\frac{1}{x}
\]

**Rule 3**  **Product Rule**:

If \( y = uv \) where \( u \) and \( v \) are differentiable functions of \( x \),
then \( \frac{dy}{dx} = u \frac{dv}{dx} + v \frac{du}{dx} \)

i.e., \[
\frac{dy}{dx} = \frac{d}{dx}(uv) = u \frac{dv}{dx} + v \frac{du}{dx}
\]

**Examples**:

1. If \( y = x^4 \log x \), find \( \frac{dy}{dx} \)

**Solution**:

Given \( y = x^4 \log x \)

\( u = x^4 \), \( v = \log x \)

\[
\frac{dy}{dx} = \frac{d}{dx}(x^4 \log x)
\]

\[
= x^4 \frac{d}{dx}(\log x) + \log x \frac{d}{dx}(x^4)
\]

\[
= x^4 \left(\frac{1}{x}\right) + \log x \left(4x^3\right)
\]

\[
= x^3 + 4x^3 \log x
\]

\[
\therefore \frac{dy}{dx} = x^3 \left[1 + 4 \log x\right]
\]

2. If \( y = x^2 e^x \), find \( \frac{dy}{dx} \)

**Solution**:

Given \( y = x^2 e^x \)

\[
\frac{dy}{dx} = \frac{d}{dx}(x^2 e^x)
\]

\[
= x^2 \frac{d}{dx}(e^x) + e^x \frac{d}{dx}(x^2)
\]

\[
= x^2 e^x + e^x(2x)
\]

\[
\therefore \frac{dy}{dx} = e^x [x^2 + 2x]
\]

**Rule 4**  **Quotient Rule**:

If \( y = \frac{u}{v} \), \( v \neq 0 \) where \( u \) and \( v \) are differentiable functions of \( x \), then
\[
\frac{dy}{dx} = \left[ \frac{v \frac{du}{dx} - u \frac{dv}{dx}}{v^2} \right]
\]

**Examples:**

(1) If \( y = x + 4 \), find \( \frac{dy}{\log x} \)

Solution: Given \( y = x + 4 \)
\( u = x + 4 \)
\( v = \log x \)

\[
\frac{dy}{dx} = \log x \frac{d((x + 4)}{dx} - (x + 4) \frac{d(\log x)}{dx}
\]
\[
= (\log x)(1+0) - (x + 4) \left(\frac{1}{4}\right)
\]
\[
\therefore \frac{dy}{dx} = (\log x) - (x + 4) \left(\frac{1}{4}\right)
\]

(2) If \( y = e^x + 5 \), find \( \frac{dy}{x^6-10} \)

Solution: Given \( y = e^x + 5 \)
\( u = e^x + 5 \)
\( v = x^6 - 10 \)

\[
\frac{dy}{dx} = (x^6 - 10) \frac{d(e^x+5)}{dx} - (e^x+5) \frac{d(x^6 - 10)}{dx}
\]
\[
= (x^6 - 10)^2
\]
\[
\therefore \frac{dy}{dx} = \frac{(x^6 - 10)(e^x) - (e^x+5)(6x^5)}{(x^6 - 10)^2}
\]

(3) If \( y = \frac{x^3 - 1}{x^3 + 1} \), find \( \frac{dy}{dx} \)

Solution: Given \( y = \frac{x^3 - 1}{x^3 + 1} \)
\( u = x^3 - 1 \)
\( v = x^3 + 1 \)

\[
\frac{dy}{dx} = \frac{(x^3 + 1) \frac{d(x^3 - 1)}{dx} - (x^3 - 1) \frac{d(x^3 + 1)}{dx}}{(x^3 + 1)^2}
\]
= \( (x^3 + 1) (3x^2) - (x^3-1) (3x^2) \)
\[ \frac{1}{(x^3+1)^2} \]
\[ = 3x^2 \left[ (x^3+1) - (x^3-1) \right] \]
\[ = 3x^2 \left[ x^3 + 1 - x^3 + 1 \right] \]
\[ = \frac{3x^2 \left[ 2 \right]}{(x^3+1)^2} \]
\[ \frac{dy}{dx} = \frac{6x^2}{(x+1)^2} \]

**Rule 5:** Scalar multiplication rule or constant multiplied by a function rule.
If \( y = cu \), \( c \) is a constant, where \( u \) is a differentiable function of \( x \), then
\[ \frac{dy}{dx} = c \frac{du}{dx} \]
i.e.,
\[ \frac{dy}{dx} = \frac{d}{dx} (cu) = c \frac{du}{dx} \]

Examples:
(1) If \( y = 5x^3 \), find \( \frac{dy}{dx} \)

Solution: Given: \( y = 5x^3 \)
\[ \therefore \frac{dy}{dx} = \frac{d}{dx} (5x^3) \]
\[ = 5 \frac{d}{dx} (x^3) = 5(3x^2) = 15x^2 \]

(2) If \( y = 10 \log x \), find \( \frac{dy}{dx} \)

Solution: Given: \( y = 10 \log x \)
\[ \therefore \frac{dy}{dx} = \frac{d}{dx} (10 \log x) \]
\[ = 10 \frac{d}{dx} (\log x) \]
\[ = 10 \left( \frac{1}{x} \right) \]
\[ = \frac{10}{x} \]

9.2.4 List of formulae:

<table>
<thead>
<tr>
<th>( y = f(x) )</th>
<th>( \frac{dy}{dx} = f'(x) )</th>
</tr>
</thead>
<tbody>
<tr>
<td>1. ( x^n )</td>
<td>( nx^{n-1} )</td>
</tr>
<tr>
<td>2. ( C )</td>
<td>0</td>
</tr>
<tr>
<td>3. ( e^x )</td>
<td>( e^x )</td>
</tr>
<tr>
<td>4. ( a^x )</td>
<td>( a^x \log a )</td>
</tr>
<tr>
<td>5. ( \log x )</td>
<td>( 1/x )</td>
</tr>
<tr>
<td>6. ( x )</td>
<td>1</td>
</tr>
</tbody>
</table>
7. \( \sqrt{x} \) \quad \quad \quad 1/2\sqrt{x} \\
8. \( 1/x \) \quad \quad \quad -1/x^2 \\

9.2.5 List of Rules:

(1) \( \frac{d}{dx} (u+v) = \frac{du}{dx} + \frac{dv}{dx} \)

(2) \( \frac{d}{dx} (u-v) = \frac{du}{dx} - \frac{dv}{dx} \)

(3) \( \frac{d}{dx} (uv) = u \frac{dv}{dx} + v \frac{du}{dx} \)

(4) \( \frac{d}{dx} \left( \frac{u}{v} \right) = \frac{v \frac{du}{dx} - u \frac{dv}{dx}}{v^2} \)

(5) \( \frac{d}{dx} (cu) = c \frac{du}{dx} \), \( c \) = constant.

9.2.6 Examples:

Find \( \frac{dy}{dx} \) for each of the following:

Ex: (1) \( y = x^6 + 4e^x + \log x + 10 \)

Solution:

\[
\frac{dy}{dx} = \frac{d}{dx} (x^6 + 4e^x + \log x + 10) \\
= \frac{d}{dx} (x^6) + 4 \frac{d}{dx} (e^x) + \frac{d}{dx} (\log x) + \frac{d}{dx} (10) \\
= 6x^5 + 4e^x + \frac{1}{x} + 0 \\
\therefore \frac{dy}{dx} = 6x^5 + 4e^x + 1/x
\]

Ex: (2) \( y = 5x^4 - 3e^{-x} + 4\sqrt{x} + 2^x \)

Solution:

\[
\frac{dy}{dx} = \frac{d}{dx} (5x^4 - 3e^{-x} + 4\sqrt{x} + 2^x) \\
= \frac{d}{dx} (5x^4) - 3 \frac{d}{dx} (e^{-x}) + 4 \frac{d}{dx} (\sqrt{x}) + \frac{d}{dx} (2^x) \\
= 5 \frac{d}{dx} (x^4) - 3 \frac{d}{dx} (e^{-x}) + 4 \frac{d}{dx} (\sqrt{x}) + \frac{d}{dx} (2^x) \\
= 5(4x^3) - 3(-xe^{-x}) + 4 \left( \frac{1}{2\sqrt{x}} \right) + \frac{d}{dx} (2^x) \\
= 20x^3 - 3e^{-x} + 4(1/2\sqrt{x}) + 2^x \log_2 \\
= 20x^3 - 3e^{-x} + 2/\sqrt{x} + 2^x \log_2
\]

Ex: (3) \( y = x^{3/2} + 4 \log x - 10x^2 + 15 \)
Solution:
\[ \frac{dy}{dx} = \frac{d}{dx} (x^{3/2} + 4 \log x - 10x^2 + 15) \]
\[ = \frac{d}{dx} (x^{3/2}) + \frac{d}{dx} (4 \log x) - \frac{d}{dx} (10x^2) + \frac{d}{dx} (15) \]
\[ = \frac{3}{2} x^{3/2-1} + \frac{4}{x} - 20x \]
\[ \frac{dy}{dx} = \frac{3}{2} x^{1/2} + 4 \cdot \frac{1}{x} - 20x. \]

Ex: (4) \( y = (x + e^x) \cdot (5 + \log x) \)

Solution: Here \( u = x + e^x \), \( v = 5 + \log x \)
\[ \frac{dy}{dx} = (x + e^x) \frac{d}{dx} (5 + \log x) + (5 + \log x) \frac{d}{dx} (x + e^x) \]
\[ = (x + e^x) \left( 0 + \frac{1}{x} \right) + (5 + \log x) \left( 1 + e^x \right) \]
\[ \therefore \frac{dy}{dx} = (x + e^x) \left( \frac{1}{x} \right) + (5 + \log x) \left( 1 + e^x \right) \]

Ex: (5) \( y = (x^{10}) \cdot (10^x) \)

Solution: Here \( u = x^{10} \), \( v = 10^x \)
\[ \therefore \frac{dy}{dx} = x^{10} \frac{d}{dx} (10^x) + 10^x \frac{d}{dx} (x^{10}) = x^{10} (10^x \log 10) + 10^x (10x^9) \]

Ex: (6) \( y = (\sqrt{x} + e^x) \cdot (2x^3 + 7) \)

Solution: Here \( u = \sqrt{x} + e^x \), \( v = 2x^3 + 7 \)
\[ \frac{dy}{dx} = (\sqrt{x} + e^x) \frac{d}{dx} (2x^3 + 7) + (2x^3 + 7) \frac{d}{dx} (\sqrt{x} + e^x) \]
\[ \therefore \frac{dy}{dx} = (\sqrt{x} + e^x) \left( 6x^2 + 21 \right) + (2x^3 + 7) \left( \frac{1}{2\sqrt{x}} + e^x \right) \]

Ex: (7) \( y = \frac{x^2 + 5x + 6}{x+7} \)

Solution: Here \( u = x^2 + 5x + 6 \)
\[ \therefore \frac{dy}{dx} = \frac{(x+7) \frac{d}{dx} (x^2 + 5x + 6) - (x^2 + 5x + 6) \frac{d}{dx} (x+7)}{(x+7)^2} \]
\[ = \frac{(x+7)(2x+5) - (x^2 + 5x + 6)(1)(x+7)}{(x+7)^2} \]
\[ = \frac{2x^2 + 19x + 35 - x^2 - 5x - 6}{(x+7)^2} \]
\[ \therefore \frac{dy}{dx} = \frac{x^2 + 14x + 29}{(x+7)^2} \]
\[
\frac{dx}{dx} \quad (x+7)^2
\]

Ex: 8 \quad y = \frac{10 e^x + 5 \log x}{x^3 + 12}

Solution: Here \( u = 10 e^x + 5 \log x \)
\( v = x^3 + 12 \)
\[
\begin{align*}
dy &= (x^3 + 12) d(10 e^x + 5 \log x) - (10 e^x + 5 \log x) d(x^3 + 12) \\
&= \frac{(x^3 + 12)(10e^x+5/x) - (10 e^x + 5 \log x) (3x^2)}{(x^3 + 12)^2}
\end{align*}
\]

Ex: 9 \quad y = \frac{4^x + 6}{2x^2 + 5}

Solution: Here \( u = 4^x + 6 \), \( v = 2x^2 + 5 \)
\[
\begin{align*}
dy &= (2x^2 + 5) d(4^x + 6) - (4^x + 6) d(2x^2 + 5) \\
&= \frac{(2x^2 + 5)(4^x \log 4) - (4^x + 6)(4x)}{(2x^2 + 5)^2}
\end{align*}
\]

Ex: 10 \quad y = 2\sqrt{x} + 16e^x + (6^x) + 20x

Solution:
\[
\begin{align*}
dy &= d(2\sqrt{x} + 16e^x + 6^x + 20x) \\
&= 2 (1/2\sqrt{x}) + 16e^x + 6^x \log 6 + 20 \\
\end{align*}
\]
\[
\begin{align*}
dy &= \frac{1}{\sqrt{x}} + 16e^x + 6^x \log 6 + 20 \\
\end{align*}
\]

Exercise: 9.2

Differentiate the following with respect to \( x \).

1. \( y = x^8 - 6e^x + 4x^{3/2} - 3x^2 + 5 \)
2. \( y = 6 \log x - 3^x + 2e^x + 10 \sqrt{x} + 2 \)
3. \( y = 5x^4 - 12.3^x + 18 e^x + 10 x^{-25} \)
4. \( y = 8^x (5x^3 + 3x + 1) \)
5. \( y = (10 x^2 + 2x + 5) (\sqrt{x} + e^x) \)
6. \( y = (2x^3 + 3x^2) (5 \log x + 14) \)
7. \( y = (x + \log x) (x^5 - 4x^2 + 10) \)
8. \( y = (8x^5 - 6x^{5/2} + 1) (40 \sqrt{x} + 2e^x) \)
(9) \( y = (e^x + 2 \log x + 2) (6x + 2x^2 + 5) \)

(10) \( y = \frac{x^2 + 1}{x^4 - 1} \)

(11) \( y = \frac{2^x + 4\sqrt{x}}{2e^x + 5} \)

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

(13) \( y = \frac{x + \sqrt{x}}{\sqrt{x} - 1} \)

(14) \( y = \frac{3\log x + 5}{x^2 + 2x} \)

(15) \( y = \frac{e^x - \sqrt{x}}{2\sqrt{x} + 1} \)

Answers:

1. \( 8x^7 - 6e^x + 6x^{1/2} - 6x \)
2. \( 6x^5 - 3x^4 \log x + 2e^x + 5/\sqrt{x} \)
3. \( 20x^3 - 36x^2 + 18e^x + 10\log 10 \)
4. \( 8^x (15x^2 + 3) + (5x^3 + 3x + 1) (8^x \log 8) \)
5. \( 10(x^2 + 2x + 5) (1/2\sqrt{x} + e^x) + (\sqrt{x} + e^x)(20x + 2) \)
6. \( (2x^3 + 3x^2) (5/x) + (5\log x + 14) (6x^2 + 6x) \)
7. \( (x + \log x) (5x^4 - 8x) + (x^5 - 4x^2 + 10) (1 + 1/x) \)
8. \( (8x^5 - 6x^{5/2} + 1) (20/\sqrt{x} + 2e^x) + (40\sqrt{x} + 2e^x) (40x^4 - 15x^{3/2}) \)
9. \( (e^x + 2\log x + 2) (6^x \log 6 + 4x) + (6^x + 2x^2 + 5) (e^x + 2/x) \)
10. \( -2x^5 - 4x^3 - 2x \)

\( (x^4 - 1)^2 \)

(11) \( \frac{(2e^x + 5) (2^x \log 2 + 2/\sqrt{x}) - (2^x + 4\sqrt{x}) (2e^x)}{(2e^x + 5)^2} \)

(12) \( x^4 - 12x^2 + 4x \)

\( (x^2 - 4)^2 \)

(13) \( \sqrt{x - 1} - (1 + x) / 2\sqrt{x} \)

\( (\sqrt{x} - 1)^2 \)

(14) \( \frac{(x^5 + 2x) (3x) - (3\log x + 2x) (5x^4 + 2)}{(x^5 + 2x)^2} \)

(15) \( \frac{(2\sqrt{x} + 1) (e^x + 1/2\sqrt{x}) - (e^x - \sqrt{x}) (1/\sqrt{x})}{(2\sqrt{x} + 1)^2} \)
9.3 SECOND ORDER DERIVATIVES

If \( y = f(x) \) is differentiable function with respect to \( x \), then
\[
\frac{dy}{dx} = f'(x) \quad \text{is called first order derivative of } y \text{ with respect to } x \quad \text{and}
\]
\[
\frac{d^2y}{dx^2} = \frac{d}{dx} \left( \frac{dy}{dx} \right) \quad \text{is called the second order derivative of } y \text{ with respect to } x
\]

Notation:
(i) First order derivative:
\[
\frac{dy}{dx} = f'(x) = y' \]
(ii) Second order derivative:
\[
\frac{d^2y}{dx^2} = f''(x) = y''
\]

9.3.1 Examples:
(1) If \( y = x^3 - 6x^2 + 19x + 100 \), find \( \frac{d^2y}{dx^2} \)

Solution:
Given: \( y = x^3 - 6x^2 + 19x + 100 \)
\[
\frac{dy}{dx} = 3x^2 - 12x + 19
\]
\[
\frac{d^2y}{dx^2} = \frac{d}{dx} \left( 3x^2 - 12x + 19 \right) = 6x - 12
\]
(2) If \( y = e^x + 2x^3 + 5x^2 + 4 \), find \( \frac{d^2y}{dx^2} \)

Solution:
Given: \( y = e^x + 2x^3 + 5x^2 + 4 \)
\[
\frac{dy}{dx} = e^x + 6x^2 + 10x
\]
\[
\frac{d^2y}{dx^2} = \frac{d}{dx} \left( e^x + 6x^2 + 10x \right) = e^x + 12x + 10
\]
(3) If \( y = x^5 - 6x^3 + 2x^2 + 10x + 5 \), find \( f''(x) \)

Solution:
Given \( f(x) = x^5 - 6x^3 + 2x^2 + 10x + 5 \)
\[ f'(x) = 5x^4 - 18x^2 + 4x + 10 \]
\[ f''(x) = 20x^3 - 36x + 4 \]

(4) If \( y = x + 1/x, \ x \neq 0 \), find \( \frac{d^2 y}{dx^2} \)

Solution: Given: \( y = x + 1/x \)
\[ \frac{dy}{dx} = 1 + (-1/x^2) = 1 - 1/x^2 \]
\[ \frac{d^2 y}{dx^2} = 0 - (-2/x^3) = 2 \]
\[ \frac{dx^3}{x^3} \]

Exercise: 9.3

Find \( \frac{d^2 y}{dx^2} \) for each of the following:
\[ \frac{dx^2}{x^2} \]
(1) \( y = 8x^5 - 16x^4 + 4x^3 + x + 2 \)
(2) \( y = 2x^3 - 5x^2 + 12x + 15 \)
(3) \( y = x + 25/x \)
(4) \( y = 2x^2 + e^x + 5x + 12 \)
(5) \( y = x^2 + x + \log x \)

Answers:
1) \( 160x^3 - 192x^2 + 24x \)
2) \( 12x - 10 \)
3) \( \frac{50}{x^3} \)
4) \( 4 + e^x \)
5) \( \frac{2 - 1}{x^2} \)

9.4 APPLICATIONS OF DERIVATIVES

Applications to Economics:

9.4.1 The Total Cost function:

(i) Total cost function \( C = f(x) \)
(ii) Average cost function \( AC = \frac{C}{x} \)
(iii) Marginal cost function:
The rate of change of cost with respect to the number of units produced is called the Marginal cost and is denoted by MC.

\[ \text{i.e, } MC = \frac{dC}{dx} \]

9.4.2 The Total Revenue Function:
9.4.3 Elasticity:
Let \( D \) be the demand and \( p \) be the price. The quantity \( -\frac{p}{D} \frac{dD}{dp} \) is called elasticity of demand with respect to the price and is denoted by \( \eta \) (\( \eta = \) eta : Greek alphabet).

i.e, \( \eta = -\frac{p}{D} \frac{dD}{dp} \)

If \( \eta = 0 \), demand \( D \) is Constant and the demand is said to be perfectly elastic.

If \( 0 < \eta < 1 \), the demand is said to be inelastic.

If \( \eta = 1 \), the demand is directly proportional to the price.

If \( \eta > 1 \), the demand is said to be elastic.

9.4.4 Relation between the Marginal Revenue and elasticity of demand.
Let \( R \) = Total revenue
\( p \) = price
\( D \) = demand

\[ R = pD \]

\( MR = \frac{dR}{dD} = \frac{d}{dD}(pD) \]

\[ = p \frac{d}{dD}(D) + D \frac{d}{dD}(p) \text{ (by product rule)} \]

\[ = p (1) + D \frac{d}{dD}p \]
\[
\text{MR} = p + D \frac{dp}{dD}
\]
\[
\eta = -p \frac{dD}{D \ dp}
\]
\[
\therefore D \frac{dp}{dD} = \frac{-p}{\eta}
\]
\[
\text{MR} = p + \left( \frac{-p}{\eta} \right)
\]
\[
\text{MR} = p \left[ 1 - \frac{1}{\eta} \right]
\]
\[
\text{MR} = \text{AR} \left[ 1 - \frac{1}{\eta} \right]
\]  \quad (\text{AR} = p)

**9.4.5 Examples:**

Ex: (1) The cost of producing \(x\) items is given by

\[2x^2 + 5x + 20\]. Find the total cost, average cost and marginal cost when \(x = 10\).

Solution:

Let \(C = f(x) = 2x^2 + 5x + 20\)

\[
\text{AC} = \frac{C}{x} = \frac{2x^2 + 5x + 20}{x}
\]

\[
\text{MC} = \frac{dC}{dx} = \frac{d}{dx} (2x^2 + 5x + 20)
\]

\[
\text{MC} = 4x + 5
\]

when \(x = 10\)

\[
C = 2(10)^2 + 5(10) + 20
\]

\[
C = 270
\]

\[
\text{AC} = \frac{C}{x} = \frac{270}{10} = 27
\]
MC = 4 ( 10 ) + 5 = 45

Ex: (2)

The demand function is given by \( \rho = 50 + 6D + 4D^2 \).
Find the total revenue, average revenue and the marginal revenue when the demand is 5 units.

**Solution:**

Given : \( \rho = 50 + 6D + 4D^2 \)

\[
R = \rho \times D
\]
\[
= ( 50 + 6D + 4D^2 ) ( D )
\]
\[
R = 50D + 6D^2 + 4D^3
\]

\[
AR = \rho = 50 + 6D + 4D^2
\]

\[
MR = \frac{dR}{dD}
\]
\[
= \frac{d}{dD} ( 50D + 6D^2 + 4D^3 )
\]

\[
MR = 50 + 12D + 12D^2
\]

When \( D = 5 \)

\[
R = 50 ( 5 ) + 6 ( 5 )^2 + 4 ( 5 )^3
\]
\[
= 250 + 150 + 500
\]
\[
R = 900
\]

\[
AR = 50 + 6 ( 5 ) + 4 ( 5 )^2
\]
\[
= 50 + 30 + 100
\]
\[
= 180
\]

\[
MR = 50 + 12 ( 5 ) + 12 ( 5 )^2
\]
\[
= 50 + 60 + 300
\]
\[
= 410
\]

Ex: (3)
The total revenue function is given by \( R = 20D - D^2 \), \( D = \) Demand. Find the demand function. Also find AR when MR = 0.

Solution:

Given \( R = 20D - D^2 \)

\[ \rho D = 20D - D^2 \quad (\therefore R = \rho D) \]

\[ \therefore \rho = \frac{20D - D^2}{D} \]

\[ \rho = 20 - D \]

\[ \therefore \] The demand function is \( \rho = 20 - D \).

Now, \( MR = \frac{dR}{dD} \)

\[ = \frac{d}{dD} (20D - D^2) \]

\[ MR = 20 - 2D \]

Given that \( MR = 0 \).

\[ \therefore 0 = 20 - 2D \]

\[ 2D = 20 \]

\[ \therefore D = 10 \]

\( AR = \rho = 20 - D \)

\[ \therefore AR = 20 - 10 \]

\[ AR = 10 \]

Ex : (4)

The demand function is given by

\[ D = 25 - 2\rho - \rho^2 \]

Find the elasticity of demand when the price is 4.

Solution:

Given : \( D = 25 - 2\rho - \rho^2 \)

\[ \therefore \frac{dD}{d\rho} = 0 - 2 - 2\rho \]
\[
\frac{dD}{dp} = -2 - 2\rho
\]

\[
\therefore \eta = \frac{-\rho}{D} \cdot \frac{dD}{dp}
\]

\[
= \frac{(-\rho)}{25 - 2\rho - \rho^2} \cdot (-2 - 2\rho)
\]

\[
\eta = \frac{p(2 + 2\rho)}{25 - 2\rho - \rho^2}
\]

When \( p = 4, \)

\[
\eta = \frac{4(2 + 8)}{25 - 8 - 16}
\]

\[
= \frac{4(10)}{1}
\]

\[
\therefore \eta = 40
\]

Ex: 5

The demand function is given by \( D = \frac{p + 3}{2p - 1} \)

where \( D = \text{Demand} \) and \( p = \text{price} \). Find the elasticity of demand when the price is 8.

**Solution:**

Given \( D = \frac{p + 3}{2p - 1} \)

\[
\therefore \frac{dD}{dp} = \frac{(2p - 1) \frac{d}{dp}(p + 3) - (p + 3) \frac{d}{dp}(2p - 1)}{(2p - 1)^2}
\]

\[
= \frac{(2p - 1)(1) - (p + 3)(2)}{(2p - 1)^2}
\]

\[
= \frac{2p - 1 - 2p - 6}{(2p - 1)^2}
\]

\[
\frac{dD}{dp} = -7
\]
\[
\eta = \frac{-p}{D} \frac{dD}{dp}
\]

\[
= \frac{-(p)}{[(p+3)/(2p-1)]} \left[ \frac{-7}{(2p-1)^2} \right]
\]

\[
= \frac{7p(2p-1)}{(p+3)(2p-1)^2}
\]

\[
\eta = \frac{7p}{(p+3)(2p-1)}
\]

When \( p = 8 \)

\[
\eta = \frac{7(8)}{(8+3)[2(8)-1]}
\]

\[
= \frac{56}{11(15)}
\]

\[
= \frac{56}{165}
\]

\[
\eta = 0.34
\]

Ex: (6) If MR = 45, AR = 75, Find \( \eta \)

Solution:

Given

\[
\begin{align*}
\text{MR} &= 45 \\
\text{AR} &= 75 \\
\eta &= ?
\end{align*}
\]

\[
.: \text{MR} = \text{AR} \ [1 - 1/\eta] \\
45 &= 75 \ [1 - 1/\eta] \\
\frac{45}{75} &= 1 - 1/\eta
\]

\[
\frac{3}{5} = 1 - 1/\eta
\]

\[
0.6 = 1 - 1/\eta
\]

\[
1/\eta = 1 - 0.6
\]

\[
1/\eta = 0.4
\]

\[
\eta = 2.5
\]
Ex : (7) 

If AR = 95 and $\eta = \frac{7}{2}$, Find MR. 

Solution : 
Given AR = 95 
$\eta = \frac{7}{2} = 3.5$ 
MR = ? 

$MR = AR \left[ 1 - \frac{1}{\eta} \right]$ 
$= 95 \left[ 1 - \frac{1}{3.5} \right]$ 
$= 95 \left[ 1 - 0.29 \right]$ 
$= 95 \left[ 0.71 \right]$ 

MR = 67.45 

Exercise : 9.4 

(1) The cost of producing x items is given by $x^3 + 4x + 15$. Find the total cost, average cost and marginal cost when x = 6. 

(2) The total cost function is given by $C = x^3 + 2x^2 + 5x + 30$. Find the total cost, average cost and marginal cost when x = 10. 

(3) The demand function is given by $p = 20 - 8D + 3D^2$. Find the total revenue, average revenue and marginal revenue when the demand is 4 units. 

(4) The total revenue function is given by $R = 30D - 2D^2 + D^3$. Find the demand function. Also find total revenue, average revenue and marginal revenue when the demand is 5 units. 

(5) The demand function is given by $D = -28 - 5p + 2p^2$. Find the elasticity of demand when the price is 3. 

(6) The demand function is given by $D = \frac{2p + 5}{p - 3}$ Where D = Demand and p = price. Find the elasticity of demand when price is 6.
(7) If AR = 65 and η = 3, find MR.

(8) If MR = 85 and η = 4.5, find AR.

(9) If MR = 55 and AR = 98, find η.

(10) If the price is 65 and the elasticity of demand is 5.2, find the marginal revenue.

Answers:

(1) \( C = 255; AC = 42.5; MC = 112 \)

(2) \( C = 1280; AC = 128; MC = 345 \)

(3) \( R = 144; AR = 36; MR = 100 \)

(4) \( p = 30 - 2D + D^2 \)
    \( R = 225; AR = 45; MR = 85 \)

(5) \( η = 0.64 \)

(6) \( η = 1.29 \)

(7) \( MR = 43.3 \)

(8) \( AR = 108.97 \)

(9) \( η = 2.28 \)

(10) \( MR = 52.5 \)

9.5 MAXIMA AND MINIMA.

Let \( y = f(x) \) be the given function. A curve \( f(x) \) is said to have a maximum or minimum point (extreme point), if \( f(x) \) attains either a maximum or minimum of that point.

Maximum at \( x = a \)
In the first figure, x = a is the point where the curve \( f(x) \) attains a maximum. In the second figure, x = b is the point where the curve \( f(x) \) attains a minimum.

### 9.5.1 Conditions for Maximum & Minimum:

1. **Condition for Maximum:**
   
   (i) \( f'(x) = 0 \)
   
   (ii) \( f''(x) < 0 \) at \( x = a \)

2. **Condition for Minimum:**

   (i) \( f'(x) = 0 \)

   (ii) \( f''(x) > 0 \) at \( x = b \)

### 9.5.2 To find the Maximum and Minimum values of \( f(x) \):

**Steps:**

(i) Find \( f'(x) \) and \( f''(x) \).

(ii) put \( f'(x) = 0 \), solve and get the values of \( x \).

(iii) Substitute the values of \( x \) in \( f''(x) \).

If \( f''(x) < 0 \), then \( f(x) \) has maximum value at \( x = a \). If \( f''(x) > 0 \), then \( f(x) \) has minimum value at \( x = b \).

(iv) To find the maximum and minimum values, put the points \( x = a \) and \( x = b \) in \( f(x) \).

**Note:** Extreme values of \( f(x) \) = Maximum and minimum values of \( f(x) \).

### 9.5.3 Examples:

Ex: (1) Find the extreme values of \( f(x) = x^3 - 3x^2 - 45x + 25 \).
Solution:

Given:

\[ f(x) = x^3 - 3x^2 - 45x + 25 \]

\[ \therefore f'(x) = 3x^2 - 6x - 45 \]

\[ f''(x) = 6x - 6. \]

Since \( f(x) \) has maximum or minimum value,

\[ f'(x) = 0 \]

\[ \therefore 3x^2 - 6x - 45 = 0 \]

\[ 3 \left( x^2 - 2x - 15 \right) = 0 \]

\[ \therefore x^2 - 2x - 15 = 0 \]

\[ x^2 - 5x + 3x - 15 = 0 \]

\[ x \left( x - 5 \right) + 3 \left( x - 5 \right) = 0 \]

\[ (x - 5)(x + 3) = 0 \]

\[ x - 5 = 0 \text{ or } x + 3 = 0 \]

\[ \therefore x = 5 \text{ or } x = -3. \]

When \( x = 5 \), \( f''(5) = 6(5) - 6 = 24 > 0 \)

\[ \therefore f(x) \text{ has minimum at } x = 5. \]

When \( x = -3 \), \( f''(-3) = 6 \left( -3 \right) - 6 \]

\[ = -24 < 0 \]

\[ \therefore f(x) \text{ has maximum at } x = -3. \]

To find the maximum and minimum values of \( f(x) \):

put \( x = 5 \) and \( x = -3 \) in \( f(x) \).

\[ \therefore f(5) = 5^3 - 3 \left( 5 \right)^2 - 45(5) + 25 \]

\[ = 125 - 75 - 225 + 25 \]

\[ = -150 \]

\[ f(-3) = (-3)^3 - 3 \left( -3 \right)^2 - 45 \left( -3 \right) + 25 \]

\[ = -27 - 27 + 135 + 25 \]
= 106

∴ Maximum value = 106 at x = -3

Minimum value = -150 at x = 5.

Ex: (2) Find the maximum and minimum values of \( f(x) = x + \frac{16}{x}, \) \( x \neq 0. \)

Solution: Given \( f(x) = x + \frac{16}{x} \)

\[
\therefore \quad f'(x) = 1 + 16 \left(-\frac{1}{x^2}\right)
\]

\[
f'(x) = 1 - \frac{16}{x^2}
\]

\[
\therefore \quad f''(x) = 0 - 16 \left(-\frac{2}{x^3}\right)
\]

\[
f''(x) = \frac{32}{x^3}
\]

Since \( f(x) \) has maximum or minimum value, \( f'(x) = 0. \)

\[
\therefore \quad 1 - \frac{16}{x^2} = 0
\]

\[
\frac{x^2 - 16}{x^2} = 0
\]

\[
x^2 - 16 = 0
\]

\[
x^2 - 16 = 0
\]

\[
x^2 - 4^2 = 0
\]

\[
(x - 4)(x + 4) = 0
\]

\( x = 4 \) or \( x = -4 \)

when \( x = 4, \) \( f''(4) = \frac{32}{(4)^3} = \frac{32}{64} = \frac{1}{2} > 0 \)

\( \therefore \) \( f(x) \) has minimum at \( x = 4. \)

when \( x = -4, \) \( f''(-4) = \frac{32}{(-4)^3} = \frac{32}{-64} = -\frac{1}{2} < 0 \)

\( \therefore \) \( f(x) \) has maximum at \( x = -4. \)

Now to find the extreme values of \( f(x) \): put \( x = 4 \) and \( x = -4 \) in \( f(x) \)
\[
\therefore f(4) = 4 + \frac{16}{4} = 4 + 4 = 8
\]
\[
f(-4) = -4 + \frac{16}{-4} = -4 - 4 = -8
\]

\therefore \text{Maximum value} = -8 \text{ at } x = -4

\text{Minimum value} = 8 \text{ at } x = 4.

Ex: (3) Divide 80 into two parts such that the sum of their squares is a minimum.

Solution:
Let \( x \) and \( 80 - x \) be the two required numbers.

\[f(x) = x^2 + (80 - x)^2 \]
\[f(x) = x^2 + 80^2 - 2(80)(x) + x^2 \]
\[f(x) = 2x^2 - 160x + 6400 \]
\[\therefore f'(x) = 4x - 160 \]
\[f''(x) = 4 \]

Since \( f(x) \) has minimum,
\[f'(x) = 0 \]
\[4x - 160 = 0 \]
\[4x = 160 \]
\[x = \frac{160}{4} = 40. \]

\[\therefore f''(x) = 4 > f''(40) = 4 > 0. \]
\[\therefore f(x) \text{ has minimum at } x = 40. \]

\[\therefore \text{The required numbers are} \]
\[40 \text{ and } 80 - 40 = 40 \]
\[\therefore \text{The required parts of 80 are} \]
\[40 \text{ and } 40. \]

Ex: (4)
A manufacturer can sell \( x \) items at a price of Rs. \( (330-x) \) each. The cost of producing \( x \) items is Rs. \( (x^2 + 10x + 12) \). Find \( x \) for which the profit is maximum.

Solution:
Given that the total cost function is
\[ C = x^2 + 10x + 12. \]

Selling price \( p = 330 - x \)

Revenue is \( R = p \times D \)
\[ = p \times x \quad (D = x) \]
\[ = (330 - x) \times x \]
\[ = 330x - x^2 \]

\[ \therefore \text{Profit} = \text{Revenue} - \text{Cost} \]
\[ \therefore \text{Profit} = (330 - x) - (x^2 + 10x + 12) \]
\[ = 330x - x^2 - 10x - 12 \]
\[ \text{Profit} = 320x - 2x^2 - 12 \]

\[ \therefore \frac{dp}{dx} = 320 - 4x \]

\[ \frac{d^2p}{dx^2} = -4 < 0 \]

\[ \therefore \text{The profit is maximum.} \]

Since the profit is maximum,
\[ \frac{dp}{dx} = 0 \]
\[ 320 - 4x = 0 \]
\[ 4x = 320 \]
\[ x = 80. \]

Hence the profit is maximum when 80 items are sold.

Ex: (5) The total cost function is
\[ C = x^3 - 9x^2 + 24x + 70. \]
Find \( x \) for which the total cost is minimum.

Solution:

Let \( C = f(x) = x^3 - 9x^2 + 24x + 70 \)
\[ \therefore C' = f' (x) = 3x^2 - 18x + 24 \]
\[ C'' = f'' (x) = 6x - 18 \]

Since \( f(x) \) has minimum,
\[ f' (x) = 0 \]
\[ 3x^2 = 18x + 24 = 0 \]
\[3(x^2 - 6x + 8) = 0\]
\[x^2 - 6x + 8 = 0\]
\[x^2 - 2x - 4x + 8 = 0\]
\[x(x - 2) - 4(x - 2) = 0\]
\[(x - 2)(x - 4) = 0\]
\[x - 2 = 0\] or \[x - 4 = 0\]
\[x = 2\] or \[x = 4\]

When \(x = 4\)

\[f''(x) = 6x - 18\]
\[f''(4) = 6(4) - 18\]
\[= 24 > 0\]
\[\therefore f(x)\] has minimum at \(x = 4\).

\[\therefore \text{The total cost is minimum at } x = 4.\]

Ex: (6) The total revenue function is given by

\[R = 4x^3 - 72x^2 + 420x + 800.\]

Find \(x\) for which the total revenue is maximum.

Solution:
Let \(R = f(x) = 4x^3 - 72x^2 + 420x + 800\)

\[\therefore R' = f'(x) = 12x^2 - 144x + 420\]

\[R'' = f''(x) = 24x - 144\]

Since \(f(x)\) has maximum,

\[f''(x) = 0\]
\[\therefore 12x^2 - 144x + 420 = 0\]
\[\therefore 12(x^2 - 12x + 35) = 0\]
\[\therefore x^2 - 12x + 35 = 0\]
\[x^2 - 5x - 7x + 35 = 0\]
\[x(x - 5) - 7(x - 5) = 0\]
\[(x - 5)(x - 7) = 0\]
\[x - 5 = 0\] or \[x - 7 = 0\]
\[\therefore x = 5\] or \[x = 7\]

When \(x = 5\)

\[R'' = f''(x) = 24x - 144\]
\[= 24(5) - 144\]
\[= -24 < 0\]

\[\therefore f(x)\] has maximum at \(x = 5\)
The total revenue is maximum at x=5

**Exercise : 9.5**

1. Find the extreme values of \( f(x) = 2x^3 - 6x^2 - 48x + 90 \).
2. Find the maximum and minimum values of \( f(x) = x + (9/x) \), \( x \neq 0 \)
3. Find the extreme values of \( f(x) = 4x^3 - 12x^2 - 36x + 25 \)
4. Find the extreme values of \( f(x) = x + (36/x) \), \( x \neq 0 \).
5. Divide 120 into two parts such that their product is maximum.
6. Divide 70 into two parts such that the sum of their squares is a minimum.
7. A manufacturer sells x items at a price of Rs. (400 - x) each. The cost of producing x items is Rs \( (x^2 + 40x + 52) \). Find x for which the profit is maximum.
8. The cost function is given by \( C = x^3 - 24x^2 + 189x + 120 \). Find x for which the cost is minimum.
9. The total revenue function is given by

\[
R = 2x^3 - 63x^2 + 648x + 250.
\]

Find x for which the total revenue is maximum.
10. The total cost of producing x units is Rs \( (x^2 + 2x + 5) \) and the price is Rs \( (30 - x) \) per unit. Find x for which the profit is maximum.

**Answers:**

1. Maximum value = 146 at \( x = -2 \)
   Minimum value = -70 at \( x = 4 \)
2. Maximum value = -6 at \( x = -3 \)
   Minimum value = 6 at \( x = 3 \)
3. Maximum value = 45 at \( x = -1 \)
   Minimum value = -83 at \( x = 3 \)
4. Maximum value = -12 at \( x = -6 \)
   Minimum value = 12 at \( x = 6 \)
5. 60, 60
6. 35, 35
7. 90
OBJECTIVES

After reading this chapter you will be able to:

- Define interest, principal, rate of interest, period.
- Find simple interest (SI), rate of S.I., period of investment.
- Find Compound Interest (CI), rate of C.I., Amount accumulated at the end of a period.
- Compound interest compounded yearly, half-yearly, quarterly or monthly.

10.1 INTRODUCTION

In every day life individuals and business firms borrow money from various sources for different reasons. This amount of money borrowed has to be returned from the lender in a stipulated time by paying some interest on the amount borrowed. In this chapter we are going to study the two types of interests viz. simple and compound interest. We start with some definitions and then proceed with the formula related to both the types of interests.

10.2 Definitions Of Terms Used In This Chapter

**Principal**: The sum borrowed by a person is called its *principal*. It is denoted by $P$.

**Period**: The time span for which money is lent is called *period*. It is denoted by $n$.

**Interest**: The amount paid by a borrower to the lender for the use of money borrowed for a certain period of time is called *Interest*. It is denoted by $I$.

**Rate of Interest**: This is the interest to be paid on the amount of Rs. 100 per annum (i.e. per year). This is denoted by $r$. 
Total Amount: The sum of the principal and interest is called as the total amount (maturity value) and is denoted by $A$. Thus, $A = P + I$. i.e. Interest paid $I = A - P$.

10.2 SIMPLE INTEREST

The interest which is payable on the principal only is called as simple interest (S.I.). For example the interest on Rs. 100 at 11% after one year is Rs.11 and the amount is $100 + 11 = Rs. 111$.

It is calculated by the formula: $I = \frac{Pnr}{100} = P \times n \times \frac{r}{100}$

<table>
<thead>
<tr>
<th>Simple Interest = Principal x period x rate of interest</th>
</tr>
</thead>
<tbody>
<tr>
<td>Amount at the end of $n^{th}$ year = $A = P + I = P + \frac{Pnr}{100} = P \left(1 + \frac{nr}{100}\right)$</td>
</tr>
</tbody>
</table>

Remark: The period $n$ is always taken in ‘years’. If the period is given in months/days, it has to be converted into years and used in the above formula. For example, if period is 4 months then we take $n = 4/12 = 1/3$ or if period is 60 days then $n = 60/365$.

Example 1: If Mr. Sagar borrows Rs. 500 for 2 years at 10% rate of interest, find (i) simple interest and (ii) total amount.

Ans: Given $P = Rs. 500$, $n = 2$ and $r = 10 \%$

(i) $I = \frac{Pnr}{100} = \frac{500 \times 2 \times 10}{100} = Rs. 100$

(ii) $A = P + I = 500 + 100 = Rs. 600$

10.2.1 Problems involving unknown factors in the formula $I = \frac{Pnr}{100}$

The formula $I = \frac{Pnr}{100}$ remaining the same, the unknown factor in the formula is taken to the LHS and its value is computed. For example, if rate of interest is unknown then the formula is rewritten as $r = \frac{I \times 100}{P \times n}$.

Example 2: If Mr. Prashant borrows Rs. 1000 for 5 years and pays an interest of Rs. 300, find rate of interest.

Ans: Given $P = 1000$, $n = 5$ and $I = Rs. 300$

Now, $I = \frac{Pnr}{100} \Rightarrow r = \frac{I \times 100}{P \times n} = \frac{300 \times 100}{1000 \times 5} = 6$

Thus, the rate of interest is 6%.
**Example 3:** Find the period for Rs. 2500 to yield Rs. 900 in simple interest at 12%.

**Ans:** Given \( P = Rs. \ 2500, \ I = 900, \ r = 12\% \)

Now, \( I = \frac{Pnr}{100} \) \( \Rightarrow \ n = \frac{I \times 100}{P \times r} = \frac{900 \times 100}{2500 \times 12} = 3 \)

Thus, the period is 3 years.

**Example 4:** Find the period for Rs. 1000 to yield Rs. 50 in simple interest at 10%.

**Ans:** Given \( P = Rs. \ 1000, \ I = 50, \ r = 10\% \)

Now, \( I = \frac{Pnr}{100} \) \( \Rightarrow \ n = \frac{I \times 100}{P \times r} = \frac{50 \times 100}{1000 \times 10} = 0.5 \)

Thus, the period is 0.5 years i.e. 6 months.

**Example 5:** Mr. Akash lent Rs. 5000 to Mr. Prashant and Rs. 4000 to Mr. Sagar for 5 years and received total simple interest of Rs. 4950. Find (i) the rate of interest and (ii) simple interest of each.

**Ans:** Let the rate of interest be \( r. \)

S.I. for Prashant = \( \frac{5000 \times 5 \times r}{100} = 250r \) \( \ldots (1) \)

and S.I. for Sagar = \( \frac{4000 \times 5 \times r}{100} = 200r \) \( \ldots (2) \)

from (1) and (2), we have,

total interest from both = \( 250r + 200r = 450r \)

But total interest received be Mr. Akash = Rs. 4950

\( \therefore 450r = 4950 \) \( \Rightarrow r = \frac{4950}{450} = 11 \)

\( \therefore \) the rate of interest = 11%

**Example 6:** The S.I. on a sum of money is one-fourth the principal. If the period is same as that of the rate of interest then find the rate of interest.

**Ans:** Given \( I = \frac{P}{4} \) and \( n = r \)

Now, we know that \( I = \frac{Pnr}{100} \)

\( \therefore \frac{P}{4} = \frac{P \times r \times r}{100} \) \( \Rightarrow \frac{100}{4} = r^2 \)

\( \therefore r^2 = 25 \) \( \Rightarrow r = 5. \)

\( \therefore \) the rate of interest = 5%

**Example 7:** If Rs. 8400 amount to Rs. 11088 in 4 years, what will Rs. 10500 amount to in 5 years at the same rate of interest?

**Ans:**
(i) Given \( n = 4, P = \text{Rs. 8400}, A = \text{Rs. 11088} \)
\[ I = A - P = 11088 - 8400 = \text{Rs. 2688} \]
Let \( r \) be the rate of interest.

Now, \( \frac{Pnr}{100} = 2688 \Rightarrow 2688 = \frac{8400 \times 4 \times r}{100} \)
\[ r = 8\% \]

(ii) To find \( A \) when \( n = 5, P = \text{Rs. 10500}, r = 8 \)
\[ A = P \left(1 + \frac{nr}{100}\right) = 10500 \times \left(1 + \frac{5 \times 8}{100}\right) = 10500 \times \frac{140}{100} = 14700 \]
\[ \therefore \text{the required amount} = \text{Rs. 14,700} \]

**Example 8:** Mr. Shirish borrowed Rs. 12,000 at 9\% interest from Mr. Girish on January 25, 2007. The interest and principal is due on August 10, 2007. Find the interest and total amount paid by Mr. Shirish.

**Ans:** Since the period is to be taken in years, we first count number of days from 25\(^{th}\) January to 10\(^{th}\) August, which is 197 days.

Now, \( I = \frac{Pnr}{100} = 12000 \times \frac{197}{365} \times \frac{9}{100} \)
\[ I = \text{Rs. 582.9} \]
Total amount = \( P + I = 12000 + 582.9 \)
\[ \therefore A = \text{Rs. 12,582.9} \]

**Check your progress 10.1**

1. Find the SI and amount for the following data giving principal, rate of interest and number of years:
   (i) 1800, 6\%, 4 years.
   (ii) 4500, 8\%, 5 years
   (iii) 7650, 5.5\%, 3 years.
   (iv) 6000, 7.5\%, 6 years
   (v) 25000, 8\%, 5 years
   (vi) 20000, 9.5\%, 10 years.

**Ans:**
(i) 432, 2232
(ii) 1800, 6300
(iii) 1262.25, 8912.25
(iv) 2700, 8700
(v) 10000, 35000
(vi) 19000, 39000

2. Find the S.I. and the total amount for a principal of Rs. 6000 for 3 years at 6\% rate of interest.

**Ans:** 1080, 7080

3. Find the S.I. and the total amount for a principal of Rs. 3300 for 6 years at 3\%\% rate of interest.

**Ans:** 693, 3993
4. Find the S.I. and the total amount for a principal of Rs. 10550 for 2 years at 10\% \text{ rate of interest. }
   \textbf{Ans:} 2162.75, 12712.75

5. Find the rate of interest if a person invests Rs. 1000 for 3 years and receives a S.I. of Rs. 150.
   \textbf{Ans:} 5\%

6. Find the rate of interest if a person invests Rs. 1200 for 2 years and receives a S.I. of Rs. 168.
   \textbf{Ans:} 7\%

7. A person invests Rs. 4050 in a bank which pays 7\% S.I. What is the balance of amount of his savings after (i) six months, (ii) one year?
   \textbf{Ans:} 141.75, 283.5

8. A person invests Rs. 3000 in a bank which offers 9\% S.I. After how many years will his balance of amount will be Rs. 3135?
   \textbf{Ans:} 6 months

9. Find the principal for which the S.I for 4 years at 8\% is 585 less than the S.I for 3\% years at 11\%.
   \textbf{Ans:} 9000

10. Find the principal for which the S.I for 5 years at 7\% is 250 less than the S.I for 4 years at 10\%.
    \textbf{Ans:} 5000

11. Find the principal for which the S.I for 8 years at 7.5\% is 825 less than the S.I for 6\% years at 10.5\%.
    \textbf{Ans:} 10000

12. Find the principal for which the S.I for 3 years at 6\% is 230 more than the S.I for 3\% years at 5\%.
    \textbf{Ans:} 46000

13. After what period of investment would a principal of Rs. 12,350 amount to Rs. 17,043 at 9.5\% rate of interest?
    \textbf{Ans:} 4 years

14. A person lent Rs. 4000 to Mr. X and Rs. 6000 to Mr. Y for a period of 10 years and received total of Rs. 3500 as S.I. Find (i) rate of interest, (ii) S.I. from Mr. X, Mr. Y.
    \textbf{Ans:} 3.5\%, 1400, 2100

15. Miss Pankaj Kansra lent Rs. 2560 to Mr. Abhishek and Rs. 3650 to Mr. Ashwin at 6\% rate of interest. After how many years should he receive a total interest of Rs. 3726?
16. If the rate of S.I. on a certain principal is same as that of the period of investment yields same interest as that of the principal, find the rate of interest.

**Ans:** 10%  

17. If the rate of S.I. on a certain principal is same as that of the period of investment yields interest equal to one-ninth of the principal, find the rate of interest.

**Ans:** $3\frac{1}{3}$ years

18. Find the principal and rate of interest if a certain principal amounts to Rs. 2250 in 1 year and to Rs. 3750 in 3 years.

**Ans:** 1500, 50%

19. Find the principal and rate of interest if a certain principal amounts to Rs. 3340 in 2 years and to Rs. 4175 in 3 years. **Ans:** 1670, 50%

20. If Rs. 2700 amount Rs. 3078 in 2 years at a certain rate of interest, what will Rs. 7200 amount to in 4 years at the same rate on interest?

**Ans:** 7%, 9216

21. At what rate on interest will certain sum of money amount to three times the principal in 20 years?

**Ans:** 15%

22. Mr. Chintan earns as interest Rs. 1020 after 3 years by lending Rs. 3000 to Mr. Bhavesh at a certain rate on interest and Rs. 2000 to Mr. Pratik at a rate on interest 2% more than that of Mr. Bhavesh. Find the rates on interest.

**Ans:** 6%, 8%

23. Mr. Chaitanya invested a certain principal for 3 years at 8% and received an interest of Rs. 2640. Mr. Chihar also invested the same amount for 6 years at 6%. Find the principal of Mr. Chaitanya and the interest received by Mr. Chihar after 6 years.

**Ans:** 11000, 3960

24. Mr. Ashfaq Khan invested some amount in a bank giving 8.5% rate of interest for 5 years and some amount in another bank at 9% for 4 years. Find the amounts invested in both the banks if his total investment was Rs. 75,000 and his total interest was Rs. 29,925.

**Ans:** 45000, 30000

25. Mrs. Prabhu lent a total of Rs. 48,000 to Mr. Diwakar at 9.5% for 5 years and to Mr. Ratnakar at 9% for 7 years. If she receives a total interest of Rs. 25,590 find the amount she lent to both.

**Ans:** 18000, 30000
10.3 COMPOUND INTEREST

The interest which is calculated on the amount in the previous year is called **compound interest**.

For example, the compound interest on Rs. 100 at 8% after one year is Rs. 8 and after two years is $108 + 8\% \text{ of } 108 = Rs. \ 116.64$

If $P$ is the principal, $r$ is the rate of interest p.a. then the amount at the end of $n^{\text{th}}$ year called as **compound amount** is given by the formula:

$$A = P \left(1 + \frac{r}{100}\right)^n$$

The **compound interest** is given by the formula:

$$CI = A - P$$

**Note:**
1. The interest may be compounded annually (yearly), semi-annually (half yearly), quarterly or monthly. Thus, the general formula to calculate the amount at the end of $n$ years is as follows:

$$A = P \left(1 + \frac{r}{p \times 100}\right)^{np}$$

Here $p$: number of times the interest is compounded in a year.

- $p = 1$ if interest is compounded **annually**
- $p = 2$ if interest is compounded **semi-annually** (**half-yearly**)  
- $p = 4$ if interest is compounded **quarterly**
- $p = 12$ if interest is compounded **monthly**

2. It is easy to calculate amount first and then the compound interest as compared with finding interest first and then the total amount in case of simple interest.

**Example 9:** Find the compound amount and compound interest of Rs. 1000 invested for 10 years at 8% if the interest is compounded annually.

**Ans:** Given $P = 1000$, $r = 8$, $n = 10$.

Since the interest is compounded annually, we have

$$A = P \left(1 + \frac{r}{100}\right)^n = 1000 \times \left(1 + \frac{8}{100}\right)^{10} = 1000 \times 2.1589 = Rs. \ 2158.9$$

**Example 10:** Find the principal which will amount to Rs. 11,236 in 2 years at 6% compound interest compounded annually.
Ans: Given \( A = \text{Rs. 11236}, n = 2, r = 6 \) and \( P = ? \)

Now, \( A = P \left( 1+ \frac{r}{100} \right)^n \)

\[ \therefore \text{11236} = P \left( 1+ \frac{6}{100} \right)^2 = P \times 1.1236 \]

\[ \therefore P = \frac{11236}{1.1236} = 10,000 \]

\[ \therefore \text{the required principal is Rs. 10,000.} \]

**Example 11:** Find the compound amount and compound interest of Rs. 1200 invested for 5 years at 5\% if the interest is compounded (i) annually, (ii) semi annually, (iii) quarterly and (iv) monthly.

**Ans:** Given \( P = \text{Rs. 1200}, r = 5, n = 5 \)

Let us recollect the formula \( A = P \left( 1+ \frac{r}{p \times 100} \right)^{np} \)

(i) If the interest is compounded annually, \( p = 1 \):

\[ A = P \left( 1+ \frac{r}{100} \right)^n = 1200 \times \left( 1+ \frac{5}{100} \right)^5 = 1200 \times 1.2763 = \text{Rs. 1531.56} \]

\[ CI = A - P = 1531.56 - 1200 = \text{Rs. 331.56} \]

(ii) If the interest is compounded semi-annually, \( p = 2 \):

\[ A = P \left( 1+ \frac{r}{2 \times 100} \right)^{2n} = 1200 \times \left( 1+ \frac{5}{200} \right)^{10} = 1200 \times 1.28 = \text{Rs. 1536} \]

\[ CI = A - P = 1536 - 1200 = \text{Rs. 336}. \]

(iii) If the interest is compounded quarterly, \( p = 4 \):

\[ A = P \left( 1+ \frac{r}{4 \times 100} \right)^{4n} = 1200 \times \left( 1+ \frac{5}{400} \right)^{20} = 1200 \times 1.2820 = \text{Rs. 1538.4} \]

\[ CI = A - P = 1538.4 - 1200 = \text{Rs. 338.4} \]

(iv) If the interest is compounded monthly, \( p = 12 \):

\[ A = P \left( 1+ \frac{r}{12 \times 100} \right)^{12n} = 1200 \times \left( 1+ \frac{5}{1200} \right)^{60} = 1200 \times 1.2834 = \text{Rs. 1540} \]

\[ CI = A - P = 1540 - 1200 = \text{Rs. 340} \]

**Example 12:** Mr. Santosh wants to invest some amount for 4 years in a bank. Bank X offers 8\% interest if compounded half yearly while bank Y offers 6\% interest if compounded monthly. Which bank should Mr. Santosh select for better benefits?

**Ans:** Given \( n = 4 \).

Let the principal Mr. Santosh wants to invest be \( P = \text{Rs. 100} \)
From Bank $X$: $r = 8$ and interest is compounded half-yearly, so $p = 2$.

\[ A = P \left( 1 + \frac{r}{2 \times 100} \right)^{2n} = 100 \times \left( 1 + \frac{8}{200} \right)^4 = 116.9858 \quad \ldots (1) \]

From Bank $Y$: $r = 6$, $p = 12$

\[ A = P \left( 1 + \frac{r}{12 \times 100} \right)^{12n} = 100 \times \left( 1 + \frac{6}{1200} \right)^{48} = 127.0489 \quad \ldots (2) \]

Comparing (1) and (2), Dr. Ashwinikumar should invest his amount in bank $Y$ as it gives more interest at the end of the period.

**Example 13:** In how many years would Rs. 75,000 amount to Rs. 1,05,794.907 at 7% compound interest compounded semi-annually?

**Ans:** Given $A = Rs. 105794.907$, $P = Rs. 75000$, $r = 7$, $p = 2$

\[ A = P \left( 1 + \frac{r}{2 \times 100} \right)^{2n} \]

\[ 105794.907 = 75000 \times \left( 1 + \frac{7}{200} \right)^{2n} \]

\[ \frac{105794.907}{75000} = (1.035)^{2n} \]

\[ 1.41059876 = (1.035)^{2n} \]

\[ (1.035)^{10} = (1.035)^{2n} \quad \Rightarrow \quad 2n = 10 \]

Thus, $n = 5$

**Example 14:** A certain principal amounts to Rs. 4410 after 2 years and to Rs.4630.50 after 3 years at a certain rate of interest compounded annually. Find the principal and the rate of interest.

**Ans:** Let the principal be $P$ and rate of interest be $r$.

Now, we know that $A = P \left( 1 + \frac{r}{100} \right)^n$

From the given data we have,

\[ 4410 = P \left( 1 + \frac{r}{100} \right)^2 \quad \text{and} \quad 4630.5 = P \left( 1 + \frac{r}{100} \right)^3 \]

\[ \therefore 4410 = P (1 + 0.01r)^2 \quad \ldots (1) \]

\[ 4630.5 = P (1 + 0.01r)^3 \quad \ldots (2) \]

Do not write ‘1 + 0.01r’ as 1.01$r$
Dividing (2) by (1), we have

\[
\frac{4630.5}{4410} = \frac{P(1+0.01r)^3}{P(1+0.01r)^2} \Rightarrow 1.05 = 1 + 0.01r
\]

\[\therefore 0.05 = 0.01r\]

Thus, \(r = 5\%\)

**Example 15:** Find the rate of interest at which a sum of Rs. 2000 amounts to Rs. 2690 in 3 years given that the interest is compounded half yearly. \((\sqrt[6]{1.345} = 1.05)\)

**Ans:** Given \(P = Rs.\ 2000, A = Rs.\ 2680, n = 3, p = 2\)

Now, \(A = P\left(1 + \frac{r}{2 \times 100}\right)^{2n}\)

\[\therefore 2690 = 2000 \times \left(1 + \frac{r}{200}\right)^6\]

\[\therefore \frac{2690}{2000} = \left(1 + \frac{r}{200}\right)^6 \Rightarrow 1.345 = \left(1 + \frac{r}{200}\right)^6\]

\[\therefore \sqrt[6]{1.345} = 1 + \frac{r}{200} \Rightarrow 1.05 = 1 + \frac{r}{200}\]

\[\therefore r = 0.05 \times 200 = 10\%\]

Thus, the rate of compound interest is \(10\%\).

**Example 16:** If the interest compounded half yearly on a certain principal at the end of one year at 8\% is Rs. 3264, find the principal.

**Ans:** Given \(CI = Rs.\ 3264, n = 1, p = 2\) and \(r = 8\)

Now, \(CI = A - P = P\left(1 + \frac{8}{200}\right)^2 - P\)

i.e. \(3264 = P\left(1.04\right)^2 - P\)

\[\therefore P = \frac{3264}{0.0816} = 40000\]

Thus, the principal is \(Rs.\ 40,000\).

**Check your progress 10.2**

1. Compute the compound amount and interest on a principal of Rs. 21,000 at 9\% p.a. after 5 years.
   **Ans:** 32,311.10, 11,311.10

2. Compute the compound amount and interest on a principal of Rs. 6000 at 11\% p.a. after 8 years.
   **Ans:** 13827.23, 7827.23
3. Compute the compound amount and compound interest of Rs. 5000 if invested at 11% for 3 years and the interest compounded (i) annually, (ii) semi annually, (iii) quarterly and (iv) monthly.  
   **Ans:** (i) 6838.16, 1838.16  
   (ii) 6894.21, 189421  
   (iii) 6923.92, 1923.92  
   (iv) 6944.39, 1944.39

4. Compute the compound amount and compound interest of Rs. 1200 if invested at 9% for 2 years and the interest compounded (i) annually, (ii) semi annually, (iii) quarterly and (iv) monthly.  
   **Ans:** (i) 1425.72, 225.72  
   (ii) 1431.02, 231.02  
   (iii) 1433.8, 233.8  
   (iv) 1435.7, 235.7

5. Miss Daizy invested Rs. 25,000 for 5 years at 7.5% with the interest compounded semi-annually. Find the compound interest at the end of 5 years.  
   **Ans:** 11,126.10

6. Mr. Dayanand borrowed a sum of Rs. 6500 from his friend at 9% interest compounded quarterly. Find the interest he has to pay at the end of 4 years?  
   **Ans:** 2779.54

7. Mr. Deepak borrowed a sum of Rs. 8000 from his friend at 8% interest compounded annually. Find the interest he has to pay at the end of 3 years?  
   **Ans:** 2077.70

8. Mr. Deshraj borrowed Rs. 1,25,000 for his business for 3 years at 25% interest compounded half yearly. Find the compound amount and interest after 3 years.  
   **Ans:** 2,53,410.82; 12,8410.82

9. Mrs. Hemlata bought a Sony Digital Camera for Rs. 15,800 from Vijay Electronics by paying a part payment of Rs. 2,800. The remaining amount was to be paid in 3 months with an interest of 9% compounded monthly on the due amount. How much amount did Mrs. Hemlata paid and also find the interest.  
   **Ans:** 13294.70, 294.70

10. Mr. Irshad bought a Kisan Vikas Patra for Rs. 10000, whose maturing value is Rs. 21,000 in 4½ years. Calculate the rate of interest if the compound interest is compounded quarterly.  
    **Ans:** 16.8%

11. What sum of money will amount to Rs. 11236 in 2 years at 6% p.a. compound interest?  
    **Ans:** 10,000
12. Find the principal which will amount to Rs. 13468.55 in 5 years at 6% interest compounded quarterly. \[ (1.015)^{20} = 1.346855 \]  
Ans: 10000

13. Find the principal which will amount to Rs. 30626.075 in 3 years at 7% interest compounded yearly.  
Ans: 25000

14. Find the principal if the compound interest payable annually at 8% p.a. for 2 years is Rs. 1664.  
Ans: 10000

15. If Mr. Sagar wants to earn Rs. 50000 after 4 years by investing a certain amount in a company at 10% rate of interest compounded annually, how much should he invest?  
Ans: 34150.67

16. Find after how many years will Rs. 4000 amount to Rs. 4494.40 at 6% rate of interest compounded yearly.  
Ans: \( n = 2 \)

17. Find after how many years Rs. 10,000 amount to Rs. 12,155 at 10% rate of interest compounded half-yearly.  
Ans: \( n = 1 \)

18. Find the rate of interest at which a principal of Rs.10000 amounts to Rs. 11236 after 2 years.  
Ans: 6%

19. Find the rate of interest at which a principal of Rs.50000 amounts to Rs. 62985.6 after 3 years. \( (\sqrt[3]{1.259712} = 1.08) \)  
Ans: 8%

20. Mrs. Manisha Lokhande deposited Rs. 20,000 in a bank for 5 years. If she received Rs.3112.50 as interest at the end of 2 years, find the rate of interest p.a. compounded annually.  
Ans: 7.5%

21. A bank \( X \) announces a super fixed deposit scheme for its customers offering 10% interest compounded half yearly for 6 years. Another bank \( Y \) offers 12% simple interest for the same period. Which bank’s scheme is more beneficial for the customers?  
Ans: Bank \( X \)

22. \( ABC \) bank offers 9% interest compounded yearly while \( XYZ \) bank offers 7% interest compounded quarterly. If Mr. Arunachalam wants to invest Rs. 18000 for 5 years, which bank should he choose?  
Ans: Bank \( ABC \)

23. Mangesh borrowed a certain amount from Manish at a rate of 9% for 4 years. He paid Rs. 360 as simple interest to Manish. This amount he invested in a bank for 3 years at 11% rate of interest compounded quarterly. Find the compound interest Mangesh
24. On a certain principal for 3 years the compound interest compounded annually is Rs. 1125.215 while the simple interest is Rs. 1050, find the principal and the rate of interest.
   \[ \text{Ans: } 5000, 7\% \]

25. On a certain principal for 4 years the compound interest compounded annually is Rs. 13923 while the simple interest is Rs. 12000, find the principal and the rate of interest.
   \[ \text{Ans: } 30000, 10\%. \]

26. Which investment is better for Mr. Hariom Sharma (i) 6% compounded half yearly or (ii) 6.2% compounded quarterly?
   \[ \text{Ans: } \]

27. Which investment is better for Mr. Suyog Apte (i) 9% compounded yearly or (ii) 8.8% compounded quarterly?
   \[ \text{Ans: } \]

28. A bank X offers 7% interest compounded semi-annually while another bank offers 7.2% interest compounded monthly. Which bank gives more interest at the end of the year?
   \[ \text{Ans: } \]

29. Mr. Nitin Tare has Rs. 10000 to be deposited in a bank. One bank offers 8% interest p.a. compounded half yearly, while the other offers 9% p.a. compounded annually. Calculate the returns in both banks after 3 years. Which bank offers maximum return after 3 years?
   \[ \text{Ans: } \]

❖❖❖
Unit VII

11

ANNUITIES AND EMI

OBJECTIVES

After reading this chapter you will be able to:

- Define annuity, future value, present value, EMI, Sinking Fund.
- Compute Future Value of annuity due, Present Value of an ordinary annuity.
- Compute EMI of a loan using reducing balance method and flat interest method.
- Compute Sinking Fund (periodic payments).

11.1 INTRODUCTION

In the previous chapter we have seen how to compute compound interest when a lump sum amount is invested at the beginning of the investment. But many a time we pay (or are paid) a certain amount not in lump sum but in periodic installments. This series of equal payments done at periodic intervals is called as annuity.

Let us start the chapter with the definition of an annuity.

11.2 ANNUITY

A series of equal periodic payments is called annuity. The payments are of equal size and at equal time interval.

The common examples of annuity are: monthly recurring deposit schemes, premiums of insurance policies, loan installments, pension installments etc. Let us understand the terms related to annuities and then begin with the chapter.

Periodic Payment:
The amount of payment made is called periodic payment.

Period of Payment:
The time interval between two payments of an annuity is called as the period of payment.

Term of an annuity:
The time of the beginning of the first payment period to the end of the last payment period is called as term of annuity. An annuity gets matured at the end of its term.

11.3 TYPES OF ANNUITIES

Though we will be discussing two types of annuities in detail, let us understand different types of annuities based on the duration of the term or on the time when the periodic payments are made. On the basis of the closing of an annuity, there are three types of annuities:

1. **Certain Annuity:**
   Here the duration of the annuity is fixed (or certain), hence called *certain annuity*. We will be learning such annuities in detail.

2. **Perpetual Annuity:**
   Here the annuity has no closing duration, i.e. it has indefinite duration. Practically there are rarely any perpetuities.

3. **Contingent Annuity:**
   Here the duration of the annuity depends on an event taking place. An example of contingent annuity is *life annuity*. Here the payments are to be done till a person is alive, like, pension, life insurance policies for children (maturing on the child turning 18 years) etc.

On the basis of when the periodic payments are made we have two types of annuities: ordinary annuity and annuity due.

11.3.1 **Immediate (Ordinary) Annuity:**
   The annuity which is paid at the *end of each period* is called as *immediate (ordinary) annuity*. The period can be monthly, quarterly or yearly etc. For example, stock dividends, salaries etc.

   Let us consider an example of an investment of Rs. 5000 each year is to be made for four years. If the investment is done at the end of each year then we have the following diagrammatic explanation for it:
11.3.2 Present Value:

The sum of all periodic payments of an annuity is called its present value. In simple words, it is that sum which if paid now will give the same amount which the periodic payments would have given at the end of the decided period. It is the one time payment of an annuity.

The formula to find the present value (PV) is as follows:

$$ PV = \frac{P}{r} \left(1 - \frac{1}{\left(1 + \frac{r}{p \times 100}\right)^{np}} \right) $$

Let \( i = \frac{r}{p \times 100} \), the rate per period, then the above formula can be rewritten as follows:

$$ PV = \frac{P}{i} \left[1 - \frac{1}{1 + i^{np}} \right] $$

11.3.3 Future Value (Accumulated value):

The sum of all periodic payments along with the interest is called the future value (accumulated amount) of the annuity.

The formula to find the future value (A) of an immediate annuity is as follows:

$$ A = \frac{P}{r} \left[\left(1 + \frac{r}{p \times 100}\right)^{np} - 1 \right] $$

Here,
- \( P \): periodic equal payment
- \( r \): rate of interest p.a.
- \( p \): period of annuity i.e. yearly, half yearly, quarterly or monthly
- \( i = \frac{r}{p \times 100} \)
**Example 1:** Find the future value after 2 years of an immediate annuity of Rs. 5000, the rate of interest being 6% p.a compounded annually.

**Ans:** Given \( n = 2, P = \text{Rs. 5000}, r = 6 \) and \( p = 1 \) \( \Rightarrow i = \frac{6}{100} = 0.06 \)

\[
A = \frac{P}{i} \left[ 1 + i^{np} - 1 \right] = \frac{5000}{0.06} \left[ 1 + 0.06^{2} - 1 \right] = 5000 \left[ \frac{1.1236 - 1}{0.06} \right]
\]

\[
A = 5000 \times 2.06 = \text{Rs. 10300}
\]

**Example 2:** Find the amount for an ordinary annuity with periodic payment of Rs. 3000, at 9% p.a. compounded semi-annually for 4 years.

**Ans:** Given \( n = 4, P = \text{Rs. 3000}, r = 9 \) and \( p = 2 \) \( \Rightarrow i = \frac{9}{2 \times 100} = 0.045 \)

Now, \( A = \frac{P}{i} \left[ 1 + i^{np} - 1 \right] = \frac{3000}{0.045} \left[ (1 + 0.045)^{2 \times 4} - 1 \right] = \frac{3000}{0.045} \times 0.4221 \)

Thus, \( A = \text{Rs. 28,140} \)

**Example 3:** Mr. Ravi invested Rs. 5000 in an annuity with quarterly payments for a period of 2 years at the rate of interest of 10%. Find the accumulated value of the annuity at the end of 2\(^{nd}\) year.

**Ans:** Given \( n = 2, P = \text{Rs.5000}, r = 10 \) and \( p = 4 \) \( \Rightarrow i = \frac{10}{4 \times 100} = 0.025 \)

Now, \( A = \frac{P}{i} \left[ 1 + i^{np} - 1 \right] = \frac{5000}{0.025} \left[ (1.025)^{2 \times 4} - 1 \right] = \frac{5000}{0.025} \times 0.2184 \)

Thus, \( A = \text{Rs. 43,680} \)

**Example 4:** Mr. Ashok Rane borrowed Rs. 20,000 at 4% p.a. compounded annually for 10 years. Find the periodic payment he has to make.

**Ans:** Given \( PV = \text{Rs. 20,000}; n = 10; p = 1 \) and \( r = 4 \) \( \Rightarrow i = 0.04 \)

Now to find the periodic payments \( P \) we use the following formula:

\[
PV = \frac{P}{i} \left[ 1 - \frac{1}{1 + i^{np}} \right]
\]

\[
\therefore 20000 = \frac{P}{0.04} \left[ 1 - \frac{1}{1 + 0.04^{10}} \right] = \frac{P}{0.04} \times 0.3244
\]

\[
\therefore P = \frac{20000 \times 0.04}{0.3244} = 2466.09
\]

Thus, Mr. Rane has to make the periodic payments of Rs. 2466.09

**Example 5:** Find the future value of an immediate annuity after 3 years with the periodic payment of Rs. 12000 at 5% p.a. if the period of payments is (i) yearly, (ii) half-yearly, (iii) quarterly and (iv) monthly.
Ans: Given $P = \text{Rs. } 1200$, $n = 3$, $r = 5$

(i) period $p = 1$ then $i = \frac{5}{100} = 0.05$

$$A = \frac{P}{i} \left[ 1 + i^{n-p} - 1 \right] = \frac{12000}{0.05} \left[ 1 + 0.05^3 - 1 \right] = \frac{12000}{0.05} \times 1.1576 - 1$$
\[ \therefore A = 12000 \times 3.1525 = \text{Rs. } 37,830 \]

(ii) period $p = 2$ then $i = \frac{5}{2 \times 100} = 0.025$

$$A = \frac{P}{i} \left[ 1 + i^{2n-1} \right] = \frac{12000}{0.025} \left[ (1 + 0.025)^6 - 1 \right] = \frac{12000}{0.025} \times 0.1597$$
\[ \therefore A = 12000 \times 6.388 = \text{Rs. } 76,656 \]

(iii) period $p = 4$ then $i = \frac{5}{4 \times 100} = 0.0125$

$$A = \frac{P}{i} \left[ 1 + i^{4n-1} \right] = \frac{12000}{0.0125} \left[ (1 + 0.0125)^{12} - 1 \right] = \frac{12000}{0.0125} \times 0.16075$$
\[ \therefore A = 12000 \times 12.86 = \text{Rs. } 1,54,320 \]

(iv) period $p = 12$ then $i = \frac{5}{12 \times 100} = 0.00417$

$$A = \frac{P}{i} \left[ 1 + i^{12n-1} \right] = \frac{12000}{0.00417} \left[ (1 + 0.00417)^{36} - 1 \right] = \frac{12000}{0.00417} \times 0.1615$$
\[ \therefore A = 12000 \times 38.729 = \text{Rs. } 4,64,748 \]

**Example 6:** Mr. Nagori invested certain principal for 3 years at 8% interest compounded half yearly. If he received Rs. 72957.5 at the end of 3rd year, find the periodic payment he made. [$(1.04)^6 = 1.2653$]

Ans: Given $n = 3$, $r = 8$, $p = 2 \Rightarrow i = \frac{8}{2 \times 100} = 0.04$

Now, $A = \frac{P}{i} \left[ 1 + i^{np} - 1 \right]$

\[ \therefore 72957.5 = \frac{P}{0.04} \left[ (1 + 0.04)^6 - 1 \right] = \frac{P}{0.04} \times 0.2653 \]
\[ \therefore 72957.5 = P[6.6325] \]
\[ \therefore P = \frac{72957.5}{6.6325} = 11000 \]

Thus, the periodic payment is Rs. 11,000

**11.4 SINKING FUND**
The fund (money) which is kept aside to accumulate a certain sum in a fixed period through periodic equal payments is called as **sinking fund**.

We can consider an example of a machine in a factory which needs to be replaced after say 10 years. The amount for buying a new machine 10 years from now may be very large, so a proportionate amount is accumulated every year so that it amounts to the required sum in 10 years. This annual amount is called as *sinking fund*. Another common example is of the *maintenance tax* collected by any Society from its members.

A sinking fund being same as an annuity, the formula to compute the terms is same as that we have learnt in section 11.3.3

**Example 7:** A company sets aside a sum of Rs. 15,000 annually to enable it to pay off a debenture issue of Rs. 1,80,000 at the end of 10 years. Assuming that the sum accumulates at 6% p.a., find the surplus after paying off the debenture stock.

**Ans:** Given \( P = \) Rs. 15000, \( n = 10 \), \( r = 6 \) \( \Rightarrow i = 0.06 \)

\[
A = \frac{P}{i} \left[ \frac{1 + i^n - 1}{1 - i} \right] = \frac{15000}{0.06} \times [(1 + 0.06)^{10} - 1] = \frac{15000}{0.06} \times 0.7908
\]

\[
A = \text{Rs. 1,97,700}
\]

Thus, the surplus amount after paying off the debenture stock is

\[
= 197712 - 180000 = \text{Rs. 17712}.
\]

**Example 8:** Shriniketan Co-op Hsg. Society has 8 members and collects Rs. 2500 as maintenance charges from every member per year. The rate of compound interest is 8% p.a. If after 4 years the society needs to do a work worth Rs. 100000, are the annual charges enough to bear the cost?

**Ans:** Since we want to verify whether Rs. 2500 yearly charges are enough or not we assume it to be \( P \) and find its value using the formula:

\[
A = \frac{P}{i} \left[ \frac{1 + i^n - 1}{1 - i} \right]
\]

Here \( A = \) Rs. 100000, \( n = 4 \), \( r = 8 \) \( \Rightarrow i = 0.08 \)

\[
\therefore P = \frac{A \times i}{1 + i^n - 1} = \frac{100000 \times 0.08}{1 + 0.08^4 - 1} = 22192
\]

Thus, the annual payment of all the members i.e. 8 members should be Rs. 22192.

\[
\therefore \text{the annual payment per member} = \frac{22192}{8} = \text{Rs. 2774}
\]
This payment is less than Rs. 2500 which the society has decided to take presently. Thus, the society should increase the annual sinking fund.

11.5 EQUATED MONTHLY INSTALLMENT (EMI)

Suppose a person takes a loan from a bank at a certain rate of interest for a fixed period. The equal payments which the person has to make to the bank per month are called as equated monthly installments in short EMI.

Thus, EMI is a kind of annuity with period of payment being monthly and the present value being the sum borrowed.

We will now study the method of finding the EMI using reducing balance method and flat interest method.

(a) Reducing balance method:

Let us recall the formula of finding the present value of an annuity.

\[ PV = \frac{P}{i} \left[ 1 - \frac{1}{1 + \frac{i}{p}}^{np} \right] \]

The equal periodic payment \( P \) is our EMI which is denoted it by \( E \).

The present value \( PV \) is same as the sum \( S \) borrowed.

Also the period being monthly \( p = 12, i = \frac{r}{1200} \) as we are interested in finding monthly installments and \( n \) is period in years.

Substituting this in the above formula we have:

\[ S = \frac{E}{i} \left[ 1 - \frac{1}{1 + \frac{i}{12n}}^{12n} \right] \]

Thus, if \( S \) is the sum borrowed for \( n \) years with rate of interest \( r \) % p.a. then the EMI is calculated by the formula:

\[ E = \frac{S \times i}{1 - \left(1 + \frac{i}{12n}\right)^{12n}} \]

(b) Flat Interest Method:

Here the amount is calculated using Simple Interest for the period and the EMI is computed by dividing the amount by total number of monthly installments.

Let \( S \) denote the sum borrowed, \( r \) denote the rate of interest and \( n \) denote the duration in years, then as we know the amount using simple interest formula is \( A = S \left(1 + \frac{nr}{100}\right)\). The total number of monthly installments for duration of \( n \) years is \( 12n \). Hence the EMI is calculated as

\[ E = \frac{A}{12n} \]
**Example 9:** Mr. Sudhir Joshi has taken a loan of Rs. 10,00,000 from a bank for 10 years at 11% p.a. Find his EMI using (a) reducing balance method and (b) Flat interest method.

**Ans:** Given \( S = Rs. \ 1000000, \ n = 10, \ r = 11 \Rightarrow i = \frac{11}{1200} = 0.0092 \)

(a) **Using flat interest method:**

\[
A = S \left(1 + \frac{nr}{100}\right) = 1000000 \left(1 + \frac{110}{100}\right) = 2100000
\]

Thus, \( E = \frac{A}{12n} = \frac{2100000}{120} = 17,500 \) … (1)

(b) **Using reducing balance method:**

Now, \( E = \frac{S \times i}{1 - \frac{1}{(1+i)^n}} = \frac{600000 \times 0.0092}{1 - \frac{1}{(1+0.0092)^{120}}} = 13797.65 \)

\( \therefore E = Rs. \ 13,798 \) approximately … (2)

Comparing (1) and (2), we can see that the EMI using flat interest method is higher than by reducing balance method.

**Example 10:** Mr. Prabhakar Naik has borrowed a sum of Rs. 60,000 from a person at 6% p.a. and is due to return it back in 4 monthly installments. Find the EMI he has to pay and also prepare the amortization table of repayment.

**Ans:** Given \( S = Rs. \ 60,000; \ n = 4 \) months;

\( r = 6\% \Rightarrow i = \frac{6}{1200} = 0.005 \)

Now, \( E = \frac{S \times i}{1 - \frac{1}{(1+i)^n}} = \frac{60000 \times 0.005}{1 - \frac{1}{(1+0.005)^4}} = \frac{300}{0.01975} \)

\( \therefore E = Rs. \ 15,187.97 \)

Now, we will prepare the **amortization table** i.e. the table of repayment of the sum borrowed using reducing balance method.

In the beginning of the 1\(^{st}\) month the outstanding principal is the sum borrowed i.e. Rs. 60000 and the EMI paid is Rs. 15187.97

The interest on the outstanding principal is \( 0.005 \times 60000 = Rs. \ 300 \) … (1)

Thus, the principal repayment is \( 15187.97 - 300 = Rs. \ 14887.97 \) … (2)

The outstanding principal (O/P) in the beginning of the 2\(^{nd}\) month is now \( 60000 - 14887.97 = 45112.03 \).

**Note:**
• (1) is called the **interest part** of the EMI and (2) is called as the **principal part** of the EMI.
• As the tenure increases the interest part reduces and the principal part increases.

This calculation can be tabulated as follows:

<table>
<thead>
<tr>
<th>Month</th>
<th>O/P</th>
<th>EMI</th>
<th>Interest Part</th>
<th>Principal Part</th>
</tr>
</thead>
<tbody>
<tr>
<td></td>
<td>(a)</td>
<td>(b)</td>
<td>(c) = (a) x i</td>
<td>(b) - (c)</td>
</tr>
<tr>
<td>1</td>
<td>60000</td>
<td>15187.97</td>
<td>300</td>
<td>14887.97</td>
</tr>
<tr>
<td>2</td>
<td>45112.03</td>
<td>15187.97</td>
<td>225.56</td>
<td>14962.45</td>
</tr>
<tr>
<td>3</td>
<td>30141.02</td>
<td>15187.97</td>
<td>150.75</td>
<td>15037.22</td>
</tr>
<tr>
<td>4</td>
<td>15111.80</td>
<td>15187.97</td>
<td>75.56</td>
<td>15112.41</td>
</tr>
</tbody>
</table>

In the beginning of the 4\(^{th}\) month the outstanding principal is Rs. 15111.80 but the actual principal repayment in that month is Rs. 15112.41. This difference is due to rounding off the values to two decimals, which leads the borrower to pay 61 paise more!!

**Example 11:** Mr. Shyam Rane has borrowed a sum of Rs. 100000 from a bank at 12\% p.a. and is due to return it back in 5 monthly installments. Find the EMI he has to pay and also prepare the amortization table of repayment.

**Ans:** Given \(S = \text{Rs. } 100000; \ n = 5 \text{ months;}\)

\[ r = 12\% \text{ p.a.} = \frac{12}{12} = 1\% \text{ p.m} \Rightarrow i = 0.01 \]

Now, 
\[
E = \frac{S \times i}{1 - \frac{1}{(1+i)^n}} = \frac{100000 \times 0.01}{1 - \frac{1}{(1+0.01)^5}} = \frac{1000}{0.0485343} = 20603.98
\]

The amortization table is as follows:

<table>
<thead>
<tr>
<th>Month</th>
<th>O/P</th>
<th>EMI</th>
<th>Interest Part</th>
<th>Principal Part</th>
</tr>
</thead>
<tbody>
<tr>
<td></td>
<td>(a)</td>
<td>(b)</td>
<td>(c) = (a) x i</td>
<td>(b) - (c)</td>
</tr>
<tr>
<td>1</td>
<td>100000</td>
<td>20603.98</td>
<td>1000</td>
<td>19603.98</td>
</tr>
<tr>
<td>2</td>
<td>80396.02</td>
<td>20603.98</td>
<td>803.96</td>
<td>19800.02</td>
</tr>
<tr>
<td>3</td>
<td>60596</td>
<td>20603.98</td>
<td>605.96</td>
<td>19998.02</td>
</tr>
</tbody>
</table>
Check your progress

1. An overdraft of Rs. 50,000 is to be paid back in equal annual installments in 20 years. Find the installments, if the interest is 12% p.a. compounded annually. \[(1.12)^{20} = 9.64629\]

2. A man borrows Rs. 30,000 at 6% p.a. compounded semi-annually for 5 years. Find the periodic payments he has to make.

3. What periodic payments Mr. Narayanan has to make if he has borrowed Rs. 1,00,000 at 12% p.a. compounded annually for 12 years? \[(1.12)^{12} = 3.896\]

4. Find the future value of an immediate annuity of Rs. 1200 at 6% p.a. compounded annually for 3 years.

5. Find the future value of an immediate annuity of Rs. 500 at 8% p.a. compounded p.m. for 5 years.

6. Find the accumulated value after 2 years if a sum of Rs. 1500 is invested at the end of every year at 10% p.a. compounded quarterly.

7. Find the accumulated amount of an immediate annuity of Rs. 1000 at 9% p.a. compounded semi-annually for 4 years.

8. Find the future value of an immediate annuity of Rs. 2800 paid at 10% p.a. compounded quarterly for 2 years. Also find the interest earned on the annuity.

9. Find the sum invested and the accumulated amount for an ordinary annuity with periodic payment of Rs. 2500, at the rate of interest of 9% p.a. for 2 years if the period of payment is (a) yearly, (b) half-yearly, (c) quarterly or (d) monthly.

10. Find the sum invested and the accumulated amount for an ordinary annuity with periodic payment of Rs. 1500, at the rate of interest of 10% p.a. for 3 years if the period of payment is (a) yearly, (b) half-yearly, (c) quarterly or (d) monthly.

11. Mr. Banerjee wants to accumulate Rs. 5,00,000 at the end of 10 years from now. How much amount should he invest every year at the rate of interest of 9% p.a. compounded annually?

12. Find the periodic payment to be made so that Rs. 25000 gets accumulated at the end of 4 years at 6% p.a. compounded annually.
13. Find the periodic payment to be made so that Rs. 30,000 gets accumulated at the end of 5 years at 8% p.a. compounded half yearly.

14. Find the periodic payment to be made so that Rs. 2000 gets accumulated at the end of 2 years at 12% p.a. compounded quarterly.

15. Find the rate of interest if a person depositing Rs. 1000 annually for 2 years receives Rs. 2070.

16. Find the rate of interest compounded p.a. if an immediate annuity of Rs. 50,000 amounts to Rs. 1,03,000 in 2 years.

17. Find the rate of interest compounded p.a. if an immediate annuity of Rs. 5000 amounts to Rs. 10400 in 2 years.

18. What is the value of the annuity at the end of 5 years, if Rs. 1000 p.m. is deposited into an account earning interest 9% p.a. compounded monthly? What is the interest paid in this amount?

19. What is the value of the annuity at the end of 3 years, if Rs. 500 p.m. is deposited into an account earning interest 6% p.a. compounded monthly? What is the interest paid in this amount?

20. Mr. Ashish Gokhale borrows Rs. 5000 from a bank at 8% compound interest. If he makes an annual payment of Rs. 1500 for 4 years, what is his remaining loan amount after 4 years?

(Hint: find the amount using compound interest formula for 4 years and then find the accumulated amount of annuity, the difference is the remaining amount.)

21. Find the present value of an immediate annuity of Rs. 10,000 for 3 years at 6% p.a. compounded annually.

22. Find the present value of an immediate annuity of Rs. 100000 for 4 years at 8% p.a. compounded half yearly.

23. Find the present value of an immediate annuity of Rs. 1600 for 2 years at 7% p.a. compounded half yearly.

24. A loan is repaid fully with interest in 5 annual installments of Rs. 15,000 at 8% p.a. Find the present value of the loan.

25. Mr. Suman borrows Rs. 50,000 from Mr. Juman and agreed to pay Rs. 14000 annually for 4 years at 10% p.a. Is this business profitable to Mr. Juman?

(Hint: Find the $PV$ of the annuity and compare with Rs. 50000)

26. Mr. Paradkar is interested in saving a certain sum which will amount to Rs. 3,50,000 in 5 years. If the rate of interest is 12% p.a., how much should he save yearly to achieve his target?
27. Mr. Kedar Pethkar invests Rs. 10000 per year for his daughter from her first birthday onwards. If he receives an interest of 8.5% p.a., what is the amount accumulated when his daughter turns 18?

28. Dr. Wakankar, a dentist has started his own dispensary. He wants to install a machine chair which costs Rs. 3,25,000. The machine chair is also available on monthly rent of Rs. 9000 at 9% p.a. for 3 years. Should Dr. Wakankar buy it in cash or rent it?

29. A sum of Rs. 50,000 is required to buy a new machine in a factory. What sinking fund should the factory accumulate at 8% p.a. compounded annually if the machine is to be replaced after 5 years?

30. The present cost of a machine is Rs. 80,000. Find the sinking fund the company has to generate so that it could buy a new machine after 10 years, whose value then would be 25% more than of today’s price. The rate of compound interest being 12% p.a. compounded annually.

31. Mr. Mistry has to options while buying a German wheel alignment machine for his garage: (a) either buy it at Rs. 1,26,000 or (b) take it on lease for 5 years at an annual rent of Rs. 30,000 at the rate of interest of 12% p.a.. Assuming no scrap value for the machine which option should Mr. Mistry exercise?

32. Regency Co-op. Hsg. Society which has 50 members require Rs. 12,60,000 at the end of 3 years from now for the society repairs. If the rate of compound interest is 10% p.a., how much fund the society should collect from every member to meet the necessary sum?

33. Mr. Lalwaney is of 40 years now and wants to create a fund of Rs. 15,00,000 when he is 60. What sum of money should he save annually so that at 13% p.a. he would achieve his target?

34. If a society accumulates Rs. 1000 p.a. from its 200 members for 5 years and receives 12% interest then find the sum accumulated at the end of the fifth year. If the society wants Rs. 13,00,000 for society maintenance after 5 years, is the annual fund of Rs. 1000 per member sufficient?

35. How much amount should a factory owner invest every year at 6% p.a. for 6 years, so that he can replace a mixture-drum (machine) costing Rs. 60,000, if the scrap value of the mixture-drum is Rs. 8,000 at the end of 6 years.

36. If a society accumulates Rs. 800 p.a. from its 100 members for 3 years and receives 9% interest then find the sum accumulated at
the end of the third year. If the society wants Rs. 2,50,000 for society maintenance after 3 years, is the annual fund of Rs. 800 per member sufficient?

37. Mr. Kanishk wants clear his loan of Rs. 10,00,000 taken at 12% p.a. in 240 monthly installments. Find his EMI using reducing balance method.

38. Using the reducing balance method find the EMI for the following:

<table>
<thead>
<tr>
<th>Loan amount (in Rs.)</th>
<th>rate of interest (in % p.a.)</th>
<th>period of loan (in yrs.)</th>
</tr>
</thead>
<tbody>
<tr>
<td>i) 1000</td>
<td>6</td>
<td>5</td>
</tr>
<tr>
<td>ii) 50000</td>
<td>6</td>
<td>10</td>
</tr>
<tr>
<td>iii) 8000</td>
<td>7</td>
<td>6</td>
</tr>
<tr>
<td>iv) 12000</td>
<td>9</td>
<td>10</td>
</tr>
<tr>
<td>v) 1000</td>
<td>9.5</td>
<td>10</td>
</tr>
<tr>
<td>vi) 1100000</td>
<td>12.5</td>
<td>20</td>
</tr>
</tbody>
</table>

39. Mr. Vilas Khopkar has taken a loan of Rs. 90,000 at 11% p.a. Find the EMI using (a) reducing balance method and (b) Flat interest method, if he has to return the loan in 4 years.

40. Find the EMI using reducing balance method on a sum of Rs. 36,000 at 9%, to be returned in 6 monthly installments.

41. Find the EMI using reducing balance method on a sum of Rs. 72,000 at 12%, to be returned in 12 installments.

42. Mr. Sachin Andhale has borrowed Rs. 10,000 from his friend at 9% p.a. and has agreed to return the amount with interest in 4 months. Find his EMI and also prepare the amortization table.

43. Mr. Arvind Kamble has borrowed Rs. 30,000 from his friend at 14% p.a. If he is to return this amount in 5 monthly installments, find the installment amount, the interest paid and prepare the amortization table for repayment.

44. Mrs. Chaphekar has taken a loan of Rs. 1,25,000 from a bank at 12% p.a. If the loan has to be returned in 3 years, find the EMI, Mrs. Chaphekar has to pay. Prepare the amortization table of repayment of loan and find the interest she has to pay.

45. A loan of Rs. 75,000 is to be returned with interest in 4 installments at 15% p.a. Find the value of the installments.

46. A loan of Rs. 60,000 is to be returned in 6 equal installments at 12% p.a. Find the amount of the installments.
47. Find the sum accumulated by paying an EMI of Rs. 11,800 for 2 years at 10% p.a.

48. Find the sum accumulated by paying an EMI of Rs. 1,800 for 2 years at 12% p.a.

49. Find the sum accumulated by paying an EMI of Rs. 12,000 for 3 years at 9% p.a.

50. Find the sum accumulated by paying an EMI of Rs. 11,000 for 8 years at 9.5% p.a.

Hints & Solutions to Check your progress

(1) 6694  (2) 3517  (3) 16,144  (4) 3820.32  
(5) 36555.65  (6) 13104  (7) 9380  (8) 24461

<table>
<thead>
<tr>
<th>Period</th>
<th>Sum Invested</th>
<th>Accumulated Amount</th>
</tr>
</thead>
<tbody>
<tr>
<td>Yearly</td>
<td>5000</td>
<td>5225</td>
</tr>
<tr>
<td>Half-yearly</td>
<td>10000</td>
<td>10695.5</td>
</tr>
<tr>
<td>Quarterly</td>
<td>20000</td>
<td>21648</td>
</tr>
<tr>
<td>Monthly</td>
<td>60000</td>
<td>65471</td>
</tr>
</tbody>
</table>

<table>
<thead>
<tr>
<th>Period</th>
<th>Sum Invested</th>
<th>Accumulated Amount</th>
</tr>
</thead>
<tbody>
<tr>
<td>Yearly</td>
<td>4500</td>
<td>4965</td>
</tr>
<tr>
<td>Half-yearly</td>
<td>9000</td>
<td>10203</td>
</tr>
<tr>
<td>Quarterly</td>
<td>18000</td>
<td>20693</td>
</tr>
<tr>
<td>Monthly</td>
<td>54000</td>
<td>62635</td>
</tr>
</tbody>
</table>

(11) 32910  (12) 5715  (13) 2498.72  (14) 225  
(15) 7%  (16) 6%  (17) 8%  (18) 75424, 15424  
(19) 19688, 1688  (20) 4719  (21) 26730  
(22) 67327.45  (23) 5877  (24) 59890.65  (25) 44378, Yes  
(26) 97093.4  (27) 393229.95  (28) 283021.25, take it on rent  
(29) 12523  (30) 17698.42  (31) 108143.28 < 126000, Mr. Mistry should use the second option.  
(32) 16245  (33) 18530  
(34) 1270569.47, not sufficient  (35) 7454.86  (36) 2.62,248; yes  
(37) 11,011

<table>
<thead>
<tr>
<th>Loan amount (in Rs.)</th>
<th>rate of interest (in % p.a.)</th>
<th>period of loan (in yrs.)</th>
<th>EMI (in Rs.)</th>
</tr>
</thead>
<tbody>
<tr>
<td>i) 1000</td>
<td>6</td>
<td>5</td>
<td>19</td>
</tr>
<tr>
<td>ii) 50000</td>
<td>6</td>
<td>10</td>
<td>555</td>
</tr>
<tr>
<td>iii) 8000</td>
<td>7</td>
<td>6</td>
<td>136</td>
</tr>
<tr>
<td>iv) 12000</td>
<td>9</td>
<td>10</td>
<td>152</td>
</tr>
<tr>
<td>Month</td>
<td>O/P</td>
<td>EMI</td>
<td>Interest Part</td>
</tr>
<tr>
<td>-------</td>
<td>-----</td>
<td>------</td>
<td>---------------</td>
</tr>
<tr>
<td>1</td>
<td>1000</td>
<td>2547.05</td>
<td>75</td>
</tr>
<tr>
<td>2</td>
<td>7527.95</td>
<td>2547.05</td>
<td>56.45</td>
</tr>
<tr>
<td>3</td>
<td>5037.35</td>
<td>2547.05</td>
<td>37.78</td>
</tr>
<tr>
<td>4</td>
<td>2528.08</td>
<td>2547.05</td>
<td>18.96</td>
</tr>
</tbody>
</table>

<table>
<thead>
<tr>
<th>Month</th>
<th>O/P</th>
<th>EMI</th>
<th>Interest Part</th>
<th>Principal Part</th>
</tr>
</thead>
<tbody>
<tr>
<td>1</td>
<td>30000</td>
<td>6212.23</td>
<td>351</td>
<td>5861.23</td>
</tr>
<tr>
<td>2</td>
<td>24138.77</td>
<td>6212.23</td>
<td>282.42</td>
<td>5929.81</td>
</tr>
<tr>
<td>3</td>
<td>18208.96</td>
<td>6212.23</td>
<td>213.04</td>
<td>5999.19</td>
</tr>
<tr>
<td>4</td>
<td>12209.77</td>
<td>6212.23</td>
<td>142.85</td>
<td>6069.38</td>
</tr>
<tr>
<td>5</td>
<td>6140.39</td>
<td>6212.23</td>
<td>71.84</td>
<td>6140.39</td>
</tr>
</tbody>
</table>

(39) 2326, 2700  (40) 6158.48  (41) 6397.11

(42) v) 1000 9.5 10 13  (38)  
vi) 110000 12.5 20 12498  

(45) 19339.57  (46) 16353  (47) 3,12,673.60  
(48) 48552.24  (49) 4,93,832.6  (50) 15,72,727
12

CORRELATION AND REGRESSION

OBJECTIVES

- To understand the relationship between two relevant characteristics of a statistical unit.
- Learn to obtain the numerical measure of the relationship between two variables.
- Use the mathematical relationship between two variables in order to estimate value of one variable from the other.
- Use the mathematical relationship to obtain the statistical constants line means and S.D.’s

12.1 INTRODUCTION

In the statistical analysis we come across the study of two or more relevant characteristics together in terms of their interrelations or interdependence. e.g. Interrelationship among production, sales and profits of a company. Inter relationship among rainfall, fertilizers, yield and profits to the farmers.

Relationship between price and demand of a commodity When we collect the information (data) on two of such characteristics it is called bivariate data. It is generally denoted by (X,Y) where X and Y are the variables representing the values on the characteristics.

Following are some examples of bivariate data.
   a) Income and Expenditure of workers.
   b) Marks of students in the two subjects of Maths and Accounts.
   c) Height of Husband and Wife in a couple.
   d) Sales and profits of a company.

   Between these variables we can note that there exist some sort of interrelationship or cause and effect relationship. i.e. change in the value of one variable brings out the change in the value of other variable also. Such relationship is called as correlation.

   Therefore, correlation analysis gives the idea about the nature and extent of relationship between two variables in the bivariate data.

12.2 TYPES OF CORRELATION:
There are two types of correlation.
   a) Positive correlation. and  b) Negative correlation.

12.2.1 Positive correlation: When the relationship between the variables $X$ and $Y$ is such that increase or decrease in $X$ brings out the increase or decrease in $Y$ also, i.e. there is direct relation between $X$ and $Y$, the correlation is said to be positive. In particular when the ‘change in $X$ equals to change in $Y$’ the correlation is perfect and positive. e.g. Sales and Profits have positive correlation.

12.2.2 Negative correlation: When the relationship between the variables $X$ and $Y$ is such that increase or decrease in $X$ brings out the decrease or increase in $Y$, i.e. there is an inverse relation between $X$ and $Y$, the correlation is said to be negative. In particular when the ‘change in $X$ equals to change in $Y$’ but in opposite direction the correlation is perfect and negative. e.g. Price and Demand have negative correlation.

12.3 MEASUREMENT OF CORRELATION

The extent of correlation can be measured by any of the following methods:
   • Scatter diagrams
   • Karl Pearson’s co-efficient of correlation
   • Spearman’s Rank correlation

12.3.1 Scatter Diagram: The Scatter diagram is a chart prepared by plotting the values of $X$ and $Y$ as the points $(X,Y)$ on the graph. The pattern of the points is used to explain the nature of correlation as follows. The following figures and the explanations would make it clearer.

(i) Perfect Positive Correlation:
   If the graph of the values of the variables is a straight line with positive slope as shown in Figure 4.1, we say there is a perfect positive correlation between $X$ and $Y$. Here $r = 1$.

(ii) Imperfect Positive Correlation:
   If the graph of the values of $X$ and $Y$ show a band of points from lower left corner to upper right corner as shown in Figure 4.2, we say that there is an imperfect positive correlation. Here $0 < r < 1$.

(iii) Perfect Negative Correlation:
   If the graph of the values of the variables is a straight line with negative slope as shown in Figure 4.3, we say there is a perfect negative correlation between $X$ and $Y$. Here $r = -1$.

(iv) Imperfect Negative Correlation:
   If the graph of the values of $X$ and $Y$ show a
band of points from upper left corner to the lower right corner as shown in Figure 4.4, then we say that there is an imperfect negative correlation. Here $-1 < r < 0$

(v) **Zero Correlation:**

If the graph of the values of $X$ and $Y$ do not show any of the above trend then we say that there is a zero correlation between $X$ and $Y$. The graph of such type can be a straight line perpendicular to the axis, as shown in Figure 4.5 and 4.6, or may be completely scattered as shown in Figure 4.7. Here $r = 0$.

The Figure 4.5 show that the increase in the values of $Y$ has no effect on the value of $X$, it remains the same, hence zero correlation. The Figure 4.6 show that the increase in the values of $X$ has no effect on the value of $Y$, it remains the same, hence zero correlation. The Figure 4.7 show that the points are completely scattered on the graph and show no particular trend, hence there is no correlation or zero correlation between $X$ and $Y$.

### 12.3.2 Karl Pearson’s co-efficient of correlation.

This co-efficient provides the numerical measure of the correlation between the variables $X$ and $Y$. It is suggested by Prof. Karl Pearson and calculated by the formula

$$r = \frac{\text{Cov}(x, y)}{\sigma_x \sigma_y}$$

Where, $\text{Cov}(x, y):$ Covariance between $x \& y$

$\sigma_x: \text{Standard deviation of } x \& \sigma_y: \text{Standard deviation of } y$

Also, $\text{Cov}(x, y) = \frac{1}{n} \sum(x - \bar{x})(y - \bar{y}) = \frac{1}{n} \sum xy - \bar{x} \bar{y}$

$$\text{S.D.}(x) = \sigma_x = \sqrt{\frac{1}{n} \sum(x - \bar{x})^2} = \sqrt{\frac{1}{n} \sum x^2 - \bar{x}^2}$$ and $$\text{S.D.}(y) = \sigma_y = \sqrt{\frac{1}{n} \sum(y - \bar{y})^2} = \sqrt{\frac{1}{n} \sum y^2 - \bar{y}^2}$$

Remark : We can also calculate this co-efficient by using the formula given by
\[
\rho = \frac{\frac{1}{n} \sum (x - \bar{x})(y - \bar{y})}{\sqrt{\frac{1}{n} \sum (x - \bar{x})^2 \sqrt{\frac{1}{n} \sum (y - \bar{y})^2}} = \frac{\frac{\sum xy}{n} - \bar{x}\bar{y}}{\left(\frac{\sum x^2}{n} - \bar{x}^2\right)^{\frac{1}{2}} \left(\frac{\sum y^2}{n} - \bar{y}^2\right)^{\frac{1}{2}}}
\]

The Pearson’s Correlation co-efficient is also called as the ‘product moment correlation co-efficient’

Properties of correlation co-efficient ‘r’

The value of ‘r’ can be positive (+) or negative(-)

The value of ‘r’ always lies between –1 & +1, i.e. –1 < r < +1

Significance of ‘r’ equals to –1, +1 & 0

When ‘r’= +1; the correlation is perfect and positive.

‘r’= -1; the correlation is perfect and negative.

and when there is no correlation ‘r’= 0

SOLVED EXAMPLES:

Example I: Calculate the Karl Pearson’s correlation coefficient from the following.

\[\begin{array}{c|c|c|c|c}
X & 12 & 10 & 20 & 13 \\
Y & 7 & 14 & 6 & 12 \\
\end{array}\]

Solution: Table of calculation,

\[
\begin{array}{c|c|c|c|c|c|c|c}
X & Y & XxY & X^2 & Y^2 \\
12 & 7 & 84 & 144 & 49 \\
10 & 14 & 140 & 100 & 196 \\
20 & 6 & 120 & 400 & 36 \\
13 & 12 & 156 & 169 & 144 \\
15 & 11 & 165 & 225 & 121 \\
\end{array}
\]

\[\Sigma x = 70 \quad \Sigma y = 50 \quad \Sigma x y = 665 \quad \Sigma x^2 = 1038 \quad \Sigma y^2 = 546 \quad \text{And n = 5} \]

The Pearson’s correlation coefficient r is given by,

\[r = \frac{\text{Cov}(x, y)}{\sigma_x \sigma_y}\]

Where,

\[\bar{x} = \frac{\Sigma x}{n} = \frac{70}{5} = 14 \quad \bar{y} = \frac{\Sigma y}{n} = \frac{50}{5} = 10\]

\[\text{Cov}(x, y) = \frac{\Sigma xy}{n} - \bar{x}\bar{y} \quad \sigma_x = \sqrt{\frac{1}{n} \sum x^2 - \bar{x}^2} \quad \sigma_y = \sqrt{\frac{1}{n} \sum y^2 - \bar{y}^2}\]

\[= \frac{665}{5} - 14 \times 10 = \sqrt{\frac{1038}{5} - 14^2} = \sqrt{\frac{546}{5} - 10^2}\]
\[ = 133 - 140 = \sqrt{11.6} = 3.40 \quad 9.2 = 3.03 = -07 \]

\[ \therefore \text{ Cov}(x,y) = -7 \quad \sigma_x = 3.40 \quad \text{and} \quad \sigma_y = 3.03 \]

Substituting the values in the formula of \( r \) we get

\[ r = \frac{-7}{3.40 \times 3.03} = -0.68 \]

\[ \therefore r = -0.68 \]

**Example 2:** Let us calculate co-efficient of correlation between Marks of students in the Subjects of Maths & Accounts. in a certain test conducted.

Table of calculation:

<table>
<thead>
<tr>
<th>Marks In Maths (X)</th>
<th>Marks In Accounts (Y)</th>
<th>XY</th>
<th>X^2</th>
<th>Y^2</th>
</tr>
</thead>
<tbody>
<tr>
<td>28</td>
<td>30</td>
<td>840</td>
<td>784</td>
<td>900</td>
</tr>
<tr>
<td>25</td>
<td>40</td>
<td>1000</td>
<td>625</td>
<td>1600</td>
</tr>
<tr>
<td>32</td>
<td>50</td>
<td>1600</td>
<td>1024</td>
<td>2500</td>
</tr>
<tr>
<td>16</td>
<td>18</td>
<td>288</td>
<td>256</td>
<td>324</td>
</tr>
<tr>
<td>20</td>
<td>25</td>
<td>500</td>
<td>400</td>
<td>625</td>
</tr>
<tr>
<td>15</td>
<td>12</td>
<td>180</td>
<td>225</td>
<td>144</td>
</tr>
<tr>
<td>19</td>
<td>11</td>
<td>209</td>
<td>361</td>
<td>121</td>
</tr>
<tr>
<td>17</td>
<td>21</td>
<td>357</td>
<td>289</td>
<td>441</td>
</tr>
<tr>
<td>40</td>
<td>45</td>
<td>1800</td>
<td>1600</td>
<td>2025</td>
</tr>
<tr>
<td>30</td>
<td>35</td>
<td>1050</td>
<td>900</td>
<td>1225</td>
</tr>
</tbody>
</table>

\[ \Sigma x = 242 \quad \Sigma y = 287 \quad \Sigma xy = 7824 \quad \Sigma x^2 = 6464 \quad \Sigma y^2 = 9905 \]

\[ n = 10 \]

Now Pearson’s co-efficient of correlation is given by the formula,

\[ r = \frac{\text{Cov}(x, y)}{\sigma_x \sigma_y} \]

Where,

\[ \bar{x} = \frac{\Sigma x}{n} = \frac{242}{10} = 24.2 \quad \bar{y} = \frac{\Sigma y}{n} = \frac{287}{10} = 28.7 \]

\[ \text{Cov}(x,y) = \frac{\Sigma xy}{n} - \bar{x} \bar{y} \quad \sigma_x = \sqrt{\frac{1}{n} \Sigma x^2 - \bar{x}^2} \quad \sigma_y = \sqrt{\frac{1}{n} \Sigma y^2 - \bar{y}^2} \]

\[ = \frac{7824}{10} - 24.2 \times 28.7 \quad \sigma_x = \sqrt{\frac{6464}{10} - 24.2^2} \quad \sigma_y = \sqrt{\frac{9905}{10} - 28.7^2} \]

\[ = 782.4 - 694.54 \quad = \sqrt{60.76} \quad = \sqrt{166.81} \]
\[ \text{Cov}(x,y) = 87.86, \quad \sigma_x = 7.79 \quad \text{and} \quad \sigma_y = 12.91 \]

\[ \therefore \text{Cov}(x,y) = 87.86, \quad \sigma_x = 7.79 \quad \text{and} \quad \sigma_y = 12.91 \]

Substituting the values in the formula of \( r \) we get

\[ r = \frac{87.86}{7.79 \times 12.91} = 0.87 \]

\[ \therefore r = 0.87 \]

### 12.3 RANK CORRELATION

In many practical situations, we do not have the scores on the characteristics, but the ranks (preference order) decided by two or more observers. Suppose, a singing competition of 10 participants is judged by two judges A and B who rank or assign scores to the participants on the basis of their performance. Then it is quite possible that the ranks or scores assigned may not be equal for all the participants. Now the difference in the ranks or scores assigned indicates that there is a difference of opinion between the judges on deciding the ranks. The rank correlation studies the association in this ranking of the observations by two or more observers. The measure of the extent of association in rank allocation by the two judges is calculated by the co-efficient of Rank correlation ‘\( R \)’. This co-efficient was developed by the British psychologist Edward Spearman in 1904.

Mathematically, Spearman’s rank correlation co-efficient is defined as,

\[ R = 1 - \frac{6 \sum d^2}{n(n^2 - 1)} \]

Where \( d \) = rank difference and \( n \) = no of pairs.

**Remarks:** We can note that, the value of ‘\( R \)’ always lies between –1 and +1.

The positive value of ‘\( R \)’ indicates the positive correlation (association) in the rank allocation. Whereas, the negative value of ‘\( R \)’ indicates the negative correlation (association) in the rank allocation.

### SOLVED EXAMPLES:

**Example 3**

a) **When ranks are given:**

Data given below read the ranks assigned by two judges to 8 participants. Calculate the co-efficient of Rank correlation.
<table>
<thead>
<tr>
<th>Participant No.</th>
<th>Ranks by Judge</th>
<th>Rank diff $d^2$</th>
</tr>
</thead>
<tbody>
<tr>
<td></td>
<td>A</td>
<td>B</td>
</tr>
<tr>
<td>1</td>
<td>5</td>
<td>4</td>
</tr>
<tr>
<td>2</td>
<td>6</td>
<td>8</td>
</tr>
<tr>
<td>3</td>
<td>7</td>
<td>1</td>
</tr>
<tr>
<td>4</td>
<td>1</td>
<td>7</td>
</tr>
<tr>
<td>5</td>
<td>8</td>
<td>5</td>
</tr>
<tr>
<td>6</td>
<td>2</td>
<td>6</td>
</tr>
<tr>
<td>7</td>
<td>3</td>
<td>2</td>
</tr>
<tr>
<td>8</td>
<td>4</td>
<td>3</td>
</tr>
<tr>
<td><strong>N = 8</strong></td>
<td></td>
<td></td>
</tr>
</tbody>
</table>

Spearman’s rank correlation co-efficient is given by

$$R=1-\frac{6\sum d^2}{n(n^2-1)}$$

Substituting the values from the table we get,

$$R= 1-\frac{6\times 104}{8(8^2-1)} = -0.23$$

The value of correlation co-efficient is - 0.23. This indicates that there is negative association in rank allocation by the two judges A and B.

b) **When scores are given:**

**Example**

<table>
<thead>
<tr>
<th>Student No</th>
<th>Marka by Examiner</th>
<th>Ranks</th>
<th>Rank difference square</th>
</tr>
</thead>
</table>

The data given below are the marks given by two Examiners to a set of 10 students in a aptitude test. Calculate the Spearman’s Rank correlation co-efficient, ‘R’
Now the Spearman’s rank correlation co-efficient is given by

\[ R = 1 - \frac{6 \sum d^2}{n(n^2 - 1)} \]

Substituting the values from the table we get,

\[ R = 1 - \frac{6 \times 5}{10(10^2 - 1)} \]

\[ = 1 - 0.03 \]
\[ = 0.97 \]

The value of correlation co-efficient is +0.97. This indicates that there is positive association in assessment of two examiners, A and B.

c) **Case of repeated values:**

It is quite possible that the two participants may be assigned the same score by the judges. In such cases Rank allocation and calculation of rank correlation can be explained as follows.

**Example:** The data given below scores assigned by two judges for 10 participants in the singing competition. Calculate the Spearman’s Rank correlation co-efficient.

<table>
<thead>
<tr>
<th>Participant No</th>
<th>Score assigned By Judges</th>
<th>Ranks</th>
<th>Rank difference square</th>
</tr>
</thead>
<tbody>
<tr>
<td>A</td>
<td>B</td>
<td></td>
<td></td>
</tr>
<tr>
<td>1</td>
<td>85</td>
<td>2</td>
<td>2</td>
</tr>
<tr>
<td>2</td>
<td>56</td>
<td>10</td>
<td>7</td>
</tr>
<tr>
<td>3</td>
<td>45</td>
<td>10</td>
<td>0</td>
</tr>
<tr>
<td>4</td>
<td>65</td>
<td>6</td>
<td>0</td>
</tr>
<tr>
<td>5</td>
<td>96</td>
<td>1</td>
<td>0</td>
</tr>
<tr>
<td>6</td>
<td>52</td>
<td>9</td>
<td>1</td>
</tr>
<tr>
<td>7</td>
<td>80</td>
<td>3</td>
<td>1</td>
</tr>
<tr>
<td>8</td>
<td>75</td>
<td>5</td>
<td>0</td>
</tr>
<tr>
<td>9</td>
<td>78</td>
<td>4</td>
<td>1</td>
</tr>
<tr>
<td>10</td>
<td>60</td>
<td>7</td>
<td>1</td>
</tr>
<tr>
<td>N = 10</td>
<td></td>
<td>Total</td>
<td>5 = ( \sum d^2 )</td>
</tr>
</tbody>
</table>
Explanation:- In the column of A and B there is repetition of scores so while assigning the ranks we first assign the ranks by treating them as different values and then for reated scores we assign the average rank. 

In col A the score 35 appears 2 times at number 4 and 5 in the order of ranking so we calculate the average rank as \((4+5)/2 = 4.5\).

Hence the ranks assigned are 4.5 each. The other repeated scores can be ranked in the same manner.

Note: In this example we can note that the ranks are in fraction e.g. 4.5, which is logically incorrect or meaningless. Therefore in the calculation of ‘R’ we add a correction factor (C.F.) to \(\Sigma d^2\) calculated as follows.

Table of correction factor (C.F.)

<table>
<thead>
<tr>
<th>Value Repeated</th>
<th>Frequency</th>
<th>(m(m^2-1))</th>
</tr>
</thead>
<tbody>
<tr>
<td>35</td>
<td>2</td>
<td>2x(2^2-1) = 6</td>
</tr>
<tr>
<td>28</td>
<td>2</td>
<td>6</td>
</tr>
<tr>
<td>26</td>
<td>2</td>
<td>6</td>
</tr>
<tr>
<td>Total</td>
<td></td>
<td>(\Sigma m(m^2-1) = 18)</td>
</tr>
</tbody>
</table>

Now \(C.F. = \frac{\Sigma(m^3 - m)}{12} = \frac{18}{12} = 1.5\)

\(\therefore \Sigma d^2 = 117.5 + 1.5 = 119\)

We use this value in the calculation of ‘R’

Now the Spearman’s rank correlation co-efficient is given by

\[
R = 1 - \frac{6\Sigma d^2}{n(n^2 - 1)}
\]

Substituting the values we get, \(R = 1 - \frac{6 \times 119}{10(10^2 - 1)} = 1 - 0.72 = 0.28\)
EXERCISE I

1. What is mean by correlation? Explain the types of correlation with suitable examples.

2. What is a scatter diagram? Draw different scattered diagrams to explain the correlation between two variables x and y.

3. State the significance of ‘r’ = +1, –1 and 0.

4. Calculate the coefficient of correlation r from the following data.
   \[ X: \quad 18 \quad 12 \quad 16 \quad 14 \quad 10 \quad 15 \quad 17 \quad 13 \]
   \[ Y: \quad 9 \quad 13 \quad 20 \quad 25 \quad 15 \quad 14 \quad 17 \quad 26 \quad 22 \]

5. The following table gives the price and demand of a certain commodity over the period of 8 months. Calculate the Pearson’s coefficient of correlation.
<table>
<thead>
<tr>
<th>Price</th>
<th>Demand</th>
</tr>
</thead>
<tbody>
<tr>
<td>15</td>
<td>45</td>
</tr>
<tr>
<td>12</td>
<td>30</td>
</tr>
<tr>
<td>23</td>
<td>60</td>
</tr>
<tr>
<td>25</td>
<td>65</td>
</tr>
<tr>
<td>18</td>
<td>48</td>
</tr>
<tr>
<td>17</td>
<td>45</td>
</tr>
<tr>
<td>11</td>
<td>28</td>
</tr>
<tr>
<td>19</td>
<td>50</td>
</tr>
</tbody>
</table>

6. Following results are obtained on a certain bivariate data.
   (i) \( n = 10 \quad \Sigma x = 75 \quad \Sigma y = 70 \quad \Sigma x^2 = 480 \)
   \( \Sigma y^2 = 600 \quad \Sigma xy = 540 \)
   (ii) \( n = 15 \quad \Sigma x = 60 \quad \Sigma y = 85 \quad \Sigma x^2 = 520 \)
   \( \Sigma y^2 = 1200 \quad \Sigma xy = 340 \)
   Calculate the Pearson’s correlation coefficient in each case.

7. Following data are available on a certain bi-variate data :
   (i) \( \Sigma (x-\bar{x}) (y-\bar{y}) = 120, \Sigma (x-\bar{x})^2 = 150 \quad \Sigma (y-\bar{y})^2 = 145 \)
   (ii) \( \Sigma (x-\bar{x}) (y-\bar{y}) = -122, \Sigma (x-\bar{x})^2 = 136 \quad \Sigma (y-\bar{y})^2 = 148 \)
   Find the correlation coefficient.

8. Calculate the Pearson’s coefficient of correlation from the given information on a bivariate series:
   No of pairs: 25
   Sum of x values:300
   Sum of y values:375
   Sum of squares of x values: 9000
   Sum of squares of y values:6500
   Sum of the product of x and y values:4000.

9. The ranks assigned to 8 participants by two judges are as follows. Calculate the Spearman’s Rank correlation coefficient ‘R’.
<table>
<thead>
<tr>
<th>Participant No:</th>
<th>1</th>
<th>2</th>
<th>3</th>
<th>4</th>
<th>5</th>
<th>6</th>
<th>7</th>
<th>8</th>
</tr>
</thead>
<tbody>
<tr>
<td>Ranks by JudgeI:</td>
<td>5</td>
<td>3</td>
<td>4</td>
<td>6</td>
<td>1</td>
<td>8</td>
<td>7</td>
<td>2</td>
</tr>
<tr>
<td>JudgeII :</td>
<td>6</td>
<td>8</td>
<td>3</td>
<td>7</td>
<td>1</td>
<td>5</td>
<td>4</td>
<td>2</td>
</tr>
</tbody>
</table>
10. Calculate the coefficient of rank correlation from the data given below.

\[ \begin{array}{cccccccc}
X: & 40 & 33 & 60 & 59 & 50 & 55 & 48 \\
Y: & 70 & 60 & 85 & 75 & 72 & 82 & 69 \\
\end{array} \]

11. Marks given by two Judges to a group of 10 participants are as follows. Calculate the coefficient of rank correlation.

Marks by Judge
A: 52 53 42 60 45 41 37 38 25 27
Judge B: 65 68 43 38 77 48 35 30 25 50.

12. An examination of 8 applicants for a clerical post was by a bank. The marks obtained by the applicants in the subjects of Mathematics and Accountancy were as follows. Calculate the rank correlation coefficient.

Applicant: A B C D E F G H
Marks in Maths: 15 20 28 12 40 60 20 80
Marks in Accounts: 40 30 50 30 20 10 25 60

### 12.4 REGRESSION ANALYSIS

As the correlation analysis studies the nature and extent of interrelationship between the two variables X and Y, regression analysis helps us to estimate or approximate the value of one variable when we know the value of other variable. Therefore we can define the ‘Regression’ as the estimation (prediction) of one variable from the other variable when they are correlated to each other. e.g. We can estimate the Demand of the commodity if we know it’s Price.

Why are there two regressions?
When the variables X and Y are correlated there are two possibilities,

(i) Variable X depends on variable y. in this case we can find the value of x if know the value of y. This is called regression of x on y.

(ii) Variable y depends on variable X. we can find the value of y if know the value of X. This is called regression of y on x. Hence there are two regressions,

(a) Regression of X on Y; (b) Regression of X on Y.

#### 8.4.1 Formulas on Regression equation,

<table>
<thead>
<tr>
<th>Regression of X on Y</th>
<th>Regression of X on Y</th>
</tr>
</thead>
<tbody>
<tr>
<td>Assumption: X depends on Y</td>
<td>Y depends on X</td>
</tr>
<tr>
<td>The regression equation is</td>
<td>The regression equation is</td>
</tr>
<tr>
<td>((x-\bar{x}) = b_{xy}(y-\bar{y}))</td>
<td>((y-\bar{y}) = b_{yx}(x-\bar{x}))</td>
</tr>
</tbody>
</table>
Maths_stats_fybcom

\[ b_{xy} = \text{Regression co-efficient of } X \text{ on } Y = \frac{\text{Cov}(x, y)}{V(y)} \]

\[ b_{yx} = \text{Regression co-efficient of } Y \text{ on } X = \frac{\text{Cov}(x, y)}{V(x)} \]

Where,

\[ \text{Cov}(x, y) = \frac{1}{n} \sum (x - \bar{x})(y - \bar{y}) = \frac{1}{n} \sum xy - \bar{x}\bar{y} \]

\[ V(x) = \frac{1}{n} \sum (x - \bar{x})^2 \quad \text{and} \quad V(y) = \frac{1}{n} \sum (y - \bar{y})^2 \]

\[ V(x) = \frac{1}{n} \sum x^2 - \bar{x}^2 \quad \text{and} \quad V(y) = \frac{1}{n} \sum y^2 - \bar{y}^2 \]

Use: To find \( x \)

Use: To find \( y \)

---

**SOLVED EXAMPLES**

**Example 1:**
Obtain the two regression equations and hence find the value of \( x \) when \( y = 25 \)

Data:-

<table>
<thead>
<tr>
<th>( x )</th>
<th>( y )</th>
<th>( x^2 )</th>
<th>( y^2 )</th>
<th>( xxY )</th>
</tr>
</thead>
<tbody>
<tr>
<td>8</td>
<td>15</td>
<td>64</td>
<td>225</td>
<td>120</td>
</tr>
<tr>
<td>10</td>
<td>20</td>
<td>100</td>
<td>400</td>
<td>200</td>
</tr>
<tr>
<td>12</td>
<td>30</td>
<td>144</td>
<td>900</td>
<td>360</td>
</tr>
<tr>
<td>15</td>
<td>40</td>
<td>225</td>
<td>1600</td>
<td>600</td>
</tr>
<tr>
<td>20</td>
<td>45</td>
<td>400</td>
<td>2025</td>
<td>900</td>
</tr>
<tr>
<td>( \Sigma x = 65 )</td>
<td>( \Sigma y = 150 )</td>
<td>( \Sigma x^2 = 933 )</td>
<td>( \Sigma y^2 = 5150 )</td>
<td>( \Sigma xy = 2180 )</td>
</tr>
</tbody>
</table>

And \( n = 5 \)

Now the two regression equations are,

\((x - \bar{x}) = b_{xy}(y - \bar{y}) \quad \text{-------} \quad \text{on } y \) (i)

\((y - \bar{y}) = b_{xy}(x - \bar{x}) \quad \text{-------} \quad \text{on } x \) (ii)

Where,

\[ \bar{x} = \frac{1}{n} \sum x = \frac{65}{5} = 13 \quad \text{and} \quad \bar{y} = \frac{1}{n} \sum y = \frac{150}{5} = 30 \]

Also,

\[ \text{Cov}(x, y) = \frac{1}{n} \sum xy - \bar{x}\bar{y} \quad \text{V}(x) = \frac{1}{n} \sum x^2 - \bar{x}^2 \quad \text{V}(y) = \frac{1}{n} \sum y^2 - \bar{y}^2 \]
\[
\begin{align*}
= \frac{2180}{5} - 13 \times 30 & \quad = \frac{933}{5} - 13^2 & \quad = \frac{5150}{5} - 30^2 \\
= 436 - 390 & \quad = 186.6 - 169 & \quad = 1030 - 900
\end{align*}
\]

\[
\therefore \text{Cov}(x, y) = 46 \quad V(x) = 17.6 \quad V(y) = 130
\]

Now we find,

\[
\begin{align*}
\text{Regression co-efficient of X on Y} & = \frac{\text{Cov}(x, y)}{V(y)} \\
\text{Regression co-efficient of X on Y} & = \frac{\text{Cov}(x, y)}{V(x)} \\
\therefore b_{xy} & = \frac{46}{130} \\
\therefore b_{yx} & = \frac{46}{17.6}
\end{align*}
\]

Now substituting the values of \(\bar{x}, \bar{y}, \) \(b_{xy}\) and \(b_{yx}\) in the regression equations we get,

\[
\begin{align*}
(x-13) = 0.35(y-30) & \quad \text{-----} x \text{ on } y \quad (i) \\
(y-30) = 2.61(x-13) & \quad \text{-----} y \text{ on } x \quad (ii)
\end{align*}
\]

as the two regression equations.

Now to estimate \(x\) when \(y = 25\), we use the regression equation of \(x\) on \(y\)

\[
\therefore (x-13) = 0.35(25-30) \\
\therefore x = 13 - 1.75 = 11.25
\]

**Remark:**

From the above example we can note some points about Regression coefficients.

- Both the regression coefficients carry the same sign (+ or -)
- Both the regression coefficients can not be greater than 1 in number
  (e.g. -1.25 and -1.32) is not possible.
- Product of both the regression coefficients \(b_{xy}\) and \(b_{yx}\) must be < 1
  i.e. \(b_{xy} \times b_{yx} < 1\)
  Here \(0.35 \times 2.61 = 0.91 < 1\) \((\text{Check this always})\)

**Example 2:**

Obtain the two regression equations and hence find the value of \(y\) when \(x = 10\)

<table>
<thead>
<tr>
<th>Data:</th>
</tr>
</thead>
<tbody>
<tr>
<td>(X)</td>
</tr>
<tr>
<td>-------</td>
</tr>
<tr>
<td>12</td>
</tr>
<tr>
<td>20</td>
</tr>
<tr>
<td>8</td>
</tr>
<tr>
<td>14</td>
</tr>
<tr>
<td>16</td>
</tr>
</tbody>
</table>

| \(\Sigma x = 70\) | \(\Sigma y = 88\) | \(\Sigma xy = 1218\) | \(\Sigma x^2 = 1060\) | \(\Sigma y^2 = 1632\) |

And \(n = 5\)

Now the two regression equations are,
\[
\begin{align*}
(x - \bar{x}) &= b_{xy}(y - \bar{y}) \quad \text{----- on } y (i) \\
(y - \bar{y}) &= b_{yx} (x - \bar{x}) \quad \text{----- on } x (ii)
\end{align*}
\]

Where,
\[
\bar{x} = \frac{1}{n} \sum x = \frac{70}{5} = 14 \quad \text{and} \quad \bar{y} = \frac{1}{n} \sum y = \frac{88}{5} = 17.6
\]

Also,
\[
\text{Cov}(x,y) = \frac{1}{n} \sum xy - \bar{x} \bar{y} \quad V(x) = \frac{1}{n} \sum x^2 - \bar{x}^2 \quad V(y) = \frac{1}{n} \sum y^2 - \bar{y}^2
\]
\[
= \frac{1218}{5} - 14 \times 17.6 \\
= 243.6 - 246.4
\]
\[
\therefore \text{Cov}(x,y) = -2.8 \\
V(x) = 16 \\
V(y) = 16.64
\]

Now we find,
\[
\text{Regression co-efficient of } X \text{ on } Y \quad \text{Regression co-efficient of } X \text{ on } Y
\]
\[
b_{xy} = \frac{\text{Cov}(x,y)}{V(y)} = \frac{2.8}{16.64} \\
\therefore b_{xy} = -0.168
\]

Now substituting the values of \( \bar{x}, \bar{y}, b_{xy} \) and \( b_{yx} \) in the regression equations we get,
\[
(x - 14) = -0.168(y - 17.6) \quad \text{----- on } y (i) \\
(y - 17.6) = -0.175(x - 14) \quad \text{----- on } x (ii)
\]
as the two regression equations.

Now to estimate y when x = 10, we use the regression equation of y on x
\[
\therefore (y - 17.6) = -0.175(10 - 14) \\
\therefore y = 17.6 + 0.7 = 24.3
\]

**Example 3:**
The following data give the experience of machine operators and their performance rating given by the number of good parts turned out per 100 pieces.

<table>
<thead>
<tr>
<th>Operator</th>
<th>1</th>
<th>2</th>
<th>3</th>
<th>4</th>
<th>5</th>
<th>6</th>
<th>7</th>
<th>8</th>
</tr>
</thead>
<tbody>
<tr>
<td>Experience:</td>
<td>16</td>
<td>12</td>
<td>18</td>
<td>4</td>
<td>3</td>
<td>10</td>
<td>5</td>
<td>12</td>
</tr>
<tr>
<td>Performance:</td>
<td>87</td>
<td>88</td>
<td>89</td>
<td>68</td>
<td>78</td>
<td>80</td>
<td>75</td>
<td>83</td>
</tr>
</tbody>
</table>

Obtain the two regression equations and estimate the permanence rating of an operator who has put 15 years in service.

**Solution:** We define the variables,
\[
X: \text{Experience} \quad y: \text{Performance rating}
\]

Table of calculations:
Now the two regression equations are,

\[(x-\bar{x}) = b_{xy}(y - \bar{y}) \]  \---- x on y (i)

\[(y - \bar{y}) = b_{yx}(x - \bar{x}) \]  \---- y on x (ii)

Where,

\[\bar{x} = \frac{1}{n} \sum x = \frac{80}{8} = 10 \] and \[\bar{y} = \frac{1}{n} \sum y = \frac{648}{8} = 81 \]

Also,

\[\text{Cov}(x,y) = \frac{1}{n} \sum xy - \bar{x} \bar{y} = \frac{6727}{8} - 10 \times 81 = 840.75 - 810 = 30.75 \]  \Rightarrow \text{Cov}(x,y) = 30.75

\[\text{V}(x) = \frac{1}{n} \sum x^2 - \bar{x}^2 = \frac{1018}{8} - 10^2 = \frac{52856}{8} - 81^2 = 6607 - 6561 = 46 \]

\[\text{V}(y) = \frac{1}{n} \sum y^2 - \bar{y}^2 = \frac{52856}{8} = 6607 \]

Now we find,

Regression co-efficient of X on Y  Regression co-efficient of X on Y

\[b_{xy} = \frac{\text{Cov}(x,y)}{\text{V}(y)} = \frac{30.75}{46} \]  \Rightarrow \text{b}_{xy} = 0.67

\[b_{yx} = \frac{\text{Cov}(x,y)}{\text{V}(x)} = \frac{30.75}{27.25} \]  \Rightarrow \text{b}_{yx} = 1.13

Now substituting the values of \[\bar{x}, \bar{y}, \text{b}_{xy} \text{ and } \text{b}_{yx} \] in the regression equations we get,

\[(x-10) = 0.67(y-81) \] \---- x on y (i)

\[(y-81) = 1.13(x-10) \] \---- y on x (ii)

as the two regression equations.

Now to estimate Performance rating \(y\) when Experience \(x\) = 15, we use the regression equation of \(y\) on \(x\)

\[\therefore (y-81) = 1.13(15-10) \]

\[\therefore y = 81 + 5.65 = 86.65 \]
Hence the estimated performance rating for the operator with 15 years of experience is approximately 86.65 i.e approximately 87

12.4.2 Regression coefficients in terms of correlation coefficient.

We can also obtain the regression coefficients \( b_{xy} \) and \( b_{yx} \) from standard deviations, \( \sigma_x \), \( \sigma_y \) and correlation coefficient ‘r’ using the formulas

\[
   b_{xy} = r \frac{\sigma_x}{\sigma_y} \quad \text{and} \quad b_{yx} = r \frac{\sigma_y}{\sigma_x}
\]

Also consider,

\[
   b_{xy} \times b_{yx} = r \frac{\sigma_x}{\sigma_y} \cdot r \frac{\sigma_y}{\sigma_x} = r^2 \quad \text{i.e.} \quad r = \sqrt{b_{xy} \times b_{yx}}
\]

Hence the correlation coefficient ‘r’ is the geometric mean of the regression coefficients, \( b_{xy} \) and \( b_{yx} \)

Example 5:
You are given the information about advertising expenditure and sales:

<table>
<thead>
<tr>
<th>Exp. on Advertisement (Rs. In Lakh)</th>
<th>Sales (Rs. In Lakh)</th>
</tr>
</thead>
<tbody>
<tr>
<td>Mean</td>
<td>10</td>
</tr>
<tr>
<td>S.D.</td>
<td>3</td>
</tr>
<tr>
<td>-------------------------------------</td>
<td>---------------------</td>
</tr>
</tbody>
</table>

Coefficient of correlation between sales and expenditure on Advertisement is 0.8. Obtain the two regression equations. Find the likely sales when advertisement budget is Rs. 15 Lakh.

Solution: We define the variables,

X: Expenditure on advertisement
Y: Sales achieved.

Therefore we have,

\( \bar{x} = 10, \quad \bar{y} = 90, \quad 6x = 3, \quad 6y = 12 \quad \text{and} \quad r = 0.8 \)

Now, using the above results we can write the two regression equations as

\[
   (x - \bar{x}) = r \frac{\sigma_x}{\sigma_y} (y - \bar{y}) \quad \text{-------} \text{on} \ y \ (i) \\
   (y - \bar{y}) = r \frac{\sigma_y}{\sigma_x} (x - \bar{x}) \quad \text{-------} \text{on} \ x \ (ii)
\]

Substituting the values in the equations we get,

\[
   (x-10) = 0.8 \frac{3}{12} (y-90)
\]

i.e \( x - 10 = 0.2 (y-90) \quad \text{-------} \text{on} \ y \ (i) \)
also \( (y-90) = 0.8 \frac{12}{3} (x-10) \)

i.e. \( y-90 = 3.2 (x-10) \) \( \quad \text{--------y on x (ii)} \)

Now when expenditure on advertisement \((x)\) is 15, we can find the sales from eqn (ii) as,

\[
y-90 = 3.2 (15-10)
\]

\[
\cdot \cdot \cdot \quad y = 90 + 16 = 106
\]

Thus the likely sales are Rs.106 Lakh.

Example 6: Compute the two regression equations on the basis of the following information:

<table>
<thead>
<tr>
<th>X</th>
<th>Y</th>
</tr>
</thead>
<tbody>
<tr>
<td>Mean</td>
<td>40</td>
</tr>
<tr>
<td>Standard deviation</td>
<td>10</td>
</tr>
</tbody>
</table>

Karl Pearson’s coefficient of correlation between \(x\) and \(y\) = 0.50.

Also estimate the value of \(x\) when \(y = 48\) using the appropriate equation.

**Solution:** We have,

\[
\bar{x} = 40, \quad \bar{y}=45, \quad \sigma x= 10, \sigma y = 9 \quad \text{and} \quad r = 0.5
\]

Now, we can write the two regression equations as

\[
(x-\bar{x}) = r \frac{\sigma x}{\sigma y} (y-\bar{y}) \quad \text{--------x on y (i)}
\]

\[
(y-\bar{y}) = r \frac{\sigma y}{\sigma x} (x-\bar{x}) \quad \text{--------y on x (ii)}
\]

Substituting the values in the equations we get,

\[
(x-40) = 0.5 \times \frac{10}{9} (y-45)
\]

i.e \( x - 40 = 0.55 (y-45) \) \( \quad \text{--------eqn of x on y (i)} \)

and \( (y-45) = 0.5 \times \frac{9}{10} (x-40) \)

i.e. \( y-45 = 0.45(x-40) \) \( \quad \text{--------eqn of y on x (ii)} \)

Now when \(y\) is 48, we can find \(x\) from eqn (i) as,

\[
x-40 = 0.55(48-45)
\]

\[
\cdot \cdot \cdot \quad x = 40 +1.65 = 41.65
\]

**Example 7:**

Find the marks of a student in the Subject of Mathematics who have scored 65 marks in Accountancy Given,

<table>
<thead>
<tr>
<th>Average marks in Mathematics</th>
<th>70</th>
</tr>
</thead>
<tbody>
<tr>
<td>Accountancy</td>
<td>80</td>
</tr>
<tr>
<td>Standard Deviation of marks in Mathematics</td>
<td>8</td>
</tr>
<tr>
<td>in Accountancy</td>
<td>10</td>
</tr>
</tbody>
</table>
Coefficient of correlation between the marks of Mathematics and marks of Accountancy is 0.64.

**Solution:** We define the variables,
- **X:** Marks in Mathematics
- **Y:** Marks in Accountancy

Therefore we have,
- $\bar{x} = 70$, $\bar{y} = 80$, $\sigma_x = 8$, $\sigma_y = 10$ and $r = 0.64$

Now we want to approximate the marks in Mathematics ($x$), we obtain the regression equation of $x$ on $y$, which is given by

$$ (x - \bar{x}) = r \frac{\sigma_x}{\sigma_y} (y - \bar{y}) \quad -------\text{on y (i)} $$

Substituting the values we get,

$$ (x-70) = 0.64 \frac{8}{10} (y-80) $$

i.e

$$ x - 70 = 0.57 (y-80) $$

Therefore, when marks in Accountancy ($Y$) = 65

$$ x - 70 = 0.57(65-80) $$

$$ \therefore x = 70 - 2.85 = 67.15 \quad \text{i.e. 67 appr.} $$

**Use of regression equations to find means $\bar{x}$, $\bar{y}$ S.D.s $\sigma_x$, $\sigma_y$ and correlation coefficient ‘r’**

As we have that, we can obtain the regression equations from the values of Means, standard deviations and correlation coefficients ‘r’, we can get back these values from the regression equations.

Now, we can note that the regression equation is a linear equation in two variables $x$ and $y$. Therefore, the linear equation of the type $Ax + By + C = 0$ or $y = a+bx$ represents a regression equation.

E.g. $3x + 5y - 15 = 0$ and $2x + 7y + 10 = 0$ represent the two regression equations.

The values of means $\bar{x}$, $\bar{y}$ can be obtain by solving the two equations as the simultaneous equations.

**Example 8:**

From the following regression equation, find means $\bar{x}$, $\bar{y}$, $\sigma_x$, $\sigma_y$ and ‘r’

$$ 3x - 2y - 10 = 0, \quad 24x - 25y + 145 = 0 $$

**Solution:** The two regression equations are,

$$ 3x - 2y - 10 = 0 \quad -------(i) $$

$$ 24x - 25y + 145 = 0 \quad ---(ii) $$

Now for $\bar{x}$ and $\bar{y}$ we solve the two equations as the simultaneous equations.

Therefore, by (i) $x \times 8$ and (ii) $x1$, we get

$$ 24x - 16y - 80 = 0 $$

$$ 24x - 25y + 145 = 0 $$

$$ \begin{array}{c}
9y - 225 = 0 \\
9 \quad \text{y} = \frac{225}{9} = 25
\end{array} $$

Putting $y = 25$ in eqn (i), we get
3x - 2(25) - 10 = 0
3x - 60 = 0
x = \frac{60}{3} = 20

Hence \( x = 20 \) and \( y = 25 \).

Now to find ‘r’ we express the equations in the form \( y = a + bx \)
So, from eqns (i) and (ii)

\[
\begin{align*}
\frac{3x}{2} - \frac{10}{2} &= \text{ and } \frac{24x}{25} + \frac{145}{25} \\
\therefore b_1 &= \frac{3}{2} = 1.5 \\
\therefore b_2 &= \frac{24}{25} = 0.96 \\
\text{Since, } b_1 > b_2 (\text{i.e. } b_2 \text{ is smaller in number irrespective of sign + or -})
\end{align*}
\]

\therefore \text{ Equation (ii) is regression of } y \text{ on } x \text{ and } b_{yx} = 0.96

Hence eqn (i) is regression of x on y and \( b_{xy} = \frac{1}{1.5} = 0.67 \)

Now we find,

\[
r = \sqrt{b_{xy} \times b_{yx}} \text{ i.e. } r = \sqrt{0.67 \times 0.96} = + 0.84
\]

(The sign of ‘r’ is same as the sign of regression coefficients)

**Example 9:**
Find the means values of \( x, y \), and \( r \) from the two regression equations.

3x + 2y - 26 = 0 and \( \sigma x + y - 31 = 0 \). Also find \( \sigma_x \) when \( \sigma_y = 3 \).

**Solution:** The two regression equations are,

\[
\begin{align*}
3x + 2y - 26 &= 0 \quad \text{(i)} \\
6x + y - 31 &= 0 \quad \text{(ii)}
\end{align*}
\]

Now for \( x \) and \( y \) we solve the two equations as the simultaneous equations.

Therefore, by (i) \( x \times 2 \) and (ii) \( x \times 1 \), we get

\[
\begin{align*}
6x + 4y - 52 &= 0 \\
6x + y - 31 &= 0
\end{align*}
\]

\[
\frac{- - +}{- - +} \quad \therefore y = \frac{21}{3} = 7
\]

Putting \( y = 7 \) in eqn (i), we get

\[
\begin{align*}
3x + 2(7) - 26 &= 0 \\
3x - 12 &= 0
\end{align*}
\]

\[
x = \frac{12}{3} = 4.
\]

Hence \( x = 4 \) and \( y = 7 \).

Now to find ‘r’ we express the equations in the form \( y = a + bx \)
So, from eqns (i) and (ii)

\[
\begin{align*}
\frac{3x}{2} - \frac{26}{2} &= \text{ and } \frac{-6x}{1} + \frac{31}{1} \\
\therefore b_1 &= -\frac{3}{2} = -1.5 \\
\therefore b_2 &= \frac{-6}{1} = -6
\end{align*}
\]

since, \( b_1 < b_2 \) (i.e. \( b_1 \) is smaller in number irrespective of sign + or -)

\therefore \text{ Equation (i) is regression of } y \text{ on } x \text{ and } b_{yx} = -1.5

Hence, eqn (ii) is regression of \( x \) on \( y \) and \( b_{xy} = -1/6 = -0.16 \)

Now we find,

\[
r = \sqrt{b_{xy} \times b_{yx}} \quad r = \sqrt{0.16 \times 1.5} = -0.16
\]
Note: The sign of ‘r’ is same as the sign of regression coefficients

Now to find 6x when 6y = 3, we use the formula,

\[ b_{yx} = r \frac{\sigma_x}{\sigma_y} \]

\[ -1.5 = -\frac{0.16 \times 3}{6x} \]

\[ \therefore 6x = \frac{0.48}{1.5} = 0.32 \]

Hence means \( \bar{x} = 4 \), \( \bar{y} = 7 \), \( r = -0.16 \) and \( 6x = 0.32 \).

EXERCISES

1. What is mean by Regression? Explain the use of regression in the statistical analysis.

2. Why are there two Regressions? Justify.

3. State the difference between Correlation and Regression.

4. Obtain the two regression equations from the data given bellow.
   
   \[
   \begin{array}{c}
   X: 7 4 6 5 8 \\
   Y: 6 5 9 8 2 \\
   \end{array}
   \]

   Hence estimate y when x = 10.

5. The data given below are the years of experience (x) and monthly wages (y) for a group of workers. Obtain the two regression equations and approximate the monthly wages of a workers who have completed 15 years of service.

<table>
<thead>
<tr>
<th>Experience: In years</th>
<th>11</th>
<th>7</th>
<th>9</th>
<th>5</th>
<th>8</th>
<th>6</th>
<th>10</th>
</tr>
</thead>
<tbody>
<tr>
<td>Monthly wages: (in ‘000Rs.)</td>
<td>10</td>
<td>8</td>
<td>6</td>
<td>8</td>
<td>9</td>
<td>7</td>
<td>11</td>
</tr>
</tbody>
</table>

6. Following results are obtained for a bivariate data. Obtain the two regression equations and find y when x = 12

\[ n = 15 \quad \Sigma x = 130 \quad \Sigma y = 220 \quad \Sigma x^2 = 2288 \quad \Sigma y^2 = 5506 \quad \Sigma xy = 3467 \]

7. Marks scored by a group of 10 students in the subjects of Maths and Stats in a class test are given below. Obtain a suitable regression equation to find the marks of a student in the subject of Stats who have scored 25 marks in Maths.

<table>
<thead>
<tr>
<th>Student no</th>
<th>1</th>
<th>2</th>
<th>3</th>
<th>4</th>
<th>5</th>
<th>6</th>
<th>7</th>
<th>8</th>
<th>9</th>
<th>10</th>
</tr>
</thead>
<tbody>
<tr>
<td>Marks</td>
<td>13</td>
<td>18</td>
<td>9</td>
<td>6</td>
<td>14</td>
<td>10</td>
<td>20</td>
<td>28</td>
<td>21</td>
<td>16</td>
</tr>
</tbody>
</table>
8. The data given below are the price and demand for a certain commodity over a period of 7 years. Find the regression equation of Price on Demand and hence obtain the most likely demand for the in the year 2008 when it’s price is Rs.23.

<table>
<thead>
<tr>
<th>Year:</th>
<th>2001</th>
<th>2002</th>
<th>2003</th>
<th>2004</th>
<th>2005</th>
<th>2006</th>
<th>2007</th>
</tr>
</thead>
<tbody>
<tr>
<td>Price(in RS):</td>
<td>15</td>
<td>12</td>
<td>18</td>
<td>22</td>
<td>19</td>
<td>21</td>
<td>25</td>
</tr>
<tr>
<td>Demand (100 units)</td>
<td>89</td>
<td>86</td>
<td>90</td>
<td>105</td>
<td>100</td>
<td>110</td>
<td>115</td>
</tr>
</tbody>
</table>

9. For a bivariate data the following results were obtained
\[
\bar{x} = 53.2, \quad \bar{y} = 27.9, \quad 6x = 4.8, \quad \sigma_y = \sigma.4 \text{ and } r = 0.75
\]
Obtain the two regression equations, find the most probable value of \(x\) when \(y = 25\).

10. A sample of 50 students in a school gave the following statistics about Marks of students in Subjects of Mathematics and Science,

<table>
<thead>
<tr>
<th>Subjects:</th>
<th>Mathematics</th>
<th>Science</th>
</tr>
</thead>
<tbody>
<tr>
<td>Mean</td>
<td>58</td>
<td>79</td>
</tr>
<tr>
<td>S.D.</td>
<td>12</td>
<td>18</td>
</tr>
</tbody>
</table>

Coefficient of correlation between the marks in Mathematics and marks in Science is 0.8. Obtain the two regression equations and approximate the marks of a student in the subject of Mathematics whose score in Science is 65.

11. It is known that the Advertisement promotes the Sales of the company. The company’s previous records give the following results.

<table>
<thead>
<tr>
<th>Expenditure on Advertisement (Rs. In Lakh)</th>
<th>Sales (Rs. In Lakh)</th>
</tr>
</thead>
<tbody>
<tr>
<td>Mean</td>
<td>15</td>
</tr>
<tr>
<td>S.D.</td>
<td>6</td>
</tr>
</tbody>
</table>

Coefficient of correlation between sales and expenditure on Advertisement is 0.6. Using the regression equation find the likely sales when advertisement budget is Rs.25 Lakh.

12. Find the values of \(x, y,\) and \(r\) from the two regression equations given bellow. \(3x+2y-26=0\) and \(6x+y-31=0.\) Also find \(6x\) when \(\sigma y = 3.\)
13. Two random variables have the regression equations:
   \[ 5x + 7y - 22 = 0 \] and \[ 6x + 2y - 20 = 0. \] Find the mean values of \( x \) and \( y \). Also find S.D. of \( x \) when S.D. of \( y = 5 \).

14. The two regression equations for a certain data were \( y = x + 5 \) and \( 16x = 9y - 94 \). Find values of \( x \), \( y \) and \( r \). Also find the S.D. of \( y \) when S.D. of \( x \) is 2.4.

✦✦✦
UNIT IX

13

TIME SERIES

OBJECTIVES

From this chapter student should learn analysis of data using various methods. Methods involve moving average method and least square method seasonal fluctuations can be studied by business for casting method.

13.1 INTRODUCTION

Every business venture needs to know their performance in the past and with the help of some predictions based on that, would like to decide their strategy for the present. By studying the past behavior of the characteristics, the nature of variation in the value can be determined. The values in the past can be compared with the present values of comparisons at different places during formulation of future plan and policies. This is applicable to economic policy makers, meteorological department, social scientists, political analysis. Forecasting thus is an important tool in Statistical analysis. The statistical data, particularly in the field of social science, are dynamic in nature. Agricultural and Industrial production increase every year or due to improved medical facilities, there is decline in the death rate over a period of time. There is increase in sales and exports of various products over a period of years. Thus, a distinct change (either increasing or decreasing) can be observed in the value of time-series.

A time series is a sequence of value of a phenomenon arranged in order of their occurrence. Mathematically it can expressed as a function, namely \( y = f(t) \) where \( t \) represents time and \( y \) represents the corresponding values. That is, the value \( y_1, y_2, y_3 \ldots \) of a phenomenon with respect to time periods \( t_1, t_2, t_3 \ldots \). Form a Time Series.

Forecasting techniques facilities prediction on the basic of data available from the past. This data from the past is called a time series. A set of observations, of a variable, taken at a regular (fixed and equal) interval of time is called time series. A time series is a bivariate data, with time as the independent variable and the other is the variable under consideration. There are various forecasting method for time series which enable us to study the variation or trends and estimate the same for the future.
13.2 IMPORTANCE OF TIME SERIES ANALYSIS

The analysis of the data in the time series using various forecasting model is called as time analysis. The importance of time series analysis is due to the following reasons:

- Understanding the past behavior
- Planning the future action
- Comparative study

13.3 COMPONENTS OF TIME SERIES

The fluctuation in a time series are due to one or more of the following factors which are called “components” of time series.

(a) Secular Trend:
The general tendency of the data, either to increase, to decrease or to remain constant is called Secular Trend. It is smooth, long term movement of the data. The changes in the values are gradual and continuous. An increasing demand for luxury items like refrigerators or colour T.V. sets reflect increasing trend. The production of steel, cement, vehicles shows a rising trend. On the other hand, decreasing in import of food grains is an example of decreasing trend. The nature of the trend may be linear or curvilinear, in practice, curvilinear trend is more common.

Trend in due to long term tendency. Hence it can be evaluation if the time series is available over a long duration.

(b) Seasonal Variation:
The regular, seasonal change in the time series are called “Seasonal Variation”. It is observed that the demand for umbrellas, raincoats reaches a peak during monsoon or the advertisement of cold drinks, ice creams get a boom in summer. The demand for greeting cards, sweets, increase during festival like Diwali, Christmas. In March, there is maximum withdrawal of bank deposits for adjustment of income-tax payment, so also variation tax-saving schemes shoot up during this period.

The causes, for these seasonal fluctuations, are thus change in weather conditions, the traditions and customs of people etc. The seasonal component is measured to isolate these change from the trend component and to study their effect, so that, in any business, future production can be
planned accordingly and necessary adjustments for seasonal change can be made

(c) Cyclical Variation:
These are changes in time series, occurring over a period which is more than a year. They are recurring and periodic in nature. The period may not be uniform. These fluctuations are due to changes in a business cycle. There are four important phases of any business activity viz. prosperity, recession, depression and recovery. During prosperity, the business flourishes and the profit reaches a maximum level. Thereafter, in recession, the profit decreases, reaching a minimum level during depression. After some time period, the business again recovers (recovery) and it is followed by period of prosperity. The variation in the time series due to these phases in a business cycle are called “Cyclical Variation”.

![Diagram of Cyclical Variation]

The knowledge of cyclic variations is important for a businessman to plan his activity or design his policy for the phase of recession or depression. But one should know that the factors affecting the cyclical variations are quite irregular, difficult to identify and measure. The cyclical variation are denoted by

(d) Irregular Variation:
The changes in the time series which can not be predicated and are erratic in nature are called “Irregular Variation”. Usually, these are short term changes having signification effect on the time series during that time interval. These are caused by unforeseen event like wars, floods, strikes, political charges, etc. During Iran-Iraq war or recent Russian revolution, prices of petrol and petroleum product soared very high. In recent budget, control on capital issued was suddenly removed. As an effect, the all Indian-Index of share market shooted very high, creates all time records. If the effect of other components of the time series is eliminated, the remaining variation are called “Irregular or
Random Variations”. No forecast of these change can be made as they do not reflect any fixed pattern.

MODELS FOR ANALYSIS OF TIME SERIES

The purpose of studying time series is to estimate or forecast the value of the variable. As there are four components of the time series, these are to be studied separately. There are two types of models which are used to express the relationship of the components of the time series. They are additive model and multiplicative model.

- **O** = Original Time Series
- **T** = Secular Trend
- **S** = Seasonal Variations
- **C** = Cyclical Variations and
- **I** = Irregular Variations

In Additive model, it is assumed that the effect of the individual components can be added to get resultant value of the time series, that is the components are independent of one another. The model can be expressed as

\[ O = T + S + C + I \]

In multiplicative model, it is assumed that the multiplication of the individual effect of the components result in the time series, that is, the components are due to different causes but they are not necessarily independent, so that changes in any one of them can affect the other components. This model is more commonly used. It is expressed as

\[ O = T \times S \times C \times I \]

If we want to estimate the value in time series, we have to first estimate the four components and them combine them to estimate the value of the time series. The irregular variations can be found. However, we will restrict ourselves, to discuss method of estimating the first components, namely Secular Trend.

### 13.4 METHODS TO FIND TREND

There are various method to find the trend. The major methods are as mentioned below:

1. Free Hand Curve.

We will study only the method of moving average and least squares.

#### 13.4.1. Method of Moving Averages

This is a simple method in which we take the arithmetic average of the given times series over a certain period of time. These
average move over period and are hence called as moving averages. The
time interval for the average is taken as 3 years, 4 years or 5 years and so
on. The average are thus called as 3 yearly, 4 yearly and 5 yearly moving
average. The moving averages are useful in smoothing the fluctuations
caued to the variable. Obviously larger the time interval of the average
more is the smoothing. We shall study the odd yearly (3 and 5) moving
average first and then the 4 yearly moving average.

Odd Yearly Moving Average

In this method the total of the value in the time series is
taken for the given time interval and is written in front of the middle
value. The average so taken is also written in front of this middle value.
This average is the trend for that middle year. The process is continued by
replacing the first value with the next value in the time series and so on till
the trend for the last middle value is calculated. Let us understand this
with example:

Example 1:
Find 3 years moving averages and draw these on a graph paper. Also
represent the original time series on the graph.

<table>
<thead>
<tr>
<th>Year</th>
<th>1999</th>
<th>2000</th>
<th>2001</th>
<th>2002</th>
<th>2003</th>
<th>2004</th>
<th>2005</th>
<th>2006</th>
<th>2007</th>
</tr>
</thead>
<tbody>
<tr>
<td>Prod. (in thousand unit)</td>
<td>12</td>
<td>15</td>
<td>20</td>
<td>18</td>
<td>25</td>
<td>32</td>
<td>30</td>
<td>40</td>
<td>44</td>
</tr>
</tbody>
</table>

Solution:
We calculate arithmetic mean of first three observations viz. 12, 15 and
20, then we delete 12 and consider the next one so that now, average of
15, 20 and 18 is calculated and so on. These averages are placed against
the middle year of each group, viz. the year 2000, 2001 and so on. Note
moving averages are not obtained for the year 1999 and 2007.

<table>
<thead>
<tr>
<th>Year</th>
<th>Production (in thousand unit)</th>
<th>3 Years Total</th>
<th>3yrly. Moving Average</th>
</tr>
</thead>
<tbody>
<tr>
<td>1999</td>
<td>12</td>
<td>47</td>
<td>15.6</td>
</tr>
<tr>
<td>2000</td>
<td>15</td>
<td>53</td>
<td>17.6</td>
</tr>
<tr>
<td>2001</td>
<td>20</td>
<td>63</td>
<td>21.0</td>
</tr>
<tr>
<td>2002</td>
<td>18</td>
<td>75</td>
<td>25.0</td>
</tr>
<tr>
<td>2003</td>
<td>25</td>
<td>87</td>
<td>29.0</td>
</tr>
<tr>
<td>2004</td>
<td>32</td>
<td>102</td>
<td>34.0</td>
</tr>
<tr>
<td>2005</td>
<td>30</td>
<td>114</td>
<td>38.0</td>
</tr>
<tr>
<td>2006</td>
<td>40</td>
<td></td>
<td></td>
</tr>
<tr>
<td>2007</td>
<td>44</td>
<td></td>
<td></td>
</tr>
</tbody>
</table>

Example 2:
Find 5 yearly moving average for the following data.

<table>
<thead>
<tr>
<th></th>
<th></th>
<th></th>
<th></th>
<th></th>
<th></th>
<th></th>
<th></th>
<th></th>
<th></th>
<th></th>
</tr>
</thead>
<tbody>
<tr>
<td>Sales (in lakhs of Rs.)</td>
<td>51</td>
<td>53</td>
<td>56</td>
<td>57</td>
<td>60</td>
<td>55</td>
<td>59</td>
<td>62</td>
<td>68</td>
<td>70</td>
</tr>
</tbody>
</table>

Solution:
We find the average of first five values, namely 51, 53, 56, 57 and 60. Then we omit the first value 51 and consider the average of next five values, that is, 53, 56, 57, 60 and 55. This process is continued till we get the average of the last five values 55, 59, 62, 68 and 70. The following table is prepared.

<table>
<thead>
<tr>
<th>Year</th>
<th>Sales (in lakhs of Rs.)</th>
<th>5 Years Total</th>
<th>Moving Average (Total / 5)</th>
</tr>
</thead>
<tbody>
<tr>
<td>1997</td>
<td>51</td>
<td>....</td>
<td>....</td>
</tr>
<tr>
<td>1998</td>
<td>53</td>
<td>....</td>
<td>....</td>
</tr>
<tr>
<td>1999</td>
<td>56</td>
<td>51 + 53 + 56 + 57 + 60 = 277</td>
<td>55.4</td>
</tr>
<tr>
<td>2000</td>
<td>57</td>
<td>53 + 56 + 57 + 60 + 55 = 281</td>
<td>56.2</td>
</tr>
<tr>
<td>2001</td>
<td>60</td>
<td>56 + 57 + 60 + 55 + 59 = 287</td>
<td>57.4</td>
</tr>
<tr>
<td>2002</td>
<td>55</td>
<td>57 + 60 + 55 + 59 + 62 = 293</td>
<td>58.6</td>
</tr>
<tr>
<td>2003</td>
<td>59</td>
<td>60 + 55 + 59 + 62 + 68 = 304</td>
<td>60.8</td>
</tr>
<tr>
<td>2004</td>
<td>62</td>
<td>55 + 59 + 62 + 68 + 70 = 314</td>
<td>62.8</td>
</tr>
<tr>
<td>2005</td>
<td>68</td>
<td>....</td>
<td>....</td>
</tr>
<tr>
<td>2006</td>
<td>70</td>
<td>....</td>
<td>....</td>
</tr>
</tbody>
</table>

Example 3:
Determine the trend of the following time series using 5 yearly moving averages.

<table>
<thead>
<tr>
<th></th>
<th></th>
<th></th>
<th></th>
<th></th>
<th></th>
<th></th>
<th></th>
<th></th>
<th></th>
<th></th>
<th></th>
</tr>
</thead>
<tbody>
<tr>
<td>Exports in '000 Rs</td>
<td>78</td>
<td>84</td>
<td>80</td>
<td>83</td>
<td>86</td>
<td>89</td>
<td>88</td>
<td>90</td>
<td>94</td>
<td>93</td>
<td>96</td>
</tr>
</tbody>
</table>
**Solution:** The time series is divided into overlapping groups of five years, their 5 yearly total and average are calculated as shown in the following table.

<table>
<thead>
<tr>
<th>Year</th>
<th>Export (Y)</th>
<th>5 – yearly total (T)</th>
<th>5 – yearly moving average: (T/5)</th>
</tr>
</thead>
<tbody>
<tr>
<td>1981</td>
<td>78</td>
<td>78+84+80+83+86 = 411</td>
<td>411 / 5 = 82.2</td>
</tr>
<tr>
<td>1982</td>
<td>84</td>
<td>84+80+83+86+89 = 422</td>
<td>422 / 5 = 84.4</td>
</tr>
<tr>
<td>1983</td>
<td>80</td>
<td>80+83+86+89+88 = 426</td>
<td>85.2</td>
</tr>
<tr>
<td>1984</td>
<td>83</td>
<td>83+86+89+88+90 = 436</td>
<td>87.2</td>
</tr>
<tr>
<td>1985</td>
<td>86</td>
<td>86+89+88+90+94 = 447</td>
<td>89.4</td>
</tr>
<tr>
<td>1986</td>
<td>89</td>
<td>89+88+90+94+93 = 454</td>
<td>90.8</td>
</tr>
<tr>
<td>1987</td>
<td>90</td>
<td>88+90+94+93+96 = 461</td>
<td>92.2</td>
</tr>
</tbody>
</table>

**Observations:**

I. In case of the 5 – yearly moving average, the total and average for the first two and the last two in the time series is not calculated. Thus, the moving average of the first two and the last two years in the series cannot be computed.

II. To find the 3 – yearly total (or 5 – yearly total) for a particular years, you can subtract the first value from the previous year’s total, and add the next value so as to save your time!

**Even yearly moving averages**

In case of even yearly moving average the method is slightly different as here we cannot find the middle year of the four years in consideration. Here we find the total for the first four years and place it between the second and the third year value of the variable. These totals are again sunned into group of two, called as centered total and is placed between the two totals. The 4 – yearly moving average is found by dividing these centered totals by 8. Let us understand this method with an example

**Example 4:** Calculate the 4 yearly moving averages for the following data.

<table>
<thead>
<tr>
<th></th>
<th></th>
<th></th>
<th></th>
<th></th>
<th></th>
<th></th>
<th></th>
<th></th>
<th></th>
</tr>
</thead>
<tbody>
<tr>
<td>Import in’000Rs</td>
<td>15</td>
<td>18</td>
<td>20</td>
<td>24</td>
<td>21</td>
<td>25</td>
<td>28</td>
<td>26</td>
<td>30</td>
</tr>
</tbody>
</table>

**Ans:** The table of calculation is show below. Student should leave one line blank after every to place the centered total in between two years.
### Example 5:
Find the moving average of length 4 for the following data. Represent the given data and the moving average on a graph paper.

<table>
<thead>
<tr>
<th>Year</th>
<th>1998</th>
<th>1999</th>
<th>2000</th>
<th>2001</th>
<th>2002</th>
<th>2003</th>
<th>2004</th>
<th>2005</th>
<th>2006</th>
<th>2007</th>
</tr>
</thead>
<tbody>
<tr>
<td>Sales (in thousand unit)</td>
<td>60</td>
<td>69</td>
<td>81</td>
<td>86</td>
<td>78</td>
<td>93</td>
<td>102</td>
<td>107</td>
<td>100</td>
<td>109</td>
</tr>
</tbody>
</table>

**Solution:** We prepare the following table.

<table>
<thead>
<tr>
<th>Year</th>
<th>Sale (in thousand unit)</th>
<th>4 Yearly Totals</th>
<th>Centred Total</th>
<th>Moving Avg. / Central = Total / 8</th>
</tr>
</thead>
<tbody>
<tr>
<td>1998</td>
<td>60</td>
<td>60 + 69 + 81 + 86 = 296</td>
<td>296 + 314 = 610</td>
<td>76.25</td>
</tr>
<tr>
<td>1999</td>
<td>69</td>
<td>69 + 81 + 86 + 78 = 314</td>
<td>314 + 338 = 652</td>
<td>81.5</td>
</tr>
<tr>
<td>2000</td>
<td>81</td>
<td>81 + 86 + 78 + 93 = 338</td>
<td></td>
<td></td>
</tr>
</tbody>
</table>
Note that 4 yearly total are written between the years 1999-2000, 2000-01, 2001-02 etc. and the central total are written against the years 2000, 2001, 2002 etc. so also the moving average are considered w.r.t. years; 2000, 2001 and so on. The moving averages are obtained by dividing the certain total by 8.

The graph of the given set of values and the moving averages against time representing the trend component are shown below. Note that the moving averages are not obtained for the years 1998, 1999, 2006 and 2007. (i.e. first and last two extreme years).

When the values in the time series are plotted, a rough idea about the type of trend whether linear or curvilinear can be obtained. Then, accordingly a linear or second degree equation can be fitted to the values. In this chapter, we will discuss linear trend only.

13.4.2. LEAST SQUARES METHOD:

Let \( y = a + bx \) be the equation of the straight line trend where \( a, b \) are constant to be determined by solving the following normal equations,

\[
\begin{align*}
\sum y &= na + b\sum x \\
\sum xy &= a\sum x + b\sum x^2
\end{align*}
\]

where \( y \) represents the given time series.
We define \( x \) from years such that \( \sum x = 0 \). So substituting \( \sum x = 0 \) in the normal equation and simplifying, we get

\[
b = \frac{\sum xy}{\sum x^2} \quad \text{and} \quad a = \frac{\sum y}{n}
\]
Using the given set of values of the time series, \( a, b \) can be calculated and the straight line trend can be determined as \( y = a + bx \). This gives the minimum sum of squares line deviations between the original data and the estimated trend values. The method provides estimates of trend values for all the years. The method has mathematical basis and so element of personal bias is not introduced in the calculation. As it is based on all the values, if any values are added, all the calculations are to be done again.

**Odd number of years in the time series**

When the number of years in the given time series is odd, for the middle year we assume the value of \( x = 0 \). For the years above the middle year the value given to \( x \) are \( \ldots, -2, -1 \) while those after the middle year are values \( 1, 2, \ldots \) and so on.

**Even number of years in the time series**

When the number of years in the time series is even, then for the upper half the value of \( x \) are assumed as \( \ldots, -5, -3, -1 \). For the lower half years, the values of \( x \) are assumed as \( 1, 3, 5, \ldots \) and so on.

**Example 6:**

Fit a straight line trend for the following data giving the annual profits (in lakhs of Rs.) of a company. Estimate the profit for the year 1999.

<table>
<thead>
<tr>
<th></th>
<th></th>
<th></th>
<th></th>
<th></th>
<th></th>
<th></th>
<th></th>
</tr>
</thead>
<tbody>
<tr>
<td>Profit</td>
<td>30</td>
<td>34</td>
<td>38</td>
<td>36</td>
<td>39</td>
<td>40</td>
<td>44</td>
</tr>
</tbody>
</table>

**Solution:** Let \( y = a + bx \) be the straight line trend.

The number of years is seven, which is odd. Thus, the value of \( x \) is taken as 0 for the middle years 1995, for upper three years as \( -3, -2, -1 \) and for lower three years as \( 1, 2, 3 \).

The table of computation is as shown below:

<table>
<thead>
<tr>
<th>Years</th>
<th>Profit (y)</th>
<th>x</th>
<th>xy</th>
<th>( x^2 )</th>
<th>Trend Value: ( Y_t = a + bx )</th>
</tr>
</thead>
<tbody>
<tr>
<td>1992</td>
<td>30</td>
<td>-3</td>
<td>-90</td>
<td>9</td>
<td>31.41</td>
</tr>
<tr>
<td>1993</td>
<td>34</td>
<td>-2</td>
<td>-68</td>
<td>4</td>
<td>33.37</td>
</tr>
<tr>
<td>1994</td>
<td>38</td>
<td>-1</td>
<td>-38</td>
<td>1</td>
<td>35.33</td>
</tr>
<tr>
<td>1995</td>
<td>36</td>
<td>0</td>
<td>0</td>
<td>0</td>
<td>37.29</td>
</tr>
<tr>
<td>1996</td>
<td>39</td>
<td>1</td>
<td>39</td>
<td>1</td>
<td>39.25</td>
</tr>
<tr>
<td>1997</td>
<td>40</td>
<td>2</td>
<td>80</td>
<td>4</td>
<td>41.21</td>
</tr>
<tr>
<td>1998</td>
<td>44</td>
<td>3</td>
<td>132</td>
<td>9</td>
<td>43.17</td>
</tr>
<tr>
<td>Total</td>
<td>( \sum y = 261 )</td>
<td>( \sum x = 0 )</td>
<td>( \sum xy = 55 )</td>
<td>( \sum x^2 = 28 )</td>
<td></td>
</tr>
</tbody>
</table>

From the table: \( n = 7, \ \sum xy = 55, \ \sum x^2 = 28, \ \sum y = 261 \)
There fore \[ b = \frac{\sum xy}{\sum x^2} = \frac{55}{28} = 1.96 \quad \text{and} \quad a = \frac{\sum y}{n} = \frac{261}{7} = 37.29 \]

Thus, the straight line trend is \[ y = 37.29 + 1.96x. \]

The trend values in the table for the respective years are calculated by substituting the corresponding value of \( x \) in the above trend line equation.

For the trend value for 1992: \( x = -3 \):

\[ y_{1992} = 37.29 + 1.96(-3) = 37.29 - 5.88 = 31.41 \]

Similarly, all the remaining trend values are calculated.

(A short-cut method in case of odd number of years to find the remaining trend values once we calculate the first one, is to add the value of \( b \) to the first trend value to get the second trend value, then to the second trend value to get the third one and so on. This is because the difference in the values of \( x \) is 1.)

To estimate the profit for the years 1999 in the trend line equation, we substitute the prospective value of \( x \), if the table was extended to 1999. i.e. we put \( x = 4 \), the next value after \( x = 3 \) for the year 1998.

\[ \therefore y_{1999} = 37.29 + 1.96(4) = 45.13 \]

Thus, the estimated profit for the year 1999 is Rs. 45.13 lakhs.

**Example 7:**
Fit straight line trend by the method of least squares for the following data representing production in thousand units. Plot the data and the trend line on a graph paper. Hence or otherwise estimate the trend for the years 2007.

<table>
<thead>
<tr>
<th>Year</th>
<th>1999</th>
<th>2000</th>
<th>2001</th>
<th>2002</th>
<th>2003</th>
<th>2004</th>
<th>2005</th>
</tr>
</thead>
<tbody>
<tr>
<td>Production (in thousand unit)</td>
<td>14</td>
<td>15</td>
<td>17</td>
<td>16</td>
<td>17</td>
<td>20</td>
<td>23</td>
</tr>
</tbody>
</table>

**Solution:**
Here, the total number of years is 7, an odd number. So we take the center as 1986 the middle-most year and define \( x \) as year 2002. The values of \( x \) will be \(-3, -2, -1, 0, 1, 2, 3\).

Prepare the following table to calculate the required summations. Note that the trend values can be written in the table only after calculation of \( a \) and \( b \).
Here, \( n = 7, \sum y = 122, \sum x^2 = 28, \sum xy = 37 \)

Now, \( a \) and \( b \) are calculated as follows:

\[
a = \frac{\sum y}{n} = \frac{122}{7} = 17.4286 \approx 17.43
\]

\[
b = \frac{\sum xy}{\sum x^2} = \frac{37}{28} = 1.3214 \approx 1.32
\]

So, the equation is used to find trend values.

\[y = a + b \times x\]

i.e. \( y = 17.43 + 1.32x\)

The equation is used to find trend values.

For the year 1999, \( x = -3 \), substituting the value of \( x \), we get,
\[y = 17.43 + 1.32 (-3) = 17.43 - 3.96 = 13.47\]

to find the remaining trend values we can make use of the property of a straight line that as all the values of \( x \) are equidistant with different of one unit (-3, -2, -1, ---- and so on), the estimated trend value will also be equidistant with a difference of \( b \) unit.

In this case as \( b = 1.32 \), the remaining trend values for \( x = -2, -1, 0, --- \) etc. are obtained by adding \( b = 1.32 \) to the previous values. So, the trend values are 13.47, 14.79, 16.11, 17.43, 18.75, 20.07 and 21.39.

Now to estimate trend for the year 2007, \( x = 5 \), substituting in the equation
\[y = 17.43 + 1.32x\]
\[= 17.43 + 1.32 (5) = 24.03\]

So, the estimated trend value for the year 2007 is 24,030 unit.

For graph of time series, all points are plotted. But for the graph of trend line, any two trend values can be plotted and the line joining these points represents the straight line trend.

<table>
<thead>
<tr>
<th>Year</th>
<th>Production (y)</th>
<th>x</th>
<th>( x^2 )</th>
<th>xy</th>
<th>Trend Values</th>
</tr>
</thead>
<tbody>
<tr>
<td>1999</td>
<td>14</td>
<td>-3</td>
<td>9</td>
<td>-42</td>
<td>13.47</td>
</tr>
<tr>
<td>2000</td>
<td>15</td>
<td>-2</td>
<td>4</td>
<td>-30</td>
<td>14.79</td>
</tr>
<tr>
<td>2001</td>
<td>17</td>
<td>-1</td>
<td>1</td>
<td>-17</td>
<td>16.11</td>
</tr>
<tr>
<td>2002</td>
<td>16</td>
<td>0</td>
<td>0</td>
<td>0</td>
<td>17.43</td>
</tr>
<tr>
<td>2003</td>
<td>17</td>
<td>1</td>
<td>1</td>
<td>17</td>
<td>18.75</td>
</tr>
<tr>
<td>2004</td>
<td>20</td>
<td>2</td>
<td>4</td>
<td>40</td>
<td>20.07</td>
</tr>
<tr>
<td>2005</td>
<td>23</td>
<td>3</td>
<td>3</td>
<td>69</td>
<td>21.39</td>
</tr>
<tr>
<td></td>
<td>122</td>
<td>28</td>
<td>37</td>
<td></td>
<td></td>
</tr>
</tbody>
</table>
For the trend line, the trend values 17.43 and 21.39 for the years 2002 and 2005 are plotted and then a straight line joining these two points is drawn and is extended on both the sides.

The estimate of trend for the year 2007 can also be obtained from the graph by drawing a perpendicular for the year 2007, from x-axis which meet the trend line at point P. From P, a perpendicular on y-axis gives the required trend estimate as 24.

Now, to find straight line trend, when number of years is even, consider the following example.

**Example 8:**

Fit a straight line trend to the following time –series, representing sales in lakhs of Rs. of a company, for the year 1998 to 2005. Plot the given data well as the trend line on a graph paper. Hence or otherwise estimate trend for the year 2006.

<table>
<thead>
<tr>
<th>Year</th>
<th>1998</th>
<th>1999</th>
<th>2000</th>
<th>2001</th>
<th>2002</th>
<th>2003</th>
<th>2004</th>
<th>2005</th>
</tr>
</thead>
<tbody>
<tr>
<td>Sales (Lakhs of Rs.)</td>
<td>31</td>
<td>33</td>
<td>30</td>
<td>34</td>
<td>38</td>
<td>40</td>
<td>45</td>
<td>49</td>
</tr>
</tbody>
</table>

**Solution:**

Here the number of years = 8, an even number, so we define 
\[ x = \frac{year - 2001.5}{0.5} \]

so that the values of x are -7, -3, -1, 1, 3, 5 and 7, to get 
\[ \sum x = 0. \]

Prepare the following table to obtain the summations \( \sum x^2, \sum y, \sum xy \).
Here, \( n = 8 \), \( \sum y = 300 \), \( \sum x^2 = 168 \), \( \sum xy = 220 \)

Now, \( a \) and \( b \) are calculated as follows:

\[
a = \frac{\sum y}{n} = \frac{300}{8} = 37.5
\]

\[
b = \frac{\sum xy}{\sum x^2} = \frac{220}{168} = 1.31
\]

So, the equation of the straight line trend is \( y = a + b x \)

\[
i.e. \ y = 37.5 + 1.31 \ x
\]

To obtain the trend values, first calculate \( y \)

for \( x = -7 \), for the year 1998

\[
y = 37.5 + 1.31 (-7) = 37.5 - 9.17 = 28.33
\]

To find the successive trend values, go on addition \( 2b = 2 \times 1.31 = 2.62 \),
to the preceding values as in this case the different between \( x \) values is of 2 units.

So, the estimated values of trend for \( x = -5, -3, -1, 1, 3, 5, 7 \) and 7 are 30.95, 33.57, 36.19, 38.81, 41.43, 44.05 and 46.67 respectively. Write down these values in the table.

Hence the estimated trend value for the year 2006 is 49.29 (in lakhs of Rs.).

Now, for the graph of trend line, note that only two trend values 30.95 and 46.67 w.r.t. years 1999 and 2005 are considered as point. The line joining these two points represents trend line.
To estimate the trend for the year 2006, draw a perpendicular from x-axis at this point meeting the line in P. Then from P, draw another perpendicular on y-axis which gives estimate of trend as 49.

**Example 9:**

Fit a straight line trend to the following data. Draw the graph of the actual time series and the trend line. Estimate the sales for the year 2007.

<table>
<thead>
<tr>
<th>Year</th>
<th>1998</th>
<th>1999</th>
<th>2000</th>
<th>2001</th>
<th>2002</th>
<th>2003</th>
<th>2004</th>
<th>2005</th>
</tr>
</thead>
<tbody>
<tr>
<td>Sales in'000Rs</td>
<td>120</td>
<td>124</td>
<td>126</td>
<td>130</td>
<td>128</td>
<td>132</td>
<td>138</td>
<td>137</td>
</tr>
</tbody>
</table>

**Solution:** let \( y = a + bx \) be the straight line trend.

The number of years in the given time series is eight, which is an even number. The upper four years are assigned the values of \( x \) as 1, 2, 3, and 7. Note that the difference between the values of \( x \) is 2, but the sum is zero.

Now, the table of computation is completed as shown below:

<table>
<thead>
<tr>
<th>Years</th>
<th>Profit (y)</th>
<th>X</th>
<th>Xy</th>
<th>X²</th>
<th>Trend Value: ( Y_t = a + bx )</th>
</tr>
</thead>
<tbody>
<tr>
<td>1998</td>
<td>120</td>
<td>-7</td>
<td>-840</td>
<td>49</td>
<td>120.84</td>
</tr>
<tr>
<td>1999</td>
<td>124</td>
<td>-5</td>
<td>-620</td>
<td>25</td>
<td>123.28</td>
</tr>
<tr>
<td>2000</td>
<td>126</td>
<td>-3</td>
<td>-378</td>
<td>9</td>
<td>125.72</td>
</tr>
<tr>
<td>2001</td>
<td>130</td>
<td>-1</td>
<td>-130</td>
<td>1</td>
<td>128.16</td>
</tr>
<tr>
<td>2002</td>
<td>128</td>
<td>1</td>
<td>128</td>
<td>1</td>
<td>130.06</td>
</tr>
<tr>
<td>2003</td>
<td>132</td>
<td>3</td>
<td>396</td>
<td>9</td>
<td>133.04</td>
</tr>
<tr>
<td>2004</td>
<td>138</td>
<td>5</td>
<td>390</td>
<td>25</td>
<td>135.48</td>
</tr>
<tr>
<td>2005</td>
<td>137</td>
<td>7</td>
<td>359</td>
<td>49</td>
<td>137.92</td>
</tr>
</tbody>
</table>
From the table : \( n = 8, \ \sum xy = 205, \ \sum x^2 = 168, \ \sum y = 1035 \)

\[ b = \frac{\sum xy}{\sum x^2} = \frac{205}{168} = 1.22 \quad \text{and} \quad a = \frac{\sum y}{n} = \frac{1035}{8} = 129.38 \]

Thus, the straight line trend is \( y = 129.38 + 1.22x \).

The trend values in the table for the respective years are calculated by substituting the corresponding value of \( x \) in the above trend line equation.

For the trend value for 1998: \( x = -7 \):

\[ y_{1998} = 129.38 + 1.22 (-7) = 129.38 - 8.54 = 120.84 \]

Similarly, all the remaining trend values are calculated.

(A short-cut method in case of even number of years to find the remaining trend values once we calculate the first one, is to add twice the value of \( b \) to the first trend value to get the second trend value, then to the second trend value to get the third one and so on. This is because the difference in the values of \( x \) is 2. In this example we add 2 x 1.22 = 2.44)

**Estimation:**

To estimate the profit for the years 2007 in the trend line equation, we substitute the prospective value of \( x \), if the table was extended to 2007. i.e. we put \( x = 11 \), the next value after \( x = 9 \) for the year 2006 and \( x = 7 \) for 2005.

\[ y_{2007} = 129.38 + 1.22 (11) = 142.8 \]

Therefore the estimated profit for the year 2007 is Rs. 1,42,800.

Now we draw the graph of actual time series by plotting the sales against the corresponding year, the period is taken on the X-axis and the sales on the Y-axis. The points are joined by straight lines. To draw the trend line it is enough to plot any two point (usually we take the first and the last trend value) and join it by straight line.

To estimate the trend value for the year 2007, we draw a line parallel to Y-axis from the period 2007 till it meet the trend line at a point say A. From this point we draw a line parallel to the X-axis till it meet the Y-axis at point say B. This point is our estimate value of sales for the year 2007. The graph and its estimate value (graphically) is shown below:
From the graph, the estimated value of the sales for the year 2007 is 142 i.e. Rs 1,42,000 (approximately)

**Example 10:**

Fit a straight line trend to the following data. Draw the graph of the actual time series and the trend line. Estimate the import for the year 1998.

<table>
<thead>
<tr>
<th></th>
<th></th>
<th></th>
<th></th>
<th></th>
<th></th>
<th></th>
</tr>
</thead>
<tbody>
<tr>
<td>Import in '000 Rs</td>
<td>40</td>
<td>44</td>
<td>48</td>
<td>50</td>
<td>46</td>
<td>52</td>
</tr>
</tbody>
</table>

**Solution:** Here again the period of years is 6 i.e. even. Proceeding similarly as in the above problem, the table of calculation and the estimation is as follows:

<table>
<thead>
<tr>
<th>Years</th>
<th>Import (y)</th>
<th>x</th>
<th>xy</th>
<th>x²</th>
<th>Trend Value: ( Y_t = a + bx )</th>
</tr>
</thead>
<tbody>
<tr>
<td>1991</td>
<td>40</td>
<td>-5</td>
<td>-200</td>
<td>25</td>
<td>41.82</td>
</tr>
<tr>
<td>1992</td>
<td>44</td>
<td>-3</td>
<td>-132</td>
<td>9</td>
<td>43.76</td>
</tr>
<tr>
<td>1993</td>
<td>48</td>
<td>-1</td>
<td>-48</td>
<td>1</td>
<td>45.7</td>
</tr>
<tr>
<td>1994</td>
<td>50</td>
<td>1</td>
<td>50</td>
<td>1</td>
<td>47.64</td>
</tr>
<tr>
<td>1995</td>
<td>46</td>
<td>3</td>
<td>138</td>
<td>9</td>
<td>49.58</td>
</tr>
<tr>
<td>1996</td>
<td>52</td>
<td>5</td>
<td>260</td>
<td>25</td>
<td>51.52</td>
</tr>
<tr>
<td>Total</td>
<td>( \sum y = 280 )</td>
<td>( \sum x = 0 )</td>
<td>( \sum xy = 68 )</td>
<td>( \sum x^2 = 70 )</td>
<td></td>
</tr>
</tbody>
</table>

From the table: \( n = 6 \), \( \sum xy = 68 \), \( \sum x^2 = 70 \), \( \sum y = 280 \)

Their four \( b = \frac{\sum xy}{\sum x^2} = \frac{68}{70} = 0.97 \) and \( \frac{\sum y}{n} = \frac{280}{8} = 46.67 \)

Thus, the straight line trend is \( y = 46.67 + 0.97x \).
All the remaining trend values are calculated as described in the above problem.

**Estimation:**
To estimate the import for the year 1998, we put \( x = 9 \) in the tried line equation. Therefore \( \hat{y}_{1997} = 46.67 + 0.97 (9) = 55.4 \)
Therefore the imports for the year 1997 is Rs. 55,400.

The graph of the actual time series and the trend values along with the graphical estimation is shown below:

From graph the estimated import are Rs. 55,000.

**Example 11:**
Fit a straight line trend to the following time –series, representing sales in lakhs of Rs. of a company, for the year 1998 to 2005. Plot the given data well as the trend line on a graph paper. Hence or otherwise estimate trend for the year 2006.

<table>
<thead>
<tr>
<th>Year</th>
<th>1998</th>
<th>1999</th>
<th>2000</th>
<th>2001</th>
<th>2002</th>
<th>2003</th>
<th>2004</th>
<th>2005</th>
</tr>
</thead>
<tbody>
<tr>
<td>Sales (Lakhs of Rs.)</td>
<td>31</td>
<td>33</td>
<td>30</td>
<td>34</td>
<td>38</td>
<td>40</td>
<td>45</td>
<td>49</td>
</tr>
</tbody>
</table>

**Solution:**
Here the number of years = 8, an even number, so we define \( x = \frac{\text{year} - 2001.5}{0.5} \), so that the values of \( x \) are -7, -3, -1, 1, 3, 5 and 7, to get \( \sum x = 0 \).
Prepare the following table to obtain the summations \( \sum x^2, \sum y, \sum xy \).
<table>
<thead>
<tr>
<th>Year</th>
<th>Sales (in Lakhs of Rs.)</th>
<th>x</th>
<th>x²</th>
<th>x y</th>
<th>Trend Values</th>
</tr>
</thead>
<tbody>
<tr>
<td>1998</td>
<td>31</td>
<td>-7</td>
<td>49</td>
<td>-217</td>
<td>28.33</td>
</tr>
<tr>
<td>1999</td>
<td>33</td>
<td>-5</td>
<td>25</td>
<td>-165</td>
<td>30.95</td>
</tr>
<tr>
<td>2000</td>
<td>30</td>
<td>-3</td>
<td>9</td>
<td>-90</td>
<td>33.57</td>
</tr>
<tr>
<td>2001</td>
<td>34</td>
<td>-1</td>
<td>1</td>
<td>-34</td>
<td>36.19</td>
</tr>
<tr>
<td>2002</td>
<td>38</td>
<td>1</td>
<td>1</td>
<td>38</td>
<td>38.81</td>
</tr>
<tr>
<td>2003</td>
<td>40</td>
<td>3</td>
<td>9</td>
<td>120</td>
<td>41.43</td>
</tr>
<tr>
<td>2004</td>
<td>45</td>
<td>5</td>
<td>25</td>
<td>225</td>
<td>44.05</td>
</tr>
<tr>
<td>2005</td>
<td>49</td>
<td>7</td>
<td>49</td>
<td>343</td>
<td>46.67</td>
</tr>
<tr>
<td></td>
<td>300</td>
<td></td>
<td>168</td>
<td>220</td>
<td></td>
</tr>
</tbody>
</table>

Here, n = 8, \( \sum y = 300, \sum x^2 = 168, \sum xy = 220 \)

Now, a and b are calculated as follows:

\[
a = \frac{\sum y}{n} = \frac{300}{8} = 37.5
\]

\[
b = \frac{\sum xy}{\sum x^2} = \frac{220}{168} = 1.31
\]

So, the equation of the straight line trend is \( y = a + b \cdot x \)

i.e. \( y = 37.5 + 1.31 \cdot x \)

To obtain the trend values, first calculate \( y \) for \( x = -7 \), for the year 1998

\[
y = 37.5 + 1.31 (-7) = 37.5 - 9.17 = 28.33
\]

To find the successive trend values, go on addition \( 2b = 2 \times 1.31 = 2.62 \),
to the preceding values as in this case the different between \( x \) values is of 2 units.

So, the estimated values of trend for \( x = -5, -3, -1, 1, 3, 5, 7 \) and 7 are

30.95, 33. 57, 36.19, 38.81, 41.43, 44.05 and 46.67 respectively. Write down these values in the table.

Hence the estimated trend value for the year 2006 is 49.29 (in lakhs of Rs.).

Now, for the graph of trend line, note that only two trend values 30.95 and 46.67 w.r.t. years 1999 and 2005 are considered as point. The line joining these two points represents trend line.
To estimate the trend for the year 2006, drawn a perpendicular from x-axis at this point meeting the line in P. then from P, draw another perpendicular on y-axis which gives estimate of trend as 49.

MEASUREMENTS OF OTHER COMPONENTS

We have studied four methods of estimation of Secular Trend. The following procedure is applied to separate the remaining components of the time series.

Using seasonal indices (s), the seasonal variations in a time series can be measured. By removing the trend and the seasonal factors, a combination of cyclical and irregular fluctuations is obtained.

If we assume, multiplicative model, represented by the equation

\[ O = T \times S \times C \times I \]

Then, to depersonalize the data, the original time series (O) divided by the seasonal indices (S), which can be express as,

\[ \frac{O}{S} = \frac{T \times S \times C \times I}{S} = T \times C \times I \]

If it is further divided by trend values (T), then we have

\[ \frac{T \times C \times I}{T} = C \times I \]

Thus a combination of cyclical and irregular variation can be obtained. Irregular fluctuations, because of their nature, can not be eliminated completely, but these can be minimized by taking short term averages and then the estimate of cyclical variation can be obtained.

METHODS TO ESTIMATE SEASONAL FLUCTUATIONS
We have seen methods to separate the trend component of Time Series. Now, let us see, how to separate the seasonal component of it.

Methods of Seasonal Index

It is used to find the effect of seasonal variations in a Time Series. The steps are as follows:

i. Find the totals for each season, as well as the grand total, say G.
ii. Find the arithmetic means of these total, and the grand total by dividing the values added.
iii. Find seasonal indices, representing the seasonal component for each season, using the formula:

\[
\text{Seasonal Index} = \frac{\text{Average for Seasonal} \times 100}{\text{Grand Average}}
\]

Where, Grand Average = \( \frac{G}{\text{Total No.of Values}} \)

Example 12:

Find the seasonal component of the time series, using method of seasonal indices.

<table>
<thead>
<tr>
<th>Seasonal / Years</th>
<th>I</th>
<th>II</th>
<th>IV</th>
<th>Grand</th>
</tr>
</thead>
<tbody>
<tr>
<td>2003</td>
<td>33</td>
<td>37</td>
<td>32</td>
<td>31</td>
</tr>
<tr>
<td>2004</td>
<td>35</td>
<td>40</td>
<td>36</td>
<td>35</td>
</tr>
<tr>
<td>2005</td>
<td>34</td>
<td>38</td>
<td>34</td>
<td>32</td>
</tr>
<tr>
<td>2006</td>
<td>36</td>
<td>41</td>
<td>35</td>
<td>36</td>
</tr>
<tr>
<td>2007</td>
<td>34</td>
<td>39</td>
<td>35</td>
<td>32</td>
</tr>
</tbody>
</table>

Solution:

<table>
<thead>
<tr>
<th></th>
<th>I</th>
<th>II</th>
<th>III</th>
<th>IV</th>
<th>Grand</th>
</tr>
</thead>
<tbody>
<tr>
<td>Total Average</td>
<td>172</td>
<td>195</td>
<td>172</td>
<td>166</td>
<td>705 (G)</td>
</tr>
<tr>
<td>Seasonal Index</td>
<td>34.4</td>
<td>39</td>
<td>34.4</td>
<td>33.2</td>
<td>35.25(G/20)</td>
</tr>
<tr>
<td></td>
<td>34.4×100</td>
<td>39×100</td>
<td>34.4×100</td>
<td>33.2×100</td>
<td></td>
</tr>
<tr>
<td></td>
<td>35.25</td>
<td>35.25 = 110.64</td>
<td>35.25 = 97.59</td>
<td>35.25 = 94.18</td>
<td></td>
</tr>
</tbody>
</table>

The time series can be deseasonalised by removing the effect of seasonal component from it. It is done using the formula.

\[
\text{Deseasonalised Value} = \frac{\text{Original Value} \times 100}{\text{Seasonal Index}}
\]

BUSINESS FORECASTING:

In this chapter, few methods of analyzing the past data and predicting the future values are already discussed. Analysis of time series plays an important role in Business Forecasting. One of the aspects of it is estimating future trend values. Now-a-days, any business or industry is governed by factors like supply of raw material, distribution network,
availability of land, labour and capital and facilitates like regular supply of power, coal, water, etc. a business has to sustain intricate government regulations, status, everchanging tastes and fashions, the latest technology, cut throat competition by other manufacturers and many other.

While making a forecast, combined effect of above factors should be considered. Scientific method are used to analyse the past business condition. The study reveals the pattern followed by the business in the past. It also bring out the relationship and interdependence of different industries which helps in interpretation of changes in the right perspective. The analysis gives an idea about the components of the time series and their movement in the past. Various indices such as index of production, prices, bank deposits, money rates, foreign exchange position etc. can provide information about short and long term variations, the general trend, the ups downs in a business.

The study of the past data and the comparison of the estimated and actual values helps in pinpointing the areas of shortcoming which can be overcome. For successful business forecasting co-ordination of all departments such as production, sales, marketing is sine-qua-nin, which result in achieving ultimate corporate goals.

There are different theories of Business Forecasting such as

i. Time lag or Sequence Theory
ii. Action and Reaction Theory
iii. Cross Cut Analysis Theory
iv. Specific Historical Analogy Theory

Of these, Time lag or Sequence Theory is most important. It is based on the fact that there is a time lag between the effect of changes at different stages but there is a sequence followed by these effect e.g. In 80’s, the invention of silicon ships brought fourth and fifth generation computers in use. The computers were introduced in various fields such as front-line and back house banking, airlines and railways reservation, new communication technique, home appliances like washing machine etc. this, in turn, increase the demand for qualified personnel in electronic filed to manufacture, handle and maintain these sophisticated machine. It has result in mad rush for admission to various branches of electronics and computer engineering in the recent past.

By applying any one of the these forecasting theories, business forecasting can be made. It should be noted that while collecting the data for analysis, utmost care has to be taken so as to increase the reliability of estimates. The information should be collected by export investigators, over a long period of time. Otherwise, it may lead to wrong conclusions.

**EXERCISE**
1. What is a time series? Describe the various components of a time series with suitable example.
2. What are seasonal variation? Explain briefly with example.
3. Describe the secular trend component of a time series.
4. What are the method of determining trend in a time series?
6. Find the trend values using the method of semi-averages for the following data expressing production in thousand unit of a company for 7 years.
7. Explain the method to calculate 3 yearly and 4 yearly moving averages.
8. What are the merits and demerits of the method of moving average?
9. Explain the simple average method to find the seasonal indices of a time series.
10. Calculate trend by considering three yearly moving average for the following time series of price indices for the years 2000-2007. Also plot on the graph the trend values.

<table>
<thead>
<tr>
<th>Year</th>
<th>2000</th>
<th>2001</th>
<th>2002</th>
<th>2003</th>
<th>2004</th>
<th>2005</th>
<th>2006</th>
<th>2007</th>
</tr>
</thead>
<tbody>
<tr>
<td>Price Index</td>
<td>111</td>
<td>115</td>
<td>116</td>
<td>118</td>
<td>119</td>
<td>120</td>
<td>122</td>
<td>124</td>
</tr>
</tbody>
</table>

11. Determine the trend for the following data using 3 yearly moving averages. Plot the graph of actual time series and the trend values.

<table>
<thead>
<tr>
<th></th>
<th></th>
<th></th>
<th></th>
<th></th>
<th></th>
<th></th>
<th></th>
<th></th>
<th></th>
</tr>
</thead>
<tbody>
<tr>
<td>Sales in‘000Rs</td>
<td>24</td>
<td>28</td>
<td>30</td>
<td>33</td>
<td>34</td>
<td>36</td>
<td>35</td>
<td>40</td>
<td>44</td>
</tr>
</tbody>
</table>

12. Determine the trend for the following data using 3 yearly moving averages. Plot the graph of actual time series and the trend values.

<table>
<thead>
<tr>
<th></th>
<th></th>
<th></th>
<th></th>
<th></th>
<th></th>
<th></th>
<th></th>
<th></th>
</tr>
</thead>
<tbody>
<tr>
<td>Sales in‘000Rs</td>
<td>46</td>
<td>54</td>
<td>52</td>
<td>56</td>
<td>58</td>
<td>62</td>
<td>59</td>
<td>63</td>
</tr>
</tbody>
</table>

13. Determine the trend for the following data using 3 yearly moving averages. Plot the graph of actual time series and the trend values.

<table>
<thead>
<tr>
<th></th>
<th></th>
<th></th>
<th></th>
<th></th>
<th></th>
<th></th>
<th></th>
<th></th>
</tr>
</thead>
<tbody>
<tr>
<td>Profit in lakhs of Rs</td>
<td>98</td>
<td>100</td>
<td>97</td>
<td>101</td>
<td>107</td>
<td>110</td>
<td>102</td>
<td>105</td>
</tr>
</tbody>
</table>

14. Determine the trend for the following data using 5 yearly moving averages. Plot the graph of actual time series and the trend values.

Values | 34 | 37 | 35 | 38 | 37 | 40 | 43 | 42 | 48 | 50 | 52
---|---|---|---|---|---|---|---|---|---|---|---

15. Determine the trend for the following data giving the production of steel in million tons, using 5 yearly moving averages. Plot the graph of actual time series and the trend values.

<table>
<thead>
<tr>
<th></th>
<th></th>
<th></th>
<th></th>
<th></th>
<th></th>
<th></th>
<th></th>
<th></th>
<th></th>
<th></th>
<th></th>
</tr>
</thead>
<tbody>
<tr>
<td>Production</td>
<td>28</td>
<td>30.5</td>
<td>32</td>
<td>36.8</td>
<td>38</td>
<td>36</td>
<td>39.4</td>
<td>40.6</td>
<td>42</td>
<td>45</td>
<td>43.5</td>
</tr>
</tbody>
</table>

16. Find five-yearly moving average for the following data which represents production in thousand unit of a small scale industry. Plot the given data as well as the moving average on a graph paper.

<table>
<thead>
<tr>
<th></th>
<th></th>
<th></th>
<th></th>
<th></th>
<th></th>
<th></th>
<th></th>
<th></th>
<th></th>
<th></th>
<th></th>
</tr>
</thead>
<tbody>
<tr>
<td>Production</td>
<td>110</td>
<td>104</td>
<td>78</td>
<td>105</td>
<td>109</td>
<td>120</td>
<td>115</td>
<td>110</td>
<td>115</td>
<td>122</td>
<td>130</td>
</tr>
</tbody>
</table>

Ans. The trend values are 101.2, 103.2, 105.4, 111.8, 113.8, 116.4 and 118.4 for the years 1982 to 1988.

17. Find the trend component of the following time series of production in thousand kilogram during 1971-1980. Plot the moving average and the original time on a graph paper.

<table>
<thead>
<tr>
<th></th>
<th></th>
<th></th>
<th></th>
<th></th>
<th></th>
<th></th>
<th></th>
<th></th>
<th></th>
<th></th>
</tr>
</thead>
<tbody>
<tr>
<td>Production</td>
<td>12</td>
<td>15</td>
<td>18</td>
<td>17</td>
<td>16</td>
<td>20</td>
<td>23</td>
<td>22</td>
<td>24</td>
<td>25</td>
</tr>
</tbody>
</table>

Ans. The trend values are 16, 17.125, 18.375, 19.625, 21.25, 22.875 for the years 1973 to 1978.

18. Fit a straight line trend to the following data representing import in million Rs. of a certain company. Also find an estimate for the year 2008.

<table>
<thead>
<tr>
<th>Year</th>
<th>2000</th>
<th>2001</th>
<th>2002</th>
<th>2003</th>
<th>2004</th>
<th>2005</th>
<th>2006</th>
</tr>
</thead>
<tbody>
<tr>
<td>Import</td>
<td>48</td>
<td>50</td>
<td>58</td>
<td>52</td>
<td>45</td>
<td>41</td>
<td>49</td>
</tr>
</tbody>
</table>

Ans. The straight line trend is \( y = 49 - x \). The trend values are 52, 51, 50, 49, 48, 47 and 46 respectively and the estimate trend for the year 2006 is 44 million Rs.

19. The production of a certain brand of television sets in thousand unit is given below. Fit a straight line trend to the data. Plot the given data and the trend line on graph find an estimate for the year 2004.

<table>
<thead>
<tr>
<th>Year</th>
<th>1997</th>
<th>1998</th>
<th>1999</th>
<th>2000</th>
<th>2001</th>
<th>2002</th>
<th>2003</th>
</tr>
</thead>
<tbody>
<tr>
<td>Production</td>
<td>865</td>
<td>882</td>
<td>910</td>
<td>925</td>
<td>965</td>
<td>1000</td>
<td>1080</td>
</tr>
</tbody>
</table>

Ans. The straight line trend is \( y = 947.71 + 33.43x \). The trend values are 846.42, 879.85, 913.28, 946.71, 1013.57 and 1047. The estimate for the year 2004 is 1080.43 thermal million.
20. The straight line trend by the method of least squares for the following data which represents the expenditure in lakhs of Rs. on advertisement of a certain company. Also find an estimate for the year 2005. Plot the given data and the trend line on a graph paper.

<table>
<thead>
<tr>
<th>Year</th>
<th>1997</th>
<th>1998</th>
<th>1999</th>
<th>2000</th>
<th>2001</th>
<th>2002</th>
<th>2003</th>
<th>2004</th>
</tr>
</thead>
<tbody>
<tr>
<td>Expenditure</td>
<td>21</td>
<td>24</td>
<td>32</td>
<td>40</td>
<td>38</td>
<td>49</td>
<td>57</td>
<td>60</td>
</tr>
</tbody>
</table>

**Ans.** The trend is \( y = 40.13 + 2.9x \). The trend values are 19.83, 25.62, 31.43, 37.23, 43.03, 48.83, 54.63 and 60.43, 2005 is 66.23.


<table>
<thead>
<tr>
<th></th>
<th></th>
<th></th>
<th></th>
<th></th>
<th></th>
<th></th>
<th></th>
<th></th>
</tr>
</thead>
<tbody>
<tr>
<td>Production</td>
<td>80</td>
<td>90</td>
<td>92</td>
<td>83</td>
<td>94</td>
<td>99</td>
<td>92</td>
<td>102</td>
</tr>
</tbody>
</table>

**Ans.** The straight line trend is \( y = 91.5 + 1.167x \). The trend values are 83.331, 85.665, 87.999, 90.333, 92.667, 95.001, 97.335 and 99.669. The estimate of trend, for the year 2003 is 102.003.

22. Calculate seasonal indices for the following data:

<table>
<thead>
<tr>
<th>Year</th>
<th>I</th>
<th>II</th>
<th>III</th>
<th>IV</th>
</tr>
</thead>
<tbody>
<tr>
<td>2003</td>
<td>55</td>
<td>53</td>
<td>57</td>
<td>51</td>
</tr>
<tr>
<td>2004</td>
<td>56</td>
<td>55</td>
<td>60</td>
<td>53</td>
</tr>
<tr>
<td>2005</td>
<td>57</td>
<td>56</td>
<td>61</td>
<td>54</td>
</tr>
</tbody>
</table>

**Ans.** 100.59, 98.2, 106.57, 94.61

23. Determine the trend for the following data giving the production of wheat in thousand tons from the years 1980 to 1990, using the 5-yearly moving averages. Plot the graph of actual time series and the trend values.

<table>
<thead>
<tr>
<th></th>
<th></th>
<th></th>
<th></th>
<th></th>
<th></th>
<th></th>
<th></th>
<th></th>
<th></th>
<th></th>
<th></th>
</tr>
</thead>
<tbody>
<tr>
<td>Production</td>
<td>13.5</td>
<td>14.7</td>
<td>17</td>
<td>16.2</td>
<td>18.1</td>
<td>20.4</td>
<td>22</td>
<td>21.2</td>
<td>24</td>
<td>25</td>
<td>26.6</td>
</tr>
</tbody>
</table>

24. Determine the trend for the following data giving the income (in million dollars) from the export of a product from the year 1988 to 1999. Use the 4-yearly moving average method and plot the graph of actual time series and trend values.

<table>
<thead>
<tr>
<th></th>
<th></th>
<th></th>
<th></th>
<th></th>
<th></th>
<th></th>
<th></th>
<th></th>
<th></th>
<th></th>
<th></th>
<th></th>
</tr>
</thead>
<tbody>
<tr>
<td>Income</td>
<td>340</td>
<td>360</td>
<td>385</td>
<td>470</td>
<td>430</td>
<td>444</td>
<td>452</td>
<td>473</td>
<td>490</td>
<td>534</td>
<td>541</td>
<td>576</td>
</tr>
</tbody>
</table>
25. Using the 4-yearly moving average method find the trend for the following data.

<table>
<thead>
<tr>
<th></th>
<th></th>
<th></th>
<th></th>
<th></th>
<th></th>
<th></th>
<th></th>
<th></th>
<th></th>
<th></th>
<th></th>
</tr>
</thead>
<tbody>
<tr>
<td>Value</td>
<td>102</td>
<td>100</td>
<td>103</td>
<td>105</td>
<td>104</td>
<td>109</td>
<td>112</td>
<td>115</td>
<td>113</td>
<td>119</td>
<td>117</td>
</tr>
</tbody>
</table>

26. Determine the trend for the following data giving the sales (in '00 Rs.) of a product per week for 20 weeks. Use appropriate moving average method.

<table>
<thead>
<tr>
<th>Week</th>
<th>1</th>
<th>2</th>
<th>3</th>
<th>4</th>
<th>5</th>
<th>6</th>
<th>7</th>
<th>8</th>
<th>9</th>
<th>10</th>
</tr>
</thead>
<tbody>
<tr>
<td>Sales</td>
<td>22</td>
<td>26</td>
<td>28</td>
<td>25</td>
<td>30</td>
<td>35</td>
<td>39</td>
<td>36</td>
<td>30</td>
<td>32</td>
</tr>
</tbody>
</table>

<table>
<thead>
<tr>
<th>Week</th>
<th>11</th>
<th>12</th>
<th>13</th>
<th>14</th>
<th>15</th>
<th>16</th>
<th>17</th>
<th>18</th>
<th>19</th>
<th>20</th>
</tr>
</thead>
<tbody>
<tr>
<td>Sales</td>
<td>29</td>
<td>34</td>
<td>36</td>
<td>35</td>
<td>35</td>
<td>39</td>
<td>43</td>
<td>48</td>
<td>52</td>
<td>49</td>
</tr>
</tbody>
</table>

27. An online marketing company works 5-days a week. The day-to-day total sales (in ‘000 Rs) of their product for 4 weeks are given below. Using a proper moving average method find the trend values.

<table>
<thead>
<tr>
<th>Days</th>
<th>1</th>
<th>2</th>
<th>3</th>
<th>4</th>
<th>5</th>
<th>6</th>
<th>7</th>
<th>8</th>
<th>9</th>
<th>10</th>
</tr>
</thead>
<tbody>
<tr>
<td>Sales</td>
<td>12</td>
<td>16</td>
<td>20</td>
<td>17</td>
<td>18</td>
<td>20</td>
<td>26</td>
<td>25</td>
<td>27</td>
<td>30</td>
</tr>
</tbody>
</table>

<table>
<thead>
<tr>
<th>Days</th>
<th>11</th>
<th>12</th>
<th>13</th>
<th>14</th>
<th>15</th>
<th>16</th>
<th>17</th>
<th>18</th>
<th>19</th>
<th>20</th>
</tr>
</thead>
<tbody>
<tr>
<td>Sales</td>
<td>35</td>
<td>32</td>
<td>32</td>
<td>38</td>
<td>36</td>
<td>35</td>
<td>34</td>
<td>38</td>
<td>40</td>
<td>41</td>
</tr>
</tbody>
</table>

28. Fit a straight line trend to the following data. Draw the graph of the actual time series and the trend line. Estimate the sales for the years 2000.

<table>
<thead>
<tr>
<th></th>
<th></th>
<th></th>
<th></th>
<th></th>
<th></th>
<th></th>
<th></th>
<th></th>
<th></th>
<th></th>
<th></th>
</tr>
</thead>
<tbody>
<tr>
<td>Sales in ‘000 Rs</td>
<td>45</td>
<td>47</td>
<td>49</td>
<td>48</td>
<td>54</td>
<td>58</td>
<td>53</td>
<td>59</td>
<td>62</td>
<td>60</td>
<td>64</td>
</tr>
</tbody>
</table>

29. Fit a straight line trend to the following data. Draw the graph of the actual time series and the trend line. Estimate the sales for the years 2001.

<table>
<thead>
<tr>
<th></th>
<th></th>
<th></th>
<th></th>
<th></th>
<th></th>
<th></th>
<th></th>
<th></th>
<th></th>
<th></th>
</tr>
</thead>
<tbody>
<tr>
<td>Profit in ‘000 Rs</td>
<td>76</td>
<td>79</td>
<td>82</td>
<td>84</td>
<td>81</td>
<td>84</td>
<td>89</td>
<td>92</td>
<td>88</td>
<td>90</td>
</tr>
</tbody>
</table>
30. Fit a straight line trend to the following data. Draw the graph of the actual time series and the trend line. Estimate the sales for the years 2007.

<table>
<thead>
<tr>
<th>Year</th>
<th>1998</th>
<th>1999</th>
<th>2001</th>
<th>2002</th>
<th>2003</th>
<th>2004</th>
<th>2005</th>
<th>2006</th>
</tr>
</thead>
<tbody>
<tr>
<td>Profit in ‘000 Rs</td>
<td>116</td>
<td>124</td>
<td>143</td>
<td>135</td>
<td>138</td>
<td>146</td>
<td>142</td>
<td>152</td>
</tr>
</tbody>
</table>

31. Fit a straight line trend to the following data giving the number of casualties (in hundred) of motorcyclists without helmet. Estimate the number for the year 1999.

<table>
<thead>
<tr>
<th></th>
<th></th>
<th></th>
<th></th>
<th></th>
<th></th>
<th></th>
<th></th>
</tr>
</thead>
<tbody>
<tr>
<td>No of casualties</td>
<td>12</td>
<td>14.2</td>
<td>15.2</td>
<td>16</td>
<td>18.8</td>
<td>19.6</td>
<td>22.1</td>
</tr>
</tbody>
</table>

32. Fit a straight line trend to the following data. Draw the graph of the actual time series and the trend line. Estimate the import for the years 2002

<table>
<thead>
<tr>
<th></th>
<th></th>
<th></th>
<th></th>
<th></th>
<th></th>
<th></th>
<th></th>
<th></th>
<th></th>
<th></th>
</tr>
</thead>
<tbody>
<tr>
<td>Import in’000 Rs.</td>
<td>55</td>
<td>52</td>
<td>50</td>
<td>53</td>
<td>54</td>
<td>56</td>
<td>58</td>
<td>60</td>
<td>57</td>
<td>59</td>
</tr>
</tbody>
</table>

33. Fit a straight line trend to the following data giving the price of crude oil per barrel in USD. Draw the graph of the actual time series and the trend line. Estimate the sales for the year 2003.

<table>
<thead>
<tr>
<th></th>
<th></th>
<th></th>
<th></th>
<th></th>
<th></th>
<th></th>
<th></th>
<th></th>
<th></th>
<th></th>
</tr>
</thead>
<tbody>
<tr>
<td>Price per barrel</td>
<td>98</td>
<td>102</td>
<td>104.5</td>
<td>108</td>
<td>105</td>
<td>109</td>
<td>112</td>
<td>118</td>
<td>115</td>
<td>120</td>
</tr>
</tbody>
</table>

34. Apply the method of least squares to find the number of student attending the library in the month of May of the academic year 2005 – 2006 from the following data.

<table>
<thead>
<tr>
<th>Month Students</th>
<th>Jun</th>
<th>Jul</th>
<th>Aug</th>
<th>Sep</th>
<th>Oct</th>
<th>Nov</th>
<th>Dec</th>
<th>Jan</th>
<th>Feb</th>
<th>Mar</th>
<th>Apr</th>
</tr>
</thead>
<tbody>
<tr>
<td></td>
<td>105</td>
<td>120</td>
<td>160</td>
<td>225</td>
<td>180</td>
<td>115</td>
<td>124</td>
<td>138</td>
<td>176</td>
<td>230</td>
<td>180</td>
</tr>
</tbody>
</table>

35. Assuming that the trend is absent, find the seasonal indices for the following data and also find the deseasonalized values.

<table>
<thead>
<tr>
<th>Quarters</th>
<th>I</th>
<th>II</th>
<th>III</th>
<th>IV</th>
</tr>
</thead>
<tbody>
<tr>
<td>1977</td>
<td>10</td>
<td>12</td>
<td>14</td>
<td>16</td>
</tr>
<tr>
<td>1978</td>
<td>12</td>
<td>15</td>
<td>18</td>
<td>22</td>
</tr>
<tr>
<td>1979</td>
<td>16</td>
<td>18</td>
<td>20</td>
<td>24</td>
</tr>
<tr>
<td>1980</td>
<td>24</td>
<td>26</td>
<td>28</td>
<td>34</td>
</tr>
</tbody>
</table>
36. Calculate seasonal indices for the following data:

<table>
<thead>
<tr>
<th>Year</th>
<th>I</th>
<th>II</th>
<th>III</th>
<th>IV</th>
</tr>
</thead>
<tbody>
<tr>
<td>2003</td>
<td>55</td>
<td>53</td>
<td>57</td>
<td>51</td>
</tr>
<tr>
<td>2004</td>
<td>56</td>
<td>55</td>
<td>60</td>
<td>53</td>
</tr>
<tr>
<td>2005</td>
<td>57</td>
<td>56</td>
<td>61</td>
<td>54</td>
</tr>
</tbody>
</table>

Ans. 100.59, 98.2, 106.57, 94.61

---

INDEX NUMBERS

**OBJECTIVES**

- To understand about the importance of Index Numbers.
- To understand different types of Index Numbers and their computations.
- To understand about *Real Income* and *Cost of Living Index Numbers*.
- To understand the problems in constructing Index Numbers.
- To know the merits and demerits of Index Numbers.

Every variable undergoes some changes over a period of time or in different regions or due to some factors affecting it. These changes are needed to be measured. In the last chapter we have seen how a time series helps in estimating the value of a variable in future. But the magnitude of the changes or variations of a variable, if known, are useful for many more reasons. For example, if the changes in prices of various household commodities are known, one can plan for a proper budget for them in advance. If a share broker is aware of the magnitude of fluctuations in the price of a particular share or about the trend of the market he can plan his course of action of buying or selling his shares. Thus, we can feel that there is a need of such a measure to describe the changes in prices, sales, profits, imports, exports etc, which are useful from a common man to a business organization.

Index number is an important statistical relative tool to measure the changes in a variable or group of variables with respect to time, geographical conditions and other characteristics of the variable(s). Index
number is a relative measure, as it is independent of the units of the variable(s) taken in to consideration. This is the advantage of index numbers over normal averages. All the averages which we studied before are absolute measures, i.e. they are expressed in units, while index numbers are percentage values which are independent of the units of the variable(s). In calculating an index number, a base period is considered for comparison and the changes in a variable are measured using various methods.

Though index numbers were initially used for measuring the changes in prices of certain variables, now it is used in almost every field of physical sciences, social sciences, government departments, economic bodies and business organizations. The gross national product (GNP), per capita income, cost of living index, production index, consumption, profit/loss etc every variable in economics uses this as a tool to measure the variations. Thus, the fluctuations, small or big, in the economy are measured by index numbers. Hence it is called as a barometer of economics.

14.2 IMPORTANCE OF INDEX NUMBERS

The important characteristics of Index numbers are as follows:

(1) **It is a relative measure**: As discussed earlier index numbers are independent of the units of the variable(s), hence it a special kind of average which can be used to compare different types of data expressed in different units at different points of time.

(2) **Economical Barometer**: A barometer is an instrument which measures the atmospheric pressure. As the index numbers measure all the ups and downs in the economy they are hence called as the economic barometers.

(3) **To generalize the characteristics of a group**: Many a time it is difficult to measure the changes in a variable in complete sense. For example, it is not possible to directly measure the changes in a business activity in a country. But instead if we measure the changes in the factors affecting the business activity, we can generalize it to the complete activity. Similarly the industrial production or the agricultural output cannot be measured directly.

(4) **To forecast trends**: Index numbers prove to be very useful in identifying trends in a variable over a period of time and hence are used to forecast the future trends.

(5) **To facilitate decision making**: Future estimations are always used for long term and short term planning and formulating a policy for the future by government and private organizations. Price Index numbers provide
the requisite for such policy decisions in economics.

(6) To measure the purchasing power of money and useful in deflating:
Index numbers help in deciding the actual purchasing power of money. We often hear from our elders saying that “In our times the salary was just Rs. 100 a month and you are paid Rs. 10,000, still you are not happy!” The answer is simple (because of index numbers!) that the money value of Rs. 100, 30 years before and now is drastically different. Calculation of real income using index numbers is an important tool to measure the actual income of an individual. This is called as deflation.

There are different types of index numbers based on their requirement like, price index, quantity index, value index etc. The price index is again classified as single price index and composite price index.

14.3 PRICE INDEX NUMBERS

14.3.1 Simple (Unweighted) Price Index Number By Aggregative Method

In this method we define the price index number as the ratio of sum of prices in current year to sum of prices in base year and express it in percentage. i.e. multiply the quotient by 100.

Symbolically, 

$$ I = \frac{\sum P_1}{\sum P_0} \times 100 \quad \ldots (1) $$

Notations:
- $P_0$: Price in Base Year
- $Q_0$: Quantity in Base Year
- $P_1$: Price in Current Year
- $Q_1$: Quantity in Current Year

The suffix ‘0’ stands for the base year and the suffix ‘1’ stands for the current year.
Steps for computation:
1. The total of all base year prices is calculated and denoted by $\Sigma P_0$.
2. The total of all current year prices is calculated and denoted by $\Sigma P_1$.
3. Using the above formula, simple price index number is computed.

Example 1
For the following data, construct the price index number by simple aggregative method:

<table>
<thead>
<tr>
<th>Commodity</th>
<th>Unit</th>
<th>Price in</th>
<th>1985</th>
<th>1986</th>
</tr>
</thead>
<tbody>
<tr>
<td>A</td>
<td>Kg</td>
<td></td>
<td>10</td>
<td>12</td>
</tr>
<tr>
<td>B</td>
<td>Kg</td>
<td></td>
<td>4</td>
<td>7</td>
</tr>
<tr>
<td>C</td>
<td>Litre</td>
<td></td>
<td>6</td>
<td>7</td>
</tr>
<tr>
<td>D</td>
<td>Litre</td>
<td></td>
<td>8</td>
<td>10</td>
</tr>
</tbody>
</table>

Solution: Following the steps for computing the index number, we find the totals of the 3rd and 4th columns as shown below:

<table>
<thead>
<tr>
<th>Commodity</th>
<th>Unit</th>
<th>1985</th>
<th>1986</th>
</tr>
</thead>
<tbody>
<tr>
<td>A</td>
<td>Kg</td>
<td>10</td>
<td>12</td>
</tr>
<tr>
<td>B</td>
<td>Kg</td>
<td>4</td>
<td>7</td>
</tr>
<tr>
<td>C</td>
<td>Litre</td>
<td>6</td>
<td>7</td>
</tr>
<tr>
<td>D</td>
<td>Litre</td>
<td>8</td>
<td>10</td>
</tr>
</tbody>
</table>

$\therefore I = \frac{\Sigma P_1}{\Sigma P_0} \times 100 = \frac{36}{28} \times 100 = 128.57$

Meaning of the value of $I$:
$I = 128.57$ means that the prices in 1986, as compared with that in 1985 have increased by 28.57%.

14.3.2 Simple (Unweighted) Price Index Number by Average of Price Relatives Method
In this method the price index is calculated for every commodity and
its arithmetic mean is taken, \( i.e. \) the sum of all price relative is divided by the total number of commodities.

Symbolically, if there are \( n \) commodities in to consideration, then the simple price index number of the group is calculated by the formula:

\[
I = \frac{1}{n} \sum \left( \frac{P_1}{P_0} \times 100 \right) \quad \text{… (2)}
\]

Steps for computation

1. The price relatives for each commodity are calculated by the formula:
\[
\frac{P_1}{P_0} \times 100
\]
2. The total of these price relatives is calculated and denoted as:
\[
\sum \left( \frac{P_1}{P_0} \times 100 \right)
\]
3. The arithmetic mean of the price relatives using the above formula no. (2) gives the required price index number.

Example 2

Construct the simple price index number for the following data using average of price relatives method:

<table>
<thead>
<tr>
<th>Commodity</th>
<th>Unit</th>
<th>Price in</th>
</tr>
</thead>
<tbody>
<tr>
<td></td>
<td></td>
<td>1997</td>
</tr>
<tr>
<td>Rice</td>
<td>Kg</td>
<td>10</td>
</tr>
<tr>
<td>Wheat</td>
<td>Kg</td>
<td>6</td>
</tr>
<tr>
<td>Milk</td>
<td>Litre</td>
<td>8</td>
</tr>
<tr>
<td>Oil</td>
<td>Litre</td>
<td>15</td>
</tr>
</tbody>
</table>

Solution: In this method we have to find price relatives for every commodity and then total these price relatives. Following the steps for computing as mentioned above, we introduce first, the column of price relatives. The table of computation is as follows:

<table>
<thead>
<tr>
<th>Commodity</th>
<th>Unit</th>
<th>Price in</th>
<th>( \frac{P_1}{P_0} \times 100 )</th>
</tr>
</thead>
<tbody>
<tr>
<td></td>
<td></td>
<td>1997((P_0 ))</td>
<td>1998((P_1 ))</td>
</tr>
<tr>
<td>Rice</td>
<td>Kg</td>
<td>10</td>
<td>13</td>
</tr>
<tr>
<td>Wheat</td>
<td>Kg</td>
<td>6</td>
<td>8</td>
</tr>
<tr>
<td>Milk</td>
<td>Litre</td>
<td>8</td>
<td>10</td>
</tr>
</tbody>
</table>
Now, \( n = 4 \) and the total of price relatives is 508.33

\[
I = \frac{1}{n} \sum \left( \frac{P_1}{P_0} \times 100 \right) = \frac{508.33}{4} = 127.08
\]

The prices in 1998 have increased by 27% as compared with in 1997.

**Remark:**

1. The simple aggregative method is calculated without taking into consideration the units of individual items in the group. This may give a misleading index number.
2. This problem is overcome in the average of price relatives method, as the individual price relatives are computed first and then their average is taken.
3. Both the methods are unreliable as they give equal weightage to all items in consideration which is not true practically.

### 14.3.3 Weighted Index Numbers by Aggregative Method

In this method weights assigned to various items are considered in the calculations. The products of the prices with the corresponding weights are computed; their totals are divided and expressed in percentages.

Symbolically, if \( W \) denotes the weights assigned and \( P_0, P_1 \) have their usual meaning, then the weighted index number using aggregative method is given by the formula:

\[
I = \frac{\sum P_1 W}{\sum P_0 W} \times 100 \tag{3}
\]

**Steps to find weighted index number using aggregative method**

1. The columns of \( P_1 W \) and \( P_0 W \) are introduced.
2. The totals of these columns are computed.
3. The formula no. (3) is used for computing the required index number.

**Example 3**

From the following data, construct the weighted price index number:

<table>
<thead>
<tr>
<th>Commodity</th>
<th>A</th>
<th>B</th>
<th>C</th>
<th>D</th>
</tr>
</thead>
<tbody>
<tr>
<td>Price in 1982</td>
<td>6</td>
<td>10</td>
<td>4</td>
<td>18</td>
</tr>
<tr>
<td>Price in 1983</td>
<td>9</td>
<td>18</td>
<td>6</td>
<td>26</td>
</tr>
<tr>
<td>Weight</td>
<td>35</td>
<td>30</td>
<td>20</td>
<td>15</td>
</tr>
</tbody>
</table>

| Commodity | Weight \((W)\) | Price in 1982 \((P_0)\) | \(P_0 W\) | Price in 1983 \((P_1)\) | \(P_1 W\) |
Solution: Following the steps mentioned above, the table of computations is as follows:

Using the totals from the table, we have

\[
\text{Weighted Index Number } I = \frac{\sum P_i W}{\sum P_0 W} \times 100 = \frac{1365}{860} \times 100 = 158.72
\]

Remark:
There are different formulae based on what to be taken as the weight while calculating the weighted index numbers. Based on the choice of the weight we are going to study here three types of weighted index numbers: (1) Laspeyre’s Index Number, (2) Paasche’s Index Number and (3) Fisher’s Index Number.

(1) Laspeyre’s Index Number:
In this method, Laspeyre assumed the base quantity \(Q_0\) as the weight in constructing the index number. Symbolically, \(P_0, P_1\) and \(Q_0\) having their usual meaning, the Laspeyre’s index number denoted by \(I_L\) is given by the formula:

\[
I_L = \frac{\sum P_1 Q_0}{\sum P_0 Q_0} \times 100 \quad \text{... (4)}
\]

Steps to compute \(I_L\):
1. The columns of the products \(P_0 Q_0\) and \(P_1 Q_0\) are introduced.
2. The totals of these columns are computed.
3. Using the above formula no. (4), \(I_L\) is computed.

Example 4
From the data given below, construct the Laspeyre’s index number:

<table>
<thead>
<tr>
<th>Commodity</th>
</tr>
</thead>
<tbody>
<tr>
<td></td>
</tr>
<tr>
<td>A</td>
</tr>
<tr>
<td>B</td>
</tr>
<tr>
<td>C</td>
</tr>
</tbody>
</table>

<table>
<thead>
<tr>
<th>Commodity</th>
<th>1965</th>
<th>1966</th>
</tr>
</thead>
<tbody>
<tr>
<td></td>
<td>Price</td>
<td>Quantity</td>
</tr>
<tr>
<td>A</td>
<td>5</td>
<td>12</td>
</tr>
<tr>
<td>B</td>
<td>7</td>
<td>12</td>
</tr>
<tr>
<td>C</td>
<td>10</td>
<td>15</td>
</tr>
</tbody>
</table>
Solution: Introducing the columns of the products $P_0Q_0$ and $P_1Q_0$, the table of computation is completed as shown below:

<table>
<thead>
<tr>
<th>Commodity</th>
<th>1965</th>
<th>1966</th>
<th>$P_0Q_0$</th>
<th>$P_1Q_0$</th>
</tr>
</thead>
<tbody>
<tr>
<td></td>
<td>Price ($P_0$)</td>
<td>Quantity ($Q_0$)</td>
<td>Price ($P_1$)</td>
<td></td>
</tr>
<tr>
<td>A</td>
<td>5</td>
<td>12</td>
<td>7</td>
<td>60</td>
</tr>
<tr>
<td>B</td>
<td>7</td>
<td>12</td>
<td>9</td>
<td>84</td>
</tr>
<tr>
<td>C</td>
<td>10</td>
<td>15</td>
<td>15</td>
<td>150</td>
</tr>
<tr>
<td>D</td>
<td>18</td>
<td>5</td>
<td>20</td>
<td>90</td>
</tr>
<tr>
<td>Total</td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
</tbody>
</table>

Using the totals from the table and substituting in the formula no. (4), we have

$$I_L = \frac{\sum P_1 Q_0}{\sum P_0 Q_0} \times 100 = \frac{517}{384} \times 100 = 134.64$$

(2) Paasche’s Index Number:

In this method, Paasch assumed the current year quantity ($Q_1$) as the weight for constructing the index number. Symbolically, $P_0$, $P_1$ and $Q_1$ having their usual meaning, the Paasche’s index number denoted by $I_p$ is given by the formula:

$$I_p = \frac{\sum P_1 Q_1}{\sum P_0 Q_0} \times 100 \tag{5}$$

The steps for computing $I_p$ are similar to that of $I_L$.

Example 5

From the data given below, construct the Paasche’s index number:

<table>
<thead>
<tr>
<th>Commodity</th>
<th>1985</th>
<th>1986</th>
</tr>
</thead>
<tbody>
<tr>
<td></td>
<td>Price</td>
<td>Price</td>
</tr>
<tr>
<td>A</td>
<td>5</td>
<td>8</td>
</tr>
<tr>
<td>B</td>
<td>10</td>
<td>14</td>
</tr>
<tr>
<td>C</td>
<td>6</td>
<td>9</td>
</tr>
<tr>
<td>D</td>
<td>8</td>
<td>10</td>
</tr>
</tbody>
</table>
Solution: Introducing the columns of the products $P_0Q_1$ and $P_1Q_1$, the table of computations is completed as shown below:

<table>
<thead>
<tr>
<th>Commodity</th>
<th>1985</th>
<th>1986</th>
<th>$P_0Q_1$</th>
<th>$P_1Q_1$</th>
</tr>
</thead>
<tbody>
<tr>
<td></td>
<td>Price ($P_0$)</td>
<td>Price ($P_1$)</td>
<td>Quantity ($Q_1$)</td>
<td></td>
</tr>
<tr>
<td>A</td>
<td>5</td>
<td>8</td>
<td>10</td>
<td>50</td>
</tr>
<tr>
<td>B</td>
<td>10</td>
<td>14</td>
<td>20</td>
<td>200</td>
</tr>
<tr>
<td>C</td>
<td>6</td>
<td>9</td>
<td>25</td>
<td>150</td>
</tr>
<tr>
<td>D</td>
<td>8</td>
<td>10</td>
<td>10</td>
<td>80</td>
</tr>
<tr>
<td>Total</td>
<td>-</td>
<td>-</td>
<td>-</td>
<td>$\Sigma P_0Q_1 = 480$</td>
</tr>
</tbody>
</table>

Using the totals from the table and substituting in the formula no. (5), we have: $I_P = \frac{\Sigma P_0Q_1}{\Sigma P_1Q_1} \times 100 = \frac{685}{480} \times 100 = 142.71$

(3) Fisher's Index Number:
Fisher developed his own method by using the formulae of Laspeyre and Paasche. He defined the index number as the geometric mean of $I_L$ and $I_P$. Symbolically, the Fisher’s Index number denoted as $I_F$ is given by the formula:

$$I_F = \sqrt{I_L \times I_P} = \sqrt{\frac{\Sigma P_0Q_1}{\Sigma P_0Q_0} \times \frac{\Sigma P_1Q_1}{\Sigma P_1Q_1}} \times 100 \quad \text{.. (6)}$$

Note:
1. The multiple 100 is outside the square root sign.
2. While computing products of the terms, care should be taken to multiply corresponding numbers properly.

Example 6
From the following data given below, construct the (i) Laspeyre’s index number, (ii) Paasche’s index number and hence (iii) Fisher’s index number.

<table>
<thead>
<tr>
<th>Item</th>
<th>1975</th>
<th>1976</th>
</tr>
</thead>
<tbody>
<tr>
<td></td>
<td>Price</td>
<td>Quantity</td>
</tr>
<tr>
<td>A</td>
<td>4</td>
<td>12</td>
</tr>
<tr>
<td>B</td>
<td>2</td>
<td>16</td>
</tr>
<tr>
<td>C</td>
<td>8</td>
<td>9</td>
</tr>
</tbody>
</table>
Solution: Introducing four columns of the products of $P_0Q_0$, $P_0Q_1$, $P_1Q_0$, and $P_1Q_1$, the table of computations is completes as shown below:

<table>
<thead>
<tr>
<th>Item</th>
<th>$P_0$</th>
<th>$Q_0$</th>
<th>$P_1$</th>
<th>$Q_1$</th>
<th>$P_0Q_0$</th>
<th>$P_0Q_1$</th>
<th>$P_1Q_0$</th>
<th>$P_1Q_1$</th>
</tr>
</thead>
<tbody>
<tr>
<td>A</td>
<td>4</td>
<td>12</td>
<td>6</td>
<td>16</td>
<td>48</td>
<td>64</td>
<td>72</td>
<td>96</td>
</tr>
<tr>
<td>B</td>
<td>2</td>
<td>16</td>
<td>3</td>
<td>20</td>
<td>32</td>
<td>40</td>
<td>48</td>
<td>60</td>
</tr>
<tr>
<td>C</td>
<td>8</td>
<td>9</td>
<td>11</td>
<td>14</td>
<td>72</td>
<td>112</td>
<td>99</td>
<td>154</td>
</tr>
<tr>
<td>Total</td>
<td></td>
<td></td>
<td></td>
<td></td>
<td>152</td>
<td>216</td>
<td>219</td>
<td>310</td>
</tr>
</tbody>
</table>

From the table, we have $\Sigma P_0Q_0 = 152$, $\Sigma P_0Q_1 = 216$, $\Sigma P_1Q_0 = 219$ and $\Sigma P_1Q_1 = 310$

\[ I_L = \frac{\Sigma P_1Q_0}{\Sigma P_0Q_0} \times 100 = \frac{219}{152} \times 100 = 144.08 \]

\[ I_P = \frac{\Sigma P_1Q_1}{\Sigma P_0Q_1} \times 100 = \frac{310}{216} \times 100 = 143.52 \]

\[ I_F = \sqrt{I_L \times I_P} = \sqrt{144.08 \times 143.52} = 143.8 \]

Remark:
1. Laspeyre’s index number though popular has a drawback that it does not consider the change in consumption over a period. (as it does not take into account the current quantity).
2. Paasche’s index number overcomes this by assigning the current year quantity as weight.
3. Fisher’s index number being the geometric mean of both these index numbers, it considers both the quantities. Hence it is called as the ideal index number.

Example 7
From the following data given below, construct the Kelly’s index number:

<table>
<thead>
<tr>
<th>Item</th>
<th>Base Year</th>
<th>Current Year</th>
</tr>
</thead>
<tbody>
<tr>
<td></td>
<td>Price</td>
<td>Quantity</td>
</tr>
<tr>
<td>A</td>
<td>18</td>
<td>20</td>
</tr>
<tr>
<td>B</td>
<td>9</td>
<td>10</td>
</tr>
<tr>
<td>C</td>
<td>10</td>
<td>15</td>
</tr>
<tr>
<td>D</td>
<td>6</td>
<td>13</td>
</tr>
<tr>
<td>E</td>
<td>32</td>
<td>14</td>
</tr>
</tbody>
</table>

Solution: Introducing the columns of $Q = \frac{Q_0 + Q_1}{2}$, $P_0Q$ and $P_1Q$, the table of computations is completed as shown blow:
From the table, we have $\sum P_0Q = 1261$ and $\sum P_1Q = 1597$

\[ I_K = \frac{\sum P_0Q}{\sum P_0Q} \times 100 = \frac{1597}{1261} \times 100 = 126.65 \]

14.3.4 Weighted Index Numbers using average of price relatives method

This is similar to what we have seen in subsection 7.3.2. Here the individual price relatives are computed first. These are multiplied with the corresponding weights. The ratio of the sum of the products and the total value of the weight is defined to be the weighted index number.

Symbolically, if $W$ denotes the weights and $I$ denote the price relatives then the weighted index number is given by the formula:

\[ IW \]

One of the important weighted index number is the *cost of living index number*, also known as the *consumer price index (CPI) number*.

14.4 Cost Of Living Index Number Or Consumer Price Index Number

There are two methods for constructing this index number:

(1) Aggregative expenditure method and (2) Family Budget Method

(1) In aggregative expenditure method we construct the index number by taking the base year quantity as the weight. In fact this index number is nothing but the Laspeyre’s index number.

(2) In family budget method, value weights are computed for each item in the group and the index number is computed using the formula:

\[ \frac{\sum IW}{\sum W} \text{, where } I = \frac{P_1}{P_0} \times 100 \text{ and } W = P_0Q_0 \]

Example 8
A survey of families in a city revealed the following information:

<table>
<thead>
<tr>
<th>Item</th>
<th>Food</th>
<th>Clothing</th>
<th>Fuel</th>
<th>House Rent</th>
<th>Misc.</th>
</tr>
</thead>
<tbody>
<tr>
<td>% Expenditure</td>
<td>30</td>
<td>20</td>
<td>15</td>
<td>20</td>
<td>15</td>
</tr>
<tr>
<td>Price in 1987</td>
<td>320</td>
<td>140</td>
<td>100</td>
<td>250</td>
<td>300</td>
</tr>
<tr>
<td>Price in 1988</td>
<td>400</td>
<td>150</td>
<td>125</td>
<td>250</td>
<td>320</td>
</tr>
</tbody>
</table>

What is the cost of living index number for 1988 as compared to that of 1987?

Solution: Here % expenditure is taken as the weight \((W)\). The table of computations is as shown below:

<table>
<thead>
<tr>
<th>Item</th>
<th>(P_0)</th>
<th>(P_1)</th>
<th>(I = \frac{P_1}{P_0} \times 100)</th>
<th>% Expenditure ((W))</th>
<th>(IW)</th>
</tr>
</thead>
<tbody>
<tr>
<td>Food</td>
<td>320</td>
<td>400</td>
<td>125</td>
<td>30</td>
<td>3750</td>
</tr>
<tr>
<td>Clothing</td>
<td>140</td>
<td>150</td>
<td>107.14</td>
<td>20</td>
<td>2142.8</td>
</tr>
<tr>
<td>Fuel</td>
<td>100</td>
<td>125</td>
<td>125</td>
<td>15</td>
<td>1875</td>
</tr>
<tr>
<td>House Rent</td>
<td>250</td>
<td>250</td>
<td>100</td>
<td>20</td>
<td>2000</td>
</tr>
<tr>
<td>Miscellaneous</td>
<td>300</td>
<td>320</td>
<td>106.67</td>
<td>15</td>
<td>1600.05</td>
</tr>
<tr>
<td><strong>Total</strong></td>
<td></td>
<td></td>
<td></td>
<td>(\sum W = 100)</td>
<td>11367.85</td>
</tr>
</tbody>
</table>

From the table, we have \(\sum W = 100\) and \(\sum IW = 11376.85\)

\[
\text{cost of living index number} = \frac{\sum IW}{\sum W} = \frac{11376.85}{100} = 113.7685
\]

14.5 Use Of Cost of Living Index Numbers

1. These index numbers reflect the effect of rise and fall in the economy or change in prices over the standard of living of the people.
2. These index numbers help in determining the purchasing power of money which is the reciprocal of the cost of living index number.
3. It is used in deflation. \(i.e.\) determining the actual income of an individual. Hence it also used by the management of government or private organizations to formulate their policies regarding the wages, allowance to their employees.

14.6 REAL INCOME
As discussed earlier in this chapter, index numbers are very useful in finding the real income of an individual or a group of them, which facilitates the different managements to decide their wage policies. The process of measuring the actual income vis-à-vis the changes in prices is called as deflation.

The formula for computing the real income is as follows:

Real Income of a year = \( \frac{\text{Money Income for the year}}{\text{Price Index of that year}} \times 100 \)

**Example 12**

Calculate the real income for the following data:

<table>
<thead>
<tr>
<th>Year</th>
<th>Income in Rs.</th>
<th>Price Index</th>
<th>Real Income</th>
</tr>
</thead>
<tbody>
<tr>
<td>1990</td>
<td>800</td>
<td>100</td>
<td>800</td>
</tr>
<tr>
<td>1991</td>
<td>1050</td>
<td>105</td>
<td>1050</td>
</tr>
<tr>
<td>1992</td>
<td>1200</td>
<td>115</td>
<td>1200</td>
</tr>
<tr>
<td>1993</td>
<td>1600</td>
<td>125</td>
<td>1600</td>
</tr>
<tr>
<td>1994</td>
<td>2500</td>
<td>130</td>
<td>2500</td>
</tr>
<tr>
<td>1995</td>
<td>2800</td>
<td>140</td>
<td>2800</td>
</tr>
</tbody>
</table>

**Solution:** The real income is calculated by the formula:

\[ \text{real income} = \frac{\text{Money Income for the year}}{\text{Price Index of that year}} \times 100 \]

The table of computation of real income’s is completed as shown below:

<table>
<thead>
<tr>
<th>Year</th>
<th>Income in Rs.</th>
<th>Price Index</th>
<th>Real Income</th>
</tr>
</thead>
<tbody>
<tr>
<td>1990</td>
<td>800</td>
<td>100</td>
<td>800</td>
</tr>
<tr>
<td>1991</td>
<td>1050</td>
<td>105</td>
<td>1050</td>
</tr>
<tr>
<td>1992</td>
<td>1200</td>
<td>115</td>
<td>1200</td>
</tr>
<tr>
<td>1993</td>
<td>1600</td>
<td>125</td>
<td>1600</td>
</tr>
<tr>
<td>1994</td>
<td>2500</td>
<td>130</td>
<td>2500</td>
</tr>
<tr>
<td>1995</td>
<td>2800</td>
<td>140</td>
<td>2800</td>
</tr>
</tbody>
</table>

14.7 Demerits Of Index Numbers

(1) There are numerous types and methods of constructing index numbers. If an appropriate method is not applied it may lead to wrong conclusions.

(2) The sample selection may not be representative of the complete series of items.

(3) The base period selection also is personalized and hence may be biased.

(4) Index number is a quantitative measure and does not take into
account the qualitative aspect of the items.

(5) Index numbers are approximations of the changes, they may not accurate.

Check Your Progress

1. Define Index Numbers.
2. Write a short note on the importance of Index Numbers.
3. "Index Numbers are the Economical barometers". Discuss this statement with examples.
4. Discuss the steps to construct Index Numbers.
5. What are the problems in constructing an Index Number?
6. Define Cost of Living Index Number and explain its importance.
7. What do you mean by (i) Chain Based Index Number and (ii) Fixed Base Index Number? Distinguish between the two.
8. Define (i) Laspeyre’s Index Number, (ii) Paasche’s Index Number and (iii) Fisher’s Index Number. What is the difference between the three? Which amongst them is called as the ideal Index Number? Why?
9. What are the demerits of Index Numbers?
10. From the following data, construct the price index number by simple aggregative method:

<table>
<thead>
<tr>
<th>Commodity</th>
<th>Unit</th>
<th>Price in 1990</th>
<th>1991</th>
</tr>
</thead>
<tbody>
<tr>
<td>A</td>
<td>Kg</td>
<td>14</td>
<td>18</td>
</tr>
<tr>
<td>B</td>
<td>Kg</td>
<td>6</td>
<td>9</td>
</tr>
<tr>
<td>C</td>
<td>Litre</td>
<td>5</td>
<td>8</td>
</tr>
<tr>
<td>D</td>
<td>Litre</td>
<td>12</td>
<td>20</td>
</tr>
</tbody>
</table>

Ans: 148.65

11. From the following data, construct the price index number for 1995, by simple aggregative method, with 1994 as the base:

<table>
<thead>
<tr>
<th>Commodity</th>
<th>Unit</th>
<th>Price in 1994</th>
<th>1995</th>
</tr>
</thead>
<tbody>
<tr>
<td>Rice</td>
<td>Kg</td>
<td>8</td>
<td>10</td>
</tr>
<tr>
<td>Wheat</td>
<td>Kg</td>
<td>5</td>
<td>6.5</td>
</tr>
<tr>
<td>Oil</td>
<td>Litre</td>
<td>10</td>
<td>13</td>
</tr>
<tr>
<td>Eggs</td>
<td>Dozen</td>
<td>4</td>
<td>6</td>
</tr>
</tbody>
</table>
12. From the following data, construct the price index number for 1986, by average of price relatives method:

<table>
<thead>
<tr>
<th>Commodity</th>
<th>Unit</th>
<th>Price in</th>
</tr>
</thead>
<tbody>
<tr>
<td></td>
<td></td>
<td>1985</td>
</tr>
<tr>
<td>Banana</td>
<td>Dozen</td>
<td>4</td>
</tr>
<tr>
<td>Rice</td>
<td>Kg</td>
<td>5</td>
</tr>
<tr>
<td>Milk</td>
<td>Litre</td>
<td>3</td>
</tr>
<tr>
<td>Slice Bread</td>
<td>One Packet</td>
<td>3</td>
</tr>
</tbody>
</table>

Ans: 131.48

13. From the following data, construct the price index number, by method of average of price relatives:

<table>
<thead>
<tr>
<th>Commodity</th>
<th>Unit</th>
<th>Price in</th>
</tr>
</thead>
<tbody>
<tr>
<td></td>
<td></td>
<td>1988</td>
</tr>
<tr>
<td>A</td>
<td>Kg</td>
<td>6</td>
</tr>
<tr>
<td>B</td>
<td>Kg</td>
<td>4</td>
</tr>
<tr>
<td>C</td>
<td>Kg</td>
<td>10</td>
</tr>
<tr>
<td>D</td>
<td>Litre</td>
<td>8</td>
</tr>
<tr>
<td>E</td>
<td>Litre</td>
<td>12</td>
</tr>
</tbody>
</table>

Ans: 132.08

14. From the following data, construct the price index number for 1998, by (i) simple aggregative method and (ii) simple average of price relatives method, with 1995 as the base:

<table>
<thead>
<tr>
<th>Commodity</th>
<th>Unit</th>
<th>Price in</th>
</tr>
</thead>
<tbody>
<tr>
<td></td>
<td></td>
<td>1995</td>
</tr>
<tr>
<td>Rice</td>
<td>Kg</td>
<td>12</td>
</tr>
<tr>
<td>Wheat</td>
<td>Kg</td>
<td>8</td>
</tr>
<tr>
<td>Jowar</td>
<td>Kg</td>
<td>7</td>
</tr>
<tr>
<td>Pulses</td>
<td>Kg</td>
<td>10</td>
</tr>
</tbody>
</table>

Ans: (i) 124.32, (ii) 125.06
15. From the following data, construct the weighted price index number:

<table>
<thead>
<tr>
<th>Commodity</th>
<th>A</th>
<th>B</th>
<th>C</th>
<th>D</th>
</tr>
</thead>
<tbody>
<tr>
<td>Price in 1985</td>
<td>10</td>
<td>18</td>
<td>36</td>
<td>8</td>
</tr>
<tr>
<td>Price in 1986</td>
<td>12</td>
<td>24</td>
<td>40</td>
<td>10</td>
</tr>
<tr>
<td>Weight</td>
<td>40</td>
<td>25</td>
<td>15</td>
<td>20</td>
</tr>
</tbody>
</table>

Ans: 121.29

16. From the following data, construct the index number using (i) simple average of price relatives and (ii) weighted average of price relatives:

<table>
<thead>
<tr>
<th>Commodity</th>
<th>Weight</th>
<th>Price in 1988</th>
<th>Price in 1990</th>
</tr>
</thead>
<tbody>
<tr>
<td>Rice</td>
<td>4</td>
<td>8</td>
<td>10</td>
</tr>
<tr>
<td>Wheat</td>
<td>2</td>
<td>6</td>
<td>8</td>
</tr>
<tr>
<td>Pulses</td>
<td>3</td>
<td>8</td>
<td>11</td>
</tr>
<tr>
<td>Oil</td>
<td>5</td>
<td>12</td>
<td>15</td>
</tr>
</tbody>
</table>

Ans: (i) 130.13, (ii) 128.82

17. From the data given below, construct the Laspeyre’s index number:

<table>
<thead>
<tr>
<th>Commodity</th>
<th>1975</th>
<th>1976</th>
</tr>
</thead>
<tbody>
<tr>
<td></td>
<td>Price</td>
<td>Quantity</td>
</tr>
<tr>
<td>A</td>
<td>5</td>
<td>10</td>
</tr>
<tr>
<td>B</td>
<td>6</td>
<td>15</td>
</tr>
<tr>
<td>C</td>
<td>2</td>
<td>20</td>
</tr>
<tr>
<td>D</td>
<td>10</td>
<td>14</td>
</tr>
</tbody>
</table>

Ans: 131.40

18. From the data given below, construct the Paasche’s index number:

<table>
<thead>
<tr>
<th>Commodity</th>
<th>1980</th>
<th>1985</th>
</tr>
</thead>
<tbody>
<tr>
<td></td>
<td>Price</td>
<td>Price</td>
</tr>
<tr>
<td>A</td>
<td>4</td>
<td>7</td>
</tr>
<tr>
<td>B</td>
<td>14</td>
<td>22</td>
</tr>
</tbody>
</table>
19. From the following data given below, construct the (i) Laspeyre’s index number, (ii) Paasche’s index number and hence (iii) Fisher’s index number.

<table>
<thead>
<tr>
<th>Commodity</th>
<th>1980</th>
<th>1990</th>
</tr>
</thead>
<tbody>
<tr>
<td></td>
<td>Price</td>
<td>Quantity</td>
</tr>
<tr>
<td>A</td>
<td>6</td>
<td>15</td>
</tr>
<tr>
<td>B</td>
<td>4</td>
<td>18</td>
</tr>
<tr>
<td>C</td>
<td>2</td>
<td>32</td>
</tr>
<tr>
<td>D</td>
<td>7</td>
<td>20</td>
</tr>
</tbody>
</table>

Ans: (i) 203.83, (ii) 203.37, (iii) 203.60

20. From the following data given below, construct the (i) Laspeyre’s index number, (ii) Paasche’s index number and (iii) Fisher’s index number.

<table>
<thead>
<tr>
<th>Commodity</th>
<th>Base Year</th>
<th>Current Year</th>
</tr>
</thead>
<tbody>
<tr>
<td></td>
<td>Price</td>
<td>Quantity</td>
</tr>
<tr>
<td>Cement</td>
<td>140</td>
<td>200</td>
</tr>
<tr>
<td>Steel</td>
<td>60</td>
<td>150</td>
</tr>
<tr>
<td>Coal</td>
<td>74</td>
<td>118</td>
</tr>
<tr>
<td>Limestone</td>
<td>35</td>
<td>50</td>
</tr>
</tbody>
</table>

Ans: (i) 103.98, (ii) 127.4, (iii) 115.09

21. From the following data given below, construct the Fisher’s index number:

<table>
<thead>
<tr>
<th>Commodity</th>
<th>Base Year</th>
<th>Current Year</th>
</tr>
</thead>
<tbody>
<tr>
<td></td>
<td>Price</td>
<td>Quantity</td>
</tr>
<tr>
<td>A</td>
<td>2</td>
<td>8</td>
</tr>
<tr>
<td>B</td>
<td>6</td>
<td>14</td>
</tr>
<tr>
<td>C</td>
<td>8.5</td>
<td>10</td>
</tr>
</tbody>
</table>
22. From the following data, construct the aggregative price index numbers by taking the average price of the three years as base.

<table>
<thead>
<tr>
<th></th>
<th></th>
<th></th>
<th></th>
</tr>
</thead>
<tbody>
<tr>
<td>A</td>
<td>10</td>
<td>12</td>
<td>16</td>
</tr>
<tr>
<td>B</td>
<td>16</td>
<td>19</td>
<td>25</td>
</tr>
<tr>
<td>C</td>
<td>5</td>
<td>7</td>
<td>10</td>
</tr>
</tbody>
</table>

Ans: 81.58, 100, 134.21

23. From the following data, construct the price index number by taking the price in 1978 as the base price using aggregative method:

<table>
<thead>
<tr>
<th>Commodity</th>
<th>Price in 1978</th>
<th>Price in 1979</th>
<th>Price in 1980</th>
</tr>
</thead>
<tbody>
<tr>
<td>A</td>
<td>16</td>
<td>18</td>
<td>24</td>
</tr>
<tr>
<td>B</td>
<td>4</td>
<td>6</td>
<td>7.5</td>
</tr>
<tr>
<td>C</td>
<td>11</td>
<td>15</td>
<td>19</td>
</tr>
<tr>
<td>D</td>
<td>20</td>
<td>28</td>
<td>30</td>
</tr>
</tbody>
</table>

Ans: 131.37, 120.15

24. From the following data, construct the price index number by taking the price in 1998 as the base price:

<table>
<thead>
<tr>
<th></th>
<th></th>
<th></th>
<th></th>
<th></th>
</tr>
</thead>
<tbody>
<tr>
<td>A</td>
<td>3</td>
<td>12</td>
<td>15</td>
<td>20</td>
</tr>
<tr>
<td>B</td>
<td>1</td>
<td>8</td>
<td>9</td>
<td>11</td>
</tr>
<tr>
<td>C</td>
<td>4</td>
<td>16</td>
<td>20</td>
<td>25</td>
</tr>
<tr>
<td>D</td>
<td>2</td>
<td>15</td>
<td>18</td>
<td>22</td>
</tr>
</tbody>
</table>

25. From the following data, construct (i) \( I_u \), (ii) \( I_p \), (iii) \( I_f \)

<table>
<thead>
<tr>
<th>Commodity</th>
<th>1969</th>
<th>1970</th>
</tr>
</thead>
<tbody>
<tr>
<td>Item</td>
<td>Quantity</td>
<td>Price in</td>
</tr>
<tr>
<td>--------</td>
<td>----------</td>
<td>----------</td>
</tr>
<tr>
<td></td>
<td></td>
<td>1975</td>
</tr>
<tr>
<td>Rice</td>
<td>2</td>
<td>10</td>
</tr>
<tr>
<td>Wheat</td>
<td>1.5</td>
<td>8</td>
</tr>
<tr>
<td>Jowar</td>
<td>1</td>
<td>6</td>
</tr>
<tr>
<td>Bajra</td>
<td>1.2</td>
<td>5</td>
</tr>
<tr>
<td>Pulses</td>
<td>4</td>
<td>14</td>
</tr>
</tbody>
</table>

Ans: 144.4, 144.56, 144.28

26. Construct the cost of living index number for 1980 using the Family Budget Method:

<table>
<thead>
<tr>
<th>Item</th>
<th>Quantity</th>
<th>Price in</th>
<th>Price in</th>
</tr>
</thead>
<tbody>
<tr>
<td></td>
<td></td>
<td>1975</td>
<td>1980</td>
</tr>
<tr>
<td>A</td>
<td>10</td>
<td>5</td>
<td>7</td>
</tr>
<tr>
<td>B</td>
<td>5</td>
<td>8</td>
<td>11</td>
</tr>
<tr>
<td>C</td>
<td>7</td>
<td>12</td>
<td>14.5</td>
</tr>
<tr>
<td>D</td>
<td>4</td>
<td>6</td>
<td>10</td>
</tr>
<tr>
<td>E</td>
<td>1</td>
<td>250</td>
<td>600</td>
</tr>
</tbody>
</table>

Ans: 192.95

27. Construct the cost of living index number for the following data with base year as 1989.

<table>
<thead>
<tr>
<th>Item</th>
<th>Weight</th>
<th>Price in</th>
<th>Price in</th>
<th>Price in</th>
</tr>
</thead>
<tbody>
<tr>
<td>Food</td>
<td>4</td>
<td>45</td>
<td>50</td>
<td>60</td>
</tr>
<tr>
<td>Clothing</td>
<td>2</td>
<td>30</td>
<td>33</td>
<td>38</td>
</tr>
<tr>
<td>Fuel</td>
<td>1</td>
<td>10</td>
<td>12</td>
<td>13</td>
</tr>
<tr>
<td>House Rent</td>
<td>3</td>
<td>40</td>
<td>42</td>
<td>45</td>
</tr>
<tr>
<td>Miscellaneous</td>
<td>1</td>
<td>5</td>
<td>8</td>
<td>10</td>
</tr>
</tbody>
</table>

Ans: for 1990: 114.49, for 1991: 132.20

28. A survey of families in a city revealed the following information:
What is the cost of living index number for 1988 as compared to that of 1987?

**Ans:** 113.65

29. Construct the consumer price index number for the following industrial data:

<table>
<thead>
<tr>
<th>Item</th>
<th>Weight</th>
<th>Price Index</th>
</tr>
</thead>
<tbody>
<tr>
<td>Industrial Production</td>
<td>30</td>
<td>180</td>
</tr>
<tr>
<td>Exports</td>
<td>15</td>
<td>145</td>
</tr>
<tr>
<td>Imports</td>
<td>10</td>
<td>150</td>
</tr>
<tr>
<td>Transportation</td>
<td>5</td>
<td>170</td>
</tr>
<tr>
<td>Other activity</td>
<td>5</td>
<td>190</td>
</tr>
</tbody>
</table>

**Ans:** 167.30

30. Calculate the real income for the following data:

<table>
<thead>
<tr>
<th></th>
<th></th>
<th></th>
<th></th>
<th></th>
<th></th>
<th></th>
</tr>
</thead>
<tbody>
<tr>
<td>Income in Rs.</td>
<td>500</td>
<td>550</td>
<td>700</td>
<td>780</td>
<td>900</td>
<td>1150</td>
</tr>
<tr>
<td>Price Index</td>
<td>100</td>
<td>110</td>
<td>115</td>
<td>130</td>
<td>140</td>
<td>155</td>
</tr>
</tbody>
</table>

31. The employees of Australian Steel Ltd. have presented the following data in support of their contention that they are entitled to a wage adjustment. Dollar amounts shown represent the average weekly take home pay of the group:

<table>
<thead>
<tr>
<th></th>
<th></th>
<th></th>
<th></th>
<th></th>
</tr>
</thead>
<tbody>
<tr>
<td>Pay in $</td>
<td>260.50</td>
<td>263.80</td>
<td>274</td>
<td>282.50</td>
</tr>
<tr>
<td>Index</td>
<td>126.8</td>
<td>129.5</td>
<td>136.2</td>
<td>141.1</td>
</tr>
</tbody>
</table>
Compute the real wages based on the take home pay and the price indices given. Also compute the amount of pay needed in 1976 to provide buying power equal to that enjoyed in 1973.

32. Calculate the real income for the following data:

<table>
<thead>
<tr>
<th></th>
<th></th>
<th></th>
<th></th>
<th></th>
<th></th>
<th></th>
</tr>
</thead>
<tbody>
<tr>
<td>Income in Rs.</td>
<td>250</td>
<td>300</td>
<td>350</td>
<td>500</td>
<td>750</td>
<td>1000</td>
</tr>
<tr>
<td>Price Index</td>
<td>100</td>
<td>105</td>
<td>110</td>
<td>120</td>
<td>125</td>
<td>140</td>
</tr>
</tbody>
</table>

33. The per capita income and the corresponding cost of living index numbers are given below. Find the per capita real income:

<table>
<thead>
<tr>
<th></th>
<th></th>
<th></th>
<th></th>
<th></th>
<th></th>
<th></th>
</tr>
</thead>
<tbody>
<tr>
<td>per capita income</td>
<td>220</td>
<td>240</td>
<td>280</td>
<td>315</td>
<td>335</td>
<td>390</td>
</tr>
<tr>
<td>cost of living l.N.</td>
<td>100</td>
<td>110</td>
<td>115</td>
<td>135</td>
<td>150</td>
<td>160</td>
</tr>
</tbody>
</table>

34. The following data gives the salaries (in ‘00 Rs.) of the employees of Hindusthan Constructions Ltd with the cost of living index number. Find the real income and suggest how much allowance should be paid to them to maintain the same standard of living.

<table>
<thead>
<tr>
<th></th>
<th></th>
<th></th>
<th></th>
<th></th>
<th></th>
<th></th>
</tr>
</thead>
<tbody>
<tr>
<td>Income</td>
<td>12</td>
<td>14</td>
<td>17</td>
<td>20</td>
<td>24</td>
<td>28</td>
</tr>
<tr>
<td>Price Index</td>
<td>100</td>
<td>120</td>
<td>135</td>
<td>155</td>
<td>180</td>
<td>225</td>
</tr>
</tbody>
</table>

35. The income of Mr. Bhushan Damle in 1999 was Rs. 8,000 per month. If he gets an increment of Rs. 1,200 in 2000 and the price index being 115 with base as 1999, can you conclude that Mr. Damle has got an increment which will maintain his standard of living as compared with the previous year?
UNIT IX

15

STATISTICAL DECISION THEORY

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.

15.1 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 $a_1, a_2, a_3, \ldots \text{etc}$. The totality of all possible actions is called action space denoted by $A$. If there are only three possible actions, we write $A = \text{action space} = (a_1, a_2, a_3, \ldots \text{etc})$. 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 $\theta_1, \theta_2, \theta_3, \ldots$ etc. Totality of all outcomes is called nature space or state space symbolised by $\Omega$. If an action leads to three outcomes $\theta_1, \theta_2$ and $\theta_3$ then we write: $\Omega = (\theta_1, \theta_2, \theta_3)$.

For example, if a product is marketed it may be highly appreciated (outcome $\theta_1$), it may not appeal to the customers (outcome $\theta_2$) or it may be liked by a certain fraction of the customers say, 25% (outcome $\theta_3$).

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**

<table>
<thead>
<tr>
<th>Acts</th>
<th>$a_1$</th>
<th>$a_2$</th>
<th>$a_3$</th>
<th>......</th>
<th>$a_k$</th>
<th>......</th>
<th>$a_m$</th>
</tr>
</thead>
<tbody>
<tr>
<td>States of nature</td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>$\theta_1$</td>
<td>$p_{11}$</td>
<td>$p_{12}$</td>
<td>$p_{13}$</td>
<td>......</td>
<td>$p_{1k}$</td>
<td>......</td>
<td>$p_{1m}$</td>
</tr>
<tr>
<td>$\theta_2$</td>
<td>$p_{21}$</td>
<td>$p_{22}$</td>
<td>$p_{23}$</td>
<td>......</td>
<td>$p_{2k}$</td>
<td>......</td>
<td>$p_{2m}$</td>
</tr>
<tr>
<td>$\theta_3$</td>
<td>$p_{31}$</td>
<td>$p_{32}$</td>
<td>$p_{33}$</td>
<td>......</td>
<td>$p_{3k}$</td>
<td>......</td>
<td>$p_{3m}$</td>
</tr>
<tr>
<td>$\theta_i$</td>
<td>$p_{i1}$</td>
<td>$p_{i2}$</td>
<td>$p_{i3}$</td>
<td>......</td>
<td>$p_{ik}$</td>
<td>......</td>
<td>$p_{im}$</td>
</tr>
<tr>
<td>$\theta_n$</td>
<td>$p_{n1}$</td>
<td>$p_{n2}$</td>
<td>$p_{n3}$</td>
<td>......</td>
<td>$p_{nk}$</td>
<td>......</td>
<td>$p_{nm}$</td>
</tr>
</tbody>
</table>

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.*,

- $\theta_1$ = demand is 78 pastries
- $\theta_2$ = demand is 79 pastries
- $\theta_3$ = 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:

<table>
<thead>
<tr>
<th>Demand x Price - Production x Cost</th>
<th>$P_{11}$ = payoff when action $a_1$ is taken but the state of nature is $\theta_1$</th>
<th>$P_{12}$ = payoff when action $a_2$ is taken but the state of nature is $\theta_1$</th>
<th>$P_{13}$ = payoff when action $a_3$ is taken but the state of nature is $\theta_1$</th>
</tr>
</thead>
<tbody>
<tr>
<td></td>
<td>$= Rs. \ [5 \times 78 - 3 \times 78] = Rs. 156.$</td>
<td>$= Rs. \ [5 \times 78 - 3 \times 79] = Rs. 153.$</td>
<td>$= Rs. \ [5 \times 78 - 3 \times 80] = Rs. 150.$</td>
</tr>
</tbody>
</table>
Similarly $P_{21}$ = payoff when action $a_1$ is taken and the state of nature is $\theta_2$

$= Rs. [5 \times 78 - 3 \times 78] = Rs. 156.$

$P_{22} = Rs. [5 \times 79-3 \times 79]= Rs. 158.$

$P_{23} = Rs. [5 \times 79 -3 \times 80] = Rs. 155.$

Also, $P_{31}$ = payoff when action $a_1$ is taken and the state of nature is $\theta_3$

$P_{31} = Rs. [5 \times 78-3 \times 78]= Rs. 156.$

$P_{32} = Rs. [5 \times 79-3 \times 79]= Rs. 158.$

$P_{33} = Rs. [5 \times 80-3 \times 80] = Rs. 160.$

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

<table>
<thead>
<tr>
<th>State of nature</th>
<th>$a_1 = 78$</th>
<th>$a_2 = 79$</th>
<th>$a_3 = 80$</th>
</tr>
</thead>
<tbody>
<tr>
<td>$\theta_1 = 78$</td>
<td>156</td>
<td>153</td>
<td>150</td>
</tr>
<tr>
<td>$\theta_2 = 79$</td>
<td>156</td>
<td>158</td>
<td>155</td>
</tr>
<tr>
<td>$\theta_3 = 80$</td>
<td>156</td>
<td>158</td>
<td>160</td>
</tr>
</tbody>
</table>

To calculate opportunity loss, we first determine maximum payoff in each state of nature.

In first state of nature $\theta_1$, maximum payoff = Rs. 156

In second state of nature $\theta_2$, maximum payoff = Rs. 158

In third state of nature $\theta_3$, maximum payoff = Rs. 160

$L_{11} = 156 - 156 = 0, L_{12} = 156 - 153 = 3, L_{13} = 156 - 150 = 6$

$L_{21} = 158 - 156 = 2, L_{22} = 158 - 158 = 0, L_{23} = 158 - 155 = 3$

$L_{31} = 160 - 156 = 4, L_{32} = 160 - 158 = 2, L_{33} = 158 - 158 = 0.$

The loss table corresponding to payoff table is given below.

<table>
<thead>
<tr>
<th>State of nature</th>
<th>$a_1$</th>
<th>$a_2$</th>
<th>$a_3$</th>
</tr>
</thead>
<tbody>
<tr>
<td>$\theta_1$</td>
<td>0</td>
<td>3</td>
<td>6</td>
</tr>
<tr>
<td>$\theta_2$</td>
<td>2</td>
<td>0</td>
<td>3</td>
</tr>
<tr>
<td>$\theta_3$</td>
<td>4</td>
<td>2</td>
<td>0</td>
</tr>
</tbody>
</table>
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,

- \( a_1 = \text{Junk the lot at a loss of Rs. 100 for the lot} \)
- \( a_2 = \text{Sell the lot at Rs. 5 each giving double money back guarantee for each defective item} \)

Thus \( A = (a_1 \ a_2) = \text{action space} \)

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 \( 0_i \) denote that there are "\( i \)" defective bulbs in the lot.

Thus state space or \( \Omega = \{0_0, 0_1, 0_2, \ldots, 0_{100}\} \)

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.

\[ P_{i1} = -100 \text{ for all } i = 0, 1, 2, \ldots, 100 \]

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

\[ P_{i2} = 100 \times 5 - i \times 10 = 500 - 10i \quad \text{for all } i = 0, 1, 2, \ldots, 100. \]

Thus his payoffs are

\[ \begin{align*}
P_{i1} &= -100 & \text{for all } i = 0, 1, 2, \ldots, 100 \\
P_{i2} &= 500 - 10i & \text{for all } i = 0, 1, 2, \ldots, 100
\end{align*} \]

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:

<table>
<thead>
<tr>
<th>Strategies</th>
<th>S₁</th>
<th>S₂</th>
<th>S₃</th>
</tr>
</thead>
<tbody>
<tr>
<td>States of Nature</td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>N₁</td>
<td>24</td>
<td>17</td>
<td>10</td>
</tr>
<tr>
<td>N₂</td>
<td>24</td>
<td>27</td>
<td>20</td>
</tr>
<tr>
<td>N₃</td>
<td>24</td>
<td>27</td>
<td>30</td>
</tr>
<tr>
<td>Maximum gains</td>
<td>24</td>
<td>27</td>
<td>(30)</td>
</tr>
<tr>
<td>Minimum gains</td>
<td>(24)</td>
<td>17</td>
<td>10</td>
</tr>
</tbody>
</table>

For maximax criterion, first of all maximum payoff of each strategy is identified. Maximum payoff for strategy S₁ is 24, for S₂ is 27 and for S₃, 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 S₃.

(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 S₁ is 24, for S₂ it is 17, and for S₃, 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 S₁ 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 $a₁$ and $a₂$:

- $Qa₁ = -30 + 50x$
- $Qa₂ = -90 + 20x$

(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 \( Q_{a1} \) and \( Q_{a2} \)

\[-30 + 50x = -90 + 20x\]

or

\[-2 = x\]

\[\therefore -2 \] is the break even point.

(ii) Substituting \( x = 10 \) in \( Q_{a1} \) and \( Q_{a2} \)

\[Q_{a1} = 470 \quad Q_{a2} = 110\]

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 \( x = -5 \)

\[Q_{a1} = -280 \quad Q_{a2} = -190\]

\( a_2 \) 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:

<table>
<thead>
<tr>
<th>State of nature</th>
<th>Pay offs (in Rs.)</th>
</tr>
</thead>
<tbody>
<tr>
<td>Strategies</td>
<td>( N_1 )</td>
</tr>
<tr>
<td>( S_1 )</td>
<td>7,00,000</td>
</tr>
<tr>
<td>( S_2 )</td>
<td>5,00,000</td>
</tr>
<tr>
<td>( S_3 )</td>
<td>3,00,000</td>
</tr>
</tbody>
</table>

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.
The executive would choose strategy $S_2$

(b) *Maximax Criterion.* In this criterion we select that strategy which gives the maximum payoffs.

<table>
<thead>
<tr>
<th>Strategy</th>
<th>Minimum Payoffs Rs</th>
</tr>
</thead>
<tbody>
<tr>
<td>$S_1$</td>
<td>1,50,000</td>
</tr>
<tr>
<td>$S_2$</td>
<td>0</td>
</tr>
<tr>
<td>$S_3$</td>
<td>3,00,000</td>
</tr>
</tbody>
</table>

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**

<table>
<thead>
<tr>
<th></th>
<th>$N_1$</th>
<th>$N_2$</th>
<th>$N_3$</th>
</tr>
</thead>
<tbody>
<tr>
<td>$S_1$</td>
<td>7,00,000 - 7,00,000 = 0</td>
<td>4,50,000 - 3,00,000 = 1,50,000</td>
<td>3,00,000 - 1,50,000 = 1,50,000</td>
</tr>
<tr>
<td>$S_2$</td>
<td>7,00,000 - 5,00,000 = 2,00,000</td>
<td>4,50,000 - 4,50,000 = 0</td>
<td>3,00,000 - 0 = 3,00,000</td>
</tr>
<tr>
<td>$S_3$</td>
<td>7,00,000 - 3,00,000 = 4,00,000</td>
<td>4,50,000 - 3,00,000 = 1,50,000</td>
<td>3,00,000 - 3,00,000 = 0</td>
</tr>
</tbody>
</table>

Maximum opportunity loss

<p>| | |</p>
<table>
<thead>
<tr>
<th></th>
<th></th>
</tr>
</thead>
<tbody>
<tr>
<td>$S_1$</td>
<td>= 1,50,000</td>
</tr>
<tr>
<td>$S_2$</td>
<td>= 3,00,000</td>
</tr>
<tr>
<td>$S_3$</td>
<td>= 4,00,000</td>
</tr>
</tbody>
</table>

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
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:

<table>
<thead>
<tr>
<th>Nature of Demand</th>
<th>Good</th>
<th>Moderate</th>
<th>Poor</th>
</tr>
</thead>
<tbody>
<tr>
<td>Product X</td>
<td>0.70</td>
<td>0.20</td>
<td>0.10</td>
</tr>
<tr>
<td>Product Y</td>
<td>0.50</td>
<td>0.30</td>
<td>0.20</td>
</tr>
<tr>
<td>Product Z</td>
<td>0.40</td>
<td>0.50</td>
<td>0.10</td>
</tr>
</tbody>
</table>

The estimated profit or loss under the three states may be taken as

- **X**:
  - Good: 30,000
  - Moderate: 20,000
  - Poor: 10,000

- **Y**:
  - Good: 60,000
  - Moderate: 30,000
  - Poor: 20,000

- **Z**:
  - Good: 40,000
  - Moderate: 10,000
  - Poor: -15,000 (loss)

Prepare the expected value table and advise the management about the choice of product.

<table>
<thead>
<tr>
<th>States of Nature</th>
<th>Expected payoff (in Rs. ‘0000) for various acts</th>
</tr>
</thead>
<tbody>
<tr>
<td></td>
<td>X</td>
</tr>
<tr>
<td>Good</td>
<td>30</td>
</tr>
<tr>
<td>Moderate</td>
<td>20</td>
</tr>
<tr>
<td>Poor</td>
<td>10</td>
</tr>
<tr>
<td>Expected Monetary value</td>
<td>26</td>
</tr>
</tbody>
</table>

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:

<table>
<thead>
<tr>
<th>Weekly Sales</th>
<th>Probability</th>
</tr>
</thead>
<tbody>
<tr>
<td>0</td>
<td>0.10</td>
</tr>
<tr>
<td>1</td>
<td>0.10</td>
</tr>
<tr>
<td>2</td>
<td>0.20</td>
</tr>
<tr>
<td>3</td>
<td>0.20</td>
</tr>
<tr>
<td>4</td>
<td>0.02</td>
</tr>
<tr>
<td>5</td>
<td>0.15</td>
</tr>
<tr>
<td>6</td>
<td>0.05</td>
</tr>
</tbody>
</table>
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.

<table>
<thead>
<tr>
<th>Demand</th>
<th>Prob.</th>
<th>Strategies (buy so many)</th>
</tr>
</thead>
<tbody>
<tr>
<td>0</td>
<td>0.10</td>
<td>0 20 40 60 80 100 120</td>
</tr>
<tr>
<td>1</td>
<td>0.10</td>
<td>50 20 40 60 80 100 120</td>
</tr>
<tr>
<td>2</td>
<td>0.20</td>
<td>100 70 40 60 80 100 120</td>
</tr>
<tr>
<td>3</td>
<td>0.20</td>
<td>150 120 90 60 80 100 120</td>
</tr>
<tr>
<td>4</td>
<td>0.20</td>
<td>200 170 140 110 80 100 120</td>
</tr>
<tr>
<td>5</td>
<td>0.15</td>
<td>250 220 190 160 130 100 120</td>
</tr>
<tr>
<td>6</td>
<td>0.05</td>
<td>300 270 240 210 180 150 120</td>
</tr>
<tr>
<td>Expected</td>
<td></td>
<td>147.5 122.5 102.5 92.5 92.5 92.5 102.5 120</td>
</tr>
</tbody>
</table>

Thus either buy 3 or 4.

**Example 8**: Following are the records of demand of an item for the past 300 days.

<table>
<thead>
<tr>
<th>$\theta$ = Demand in units</th>
<th>No. of days</th>
<th>Prob ($\theta$)</th>
</tr>
</thead>
<tbody>
<tr>
<td>10,000</td>
<td>18</td>
<td>0.06</td>
</tr>
<tr>
<td>11,000</td>
<td>90</td>
<td>0.30</td>
</tr>
<tr>
<td>12,000</td>
<td>120</td>
<td>0.40</td>
</tr>
<tr>
<td>13,000</td>
<td>60</td>
<td>0.20</td>
</tr>
<tr>
<td>14,000</td>
<td>12</td>
<td>0.04</td>
</tr>
<tr>
<td></td>
<td>300</td>
<td>1.00</td>
</tr>
</tbody>
</table>

(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 $= 10 \times 0.06 = 0.60$.

$= 11 \times 0.30 = 3.30$

$= 12 \times 0.40 = 4.80$

$= 13 \times 0.20 = 2.60$

$= 14 \times 0.04 = 0.56$

$= 11.86$

Expected demand = 11,860 (Answer)

The daily pay-off matrix is constructed below:

<table>
<thead>
<tr>
<th>P($\theta$)</th>
<th>states of nature</th>
<th>10</th>
<th>11</th>
<th>12</th>
<th>13</th>
<th>14</th>
</tr>
</thead>
<tbody>
<tr>
<td>0.06</td>
<td>10</td>
<td>50</td>
<td>45</td>
<td>40</td>
<td>35</td>
<td>30</td>
</tr>
<tr>
<td>0.30</td>
<td>11</td>
<td>50</td>
<td>55</td>
<td>50</td>
<td>45</td>
<td>40</td>
</tr>
<tr>
<td>0.40</td>
<td>12</td>
<td>50</td>
<td>55</td>
<td>60</td>
<td>55</td>
<td>50</td>
</tr>
<tr>
<td>0.20</td>
<td>13</td>
<td>50</td>
<td>55</td>
<td>60</td>
<td>65</td>
<td>60</td>
</tr>
<tr>
<td>0.04</td>
<td>14</td>
<td>50</td>
<td>55</td>
<td>60</td>
<td>65</td>
<td>70</td>
</tr>
<tr>
<td>Expected Pay off</td>
<td></td>
<td>50</td>
<td>54.4</td>
<td>55.8</td>
<td>53.2</td>
<td>48.6</td>
</tr>
</tbody>
</table>

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 80000 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:

\[
\text{Prob. (50,000)} = \frac{2}{5} = 0.40
\]

Prob. (less than or more than 50,000) = 0.60

Prob. (less than 50,000): Prob. (More than 50,000): 1:4

Thus \( \text{Prob. (less than 50,000)} = 0.60 \times \frac{1}{1+4} = 0.12 \)

\( \text{Prob. (More than 50,000)} = 0.60 \times \frac{4}{1+4} = 0.48 \)

Prob. (60,000): Prob. (70,000): Prob. (80,000):: 1:4:1

Thus \( \text{Prob. (60,000)} = 0.48 \times \frac{1}{1+4+1} = 0.08 \)

\( \therefore \text{Prob. (80,000)} = 0.08 \)

and \( \text{Prob. (70,000)} = \frac{4}{1+4+1} \times 0.48 = 0.32 \)

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.

<table>
<thead>
<tr>
<th>Demand</th>
<th>Prob</th>
<th>&lt; 50,000 (take it 25,000)</th>
<th>50,000</th>
<th>60,000</th>
<th>70,000</th>
<th>80,000</th>
</tr>
</thead>
<tbody>
<tr>
<td>&lt; 50,000 (worst 25,000)</td>
<td>0.12</td>
<td>500</td>
<td>-250</td>
<td>-550</td>
<td>-850</td>
<td>-1150</td>
</tr>
<tr>
<td>50,000</td>
<td>0.40</td>
<td>500</td>
<td>1,000</td>
<td>700</td>
<td>400</td>
<td>100</td>
</tr>
<tr>
<td>60,000</td>
<td>0.08</td>
<td>500</td>
<td>1,000</td>
<td>1200</td>
<td>900</td>
<td>600</td>
</tr>
<tr>
<td>70,000</td>
<td>0.32</td>
<td>500</td>
<td>1,000</td>
<td>1200</td>
<td>1400</td>
<td>1100</td>
</tr>
<tr>
<td>80,000</td>
<td>0.08</td>
<td>500</td>
<td>1,000</td>
<td>1200</td>
<td>1400</td>
<td>1600</td>
</tr>
<tr>
<td>Expected Payoff</td>
<td>500</td>
<td>850</td>
<td>790</td>
<td></td>
<td></td>
<td></td>
</tr>
</tbody>
</table>
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.

EMV (at C) = .3 x Rs. 1,40,000 + .4 x Rs. 70,000 + .3 x (- Rs.10,000)
           = Rs. 42,000+ Rs. 28,000 - Rs. 3,000 = Rs. 67,000

EMV (at D) = .4 x Rs. 1,50,000 + .3 x Rs. 80,000 + .3 x (- Rs. 15,000)
           = Rs. 60,000 + Rs.24,000-Rs. 4,500 = Rs. 79,500

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:

The decision tree is drawn in the above figure. Calculations of EMV proceed from right to left.

Go

\[
30,000 \times 0.6 = 18,000 \\
15,000 \times 1 = 15,000
\]

\[
15,000 \times 0.4 = 6000 \\
24,000 \times 0.6 = 14,400
\]

\[
24,000 \times 0.4 = 9600
\]

Not Go

\[
\text{(where 1 is the probability)}
\]

\[
\text{Hence Mr. X should proceed to Sri Lanka}
\]

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(R₁) 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 x 0.6 + 14,000 x 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 x 0.9 + 0.1 x 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 x 25,000 + 0.1 X 14,000 = Rs. 23,900
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.

<table>
<thead>
<tr>
<th>Type of perfume</th>
<th>Estimated sales (units)</th>
</tr>
</thead>
<tbody>
<tr>
<td></td>
<td>25,000</td>
</tr>
<tr>
<td>I</td>
<td>200</td>
</tr>
<tr>
<td>II</td>
<td>30</td>
</tr>
<tr>
<td>III</td>
<td>50</td>
</tr>
</tbody>
</table>

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?

<table>
<thead>
<tr>
<th>States of Nature</th>
<th>Strategies</th>
</tr>
</thead>
<tbody>
<tr>
<td></td>
<td>$S_1$</td>
</tr>
<tr>
<td>$E_1$</td>
<td>30,000</td>
</tr>
<tr>
<td>$E_2$</td>
<td>50,000</td>
</tr>
<tr>
<td>$E_3$</td>
<td>40,000</td>
</tr>
</tbody>
</table>

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.

<table>
<thead>
<tr>
<th>Product</th>
<th>Fixed Cost Rs.</th>
<th>Variable cost unit in Rs.</th>
</tr>
</thead>
<tbody>
<tr>
<td>X</td>
<td>250</td>
<td>12</td>
</tr>
<tr>
<td>Y</td>
<td>350</td>
<td>10</td>
</tr>
<tr>
<td>Z</td>
<td>500</td>
<td>5</td>
</tr>
</tbody>
</table>

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:

<table>
<thead>
<tr>
<th>Product</th>
<th>Nature of demand</th>
</tr>
</thead>
<tbody>
<tr>
<td></td>
<td>Good</td>
</tr>
<tr>
<td>A</td>
<td>0.7</td>
</tr>
<tr>
<td>B</td>
<td>0.4</td>
</tr>
<tr>
<td>C</td>
<td>0.5</td>
</tr>
</tbody>
</table>
The estimated profit/loss under the three states of nature are given as follows:

<table>
<thead>
<tr>
<th>Product</th>
<th>Profit in 10,000 Rs.</th>
</tr>
</thead>
<tbody>
<tr>
<td></td>
<td>Good</td>
</tr>
<tr>
<td>A</td>
<td>30</td>
</tr>
<tr>
<td>B</td>
<td>40</td>
</tr>
<tr>
<td>C</td>
<td>60</td>
</tr>
</tbody>
</table>

Using EMV criterion, advice the management the optimum decision.

**Ans:** Select product C

5. You are given, the following payoff table for three acts A₁, A₂, A₃ and the states of nature S₁, S₂, S₃.

<table>
<thead>
<tr>
<th>States of nature</th>
<th>A₁</th>
<th>A₂</th>
<th>A₃</th>
</tr>
</thead>
<tbody>
<tr>
<td>E₁</td>
<td>25</td>
<td>-10</td>
<td>-125</td>
</tr>
<tr>
<td>E₂</td>
<td>400</td>
<td>440</td>
<td>400</td>
</tr>
<tr>
<td>E₃</td>
<td>650</td>
<td>740</td>
<td>750</td>
</tr>
</tbody>
</table>

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 A₂

6. An investor is given the following investment alternatives and percentage rate of return.

<table>
<thead>
<tr>
<th>State of Nature</th>
<th>Course of Action</th>
</tr>
</thead>
<tbody>
<tr>
<td></td>
<td>Regular Share</td>
</tr>
<tr>
<td>Good</td>
<td>7%</td>
</tr>
<tr>
<td>Better</td>
<td>10%</td>
</tr>
<tr>
<td>Best</td>
<td>15%</td>
</tr>
</tbody>
</table>

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.

<table>
<thead>
<tr>
<th>Cases demanded</th>
<th>15</th>
<th>16</th>
<th>17</th>
<th>18</th>
</tr>
</thead>
<tbody>
<tr>
<td>Probability</td>
<td>0.1</td>
<td>0.2</td>
<td>0.4</td>
<td>0.3</td>
</tr>
</tbody>
</table>

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 $A_1, A_2, A_3, A_4$.

<table>
<thead>
<tr>
<th>State of Nature</th>
<th>Course of Action</th>
<th>Probability of State</th>
</tr>
</thead>
<tbody>
<tr>
<td></td>
<td>$A_1$</td>
<td>$A_2$</td>
</tr>
<tr>
<td>$s_1$</td>
<td>50</td>
<td>400</td>
</tr>
<tr>
<td>$S_2$</td>
<td>300</td>
<td>0</td>
</tr>
<tr>
<td>$S_3$</td>
<td>-150</td>
<td>100</td>
</tr>
<tr>
<td>$s_4$</td>
<td>50</td>
<td>0</td>
</tr>
</tbody>
</table>

(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.

<table>
<thead>
<tr>
<th>Acts</th>
<th>Events</th>
</tr>
</thead>
<tbody>
<tr>
<td></td>
<td>$E_1$</td>
</tr>
<tr>
<td>$A_1$</td>
<td>40</td>
</tr>
<tr>
<td>$A_2$</td>
<td>300</td>
</tr>
<tr>
<td>$A_3$</td>
<td>50</td>
</tr>
<tr>
<td>$A_4$</td>
<td>300</td>
</tr>
</tbody>
</table>

**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.

<table>
<thead>
<tr>
<th>Dozens of roses</th>
<th>7</th>
<th>8</th>
<th>9</th>
<th>10</th>
</tr>
</thead>
<tbody>
<tr>
<td>Probability</td>
<td>0.1</td>
<td>0.2</td>
<td>0.4</td>
<td>0.3</td>
</tr>
</tbody>
</table>

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:

<table>
<thead>
<tr>
<th>No. of cars demanded</th>
<th>0</th>
<th>1</th>
<th>2</th>
<th>3</th>
<th>4</th>
</tr>
</thead>
<tbody>
<tr>
<td>Probability</td>
<td>0.1</td>
<td>0.2</td>
<td>0.3</td>
<td>0.2</td>
<td>0.2</td>
</tr>
</tbody>
</table>

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:

<table>
<thead>
<tr>
<th>Decision Di</th>
<th>Probability of decision Di given research R P(Di \ R)</th>
<th>Outcome number</th>
<th>Probability of outcome Xi given Di P(Xi \ Di)</th>
<th>Payoff value of outcome, Xi (Rs ‘000)</th>
</tr>
</thead>
<tbody>
<tr>
<td>Develop</td>
<td>0.5</td>
<td>1</td>
<td>0.6</td>
<td>600</td>
</tr>
<tr>
<td></td>
<td></td>
<td>2</td>
<td>0.3</td>
<td>-100</td>
</tr>
<tr>
<td></td>
<td></td>
<td>3</td>
<td>0.1</td>
<td>0</td>
</tr>
<tr>
<td>Do not develop</td>
<td>0.5</td>
<td>1</td>
<td>0.0</td>
<td>600</td>
</tr>
<tr>
<td></td>
<td></td>
<td>2</td>
<td>0.0</td>
<td>-100</td>
</tr>
<tr>
<td></td>
<td></td>
<td>3</td>
<td>1.0</td>
<td>0</td>
</tr>
</tbody>
</table>

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 action—subcontracting (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 analyses the tree we start working backward from the end branches.

He 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.
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.

✦✦✦
F. Y. B. Com
MATHEMATICAL AND STATISTICAL TECHNIQUES

Unit I
Commission, Brokerage, Discount and Partnership:
Commission and Brokerage : Simple examples on calculation of commission and brokerage.
Problems involving mixture of discount and profit are expected.

Unit II
Shares and Mutual Funds:
Concepts of shares, face value, market value, dividend, equity shares preferential shares, bonus shares, Simple examples.
Mutual Funds, Simple problems on calculation of Net Income after considering entry load, dividend, change in Net Asset Value (N.A.V) and exit load. Averaging of price under the ‘Systematic Investment Plan (S.I.P)’

Linear Programming Problems:
Sketching or graphs of (i) linear equation $Ax + By + C = 0$ (ii) linear inequalities. Mathematical Formulation of Linear Programming Problems upto 2 variables.

Unit III
Introduction
Meaning, Scope and Limitations of Statistics:

Collection of Data:
Primary and Secondary, Sample and Census, Survey (concept only), Tabulation of data upto 3 characteristics (Simple examples)

Diagrams and graphs:
Given a diagram, interpretation of it. Simple bar diagram, Multiple bar diagram, Percentage bar diagram, Pie diagram.
Drawing of frequency curve, frequency polygon, Histogram (class – intervals of equal lengths only) and ogives.

Unit IV
Measures of Central Tendency: Arithmetic mean, Weighted mean, Combined mean, Median, Mode-without grouping, Quartiles (No Example on missing frequency)

Measures of Dispersion: Range, Quartile deviation, Mean deviation from mean Standard deviation and their relative measures. (Concepts of shift of origin and change or scale are not to be done)
Unit V

**Elementary Probability Theory:**
Concept of Random experiment/trial and possible outcomes; Sample Space and Discrete Sample Space; Events and their types, Algebra of Events, Mutually Exclusive and Exhaustive Events, concept of \(^nC_r\).
Classical definition of Probability, Addition theorem (without proof);
Independence of Events: \(P(A \cap B) = P(A)P(B)\)
Simple examples.

**Random Variable:** Probability distribution of a discrete random variable;
Expectation and Variance; Simple examples.
Concept of Normal distribution and Standard Normal Variate (SNV), simple examples.

Unit VI

**Functions Derivatives and Their Applications**
Concept of real functions: constant function, linear function, \(x^2\), \(e^x\), \(a^x\), \(\log x\).
Demand, Supply, Total Revenue, Average Revenue, Total Cost, Average Cost and Profit function. Equilibrium Point.
Derivative as rate measure.
Derivatives of functions: Constant function, \(x^n\), \(e^x\), \(a^x\), \(\log x\)
Second Order derivatives.
Applications: Marginal Cost, Marginal Revenue, Elasticity of Demand.
Maxima and Minima for functions in Economics and Commerce.

Unit VII

**Interest and Annuity**
Simple Interest and Compound Interest
Interest Compounded more than once a year. Calculations involving upto 4 time periods.
Equated Monthly Instalments (EMI) using reducing & flat interest system.
Present value, Future value.

Annuity, Immediate and due: Simple problems with
\[
A = P \left(1 + \frac{r}{100}\right)^n \quad \text{with} \quad n \leq 4
\]

Unit VIII

**Bivariate Linear Correlation:** Scatter Diagram, Computation of Karl Pearson’s Coefficient of Correlation (Case of Bivariate Frequency Table to be excluded), Computation of Spearman’s Rank Correlation Coefficient (case of repeated ranks upto 2 repetition only)
Bivariate Linear Regression: Finding Regression lines by method of least squares.
Properties of Regression Coefficients – i) \(r = \pm \sqrt{b_{xy} \cdot b_{yx}}\) ii) \((x, y)\) is a point of intersection of two regression lines.
Unit IX

**Times Series** : Concept and Components of time series. Estimation of Trend using Moving Average Method & Least Squares Method (only Linear Trend)
Estimation of Seasonal Component using Simple Arithmetic Mean. (For Trend free data only)
Concept of Forecasting using Least Squares Method.

**Index Numbers** : Concept and uses. Simple and Composite Index Nos. (unweighted, weighted) Laspeyre’s Price Index No., Paasche’s Price Index No. Fisher’s Price Index No., Cost of Living Index No., Real Income, Simple Examples of Wholesale price Index no. (Examples on missing values should not be done)

Unit X

**Decision Theory** : Decision making situation; Decision maker, Courses of Action, States of Nature, Pay-off and Pay-off matrix; Decision making under Uncertainty, Maximum, Maximax and Laplace criteria, simple examples to find optimum decision.
Decision making under Risk Expected Monetary Value (EMV)l Decision tree, simple examples based on EMV

For IDOL Examination : 100 marks

**Section I** 
Attempt any two questions out of four.
Q. 1 Based on Unit I
Q. 2 Based on Unit II
Q. 3 Based on Unit VI
Q. 4 Based on Unit VII

**Section II** 
Attempt any three questions out of six.
Q. 5 Based on Unit III
Q. 6 Based on Unit IV
Q. 7 Based on Unit V
Q. 8 Based on Unit VIII
Q. 9 Based on Unit IX
Q. 10 Based on Unit X

 okhttp