saransh

Last Mile Referencer for

The Institute of Chartered Accountants of India
(Set up by an Act of Parliament)

Board of Studies

FOUNDATION

PAPER 3

QUANTITATIVE APTITUDE

www.icai.org I https://boslive.icai.org

Saransh –
QUANTITATIVE APTITUDE

Published in May 2024 by:
Board of Studies
The Institute of Chartered Accountants of India
''ICAI Bhawan" A-29, Sector-62,
Noida 201 309

PREFACE

The Board of Studies (BoS) is dedicated to delivering outstanding services to its students, tirelessly striving for the highest standards of education and support. It imparts quality academic education through its value added study materials, that explain concepts clearly and in simple language. Illustrations and Test Your Knowledge Questions contained therein facilitate enhanced understanding and application of concepts learnt. Revision Test Papers provides Questions & Answers to help students to update themselves to revise the concepts by solving questions contained therein. Suggested Answers containing the ideal manner of answering questions set at examination which also helps students to revise for the forthcoming examination. Mock Test Papers empower students to gauge their readiness ahead of each examination, ensuring confidence and clarity. Additionally, BoS offers engaging live virtual classes led by distinguished faculty, reaching students far and wide across the nation.

To effectively engage with its students, the Board of Studies (BoS) has been publishing subject specific capsules in its monthly Students' Journal, "The Chartered Accountant Student," since 2017. These capsules are aimed at facilitating efficient revision of concepts covered across various topics at the Foundation, Intermediate, and Final levels of the Chartered Accountancy Course. This initiative underscores the BoS's commitment to enhancing learning and comprehension among its students through accessible and attractive educational materials. Each issue of the journal includes a capsule relating to specific topic(s) in one subject at each of the three levels. In these capsules, the concepts and provisions are presented in attractive colours in the form of tables, diagrams and flow charts for facilitating easy retention and quick revision of topics.

The Board of Studies (BoS) is now introducing a comprehensive booklet titled 'Saransh - Last Mile Referencer for Foundation Paper 3–Quantitative Aptitude'. This booklet aims to consolidate significant concepts across various topics in Business mathematics, Logical Reasoning and Statistics, particularly those included in the syllabi at the Foundation level of CA Course. It features diagrams, flow charts, tables, and illustrated journal entries, providing a one-stop repository for key Quantitative Aptitude topics.

However, the readers are advised to refer the study material for comprehensive study and revision. Under no circumstances, this booklet substitutes the detailed study of the material provided by the Board of Studies. Further, the readers are advised to enhance their ability to address the issues and solve the problems based on fundamentals of Accounting, illustrations and questions given in the study material, revision test papers and mock test papers. By capturing essential points in a concise format, 'Saransh' will facilitate better understanding and retention of critical accounting principles, enhancing the learning experience for students preparing for their examinations and beyond. It will indeed serve as a valuable ready reckoner for readers, enabling them to grasp the essence of the subject comprehensively.

Happy Reading!

President Message

It is with immense pride that I introduce the Saransh booklets, a meticulously curated resource available across the Foundation, Intermediate, and Final levels of the Chartered Accountancy course. ICAI has always been dedicated to providing our students with the best possible resources to succeed in their studies and careers, and Saransh is a demonstration of this commitment.

The Saransh — Last Mile Referencers have been thoughtfully designed by the Board of Studies (BoS) to serve as an invaluable companion for your studies and exam preparation. Our aim is to simplify complex concepts and provisions, making them easier to understand, memorize, and revise. However, Saransh is not a substitute for the detailed BoS study material but a supplementary tool to complement your in-depth study.

The newly revamped Saransh booklets have been updated not only in content but also in their presentation. With a more logical and organized structure, enhanced visual appeal, and a userfriendly layout, these booklets are now more effective in aiding your studies.

We have extended the Saransh series to cover all core areas of the Chartered Accountancy course. Whether you are studying Direct Tax Laws and International Taxation, Indirect Tax Laws, Accounting Standards, Indian Accounting Standards, Auditing, Cost and Management Accounting, Strategic Cost Management and Performance Evaluation, Company Law, or Financial Management and Strategic Management, you will find a Saransh booklet for each subject.

Saransh is designed not only to help you grasp and recall essential concepts but also to guide you in approaching each subject strategically. The insights provided in these booklets will help you develop a structured approach to your studies, ensuring that you are well-prepared for your examinations.

I urge you to make the most of the Saransh booklets. While these booklets will support you, it is your dedication, perseverance, and hard work that will ultimately determine your success.

I wish each of you the very best in your studies and future careers.

Warm regards,

CA. Ranjeet Kumar Agarwal
President, ICAI

The Institute of Chartered Accountants of India

(Set up by an Act of Parliament)

Board of Studies

Saransh
Foundation
Paper 3: Quantitative Aptitude

INDEX

The Institute of Chartered Accountants of India
(Set up by an Act of Parliament)
Board of Studies

Saransh
Foundation
Paper 3: Quantitative Aptitude

INDEX

Before We Begin

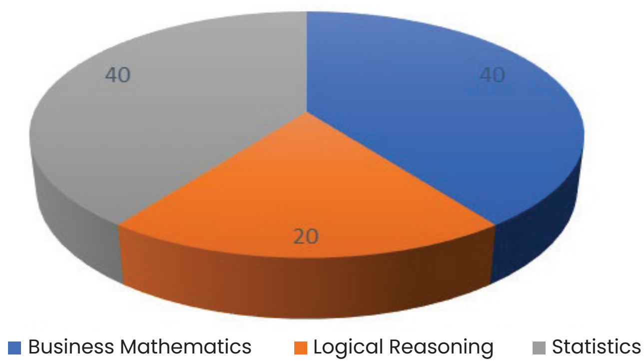
The subject "Quantitative Aptitude" has been designed specifically for the students who are aiming to pursue CA course, keeping in view the relevance of subjects after they become full-fledged professional.

Mathematics and Statistics applications are very important for the students of Chartered Accountancy Course. As professional work in future will demand quantitative and analytical skills. Logical Reasoning has been included in the syllabus to test analytical and mental ability skills which will help them in honing their interpretative skills while pursuing and thereafter CA course.

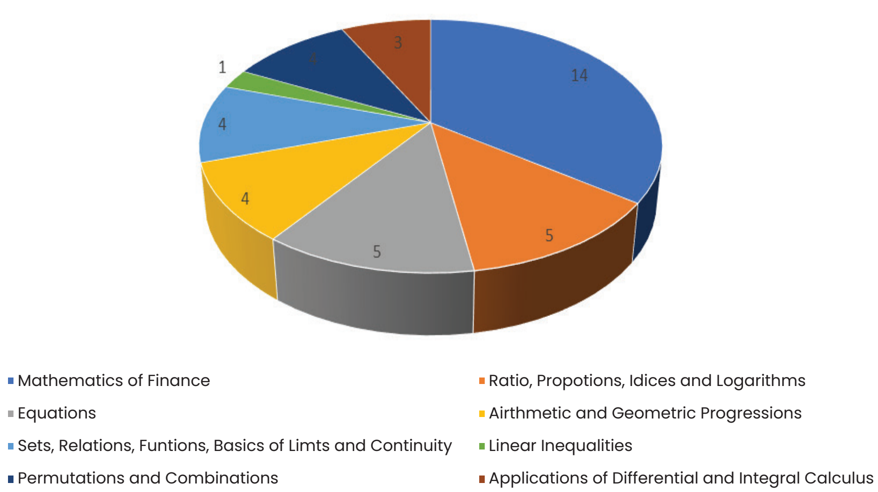
Students will be equipped with the knowledge to absorb various concepts of other subjects of the chartered accountancy course like accounting, auditing and assurance, financial management, cost and management accounting, strategic cost management, etc.

Understand the Syllabus of Quantitative Aptitude

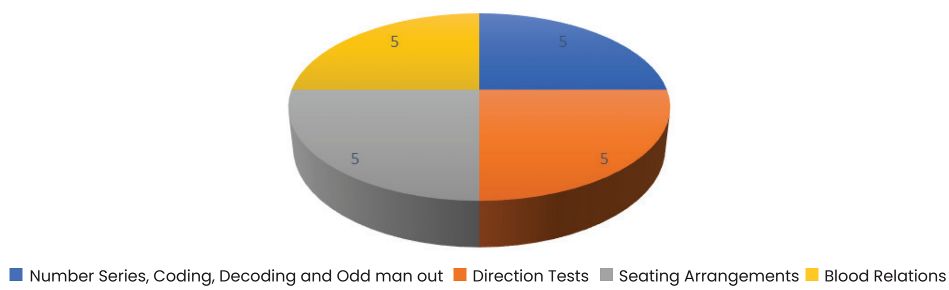
Quantitative Aptitude (100 Marks)



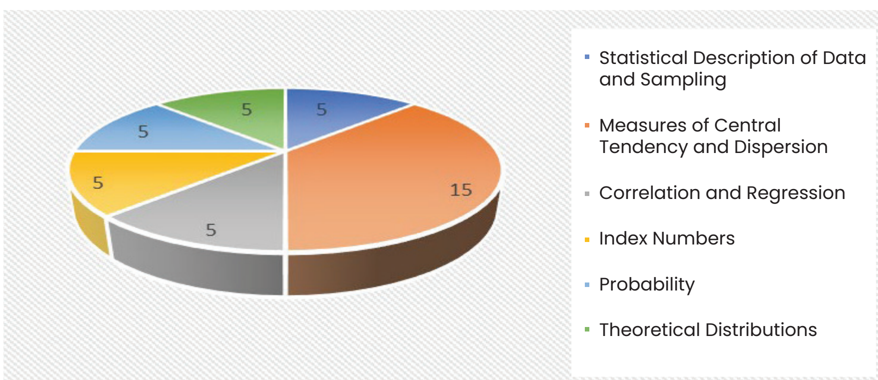
Business Mathematics (40 Marks):



Logical Reasoning (20 Marks):



Statistics (40 Marks):



The Saransh Quantitative Aptitude has been developed in a simple and easy to understand language. Every concept has been explained with the help of solved examples. A number of illustrations have been incorporated in each chapter to explain various concepts and related computational techniques. The diagrams have been drawn neatly in a such way that the students have the complete understanding of the problem by perusing them.

The Quantitative Aptitude paper is tested on multiple choice questions or objective type of questions pattern only.

Quantitative Aptitude is a crucial and scoring paper in the CA Foundation course. To excel in this paper, it's essential to have a clear understanding of the concepts, regular practice, and a disciplined study routine. By following tips such as focusing on conceptual clarity, practicing a variety of problems, managing time effectively during the exam, and staying consistent with preparation by focusing on these strategies student will crack the CA Foundation Paper 3 Quantitative Aptitude.

Happy Reading and Best Wishes!

Chapter -1:

Ratio and Portion, Indices and Logarithms

Unit 1: Ratio

Ratio Definition

If a and b are two non-zero number of same kind then fraction $\frac{a}{b}$ is called RATIO of a to b. It is denoted by a : b . Here a is called ANTECEDENT and b is called CONSEQUENT.

For example, Ratio of two numbers 4 and 6 is 2 : 3.

(i) the ratio between $\sqrt{5}$,and $\sqrt{6}$ is $\sqrt{5} : \sqrt{6}$

(ii) the ratio between 150gm and 2 kg. is 3 : 40

(iii) the ratio between 25 minutes and 45 Seconds is 100 : 3

Inverse Ratio

If a:b then b:a is called inverse ratio of a:b, **Example :** The inverse ratio of 11: 15 is 15: 11

Duplicate Ratio:

If a : b is a given ratio then Duplicate Ratio is $a^2:b^2$, **Example :** The duplicate ratio of 2: 3 4: 9

Sub Duplicate Ratio

If a : b is given ratio then $\sqrt{a} : \sqrt{b}$ is called Sub Duplicate Ratio, **Example :** The Sub Duplicate Ratio of 25 : 16 is 5 : 4

Triplicate Ratio

If a : b is given ratio then $a^3:b^3$ is called Triplicate Ratio, **Example** : Triplicate Ratio of 3 : 2 is 27 : 8

Sub-Triplicate Ratio

If a : b is a given ratio then $\sqrt[3]{a} : \sqrt[3]{b}$ is called Sub Triplicate Ratio of a : b **Example**: Triplicate Ratio of 125 : 27 is 5 : 3

Compounded Ratio

The Ratio Compound of a : b and c : d is ac : bd. Compounded Ratio of (a) 5 : 6 and 2 : 3 is 5 : 9
(b) 2 : 5, 5 : 7, and 6 : is 3 : 14

Question 1: Ratio of 5 kg and 15 kg is

Solution Ratio $= \frac{5kg}{15kg} = \frac{1}{3} = 1:3$

They have same units, so it is a ratio.

Question 2 Ratio of 5m and 200 cm; is

Solution Since 1 metre = 100 cm, Ratio of 5 m & 200 cm $= \frac{5 \times 100\text{cm}}{200\text{cm}} = \frac{5}{2} = 5:2$

Question 3 Ratio of the height 170 cm and weight 60 kg of a person is

Solution Cm is unit of length and kg is unit of mass. So, they cannot be converted in same units. 170 cm : 60 kg is not a ratio.

Question 4 Which ratio out of the following ratios is greater. $2\frac{1}{3} : 3\frac{1}{3}$ and 3.6:4.8

Solution $2\frac{1}{3} : 3\frac{1}{3} : = \frac{7}{3} : \frac{10}{3} = 7 : 10$ $7:10 = 0.70$ $\frac{3.6}{4.8} = 0.75$ Clearly 3.6 : 4.8 is greater ratio.

Unit II Proportion

Definition

AN EQUALITY OF TWO RATIO IS CALLED A PROPORTION

If a, b, c, d are four numbers then a : b = c : d [Also written as a : b : : c : d] is called Proportion of four numbers.

a : b = c : d

$\therefore \frac{a}{b} = \frac{c}{d}$

$\therefore$ ad = bc

[Product of 1st and 4th] = [Product of 2nd and 3rd]

Note : If a : b = c : d then 'd' is called 'Fourth Proportion'

Mean Proportion : If a, b, c are in Continuous Proportion Then Middle Term b is called The Mean Proportion.

$b^2 = ac$ **then** $b = \sqrt{ac}$ Mean Proportion

Note : Mean Proportion of two number x and y is $\sqrt{xy}$

Types Of Proportion

1. INVERTENDO PROPORTION

If a : b=c : d then b : a=d : c is called INVERTENDO PROPORTION

2. ALTERNENDO PROPORTION

If a : b=c : d then a : c=b : d is called ALTERNENDO PROPORTION

3. COMPONENDO PROPORTION

If a : b=c : d then (a+b): b=(c+d):d is called COMPOUND PROPORTION

4. DIVIDENDO PROPORTION

If a : b=c : d then (a-b): b=(c-d): d is called DIVIDENDO PROPORTION

5. COMPOUNDED & DIVIDENDO PROPORTION

If a : b=c : d then (a+b):(a-b)=(c+d):(c-d) is CALLED COMPOUNDED and DIVIDENDO PROPORTION

6. ADDENDO PROPORTION

If a : b = c : d=e : f .... then (a+c+e ....) : (b+d+f+ ....) is called ADDENDO PROPORTION

Continuous Proportion

If a, b, c are Continuous Proportion then $\frac{a}{b} = \frac{c}{d}$ or $b^2 = ac$

i.e. Ratio of 1st of 2nd = Ratio of 2nd to 3rd

i.e. Square of middle term = Product of 1st and last

Note : If x, y, z a, b, c .... are Continuous Proportion then

$$\frac{x}{y} = \frac{y}{z} = \frac{z}{a} = \frac{a}{b} = \frac{b}{c} = \text{..................................}$$

Question 5 If A: B= 2 : 3 ; B : C = 4 : 5 and C : D = 6 : 7 Then A : D is equal to

Solution $\frac{A}{B} \times \frac{B}{C} \times \frac{C}{D} = \frac{2}{3} \times \frac{4}{5} \times \frac{6}{7}$ A : D = 16 : 35

Question 6 Two vessels contain equal quantity of mixtures of milk and water in the ratio 9 : 5 and 4 : 3 respectively. Both the mixtures are mixed find the ratio of milk to water in the new mixture.

Tricks Milk : Water sum

Vessel I	9	:	5] 14 × 1 = 14
Vessel II	4	:	3] 7 ×2 = 14
Vessel III	(9 + 4×2)	:	5 + 3x2]

(Adding corresponding terms of the ratio)

Take LCM of 14 & 7 = 14.

Then make it of equal quantity (i.e. = 14)

∴ Ratio of Milk to Water in the new vessel = 17 : 11.

Question 7 Two vessels containing water and milk in the ratio 2 : 3 and 4 : 5 are mixed in the ratio 1 :2. The ratio of milk and water in the resulting mixture.

Tricks

Step 1. Milk : Water sum

Vessel I 3:2]5×9=45

Vessel II 5 : 4] 9 x 5 = 45

Step 2.

∴Vessel I 27: 18] 45x1

Vessel II 25 : 20] 45 x 2

Step 3.

Mixing Vessel I & Vessel II mixtures, we get Vessel III (27 + 50) : (18+40) = 58 : 77

Take LCM of 5 & 9 = 45.

Making same quantities of Vessel I & II i.e. 45. We have to mix the mixtures of vessel I & II in the ratio 1 : 2

∴ Multiply I by 1 and II by 2 ,Then we add. We get the answer.

Rules of Alligation

It is used in mixing of two varieties of same kind. It is derived from the weighted mean method.

Rule

If two varieties of rices with rate Rs. x per kg and Rs. y per kg are mixed to make a third variety of rice with rate Rs. z per kg. The ratio in which these two varieties are mixed is

Remember

(i) If x represents cost then y and z must be cost.

(ii) If x represents selling price then y and z must be selling price.

(iii) If x represents profit then y and z must be profit or loss.

(iv) If x represent milk of 1st mixture then y and z must represent milk of 2nd mixture and mixed mixture.

Question 8 Sugar at rate Rs. 15 per kg is mixed with sugar at rate Rs. 20 per kg in the ratio 2:3. Find the price per kg of the mixture.

Solution : Tricks $\frac{20 - x}{x - 15} = \frac{2}{3}$ x = Rs. 18.

Question 9 Varun mixes milk costing Rs.6.92 per litre with milk costing Rs. 7.77 per litre and sells the mixture at Rs. 8.80 per litre and earns a profit of 17.5% on his sales price. In what proportion does he mix them?

Solution Selling price per litre = Rs. 8.80 Less profit 17.5% of 8.80 = Rs. 1.54 Cost per litre **Rs. 7.26 = 3:2**

Question 10 Two numbers are in the ratio 2 : 3 and the difference of their squares is 320. The number are :

(a) 12,18 (b) 16,24 (c) 14,21 (d) None

Solution : Tricks Go by choices (a); (b) & (c) all are in ratio 2:3 But For option

(a) 182 -122 ≠ 320

(b) 242 -162 =(24 + 16) (24 - 16)

= 40 x 8 = 320

Question 11 If p : q is the sub-duplicate ratio of $p - x^2 : q - x^2$, then x^2 is :

Solution Detail Method = $\frac{\sqrt{p-x^2}}{\sqrt{q-x^2}} = \frac{p}{q}$

Squaring on both side; we get : $\frac{p-x^2}{q-x^2} = \frac{p^2}{q^2}$

or $pq^2-q^2 x^2 = p^2 q-p^2 x^2$

or $p^2 x^2-q^2 x^2 = p^2 q-pq^2$

or $x^2 (p^2-q^2) = Pq(P-q)$

or $x^2 (p+q)(p-q)=pq(p - q)$

or $x^2 = \frac{pq}{p+q}$

Question 12 An alloy is to contain copper and zinc in the ratio 9:4. The zinc required to melt with 24 kg of copper is:

Solution Let Zinc = x kg $\therefore \frac{9}{4} = \frac{24}{x} \therefore x=\frac{4\times24}{9} = \frac{32}{3} = 10\frac{2}{3}$ kg

Question 13 Two numbers are in the ratio 7: 8. If 3 is added to each of them, their ratio becomes 8 : 9. The numbers are:

Solution Let x is common in the ratio

$\therefore$ Numbers are 7x & 8x

Now $\frac{7x+3}{8x+3} = \frac{8}{9}$

or 64x + 24 = 63x + 27

or 64x - 63x = 27-24

or x = 3

1st number = 7x = 7×3 = 21

2nd number = 8x = 8x3 = 24

Question 14 A box contains Rs. 56 in the form of coins of one rupee, 50 paise and 25 paise. The number of 50 paise coin is double the number of 25 paise coins and four times the numbers of one rupee coins. The numbers of 50 paise coins in the box is :

Solution Let No. of 50 Paise coins = x

$\therefore$ No. of Re. 1 coins = $\frac{x}{4}$ and No. of 25 Paise coins = $\frac{x}{2}$

$\therefore$ Total Value = $\frac{x}{4}$ x 1 + x 0.50 + $\frac{x}{2}$ x 0.25 = 56

or 0.25x + 0.50x + 0.125x

or 0.875x = 56

or $x = \frac{56}{0.875} = 64$

Question 15 A bag contains Rs. 187 in the form ofl rupee, 50 paise and 10 paise coins in the ratio 3:4:5. Find the number of each type of coins:

Solution Let x is common in the ratio

No. of 1 Rupee ; 50 Paise and 10 Paise

Coins are 3x ; 4x and 5x

3x×1+4x×0.50+5x×0.10=187

$x = \frac{187}{5.50} = 34$

∴ No. of 1 Rupee coins = 3x = 3 × 34 = 102

No. of 50 Paise coins = 4x= 4x34 = 136

No. of 10 Paise coins = 5x = 5x34 = 170

Question 16 Ratio of earnings of A and B is **4** : 7. If the earnings of A increase by 50% and those of B decrease by 25%, the new ratio of their earning becomes 8 : 7. What is A's earning ?

Solution Let x is common in the ratio

∴ A's and B's present earnings are 4x and 7x respectively From question

$x = \frac{4x + 4x \text{ x } 0.50}{7x - 7x \text{ x } 0.25} = \frac{8}{7}$

OR; $\frac{6}{5.25} = \frac{8}{7}$

x cannot be found. ∴ Data is inadequate

Question 17 P, Q and R are three cities. The ratio of average temperature between P and Q is 11 : 12 and that between P and R is 9: 8. The ratio between the average temperature of Q and R is :

Solution P : Q = 11 : 12 Q : P = 12 : 11

$\frac{Q}{P} \text{ x } \frac{P}{R} = \frac{12}{11} \text{ x } \frac{9}{8}$

$\frac{Q}{R} = \frac{27_1}{22_1}$ Q : R = 27.22

Question 20 Rs. 407 are to be divided among A, B and C so that their shares are in the ratio 1/4:1/5:1/6 The respective shares of A, B, C are:

Solution A : B : C = $\frac{1}{4}$; $\frac{1}{5}$; $\frac{1}{6}$] x LCM of denominator = 60 = 15 : 12 :10

A's share = $\frac{407}{15+12+10}$ = Rs. 165

B's share = $\frac{407}{37}$ x 12 = Rs. 132

C's share = $\frac{407}{37}$ x 10 = Rs,110

Question 21 The incomes of A and B are in the ratio 3 :2 and their expenditures in the ratio 5 : 3. If each saves Rs. 1,500, then B's income is:

Solution Detail Method Let x is common in the ratio.

∴ A's income = 3x

B's income = 2x

$$\frac{3x - 1500}{2x - 1500} = \frac{5}{3}$$

or 10x - 7500 = **9x** - 4500 or 10x - 9x = 7500 - 4500 or x = 3000

:- B's income = 2x = 2x3000 = Rs 6000.

Question 22 In 40 litres mixture of glycerine and water, the ratio of glycerine and water is 3:1. The quantity of water added in the mixture in order to make this ratio 2:1 is:

Solution Glycerine = $\frac{40}{3+1} \times 3 = 30$ litres.

Water = $\frac{40}{4} \times 1 = 10$ litres.

Let x liters of water is added to the mixture

Then $\frac{30}{10 + x} = \frac{2}{1}$ or 2x + 20 = 30 or x = 5

Question 23 The third proportional between (a^2-b^2) and $(a+b)^2$ is :

Solution 3rd Proportion = $\frac{(\text{Meanprop.})^2}{\text{IstProp}}$

$$= \frac{\{(a+b)^2\}^2}{a^2-b^2} = \frac{(a+b)^4}{3(a+b)(a+b)} = \frac{(a+b)^3}{a-b}$$

Question 24 If A, B and C started a business by investing Rs. 1,26,000, Rs. 84,000 and Rs. 2,10,000. If at the end of the year profit is Rs. 2,42,000 then the share of each is :

Solution Investment ratio is

A : B : C = 126,000 : 84,000 : 2,10,000] ÷14,000

= 9:6:15] +3

= 3: 2 : 5

A's share = $\frac{\text{Rs}242{,}000}{3+2+5} \times 3 = \text{Rs}72{,}600$

B's share = $\frac{242{,}000}{20} \times 2$ = RS. 48,400

C's share = $\frac{242{,}000}{10} \times 5$ = Rs. 1,21,000

Question 24 If $\frac{p}{q} = -\frac{2}{3}$ then the value of $\frac{2p+q}{2p-q}$ is:

Solution $\because \frac{p}{q} = \frac{-2}{3}$ Tricks $\frac{2p+q}{20-q} = \frac{2(-2)+3}{2(-2)-3} = \frac{-4+3}{-4-3} = \frac{-1}{-7} = \frac{1}{7}$

Question 26 Fourth proportional to x, 2x, (x +1) is:

Solution Let Fourth Proportional is K. $\therefore \frac{x}{2x} = \frac{x+1}{K}$

or k.x = 2x (x + 1)

or k = 2 (x + 1) = 2x +2

Question 27 What must be added to each term of the ratio 49 : 68 so that it becomes 3.: 4 ?

Solution Let Fourth Proportional is K. Then $= \frac{49+x}{68+8} = \frac{3}{4}$

or 196 +4x = 204 + 3x

or 4x - 3x = 204 - 196

or x = 8

Question 28 If A : B= 2:5, then (10A + 3B): (5A+2B) is equal to

Solution It A : B = 2 : 5 Then

$\frac{10A+3B}{5A+2B} = \frac{10\times2+3\times5}{5\times2+2\times5} = \frac{35}{20} = \frac{7}{4} = 7:4$

Question 29 In a film shooting, A and B received money in a certain ratio and B and C also received the money in the same ratio. If A gets Rs. 1,60,000 and C gets Rs.2,50,000. Find the amount received by B ?

Solution A : B = B : C

So, $B^2 = AC$;

So, $B = \sqrt{AC} = \sqrt{1,60,000\times2,50,000}$

= 400 × 500 = 2,00,000

Question 30 The ratio compounded of 4:5 and sub-duplicate of "a" : 9 is 8:15. Then value of "a" is

Solution $\frac{4}{5} \times \sqrt{\frac{a}{9}} = \frac{8}{15}$ or $\frac{4}{5} \times \frac{\sqrt{a}}{9} = \frac{8}{15}$ $\sqrt{a} = 2 \Rightarrow a=4$

Question 31 The ratio compounded of 4:5 and sub-duplicate of "a" : 9 is 8:15. Then value of "a" is

Solution $\frac{4}{5} \times \sqrt{\frac{a}{9}} = \frac{8}{15}$

or $\frac{4}{5} \times \sqrt{\frac{a}{3}} = \frac{8}{15}$

$\therefore \sqrt{a} = 2 \Rightarrow a=4$

Question 32 If X varies inversely as square of Y and given that Y =2 for X = 1, then the value of X for Y =6 will be

Solution $x = \frac{1}{y^2} = X = K\frac{1}{y^2} \Rightarrow X = \frac{k}{y^2}$

When x = 1 Then y = 2

$x = \frac{4}{y^2}$

When y = 6 Then $x = \frac{4}{6^2} = \frac{1}{9}$

$\therefore x = \frac{1}{9}$

Question 33 Triplicate ratio of 4 : 5 is

(a) 125 : 64 (b) 16 : 25

(c) 64 : 125 (d) 120 : 46

Solution (c) Triplicate ratio of 4:5 = $4^3 : 5^3$ = 54:125

Question 34 The mean proportion between 24 and 54 is _______

(a) 33 (b) 34

(c) 35 (d) 36

Solution (d) Mean - Proportion = $\sqrt{24 \times 54}$ = 36

Question 35 If P is 25% less than Q and R is 20% higher than Q the Ratio of R and P

(a) 5:8 (b) 8:5

(c) 5:3 (d) 3:5

Solution (b) is correct

Let Q = 100, So, P = 100-025 = 75

&R= 100 + 20= 120

$\frac{R}{P} = \frac{120}{75} = \frac{8}{5}$

Question 36 A person has assets worth Rs. 1,48,200. He wish to divide it amongst his wife, son and daughter in the ratio 3 : 2 : 1 respectively. From this assets the share of his son will be

Solution Share of son = $\frac{2}{3+2+1}$ X 1,48,200
= Rs.49,400

Question 37 If x : y = 2 : 3 then (5x+2y): (3x -y) =

Solution $\frac{5x + 2y}{3y - y} = \frac{5 \times 2 + 2 \times 3}{3 \times 2 - 3} = \frac{16}{3}$

Question 38 The first, second and third month salaries of a person are in the ratio 2:4:5. The difference between the product of the salaries of first 2 months & last 2 months is Rs. 4,80,00,000. Find the salary of the second month

Solution Let x be common in the ratio.
1^{st}, 2^{nd}, and 3^{rd} month salaries of a person = 2x ; 4x ; 5x :- From this question
4x×5x-2x×^x = 4,80,00,000.
or,$12x^2$ -4,80,00,000.
or,x^2 = 4000000 x = 2000.
2^{nd} month salary = 4x = 4x 2000
= Rs. 8000

Question 39 $(2p^2 - q^2) = 7pq$, where p, q are positive then p : q
(a) 5:6 (b) 5:7
(c) 3:5 (d) 3:7

Solution (a) is correct
$15(2p^2 - q^2) = 7pq$

Tricks: Go by choices For (a) put p = 5; q = 6 we get
$15|2\times5^2 - 6^2| = 3\times5\times6$
or 15x14 = 210
or 210 = 210

Question 40 Find the ratio of third proportional of 12,30 and mean proportional of 9, 25:

(a) 7:2 (b) 5 : 1

(c) 9:4 (d) None of these

Solution 3rd proportional = $\frac{30^2}{12}$ = 75

Mean Proportional = $\sqrt{9 \text{ X } 25}$ = 15

Ratio = $\frac{75}{15}$ = 5:1

$\therefore$ (b) is correct

Question 41 What must be added to each of the numbers 10, 18,22, 38 to make them proportional:

(a) 5 (b) 2

(c) 3 (d) 9

Solution (b) is correct

let x be added.

$\frac{10 + x}{18 + x} = \frac{22 + x}{38 + x}$

Tricks: Go by choices.

$\therefore$ x = 2 satisfies it.

Question 42 x, y, z together starts a business. If x invests 3 times as much as y invests and y invests two third of what z invests, then the ratio of capitals of x, y, z is

(a) 3:9:2 (b) 6:3:2

(c) 3:6:2 (d) 6 : 2 : 3

Solution (d)

Tricks: Go by choices

6 = 3×2 and 2 = 3 x $\frac{2}{3}$

Question 43 If a:b = 2:3,b:c = 4:5,c:d=6:7 then a: d is _____

Solution Multiply all ratios. = $\frac{2}{3}$ x $\frac{4}{5}$ x $\frac{6}{7}$ = $\frac{16}{35}$

Question 43 If $\frac{1}{2}, \frac{1}{3}, \frac{1}{5}, \frac{1}{x}$ are in proportion then x =

Solution Product of middle two terms = Product of extremes

$\frac{1}{2x} = \frac{1}{15}$; x = 15/2

Unit III Indices

Definition

If n is a positive integer, and 'a' is a real number, i.e. $n \in N$ and $a \in R$ (where N is the set of positive integers and R is the set of real numbers), 'a' is used to denote the continued product of n factors each equal to 'a' as shown below:

$a^n = a \times a \times a$ ....to n factors.

Here a^n is a power of "a" whose base is "a" and the index or power is "n". For example, in $3 \times 3 \times 3 \times 3 = 3^4$, 3 is base and 4 is index or power.

Laws of Indices

1. $a^m a^n = (a)^{m+n}$
2. $\frac{a^m}{a^n} = a^{m-n}$
3. $\frac{a^m}{a^{-n}} = (a)^{m+n}$
4. $a^0 = 1$
5. $[a^m]^n = a^{mn}$
6. $[a^m]^{\frac{1}{n}} = (a)^{\frac{m}{n}}$
7. $\sqrt[n]{a^m} = (a)^{\frac{m}{n}}$
8. $[ab]^m = (a)^m (b)^m$
9. $\left[\frac{a}{b}\right] = \frac{a^m}{b^m}$
10. $a^{-m} = \frac{1}{a^m}$
11. If $a^x = a^y \therefore x = y$
12. If$(a)^x = (b)^x \therefore x = y$
13. If $a^x = 1 \therefore x = 0$

Solved Questions

Question 1: Value of $(a^{1/8} + a^{-1/8})(a^{1/8} - a^{-1/8})$

Solution $[a^{1/8} + a^{-1/8}][a^{1/8} - a^{-1/8}][a^{1/4} + a^{-1/4}][a^{1/2} + a^{-1/2}]$

[Use Formula $(a+b)(a-b) = a^2 - b^2$]

$= [(a^{1/8})^2-(a^{1/8})^2][a^{1/4}+a^{-1/4}][a^{1/2}+a^{-1/2}]$

$= (a^{1/4}-a^{-1/4})(a^{1/4}+a^{-1/4})(a^{1/2}+a^{-1/2})$

$= [(a^{1/4})^2-(a^{-1/4})^2][a^{1/2}+a^{-1/2}]$

$= (a^{1/2}-a^{-1/2})(a^{1/2}+a^{-1/2})$

$= (a^{1/2})^2-(a^{-1/2})^2=a-a^{-1}=a-\frac{1}{a}$

Question 2 If $4^x=5^y=20^z$ then z is equal to:

Solution Let $4^x=5^y=20^z=k$

or $4=k^{1/x}; 5=k^{1/y}; 20=k1/8$

$\therefore 20=4\times5$

or $k^{1/z} = k^{1/x}k^{\frac{1}{y}}$

$k^{1/z} = k^{\frac{1}{k}+\frac{1}{y}}$

$\frac{1}{y}=\frac{1}{x}+\frac{1}{y}=\frac{1}{z}=\frac{y+x}{xy}$

$z = \frac{xy}{x+y}$

Question 3 $\left(\frac{\sqrt{3}}{9}\right)^{5/2}\left(\frac{9}{3\sqrt{3}}\right)^{7/2}$ x 9 is equal to.....

Solution $\left(\frac{\sqrt{3}}{9}\right)^{5/2}\left(\frac{9}{3\sqrt{3}}\right)^{7/2}$ x 9

$=[(\frac{3^{1/2}}{3^2}\ 1^5\ (\frac{3^2}{33^{1/2}})^7]^{1/2} \times 3^2$

$=[3(\frac{1}{2}-2)\ 5 \times 3(2-1-\frac{1}{2})\ 7]^{1/2}.3^2$

$=[3\ \frac{-15}{2}.\ 3^{7/2}]^{1/2} \times 3^2 = (3\ \frac{-15}{2}+\frac{7}{2})^{1/2}.3^2$

$=(3^{-4})^{112}.3^2=3^{-2}.3^2=3^{-2+2}=3^0=i$

Question 4 If $2^x-2^{x-1} = 4$, then the value of x^x is:

Solution $2^x-2^{x-1} = 4$

or $2^{x-1}\langle 2-1)=4$

or $2^{x-1} \times 1=2^2$

$: x - 1 = 2$

$: x = 3$

$\therefore x^x=3^3=27$

Question 5 If $x = 3^{1/3}+3^{-1/3}$ then fmd value of $3x^3-9x$

Solution If $x=3^{1/3}+3^{-1/3}$

By Cubing on both sides; we get

$x^3=(3^{1/3}\rangle 3+(3^{-1/3}\rangle 3+3.3^{1/3}.3^{-1/3}\langle 3^{1/3}+3^{-1/3})$

$= 3 + 3^{-1} + 3 \times 1/xx$ (From (1))

or $x^3 = 3 + \frac{1}{3} + 3x$

or $x^3 - 3x = \frac{9+1}{3}$

or $3x^3-9x=10$

Question 6 Find the value of. $[1-\{1-(1-x^2)^{-1}\}1]^{\frac{1}{3}}$ is

Solution
$[1-\{1-(1-x^2)^{-1}\}-1]^{1/2}$

$[1-\{1-\frac{1}{1-x^2}\}^{-1}]^{-1/2}$

$= [1-\{\frac{1-x^2-1}{1-x^2}\}^{-1}]^{-1/2}$

$= [1-\{\frac{-x^2}{1-x^2}\}^{-1}]^{-1/2}$

$= [1-\frac{1-x^2}{-x^2}]^{-1/2} = [1+\frac{1-x^2}{x^2}]^{1/2}$

$= [\frac{x^2+1-x^2}{x^2}]^{-1/2}$

$= (\frac{1}{x^2})^{-1/2} = (x^{-2})^{\frac{1}{2}} = x$

Question 7 $\frac{2^n+2^{n-1}}{2^{n+1}-2^n}$

Solution $\frac{2^n+2^{n-1}}{2^{n+1}-2^n} = \frac{2^n(2^0+2^{-1})}{2^n(2^1-2^0)} = \frac{1+\frac{1}{2}}{2-1} = \frac{3}{2}$

Question 8 Simplify $2x^{1/2}\,3x^{-1}$ if $x = 4$

Solution We have $2x^{1/2}\,3x^{-1} = 6x^{1/2}\,x^{-1} = 6x^{1/2-1} = 6x^{-1/2}$

$= \frac{6}{x^{1/2}} = \frac{6}{4^{1/2}} = \frac{6}{[2^2]^{1/2}} = \frac{6}{2} = 3$

Question 9 Simplify $6ab^2c^3 \times 4b^{-2}c^{-3}d$.

Solution
$6ab^2c^3 \times 4b^{-2}c^{-3}d$
$= 24 \times a \times b^2 \times b^{-2} \times c^3 \times c^{-3} \times d$
$= 24 \times a \times b^{2+(-2)} \times c^{3+(-3)} \times d$
$= 24 \times a \times b^{2-2} \times c^{3-3} \times d$
$= 24a\, b^0 \times c^0 \times d = 24ad$

Question 10 Simplify $(x^a.y^{-b})^3 . (x^3 y^2)^{-a}$

Solution
$(x^a.y^{-b})^3 . (x^3 y^2)^{-a}$

$= (x^a)^3 . (y^{-b})^3 . (x^3)^{-a} . (y^2)^{-a}$

$= x^{3a-3a} . y^{-3b-2a}.$

$= x^0 . y^{-3b-2a}. = 1/y^{3b+2a}$

Question 11 Find x, if $\sqrt{x} = (x\sqrt{x})^x$

Solution $x\,(x)^{1/2} = x^x . x^{x/2}$

or, $x^{1+1/2} = x^{x+x/2}$

or, $x^{3/2} = x^{3x/2}$

[If base is equal, then power is also equal]

i.e $\frac{3}{2} = \frac{3x}{2}$ or $\frac{3}{2} \times \frac{2}{2} = 1$

$\therefore x = 1$

Question 12 Find the value of k from $(\sqrt{9})^{-7} \times (\sqrt{3})^{-5} = 3^k$

Solution $(\sqrt{9})^{-7} \times (\sqrt{3})^{-5} = 3^k$

or, $(3^{2 \times 1/2})^{-7} \times (3^{1/2})^{-5} = 3^k$

$3^{-7} \times 3^{-5/2} = 3^k$ or $3^{-7-5/2} = 3^k$

or, $3^{-19/2} = 3^k$ or, $k = -19/2$

Unit IV Logarithms

At the foundation level the concepts Logarithms is used in accounting and finance. Here in this Unit Logarithms an attempt is made for solving and understanding the concepts of with the help of following questions with solutions.

if $a^b = c$; Where $a \neq 1$ and a; $c > 0$ (positive)
Then b is said to be the logarithm of the number c to the base "a" and expressed as
$\text{Log}_a c = b$; Where $a \neq 1$.

Types of Logarithms

(I) Natural Logarithm:

The Logarithm of a number to base "e" is called Natural Logarithm.

i.e. $\text{Log}_e x$

where x = a number e = 2.7183

(ii) Common Logarithm:

Logarithm of a number to the base 10 is called common Logarithm.

i.e. $\text{Log}_{10} x$

where x = A number

Note; If base is not given then in arithmetical or commercial work; base is always taken as 10.

Remember Some Formulae

1. If $a^b = c \iff \text{Log}_a c = b$; Where $a \neq 1$.

2. $a^{x\log ab} = bx$

3. $\log_a \alpha = l$

4. $\log_a 1 = 0$

5. $\log_b a = \dfrac{1}{(\log_a b)} \Rightarrow \log_b a \cdot \log_a b = 1$

6. (i) $\log_b a = \log_b x \log_x a = \log_x a \log_b x$

   (ii) $\log_b a = \log_x a.\log_y x.\log_z y....\text{iog}_b k$

   $\log_b a = \log_b x.\log_x y.\text{logyz} \log_{l\varsigma} a$

7. (i) $\log_b a = \dfrac{\log_x a}{\log_x b}$

   (i) $\log_b a = \dfrac{\log_b a}{\log_l b}$

8. if logb a =x

   Then (i) $\log_{\frac{1}{b}} a = \_X$

   (ii) $\log_b \frac{1}{a} = -x$

   (iii) $3^{\log 1} \underline{1} + a$

9. (i) $\log_a (mn) = \log am + \log an$

(ii) $\log a(mnr...) = \log am + \log_a n + \log_a r + ...$

10. $\log_a\left(\frac{m}{n}\right) = \log_a m - \log_a n$

11. (i) $\log(a^b)n^n = \frac{n}{b} \log a^m$

(ii) $\log_a (m^n) = n\log_a m$

(iii) $\log_{a^b} m = \frac{1}{b} \log_a m$

12. (i) If $\log_a m = \log_b m \Rightarrow a=b$

(ii) If $\log_a m = \log_a n \Rightarrow m = n$

Question 1 $\log_2 16 = ?$

(a) 3 (b) 4 (c) 5 (d) None

Solution (b) is correct Let $\log_2 16 = x$

$\therefore 2^x = 16 = *$ $\quad 2^x = 2^4$

$\therefore x = 4.$

Question 2 If $\log_a 16 = 4$; then a equals.

(a) 4 (b) 16 (c) 2 (d) None

Solution (c) is correct Go by Choices $a^4 = 16$

Question 3 $16^{\log 4} 5 = ?$

(a) 5 (b) 16 (c) 25 (d) None

Solution (c) is correct

$16^{\log 4} 5 = (4^2)^{\log 4} 5$

$= 4^{2\log_4 5} = 5^2 = 25.$

Question 4 If $\log_{9} x + \log_4 x + \log_2 x = 11$, The value of x is

Solution $\log_2 3^x + \log_2 2^x + \log_2 x = 11$

$\Rightarrow \frac{1}{3} \log_2 X + \frac{1}{2} \log_2 X + \log_2 X = 11$

$\Rightarrow \log_2 X \left[\frac{1}{3} + \frac{1}{2} + 1\right] = 11$

$\Rightarrow \log_2 X \left[\frac{2 + 3 + 6}{6}\right] = 11$

$\Rightarrow \log_2 X . \frac{11}{6} = 11$

or$\log_2 x = \frac{11 \times 6}{11} = 6$

or $2^6 = x$

$= x = 64$

Question 5 The value of $\log_5 \sqrt{5\sqrt{5\sqrt{5.....10}}}$ to ∞ is

(a) 0 (b) 1 (c) 2 (d) None

Solution (b) is correct

$\therefore \sqrt{5\sqrt{5\sqrt{5... ...+}}}$ upto ∞ = 5 is

$\therefore \log_5 5 = 1$

Question 6 The value of $\log_8 25$ given log 2 = 0.3010 is
(a) 1 (b) 2 (c) 1.5482 (d) None

Solution (c) is correct

$\log_8 25 = \log_2 a 5^2 = \frac{2}{3} \log_2 5$

$= \frac{2}{3} \log_2 \left(\frac{10}{2}\right) = \frac{2}{3} [\log_2 10 - \log_2 2]$

$= \frac{2}{3} \left[\frac{1}{\log_{10} 2} - 1\right]$

$= \frac{2}{3} \left[\frac{1}{0.3010} - 1\right] = 1.5482$

$\therefore$ (c) is correct.

Question 7 $\frac{1}{1 + \log_a(bc)} + \frac{1}{1 + \log_b(ca)} + \frac{1}{1 + \log_c(ab)}$ is equal to

(a) 0 (b) 1 (c) 3 (d) -1

Solution (b) is correct

$$\frac{1}{\log_a a + \log_a(bc)} + \frac{1}{\log_b b + \log_b(ca)} + \frac{1}{\log_c c + \log_c(ab)}$$

$$= \frac{1}{\log_a (abc)} + \frac{1}{\log_b (abc)} + \frac{1}{\log_c (abc)}$$

$= \log_{abc} a + \log_{abc} b + \log_{abc} c$

$= \log_{abc} abc = 1$

Question 8 If $\frac{\log a}{y-z} = \frac{\log b}{z-x} = \frac{\log c}{x-y}$ the value of $a^{y+z}. b^{z+x}. c^{x+y}$ is given by

(a) 0 (b) 1 (c) -1 (d) None

Solution (b) is correct

Let $\frac{\log a}{y-z} = \frac{\log b}{z-x} = \frac{\log c}{x-y} = k$

$\log a=k(y-z) \Rightarrow (y+z)\log a=k(y-z)(y+z)$

$\Rightarrow \log a^{y+z} = k(y^2-z^2)$

$\log b = k(z-x) \Rightarrow (z+x)\log b=k(z-x)(z+x)$

$\Rightarrow \log a^{z+x} = k(z^2-x^2)$

Similarly

$\log c = k\ (x - y) => \log c^{x+y} = k(x^2 - y^2)$

Adding Then

$\log a^{y+z} + \log b^{z+x} + \log c^{x+y} = k\ (y^2 - z^2 + z^2 - x^2 + z^2 - x^2 + x^2 - y^2)$

or $\log\ (a^{y+z}. b^{z+x}. c^{x+y}) = k\text{x}\ 0 = 0 = \log 1,$

$\therefore a^{y+z}. b^{z+x} \cdot c^{x+y} = 1$

Question 9 The value of the expression ${}_a\log_a b \cdot \log_b c \cdot \log_c{}^d \cdot \log_d{}^t)$

(a) t (b) abcdt c) (a + b + c + d +1) (d) None.

Solution ${}_a\log_a{}^b \cdot \log_b c \cdot \log_c d \cdot \log_d t$

$= a^{1\log_a t}) = t^1 = t$

$\therefore$ (a) is correct

Question 10 If $\log 10000^x = \frac{-1}{4}$ then x is given by :

(a) 1/100 (b) 1/10 (c) 1/20 (d) None of these.

Solution $\log_{10000} x = \frac{1}{4}$

$\therefore (10000)^{\frac{1}{4}} = x$

or $(10^4)^{\frac{1}{4}} = x$

or $x = 10^{-1} = \frac{1}{10}$

$\therefore$ (b) is correct

Question 11 If log (2a- 3b) = log a- log b, then a=?

(a) $\frac{3b^2}{2b-1}$ (b) $\frac{3b}{2b-1}$ (c) $\frac{b^2}{2b+1}$ (d) $\frac{3b^2}{2b+1}$

Solution $\log(2a-3b) = \log\frac{a}{b}$

or $2a-3b = \frac{a}{b}$

or $2ab-3b^2=a$

or $2ab-a=3b^2$

or $a(2b-1)=3b^2$

or $a = \frac{3b^2}{2b-1}$

$\therefore$ (a) is correct

Question 12 $\frac{1}{\log\varphi\langle abc\rangle} + \frac{1}{\log_{bc}(abc)} + \frac{1}{\log_{co}(abc)}$ is equal to :

(a) 0 (b) 1 (c) 2 (d) -1

Solution $\frac{1}{\log_{ab}\langle abc\rangle} + \frac{1}{\log_{bc}(abc)} + \frac{1}{\log_{ca}(abc)}$

$= \log_{abc}(abbcca) = \log_{abc}(abc)2$

$=2\log_{abc}(abc) = 2\times1=2$

$\therefore$ (c) is correct

Question 13 Number of digits in the numeral for 2^{64} [Given log 2 = 0.30103] :

(a) 18 digits (b) 19 digits (c) 20 digits (d) 21 digits

Solution (c) Let $x = 2^{64}$

log x = 64 log 2

= 64x0.30103 = 19.26592

x = Anti Log (19.26592)

characteristics =19

No. of digits in the number =19 + 1

$\therefore$ (c) is correct = 20

Question 14 The value $\frac{\log_3 8}{\log_9 16\cdot\log_4 10}$

(a) $3\log_{10} 2$ (b) $7\log_{10}3$ (c) $3\log_e z$ (d) None

Solution $\frac{\log_3 8}{\log_9 16\cdot\log_4 10} = \frac{\log_3 2^3}{\log_3 2 2^4\times\log_2 2^{10}}$

$\frac{3\log_3{}^2}{\frac{4}{2}\log_3{}^2 x \frac{4}{2}\log_2 10} = \frac{3}{\log_2 10} = 3\log_{10}2$ $\therefore$ (a) is correct

Question 15 If $x = \frac{e^n - e^{-n}}{e^n + e^{-n}}$ then the value of n is :

Solution $\frac{x}{1} = \frac{e^n - e^{-n}}{e^n + e^{-n}}$

By Componendo and dividendo; we get

$\frac{1+x}{1-x} = \frac{e^n + e^{-n} + e^n - e^{-n}}{e^n + e^{-n} - e^n - e^{-n}} = \frac{2e^n}{2e^{-n}}$

or $\frac{1+x}{1-x} = e^{n+n}$

or $e^{2n} = \frac{1+x}{1-x}$

Taking log on both sides; we ge

$\log_e e^{2n} = \log_e \left(\frac{1+x}{1-x}\right)$

or $2n\log^e_e = \log_\varepsilon \left(\frac{1+x}{1-x}\right)$

or $2n \times 1 = \log \left(\frac{1+x}{1-x}\right)$

$n = \frac{1}{2} \log \left(\frac{1+x}{1-x}\right)$

Question 16 log 144 is equal to:

Solution log l44= log (16×9)

= log 16+ log 9

= log 2^4+ log 3^2

=4 log 2+2 log 3

Question 17 If log2 [$\log_3$ ($\log_2$ x)]=1, then x equals;

Solution $\log_2[\log_3 (\log2\ x)] = 1$

Tricks : $2^{321} = x$

or $x = 2^{32} = 2^{3\times3} = 2^9 = 512$

Detail Method:

$\log_2[\log3 (\log2\ x)] = 1$

or $\log_3 (\log_2 x) = 2^1 = 2$

or $\log_2 x = 3^2$

or $\log_2 x = 9$

or $x = 2^9 = 512$

Question 18 If $\log\left(\frac{a+b}{4}\right) = \frac{1}{2}(\log a + \log b)$ then: $\frac{a}{b} + \frac{b}{a}$

Solution $\log\left(\frac{a+b}{4}\right) = \frac{1}{2}(\log a + \log b)$

or $\log\left(\frac{a+b}{4}\right) = \log (ab)^{1/2}$

or $\frac{a+b}{4} = \sqrt{ab}$

or $a+b=4\sqrt{ab}$

Squaring on both sides; we get

$(a+b)^2=16ab$

or $a^2+b^2+2ab=16ab.$

or $a^2+b^2=14ab$

or $\frac{a^2}{ab} + \frac{b^2}{ab} = \frac{14ab}{ab}$ [Dividing by ab on both sides]

or $\frac{a}{b} + \frac{b}{a} = 14$

Question 19 $\log (m+n)= \log m+ \log n$, m can be expressed as:

Solution $\log(m+n) = \log m+ \log n$

or $\log (m+n)= \log (mn)$

or $m+n=mn$

or $m-mn =-n$

or $m(1-n) =-n$

or $m = \frac{-n}{1-n} = \frac{n}{n-1}$

Question 20 $\log_4 (x^2+x)-\log_4 (x+1)=2$. find x

Solution $\log_4 (x^2+x) -\log_4 (x+1)=2$

or $\log_4 \frac{x^2+x}{x+1} = 2$

or $\frac{x(x+1)}{x+1} =4^2$

or $x=16$

Question 21 $2\log x + 2\log x^2 + 2\log x^3 + \ldots\ldots + 2\log x^n$ will be : -

(a) $\frac{n(n+1)\log x}{2}$ (b) $n(n+1)\log x$

(c) $n^2 \log x$ (d) None of these

Solution Detail Method

$2\log x + 2\log x^2 + 2\log x^3 + ---- + 2\log x^n$

$= 2\log x + 22\log x + 2.3\log x + -- + 2n.\log x$

$= 2\log x.[1+2+3+\_--+n]$

$= 2\log x.\frac{n(n+1)}{2} = n(n+1)\log x$

=(b) is correct

Tricks :-Put n=2 in options directly This should be equal to sum of 1st 2 terms =2 log +2.2logx =6 logx

Which gives option (*b*)

Question 22 Solve: $\frac{(\log_x 10 - 3)}{2} + \frac{11 - \log_x 10}{3} = 20$

(a) 10^{-1} (b) 10^2 (c) 10 (d) 10^3

Solution $\frac{\log_x 10 - 3}{2} + \frac{11 - \log_x 10}{3} = 2$

Tricks: - Go by choices [Do Mentally]

For (a) $x = 10^{-1}$

L.H. S $\frac{\log(10^1)^{10-3}}{2} + \frac{11 - \log_{10^{-1}} 10}{3}$

$= \frac{-1-3}{2} + \frac{11-(-1)}{3}$

$= -2+4 = 2$ (R.H.S) $\therefore$ (a) is correct

Question 23 If n =m! where (' m' is a positive integer >2) then the value of:

$\frac{1}{\log_2 n} + \frac{1}{\log^n{}_3} + \frac{1}{\log_{4n}} + \ldots \ldots + \frac{1}{\log_m{}^n}$

Solution Given n=m !

$= \frac{1}{\log_2 n} + \frac{1}{\log_3 n} + \frac{1}{\log_4 n} + \ldots \ldots + \frac{1}{\log_m{}^n}$

$= \log_n{}^2 + \log_n{}^3 + \log_n{}^4 + \ldots\ldots\ldots + \log_n{}^m$

$= \log_n (2.3.4\ldots\ldots\ldots m)$

$= \log_n (1.2.3.4\ldots\ldots m)$

$\log_{(m!)}(m!) = 1$

Question 24 If $\log 2\ x + \log_4 x = 6$, then the value ofx is

(a) 16 (b) 32 (c) 64 (d) 128

Solution (a) is correct

Tricks: Go by choices for (a) if x=16

L.H. S=$\log_2 16 + \log_4 16 = 4+2 = 6$(RHS)

$\therefore$ (a) is correct

Detail Method : $\log_2 x + \log_4 x = 6$

or $\log_2 x + \log_2 2$) x=6

or $\log_2 x + \frac{1}{2} \log_2 2^x = 6$

or $(1+\frac{1}{2}) \log_2 x = 6$

or $\log_2 x \ \frac{6\times2}{3} = 4$

$x=2^4=16$

Question 25 If log KY=100 and $\log_2 x = 10$, then the value of 'Y' is:

(a) 2^{10} (b) 2^{100} (c) $2^{1,000}$ (d) 2^{10000}

Solution (c) $\log_2 x = 10$ $x=2^{10}$

Now $\log_x y = 100. = y = x^{100}$

.: $y=(2^{10})^{100}=2^{1000}$

$\therefore$ (c) is correct

Question 26 Which of the following is true.

If $\frac{1}{ab}+\frac{1}{bc}+\frac{1}{ca}=\frac{1}{abc}$

(a) log (ab+bc+ca)=abc (b) $\log(\frac{1}{a}+\frac{1}{b}+\frac{1}{c})=abc$

(c) log (abc) =0 (d) log (a+b+c)=0

Solution (d) is correct

$=\frac{1}{ab}+\frac{1}{bc}+\frac{1}{ca}=\frac{1}{abc}$

Multiplying both sides by abc;

$\frac{abc}{ab}+\frac{abc}{bc}+\frac{abc}{ca}+\frac{abc}{abc}$

or c+a+b=1

or a+b+c=1

Taking log on both sides; we get

log (a+b+c)= logl

=0

Question 27 If $(\log_{\sqrt{x}} 2)^2 = \log_x 2$ then x=

(a) 16 (b) 32nn (c) 8 (d) 4

Solution (a) is correct

$(\log_{fk} 2)^2 = \log_k 2$

or $(\log_x^{12} 2)^z = \log_x 2$

or $[\frac{1}{\frac{1}{2}} \log_k 2)^2 = \log_k 2$

or $4(\log K^2)^2 - \log_k 2 = 0$

or logx $2[4\log\lambda 2-1]=0$

If $\log K^2=0$ (Invalid)

$4\log_k 2-1=0$

or $4\log_k 2=1$

or $\log\lambda 2 = \frac{1}{4}$

or $x^{l/4}=2=x=2^4=16$

Tricks :- Go by choices

For (a) LHS 1 $(\log_{\sqrt{16}} 2)^2 = (\log_4 2)^2$

$= (\frac{1}{2}\log_2 2)^2 = \frac{1}{4}$

RHS $=\log_{16} 2 = \log_2 4^2 = \frac{1}{4}\log_2 2 = \frac{1}{4}$

Question 28 Find Value of $[\log_y x.\log_z y.\log_x z]^3$=

(a) 0 (b) -1 (c) (d) 3

Solution (c) is correct

$[\log_y x.\log_z y.\log_x z]^3$

$= [\log_x x]^3 = [1]^3 = 1$

Question 29 Find the value of $\text{Log}_4 9\text{Log}_3 2$=

(a) 3 (b) 9 (c) 2 (d) 1

Solution (d) is correct

Log4 $91xg_3 2$=

$\log_{(2l)} (3^2).\log_3 2$

$=\frac{2}{2}\log_2 3.\log_3 2$

$=1\times1=1$

Question 30 If $X=\log_{24} 12 : y=\log_{36} 24 ; z=\log_{48} 36$ then xyz +1=?

(a) 2xy (b) 2zx (c) 2yz (d) 2

Solution (c) is correct

$xyz+1=\log_{24} 12.\log_{36} 24.\log_{48} 36+1$

$= \log_{48} 12+\log_{-48} 48$

$= \log_{48} (12\times48)=\log_{48} (12\times2)^2$

$= 2\log_{48} 24=2\log_{36} 24.\log_{48}36$

$= 2yz$

Question 31 If $x^2+y^2=7_{xy}$ ₜhell $\log \frac{1}{3}(x+y)=$

(a) logx + logy (b) $\frac{1}{2}$(log x+ log y)

(c) $\frac{1}{3}$(log x+ log y) (d) $\frac{1}{3}$ (logx.log y)

Solution $\log \frac{1}{3}(x+y) = \frac{1}{2}\ 2 \log \{\frac{1}{3}(x + y)\}$

$= \frac{1}{2} \log \{\frac{1}{3}(x + y)\}^2 = \frac{1}{2} \log (\frac{x^2+y^2+2xy}{9})$

$= \frac{1}{2} 10 \log (\frac{7xy+2xy}{9}) \frac{1}{2} \log (xy)$

(b) is correct

Question 32 If log x=a-b;logy =a+b then log ( )

(a) 1-a-3b (b) a-1+3b (c) a+3b-1 (d) 1-b+3a

Solution (a) is correct

: - log x =a+b; logy =a-b.

$\log (\frac{10x}{y^2}) = \log_{10}10+ \log x-\log y^2$

=1+a+b-2 log y=1+a+b-2 (a-b)

=1+a+b-2a+2b=1-a+3b

Question 33 If x=1+$\log_p$ qr, y=1+$\log_q$ rp and z=] +$\log_r$ pq men me value or $\frac{1}{x}+\frac{1}{y}+\frac{1}{z}=$

(a) 0 (b) 1 (c) -1 (d) 3

Solution $\frac{1}{x}+\frac{1}{y}+\frac{1}{z}$

$=\frac{1}{1+\log_p qr}+\frac{1}{1+\log_q rp}+\frac{1}{1+\log_r pq}$

$=\frac{1}{\log_p p+\log_p qr}+\frac{1}{\log_q q+\log_q rp}$

$+\frac{1}{\log_r r+\log_r pq}$

$=\frac{1}{\log_p pqr}+\frac{1}{\log_q pqr}+\frac{1}{\log_r pqr}$

$=\log_{pqr} p+\log_{pq,}q+\log_{p,,r}r$

$=\log_{pqr}Pqr=1$

(b) is correct

Question 34 If log x=m+n; logy =m-n theti log $\left(\frac{10x}{y^2}\right)$ =

(a) 0 (b) 1 (c) -1 (d) 3

Solution If log x=m+n; logy =m-n Then log $\left(\frac{10x}{y^2}\right)$

= log 10+ log x-log y^2

= 1+ logx -2 log y

= 1+(m+n)-2(m-n)

= 1+m+n-2m+2n

= 1-m+3n

Question 35 log35. xlog5 4×$\log_2$ 3:

(a) 2 (b) 5 (c) -2 (d) None ofthese

Solution $\log_3 5.\log_5 4.\log_2 3$

$=\log_3 4.\log_2 3=\log_2 4=2$

(a) is correct

Question 36 The integral part of a logarithm is called , and the decimal part of a logarithm is called

(a) Mantissa, Characteristic (b) Characteristic, Mantissa

(c) Whole, Decimal (d) None of these

Solution (b) is correct.

Question 37 The value of $\frac{1}{\log_3 60} + \frac{1}{\log_4 60} + \frac{1}{\log_5 60} = -$

(a) 0 (b) 1 (c) 5 (d) 60

Solution (b) is correct

$\log\omega\ 3+\log_{60}4+\log_{60}5$

$=\log_{60}(3\times4\times5)+\log_{60}60=1$

Question 38 If $\log 4k^2+x+\log_4(x+1)=2$ then the value of x is

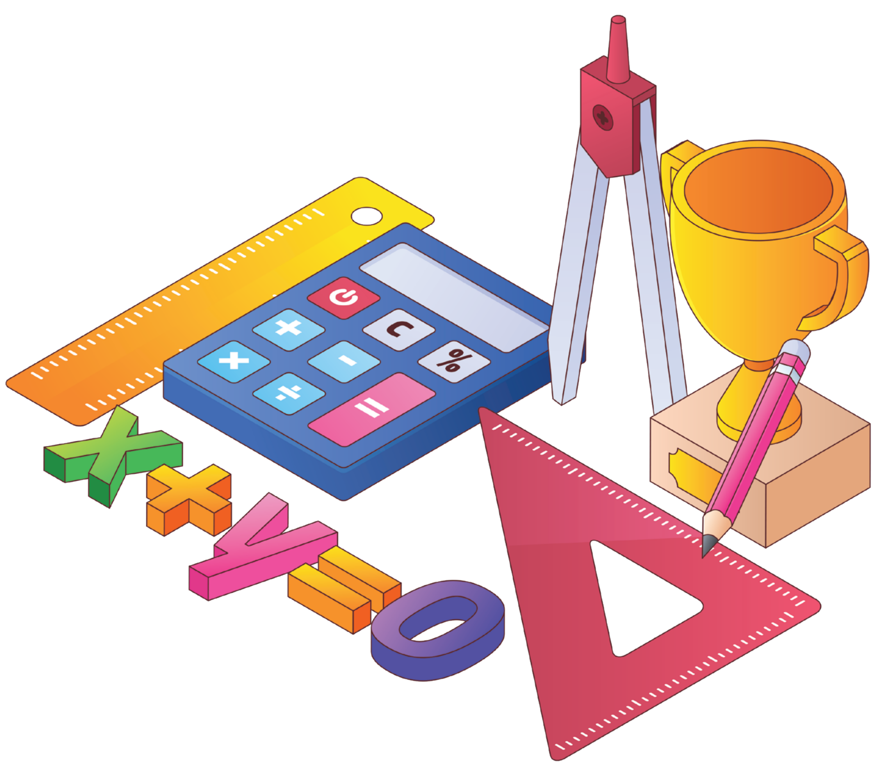
(a) 2 (b) 3 (c) 16 (d) 8

Solution (c) is correct

$\log_4 \frac{(x^2+x)}{(x+1)} = 2$

or $\log_4\{\frac{x(x+1)}{(x+1)}\}=2$

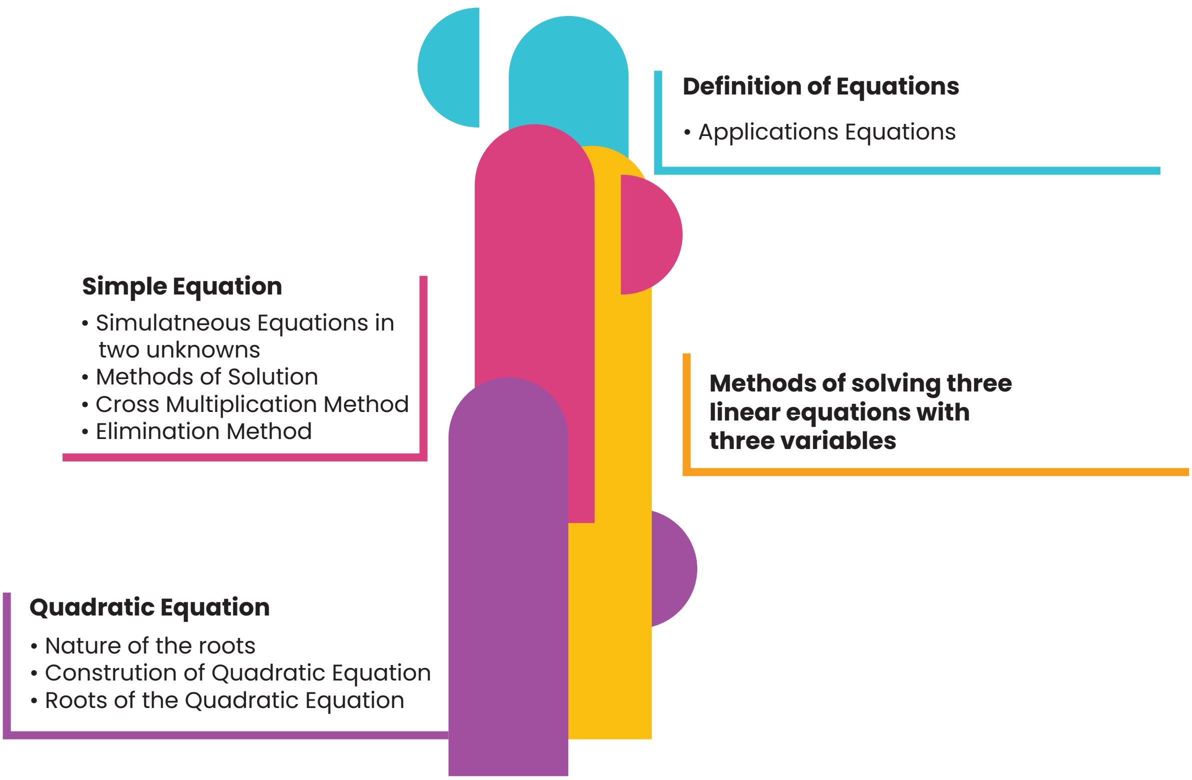
or $\log_4 x=2 \Rightarrow x=4^2=16$



Chapter – 2:

Equations

In this chapter, students are able to understand the concepts of equations such as linear, simultaneous, quadratic, and cubic. They will also know how to solve the different equations using different solution methods.



Introduction

- Equation is defined to be a mathematical statement of equality. If the equality is true for certain value of the variable involved, the equation is often called a conditional equation and equality sign '=' is used; while if the equality is true for all values of the variable involved, the equation is called an identity.

Determination of value of the variable which satisfies an equation is called solution of the equation or root of the equation.

For example:

$\frac{x+2}{3} + \frac{x+3}{2} = 3$ holds true only for x = 1.

So it is a conditional. On the other hand,

$$\frac{x+2}{3} + \frac{x+3}{2} = \frac{5x+13}{6}$$

is an identity since it holds for all values of the variable x.

Equations

An equation in which highest power of the variable is 1 is called a Linear (or a simple) equation. This is also called the equation of degree 1.

Two or more linear equations involving two or more variables are called Simultaneous Linear Equations. An equation of degree 2 (Highest Power of the variable is 2) is called Quadratic equation and the equation of degree 3 is called Cubic Equation

For Example: $8x+17(x-3) = 4(4x-9) + 12$ is a Linear equation.
$3x^2 + 5x + 6 = 0$ is a Quadratic equation.
$4x^3 + 3x^2 + x-7 = 1$ is a Cubic equation.
$x + 2y = 1, 2x + 3y = 2$ are jointly called Simultaneous equations.

Simple equation in one unknown x is in the form $ax + b = 0$.
Where a, b are known constants and $a \neq 0$
Note: A simple equation has only one root.

Example 1: $\frac{4x}{3} - 1 = \frac{14}{15}x + \frac{19}{5}$

Solution: By transposing the variables in one side and the constants in other side we have

$$\frac{4x}{3} - \frac{14x}{15} = \frac{19}{5} + 1 \text{ or } \frac{(20-14)x}{15} = \frac{19+5}{5} \text{ or } \frac{6x}{15} = \frac{24}{5}$$

$$x = \frac{24 \times 15}{5 \times 6} = 12$$

Example 2: The denominator of a fraction exceeds the numerator by 5 and if 3 be added to both the fraction becomes $\frac{3}{4}$. Find the fraction.

Solution:

Let x be the numerator and the fraction be $\frac{x}{x+5}$

By the question $\frac{x+3}{x+5+3} = \frac{3}{4}$ or $4x + 12 = 3x + 24$ or $x = 12$

The required fraction is $\frac{12}{17}$

Example 3: If thrice of A's age 6 years ago be subtracted from twice his present age, the result would be equal to his present age. Find A's present age.

Solution:
Let x years be A's present age. By the question
$2x-3(x-6) = x$
or $2x-3x + 18 = x$
or $-x + 18 = x$
or $2x = 18$
or $x=9$
$\therefore$ A's present age is 9 years

Example 4: A number consists of two digits the digit in the ten's place is twice the digit in the unit's place. If 18 be subtracted from the number the digits are reversed. Find the number.

Solution: Let x be the digit in the unit's place. So the digit in the ten's place is 2x. Thus the number becomes 10(2x) + x.

By the question 20x + x– 18 = 10x + 2x or 21x – 18 = 12x or 9x = 18 or x = 2

So the required number is 10 (2 × 2) + 2 = 42.

Example 5: For a certain commodity the demand equation giving demand 'd' in kg, for a price 'p' in rupees per kg. is d = 100 (10 – p). The supply equation giving the supply s in kg. for a price p in rupees per kg. is s = 75(p – 3). The market price is such at which demand equals supply. Find the market price and quantity that will be bought and sold.

Solution: Given d = 100(10 – p) and s = 75(p – 3).

Since the market price is such that demand (d) = supply(s) we have

100 (10 – p) = 75 (p – 3) or 1000 – 100p = 75p – 225

or – 175p = – 1225.

$$\therefore \quad p = \frac{-1225}{-175} = 7$$

So market price of the commodity is ₹ 7 per kg.

∴ the required quantity bought = 100 (10 – 7) = 300 kg.

and the quantity sold = 75 (7 – 3) = 300 kg.

Simultaneous linear equation in two unknowns

The general form of a linear equations in two unknowns x and y is ax + by + c = 0 where a, b are non-zero coefficients and c is a constant. Two such equations $a_1x + b_1y + c_1 = 0$ and $a_2 x + b_2 y + c_2 = 0$ form a pair of simultaneous equations in x and y. A value for each unknown which satisfies simultaneously both the equations will give the roots of the equations.

Method of Solution

1. Elimination Method: In this method two given linear equations are reduced to a linear equation in one unknown by eliminating one of the unknowns and then solving for the other unknown.

Example 1: Solve: 2x + 5y = 9 and 3x – y = 5.

Solution: 2x + 5y = 9 (i)

3x – y = 5 (ii)

By making (i) x 1, 2x + 5y = 9

and by making (ii) x 5, 15x – 5y = 25

Adding 17x = 34 or x = 2. Substituting this values of x in (i) i.e.

5y = 9 – 2x we find; 5y = 9 – 4 = 5 ∴ y = 1 ∴ x = 2, y = 1.

2. Cross Multiplication Method: Let two equations be:

$a_1x + b_1y + c_1 = 0$, $a_2x + b_2y + c_2 = 0$

We write the coefficients of x, y and constant terms and two more columns by repeating the coefficients of x and y as follows:

$$\begin{array}{ccccccc} 1 & \underline{\quad x \quad} & 2 & \underline{\quad y \quad} & 3 & \underline{\quad 1 \quad} & 4 \\ b_1 & \times & c_1 & \times & a_1 & \times & b_1 \\ b_2 & & c_2 & & a_2 & & b_2 \end{array}$$

and the result is given by:

$$\frac{x}{(b_1c_2 - b_2c_1)} = \frac{y}{(c_1a_2 - c_2a_1)} = \frac{1}{(a_1b_2 - a_2b_1)}$$

so the solution is : $x = \frac{b_1c_2 - b_2c_1}{a_1b_2 - a_2b_1}$ $\quad y = \frac{c_1a_2 - c_2a_1}{a_1b_2 - a_2b_1}$

Example 2: Solve 3x + 2y + 17 = 0, 5x – 6y – 9 = 0

Solution: 3x + 2y + 17 = 0 (i)
5x – 6y – 9 = 0 (ii)

Method of elimination: By (i) x 3 we get 9x + 6y + 51 = 0 (iii)
Adding (ii) & (iii) we get 14x + 42 = 0

or $x = -\frac{42}{14} = -3$

Putting x = –3 in (i) we get 3(–3) + 2y + 17= 0

or, 2y + 8= 0 or, $y = -\frac{8}{2} = -4$

So x = –3 and y = –4

Method of cross-multiplication:
3x + 2y + 17 = 0
5x – 6y – 9 = 0

$$\frac{x}{2(-9)-17(-6)} = \frac{y}{17x(5)-3(-9)} = \frac{1}{3(-6)-5x(2)}$$

or, $\frac{x}{84} = \frac{y}{112} = \frac{1}{-28}$

or, $\frac{x}{3} = \frac{y}{4} = \frac{1}{-1}$

or x = –3, y = –4

Methos of Solving Simultaneous Linear Equation with three variables

Example 1: Solve for x, y and z:

$2x - y + z = 3, x + 3y - 2z = 11, 3x - 2y + 4z = 1$

Solution:

(a) Method of elimination

$2x - y + z = 3$...(i)

$x + 3y - 2z = 11$...(ii)

$3x - 2y + 4z = 1$...(iii)

By (i) × 2 we get

$4x - 2y + 2z = 6$...(iv)

By (ii) + (iv), $5x + y = 17$...(v)

[the variable z is thus eliminated]

By (ii) × 2, $2x + 6y - 4z = 22$..(vi)

By (iii) + (vi), $5x + 4y = 23$..(vii)

By (v) – (vii), $-3y = -6$ or $y = 2$

Putting $y = 2$ in (v) $5x + 2 = 17$, or $5x = 15$ or, $x = 3$

Putting $x = 3$ and $y = 2$ in (i)

$2 \times 3 - 2 + z = 3$

or $6 - 2 + z = 3$

or $4 + z = 3$ or $z = -1$

So $x = 3, y = 2, z = -1$ is the required solution.

(Any two of 3 equations can be chosen for elimination of one of the variables)

(b) Method of cross multiplication

We write the equations as follows:

$2x - y + (z - 3) = 0$

$x + 3y + (-2z - 11) = 0$

By cross multiplication

$$\frac{x}{-1(-2z-11)-3(z-3)} = \frac{y}{(z-3)-2(-2z-11)} = \frac{1}{2'3-1(-1)}$$

$$\frac{x}{20-z} = \frac{y}{5z+19} = \frac{1}{7}$$

$$x=\frac{20-z}{7}, \quad y=\frac{5z+19}{7}$$

Substituting above values for x and y in equation (iii) i.e. $3x - 2y + yz = 1$, we have

$$3\left(\frac{20-z}{7}\right) -2\left(\frac{5z+19}{7}\right) + 4z=1$$

or $60 - 3z - 10z - 38 + 28z = 7$

or $15z = 7 - 22$ or $15z = -15$ or $z = -1$

$$x=\frac{20-(-1)}{7} = \frac{21}{7} = 3, \qquad y=\frac{5-(-1)+19}{7} = \frac{14}{7} = 2$$

Thus $x = 3, y = 2, z = -1$

Example 2: Solve for x, y and z :

$$\frac{1}{x}+\frac{1}{y}+\frac{1}{z}=5,\quad \frac{2}{x}-\frac{3}{y}-\frac{4}{z}=11,\quad \frac{3}{x}+\frac{2}{y}-\frac{1}{z}=-6$$

Solution: We put $u=\frac{1}{x}$; $v=\frac{1}{y}$; $w=\frac{1}{z}$ and get

$u + v + w = 5$..........(i)

$2u - 3v - 4w = -11$..........(ii)

$3u + 2v - w = -6$..........(iii)

By (i) + (iii) $4u + 3v = -1$..........(iv)

By (iii) x 4 $12u + 8v - 4w = -24$..........(v)

By (ii) − (v) $-10u - 11v = 13$ or $10u + 11v = -13$..........(vi)

By (iv) × 11 $44x + 33v = -11$..........(vii)

By (vi) × 3 $30u + 33v = -39$..........(viii)

By (vii) − (viii) $14u = 28$ or $u = 2$

Putting u = 2 in (iv) $4 \times 2 + 3v = -1$

or $8 + 3v = -1$

or $3v = -9$ or $v = -3$

Putting $u = 2, v = -3$ in (i) or $2-3 + w = 5$

or $-1 + w = 5$ or $w = 5 + 1$ or $w = 6$

Thus $x=\frac{1}{u}=\frac{1}{2}$, $y=\frac{1}{y}=\frac{1}{3}$, $z=\frac{1}{w}=\frac{1}{6}$ is the solution.

Example 3: Solve for x, y and z:

$$\frac{xy}{x+y}=70,\quad \frac{xy}{x+z}=84,\quad \frac{yz}{y+z}=140$$

Solution: We can write as

$\frac{x+y}{xy}=\frac{1}{70}$ or $\frac{1}{x}+\frac{1}{y}=\frac{1}{70}$..........(i)

$\frac{x+z}{xz}=\frac{1}{84}$ or $\frac{1}{z}+\frac{1}{x}=\frac{1}{84}$..........(ii)

$\frac{y+z}{yz}=\frac{1}{140}$ or $\frac{1}{y}+\frac{1}{z}=\frac{1}{140}$..........(iii)

By (i) + (ii) + (iii), we get

$$2\left(\frac{1}{x}+\frac{1}{y}+\frac{1}{z}\right)=\frac{1}{70}+\frac{1}{84}+\frac{1}{140}=\frac{14}{420}$$

or $\frac{1}{x}+\frac{1}{y}+\frac{1}{z}=\frac{7}{420}=\frac{1}{60}$..........(iv)

By (iv)−(iii) $\frac{1}{x} = \frac{1}{60} - \frac{1}{140} = \frac{4}{420}$ or $x = 105$

By (iv)−(ii) $\frac{1}{y} = \frac{1}{60} - \frac{1}{84} = \frac{2}{420}$ or $y = 210$

By (iv)−(i) $\frac{1}{z} = \frac{1}{60} - \frac{1}{70}$ or $z = 420$

Required solution is $x = 105$, $y = 210$, $z = 420$

1. If the numerator of a fraction is increased by 2 and the denominator by 1 it becomes 1. Again if the numerator is decreased by 4 and the denominator by 2 it becomes 1/2. Find the fraction.

Solution: Let x/y be the required fraction.

By the question $\frac{x+2y}{y+1} = 1, \quad \frac{x-4}{y-2} = \frac{1}{2}$

Thus $x + 2 = y + 1$ or $x - y = -1$ .... (i)
and $2x - 8 = y - 2$ or $2x - y = 6$ .... (ii)
By (i) − (ii) $-x = -7$ or $x = 7$
from (i) $7 - y = -1$ or $y = 8$
So the required fraction is 7/8

2. The age of a man is three times the sum of the ages of his two sons and 5 years hence his age will be double the sum of their ages. Find the present age of the man?

Solution: Let x years be the present age of the man and sum of

the present ages of the two sons be y years.

By the condition $x = 3y$ ....(i)

and $x + 5 = 2(y + 5 + 5)$..............................(ii)

From (i) & (ii) $3y + 5 = 2(y + 10)$ or $3y + 5 = 2y + 20$

or $3y - 2y = 20 - 5$ or $y = 15$

$\therefore x = 3 \times y = 3 \times 15 = 45$

Hence the present age of the man is 45 years

3. A number consist of three digit of which the middle one is zero and the sum of the other digits is 9. The number formed by interchanging the first and third digits is more than the original number by 297 find the number.

Solution: Let the number be $100x + y$. we have $x + y = 9$.........(i)

Also $100y + x = 100x + y + 297$...(ii)

From (ii) $99(x - y) = -297$

or $x - y = -3$ ....(iii)

Adding (i) and (iii) $2x = 6$ or $x = 3$ ⊠ from (i) $y = 6$

$\therefore$ Hence the number is 306.

Quadratic Equation: An equation of the form $ax^2 + bx + c = 0$ where x is a variable and a, b, c are constants with $a \neq 0$ is called a quadratic equation or equation of the second degree.

When b=0 the equation is called a pure quadratic equation, when $b \neq 0$ the equation is called an affected quadratic.

Examples: i) $2x^2 + 3x + 5 = 0$

ii) $x^2 - x = 0$

iii) $5x^2 - 6x - 3 = 0$

The value of the variable say x is called the root of the equation. A quadratic equation has got two roots.

Roots of a quadratic equation:

$ax^2 + bx + c = 0$ $(a \neq 0)$

$$\text{or } x = \frac{-b \pm \sqrt{b^2 - 4ac}}{2a}$$

Sum and Product of the Roots:

Let one root be α and the other root be β

$$\alpha + \beta = \frac{-b}{a} = \frac{-\text{coefficient of } x}{\text{coeffient of } x^2}$$

$$\text{product of the roots} = \frac{c}{a} = \frac{\text{constant term}}{\text{coefficient of } x^2}$$

How to construct a Quadratic Equation

For the equation $ax^2 + bx + c = 0$ we have

$$\text{Or } x^2 + = \frac{-b}{a} x + \frac{c}{a} 0$$

$$\text{Or } x^2 \left(- \frac{-b}{a}\right) x + \frac{c}{a} = 0$$

Or x^2 – (Sum of the roots) x + Product of the roots = 0

Nature of the Roots $x = \dfrac{-b \pm \sqrt{b^2 - 4ac}}{2a}$

i) If $b^2 - 4ac = 0$ the roots are real and equal;

ii) If $b^2 - 4ac > 0$ then the roots are real and unequal (or distinct);

iii) If $b^2 - 4ac < 0$ then the roots are imaginary;

iv) If $b^2 - 4ac$ is a perfect square ($\neq 0$) the roots are real, rational and unequal (distinct);

v) If $b^2 - 4ac > 0$ but not a perfect square the roots are real, irrational and unequal.

Since $b^2 - 4ac$ discriminates the roots b2 – 4ac is called the discriminant in the equation $ax^2 + bx + c = 0$ as it actually discriminates between the roots.

Note: (a) Irrational roots occur in conjugate pairs that is if $(m+\sqrt{n})$ is a root then $(m-\sqrt{n})$ is the other root of the same equation.

If one root is reciprocal to the other root then their product is 1 and so

$$\frac{c}{a} = \text{ i.e. } c = a$$

If one root is equal to other root but opposite in sign then.

$$\text{their sum} = 0 \text{ and so } \frac{b}{a} = 0. \text{ i.e. } b = 0$$

1. Solve $x^2 - 5x + 6 = 0$

Solution: 1st method : $x^2 - 5x + 6 = 0$

or $x^2 - 2x - 3x + 6 = 0$ or $x(x-2) - 3(x-2) = 0$

or $(x-2)(x-3) = 0$

or x = 2 or 3

2nd method (By formula) $x^2 - 5x + 6 = 0$

Here a = 1, b = −5 , c = 6 (comparing the equation with $ax^2 + bx + c = 0$)

$$x = \frac{-b \pm \sqrt{b^2-4ac}}{2a} = \frac{-(5) \pm \sqrt{25-24}}{2}$$

$$\frac{5 \pm 1}{2} = \frac{6}{2} \text{ and } \frac{4}{2}; \quad \therefore x = 3 \text{ and } 2$$

2. Examine the nature of the roots of the following equations.

i) $x^2 - 8x + 16 = 0$ ii) $3x^2 - 8x + 4 = 0$

iii) $5x^2 - 4x + 2 = 0$ iv) $2x^2 - 6x - 3 = 0$

Solution:

(i) a = 1, b = −8, c = 16

$b^2 - 4ac = (-8)^2 - 4 \times 1 \times 16 = 64 - 64 = 0$

The roots are real and equal.

(ii) $3x^2 - 8x + 4 = 0$

a = 3, b = −8, c = 4

$b^2 - 4ac = (-8)^2 - 4 \times 3 \times 4 = 64 - 48 = 16 > 0$ and a perfect square

The roots are real, rational and unequal

(iii) $5x^2 - 4x + 2 = 0$

$b^2 - 4ac = (-4)^2 - 4 \times 5 \times 2 = 16 - 40 = -24 < 0$

The roots are imaginary and unequal

(iv) $2x^2 - 6x - 3 = 0$

$b^2 - 4ac = (-6)^2 - 4 \times 2(-3)$

$= 36 + 24 = 60 > 0$

The roots are real and unequal. Since $b^2 - 4ac$ is not a perfect square the roots are real irrational and unequal.

3. If α and β be the roots of $x^2 + 7x + 12 = 0$ find the equation whose roots are $(\alpha-\beta)^2$.

Solution

Now sum of the roots of the required equation

$= (\alpha + \beta)^2 + (\alpha - \beta)^2 = (-7)^2 + (\alpha + \beta)^2 - 4\alpha\beta$

$= 49 + (-7)^2 - 4x12$

$= 49 + 49 - 48 = 50$

Product of the roots of the required equation = $(\alpha+ß)^2$

$((\alpha - \beta)^2 = 49 (49-48) = 49$

Hence the required equation is x^2 – (sum of the roots)

x + product of the roots = 0

or $x^2 - 50x + 49 = 0$

4. If α, β be the roots of $2x^2 - 4x - 1 = 0$ find the $\frac{\alpha^2}{\beta} + \frac{\beta^2}{\alpha}$ value of

Solution: $\alpha+\beta = \frac{-(-4)}{2} = 2$, $\alpha\beta = \frac{-1}{2}$

$$\frac{\alpha^2}{\beta} + \frac{\beta^2}{\alpha} = \frac{\alpha^3+\beta^3}{\alpha\beta} = \frac{(\alpha+\beta)^3 - 3\,\alpha\beta\,(\alpha+\beta)}{\alpha\beta}$$

$$\frac{2^3-3\left(-\frac{1}{1}\right).2}{\left(-\frac{1}{1}\right)}$$

5. Solve x : $4^x - 3.2x+2 + 25 = 0$

Solution: $4^x - 3.2^x+2 + 25 = 0$ or $(2^x)^2 - 3.2^x. 22 + 32 = 0$

or $(2^x)2 - 12.\ 2^x + 32 = 0$

or $y^2 - 12y + 32 = 0$ (taking $y = 2^x$) or $y^2 - 8y - 4y + 32 = 0$

or $y(y - 8) - 4(y - 8) = 0$ $\therefore (y - 8)(y - 4) = 0$

either $y - 8 = 0$ or $y - 4 = 0$ $\therefore y = 8$ or $y = 4$.

$\Rightarrow 2^x = 8 = 2^3$ or $2^x = 4 = 2^2$ Therefore $x = 3$ or $x = 2$.

6. Solve $\left(x - \frac{1}{x}\right)^2 + 2\left(x + \frac{1}{x}\right) = 7\frac{1}{4}$

$$\left(x - \frac{1}{x}\right)^2 + 2\left(x + \frac{1}{x}\right) = \frac{29}{4}$$

or $\left(x - \frac{1}{x}\right)^2 -4+2\left(x + \frac{1}{x}\right) = \frac{29}{4}$

[as $(a-b)^2 = (a+b)^2 - 4ab$)]

or $p^2+2p - \frac{45}{4} = 0$ Taking $p = x + \frac{1}{x}$

or $4p^2 + 8p - 45 = 0$

or $4p^2 + 18p - 10p - 45 = 0$

or $2p(2p + 9) - 5(2p + 9) = 0$

or $(2p - 5)(2p + 9) = 0$.

$\therefore$ Either $2p + 9 = 0$ or $2p - 5 = 0$ $\Rightarrow p = -\frac{9}{2}$ or $p - \frac{5}{2}$

$\therefore$ Either $x + \frac{1}{x} = -\frac{9}{2}$ or $x + \frac{1}{x} = \frac{5}{2}$

i.e. Either $2x^2 + 9x +2 = 0$ or $2x^2 - 5x + 2 = 0$

i.e. Either $x = \frac{-9\pm\sqrt{81-16}}{4}$ or, $x = \frac{5\pm\sqrt{25-16}}{4}$

i.e. Either $x = \frac{-9\pm\sqrt{65}}{4}$ or $x = 2$ or $\frac{1}{2}$

7. Solve $2^{x-2} + 2^{3-x} = 3$

Solution:

$2^{x-2} + 2^{3-x} = 3$

or $2^x . 2^{-2} + 2^3 . 2^{-x} = 3$

or $\frac{2^x}{2^2} + \frac{2^3}{2^x} = 3$

or $\frac{t}{4} + \frac{8}{t} = 3$ when $t = 2^x$

or $t^2 + 32 = 12t$

or $t^2 - 12t + 32 = 0$

or $t^2 - 8t - 4t + 32 = 0$ or $t(t-8) - 4(t-8) = 0$

or $(t-4)(t-8) = 0$

∴ $t = 4, 8$

For t = 4, $2^x = 4 = 22$ i.e. x = 2

For t = 8, $2^x = 8 = 22$ i.e. x = 3

8. If one root of the equation is $2-\sqrt{3}$ form the equation given that the roots are irrational

Solution: Other root is $2+\sqrt{3}$

∴ sum of two roots $=2-\sqrt{3}+2+\sqrt{3}=4$

Product of roots $= (2-\sqrt{3})(2+\sqrt{3})=4-3=1$

∴ Required equation is : $x^2 -$ (sum of roots)x + (product of roots) = 0 or $x^2 - 4x + 1 = 0$.

9. If the roots of the equation $p(q-r)x^2 + q(r-p)x + r(p-q) = 0$

are equal show that $\frac{2}{q} = \frac{1}{p} + \frac{1}{r}$

Solution: Since the roots of the given equation are equal the discriminant must be zero

ie. $q^2(r-p)^2 - 4.\ p(q-r)\ r(p-q) = 0$

or $q^2 r^2 + q^2 p^2 - 2q^2 rp - 4pr(pq - pr - q^2 + qr) = 0$

or $p^2q^2 + q^2r^2 + 4p^2r^2 + 2q^2pr - 4p^2qr - 4pqr^2 = 0$

or $(pq + qr - 2rp)^2 = 0$

∴ $pq + qr = 2pr$

or $\frac{pq+qr}{2pr} =1$ or $\frac{q}{2} - \frac{(p+r)}{pr} =1$ or $\frac{1}{r} + \frac{1}{p} = \frac{2}{q}$

10. If $\alpha\ \beta$ are the two roots of the equation $x^2 - px + q = 0$ form the equation whose roots are $\frac{\alpha}{\beta}$ and $\frac{\beta}{\alpha}$

Solution: As α, β are the roots of the equation $x^2 - px + q = 0$

$\alpha + \beta = -(-p) = p$ and $\alpha \times \beta = q$.

Now $\frac{\alpha}{\beta} + \frac{\beta}{\alpha} = \frac{\alpha^2+\beta^2}{\alpha\beta} = \frac{(\alpha+\beta)^2 - 2\alpha\beta}{\alpha\beta} = \frac{p^2 - 2q}{q}$ and $\frac{\alpha}{\beta} \times \frac{\beta}{\alpha} = 1$

∴ Required equation is $x^2 \left(\frac{p^2 - 2q}{q}\right) x + 1 = 0$

or $q x^2 - (p^2 - 2q) x + q = 0$

11. Difference between a number and its positive square root is 12; find the numbers?

Solution: Let the number be x.

Then $x - \sqrt{x} = 12$ ....(1)

$(\sqrt{x})^2 - \sqrt{x} - 12 = 0.$ Taking $y = \sqrt{x}$, $y^2 - y - 12 = 0$

Or $(y - 4)(y + 3) = 0$

$\therefore$ Either $y = 4$ or $y = -3$ i.e. Either $\sqrt{x} = 4$ or $\sqrt{x} = -3$

If $\sqrt{x} = -3$ $x = 9$ if does not satisfy equation (i) so $\sqrt{x} = 4$ or $x = 16$.

12. A piece of iron rod costs ₹ 60. If the rod was 2 metre shorter and each metre costs ₹ 1.00 more, the cost would remain unchanged. What is the length of the rod?

Solution: Let the length of the rod be x metres. The rate per meter is ₹ 60/X .

New Length = (x – 2); as the cost remain the same the new rate per meter is

$$\frac{60}{x-2}$$

As given $\frac{60}{x-2} = \frac{60}{x} + 1$

or $\frac{60}{x-2} - \frac{60}{x} = 1$

or $\frac{120}{x(x-2)} = 1$

or $x^2 - 2x = 120$

or $x^2 - 2x - 120 = 0$ or $(x - 12)(x + 10) = 0$.

Either $x = 12$ or $x = -10$ (not possible)

$\therefore$ Hence the required length = 12m.

13. Divide 25 into two parts so that sum of their reciprocals is 1/6.

Solution: let the parts be x and 25 – x

By the question $\frac{1}{x} + \frac{1}{25 - x} = \frac{1}{6}$

or $150 = 25x - x^2$

or $x^2 - 25x + 150 = 0$

or $x^2 - 15x - 10x + 150 = 0$ or $x(x-15) - 10(x-15) = 0$

or $(x-15)(x-10) = 0$

or $x = 10, 15$

So the parts of 25 are 10 and 15.

SOLUTION OF CUBIC EQUATION

On trial basis putting if some value of x stratifies the equation then we get a factor. This is a trial and error method. With this factor to factorise the LHS and then other get values of x.

1. Solve $x^3 - 7x + 6 = 0$

Putting x = 1 L.H.S is Zero. So (x–1) is a factor of $x^3 - 7x + 6$ We write $x^3 - 7x + 6 = 0$ in such a way that (x–1) becomes its factor. This can be achieved by writing the equation in the following form.

or $x^3-x^2+x^2-x-6x+6 = 0$

or $x^2(x-1) + x(x-1) - 6(x-1) = 0$ or $(x-1)(x^2+x-6) = 0$

or $(x-1)(x^2+3x-2x-6) = 0$

or $(x-1)\{ x(x+3) - 2(x+3) \} = 0$

or $(x-1)(x-2)(x+3) = 0$

∴ or x = 1, 2, −3

2. Solve for real x: $x^3 + x + 2 = 0$

Solution: By trial we find that x = −1 makes the LHS zero. So (x + 1) is a factor of $x^3 + x + 2$

We write $x^3 + x + 2 = 0$ as $x^3 + x^2 - x^2 - x + 2x + 2 = 0$ or $x^2 (x + 1) - x(x + 1) + 2(x + 1) = 0$

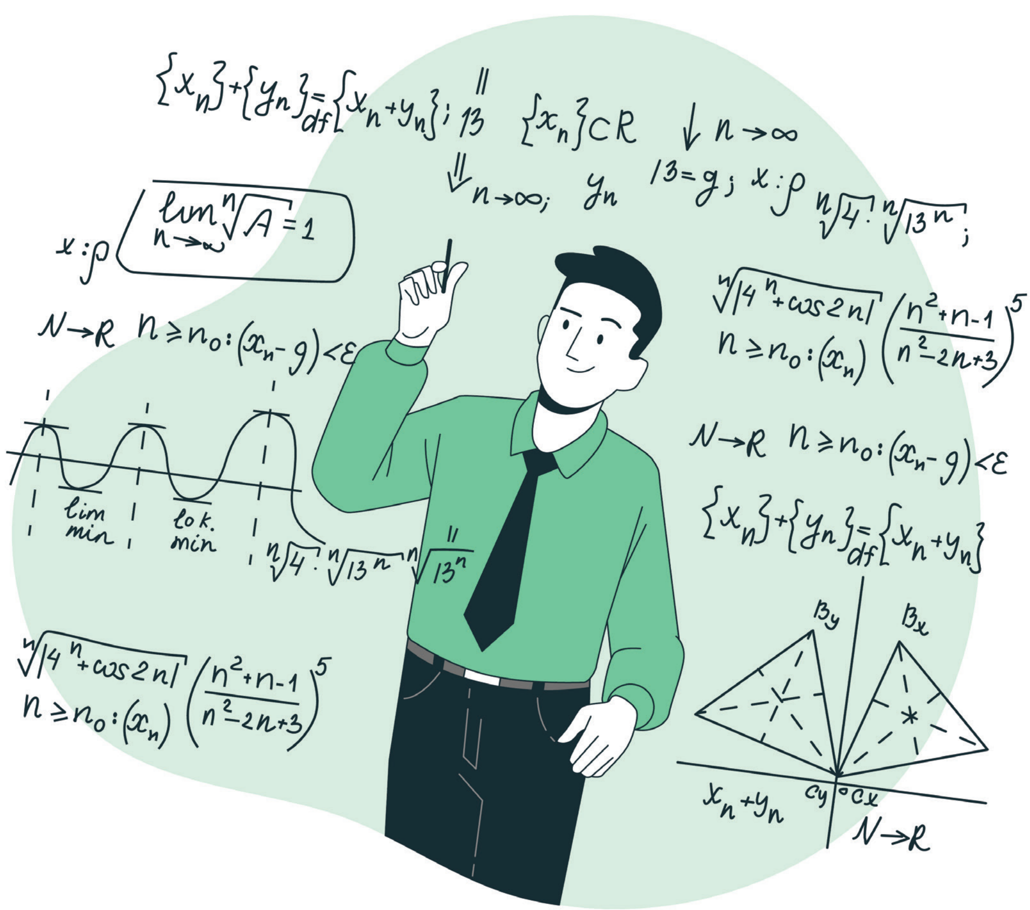
or $(x + 1) (x^2 - x + 2) = 0$.

Either x + 1 = 0; x = -1

or $x^2 - x + 2 = 0$ i.e. x = −1

i.e. $x = \dfrac{1 \pm \sqrt{1-8}}{2} = \dfrac{1 \pm \sqrt{1-7}}{2}$

As $x = \dfrac{1 \pm \sqrt{-7}}{2}$ is not real, x = −1 is the required solution.



Chapter – 3:

Linear Inequalities

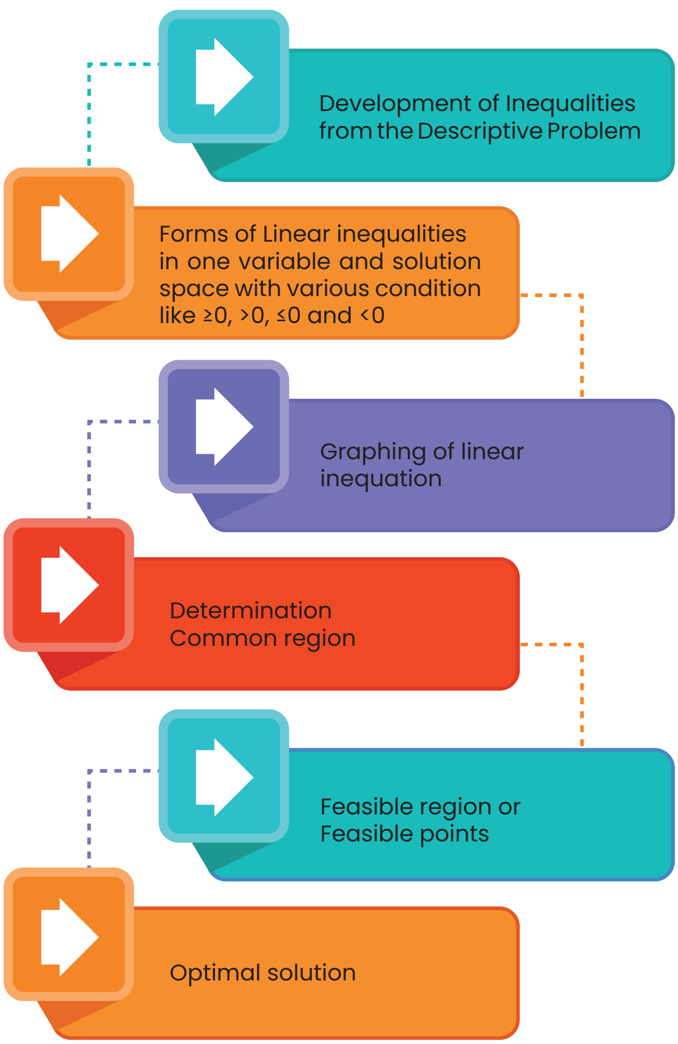
Learning Objectives

One of the widely used decision making problems, nowadays, is to decide on the optimal mix of scarce resources in meeting the desired goal. In simplest form, it uses several linear inequations in two variables derived from the description of the problem.

The objective in this section is to make a foundation of the working methodology for the above by way of introduction of the idea of :

- development of inequations from the descriptive problem;
- graphing of linear inequations; and
- determination of common region satisfying the inequations.

Chapter Overview



3.1 Inequalites

Inequalities are statements where two quantities are unequal but a relationship exists between them. These type of inequalities occur in business whenever there is a limit on supply, demand, sales etc. For example, if a producer requires a certain type of raw material for his factory and there is an upper limit in the availability of that raw material, then any decision which he takes about production should involve this constraint also. We will see in this chapter more about such situations.

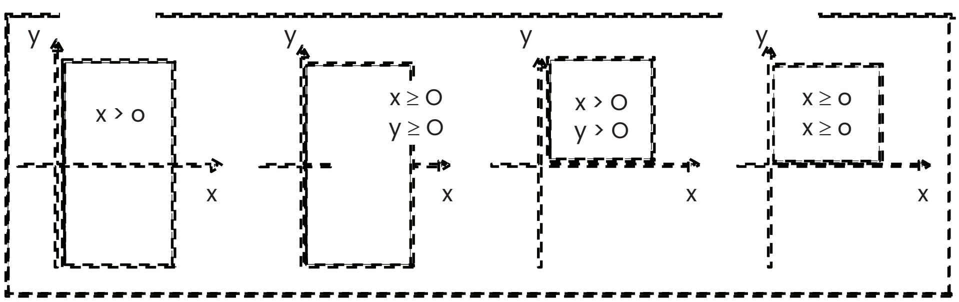
3.2 Linear inequalities in one variable and the solution space

Any linear function that involves an inequality sign is a linear inequality. It may be of one variable, or, of more than one variable. Simple example of linear inequalities are those of one variable only; viz., x > 0, x < 0 etc.

The values of the variables that satisfy an inequality are called the solution space, and is abbreviated as S.S. The solution spaces for (i) x > 0, (ii) x ⍰ 0 are shaded in the above diagrams, by using deep lines.

Linear inequalities in two variables: Now we turn to linear inequalities in two variables x and y and shade a few S.S.

Let us now consider a linear inequality in two variables given by 3x + y < 6

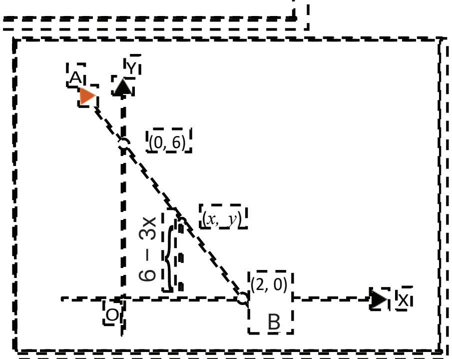


The inequality mentioned above is true for certain pairs of numbers (x, y) that satisfy 3x + y < 6. By trial, we may arbitrarily find such a pair to be (1,1) because 3 ⍰ 1 + 1 = 4, and 4 < 6.

Linear inequalities in two variables may be solved easily by extending our knowledge of straight lines.

For this purpose, we replace the inequality by an equality and seek the pairs of number that satisfy 3x + y = 6. We may write 3x + y = 6 as y = 6 – 3x, and draw the graph of this linear function.

Let x = 0 so that y = 6. Let y = 0, so that x = 2.



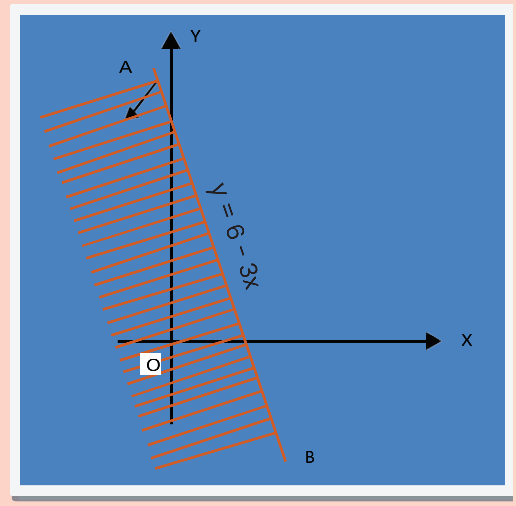
Any pair of numbers (x, y) that satisfies the equation y = 6 – 3x falls on the line AB.

Note: The pair of inequalities x 0, y 0 play an important role in linear programming problems.

Therefore, if y is to be less than 6 – 3x for the same value of x, it must assume a value that is less than the ordinate of length 6 – 3x.

All such points (x, y) for which the ordinate is less than 6 – 3x lie below the line AB.

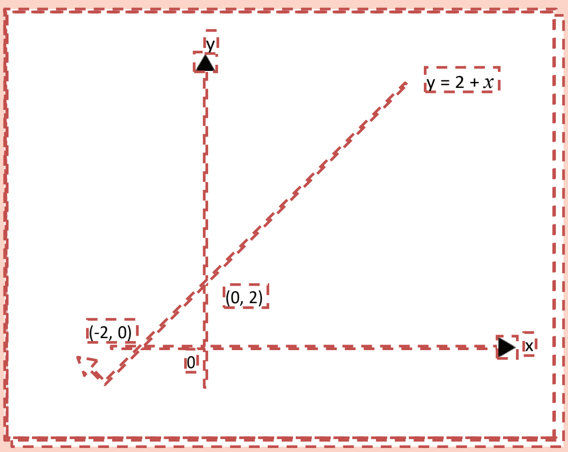
The region where these points fall is indicated by an arrow and is shaded too in the adjoining diagram. Now we consider two inequalities $3x + y \geq 6$ and $x - y \geq -2$ being satisfied simultaneously by x and y. The pairs of numbers (x, y) that satisfy both the inequalities may be found by drawing the graphs of the two lines $y = 6 - 3x$ and $y = 2 + x$, and determining the region where both the inequalities hold. It is convenient to express each equality with y on the left-side and the remaining terms in the right side. The first inequality $3x + y \leq 6$ is equivalent to $y \leq 6 - 3x$ and it requires the value of y for each x to be less than or equal to that of and on $6 - 3x$. The inequality is therefore satisfied by all points lying below the line $y = 6 - 3x$. The region where these points fall has been shaded in the adjoining diagram.



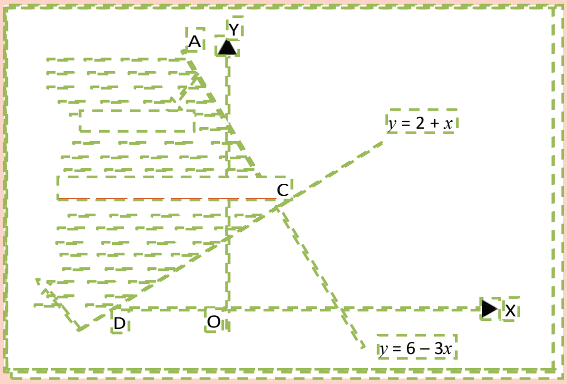
We consider the second inequality $x - y \leq -2$, and note that this is equivalent to $y \geq 2 + x$. It requires the value of y for each x to be larger than or equal to that of $2 + x$. The inequality is, therefore, satisfied by all points lying on and above the line $y = 2 + x$.

The region of interest is indicated by an arrow on the line $y = 2 + x$ in the diagram below. For $x = 0$, $y = 2 + 0 = 2$;

For $y = 0$, $0 = 2 + x$ i.e, $x = -2$.



By superimposing the above two graphs we determine the common region ACD in which the pairs (x, y) satisfy both inequalities.



Example 1: We now consider the problem of drawing graphs of the following inequalities $x \geq 0$, $y \geq 0$, $x \leq 6$, $y \leq 7$, $x + y \leq 12$ and shading the common region.

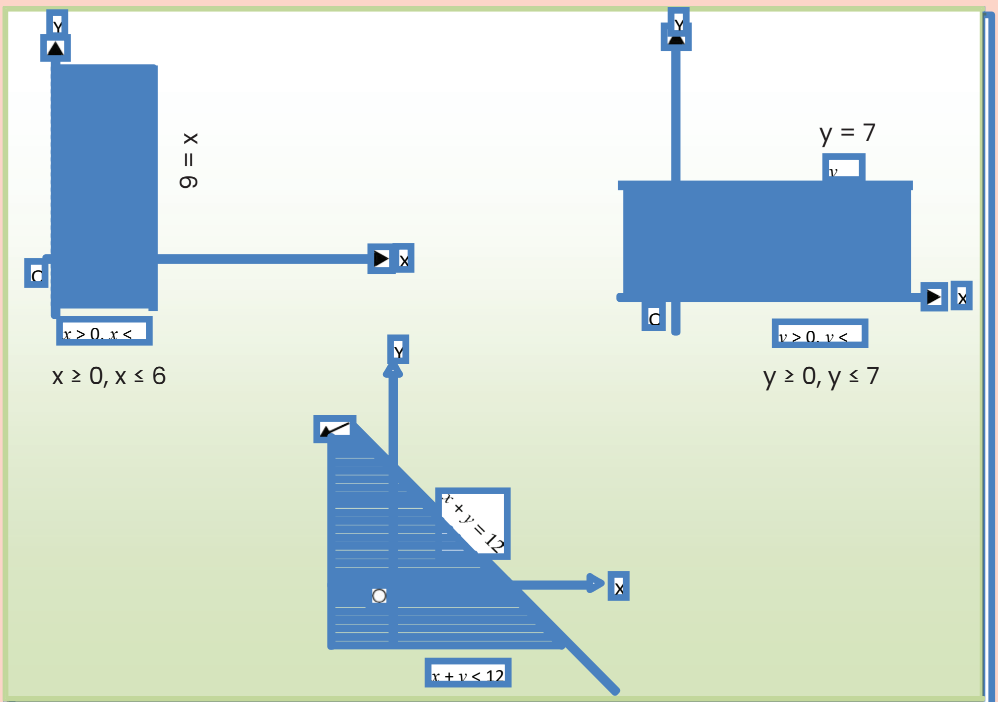
Note:

[1] The inequalities $3x + y \leq 6$ and $x - y \leq 2$ differ from the preceding ones in that these also include equality signs. It means that the points lying on the corresponding lines are also included in the region.

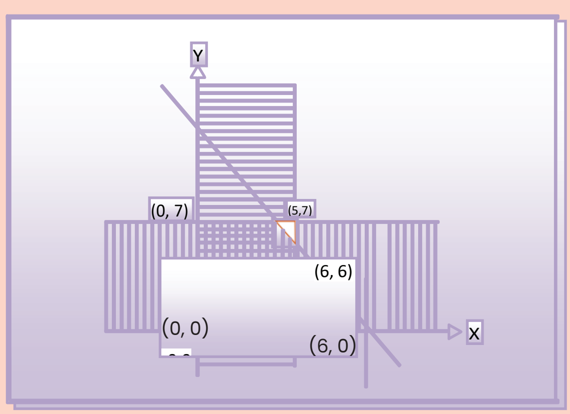
[2] The procedure may be extended to any number of inequalities. We note that the given inequalities may be grouped as follows :

$x \geq 0, y \geq 0$

$x \leq 6, y \leq 7, x + y \leq 12$



By superimposing the above three graphs, we determine the common region in the xy plane where all the five inequalities are simultaneously satisfied.



This common region is known as feasible region or the solution set (or the polygonal convex sets).

A region is said to be bounded if it can be totally included within a (very large) circle. The shaded region enclosed by deep lines in the previous diagram is bounded, since it can be included within a circle.

The objective function attains a maximum or a minimum value at one of the corner points of the feasible solution known as extreme points of the solution set. Once these extreme points (the points of intersection of lines bounding the region) are known, a compact matrix representation of these points is possible. We shall denote the matrix of the extreme points by E.

The coefficients of the objective function may also be represented by a column vector. We shall represent this column vector by C.

The elements in the product matrix EC shows different values, which the objective function attains at the various extreme points. The largest and the smallest elements in matrix EC are respectively the maximum and the minimum values of the objective function. The row in matrix EC in which this happens is noted and the elements in that row indicate the appropriate pairing and is known as the optimal solution.

In the context of the problem under consideration.

$$\begin{array}{c} \begin{array}{cc} X & Y \end{array} \\ \begin{bmatrix} 0 & 0 \\ 0 & 7 \\ 5 & 7 \\ 6 & 0 \\ 6 & 6 \end{bmatrix} \end{array}, C = \begin{bmatrix} 1 \\ 2 \end{bmatrix}$$

$$EC = \begin{bmatrix} 0 & 0 \\ 0 & 7 \\ 5 & 7 \\ 6 & 0 \\ 6 & 6 \end{bmatrix} \begin{bmatrix} 1 \\ 2 \end{bmatrix} = \begin{bmatrix} 0 & 0 \\ 0 & 7 \\ 5 & 7 \\ 6 & 0 \\ 6 & 0 \end{bmatrix} \begin{bmatrix} 1 \\ 2 \end{bmatrix} = \begin{bmatrix} 0 \text{ x } 1 + 0 \text{ x } 2 \\ 0 \text{ x } 1 + 7 \text{ x } 2 \\ 5 \text{ x } 1 + 7 \text{ x } 2 \\ 6 \text{ x } 1 + 0 \text{ x } 2 \\ 6 \text{ x } 1 + 6 \text{ x } 2 \end{bmatrix} = \begin{bmatrix} 0 \\ 14 \\ 19 \\ 6 \\ 18 \end{bmatrix}$$

The given objective function viz. Z = x + 2y is maximum at the points (5, 7) present in the third row of the matrix E. Thus the optimal solution is x = 5, y = 7, and the maximum value of the objective function is 19.

We now list the steps to be followed under graphical solution to a linear programming problem.

Step 1 Determine the region that satisfies the set of given inequalities.

Step 2 Ensure that the region is bounded*. If the region is not bounded, either there are additional hidden conditions which can be used to bound the region or there is no solution to the problem.

Step 3 Construct the matrix E of the extreme points, and the column vector C of the objective function.

Step 4 Find the matrix product EC. For maximization, determine the row in EC where the largest element appears; while for minimization, determine the row in EC where the smallest element appears.

Step 5 The objective function is optimized corresponding to the same row elements of the extreme point matrix E.

Note: If the slope of the objective function be same as that of one side of feasible region, there are multiple solutions to the problem. However, the optimized value of the objective function remains the same.

Example 2:

A manufacturer produces two products A and B, and has his machines in operation for 24 hours a day. Production of A requires 2 hours of processing in machine M1 and 6 hours in machine M2. Production of B requires 6 hours of processing in machine M1 and 2 hours in machine M2. The manufacturer earns a profit of ` 5 on each unit of A and ` 2 on each unit of B. How many units of each product should be produced in a day in order to achieve maximum profit?

Solution:

Let x1 be the number of units of type A product to be produced, and x2 is that of type B product to be produced. The formulation of the L.P.P. in this case is as below:

Maximize Z = $5x_1 + 2x_2$

* It is inconceivable for a practical problem to have an unbounded solution. subject to the constraints.

$2x_1 + 6x_2 < 24$

$6x_1 + 2x_2 < 24$

$x_1 \geq 0, x_2 \geq 0$

For the line 2x1 + 6x2 = 24

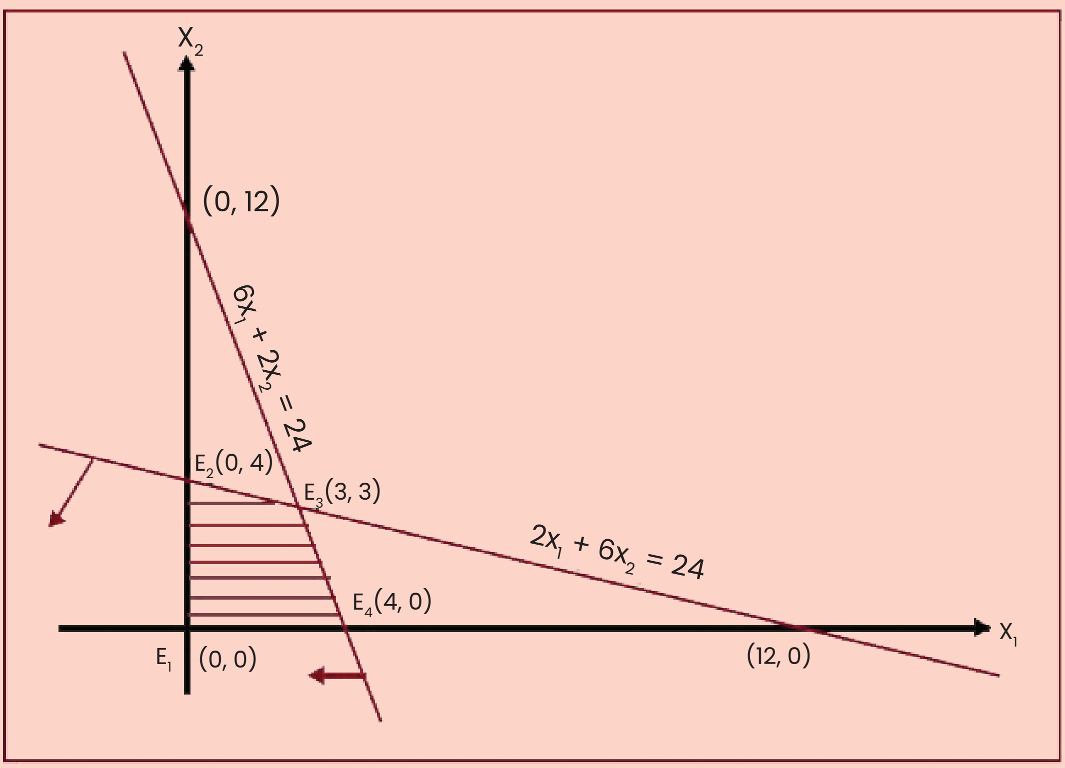
Let $x_1 = 0$, so that $x_2 = 4$

Let $x_2 = 0$, so that $x_1 = 12$

For the line $6x_1 + 2x_2 = 24$

Let $x_1 = 0$, so that $x_2 = 1$

Let $x_2 = 0$, so that $x_1 = 4$



The shaded portion in the diagram is the feasible region and the matrix of the extreme points E1, E2, E3 and E4 is

$$E = \begin{array}{c} \begin{array}{cc} X1 & Y2 \end{array} \\ \begin{bmatrix} 0 & 0 \\ 0 & 4 \\ 3 & 3 \\ 4 & 0 \end{bmatrix} \end{array} \begin{array}{l} \\ E1 \\ E2 \\ E3 \\ E4 \end{array}$$

The column vector for the objective function is $C = \begin{array}{c} \begin{bmatrix} 5 \\ 2 \end{bmatrix} \end{array}\begin{array}{l} x_1 \\ x_2 \end{array}$

The column vector the values of the objective function is given by

$$EC = \begin{bmatrix} 00 \\ 04 \\ 33 \\ 40 \end{bmatrix} \begin{bmatrix} 5 \\ 2 \end{bmatrix} = \begin{bmatrix} 0 \times 5 + 0 \times 2 \\ 0 \times 5 + 4 \times 2 \\ 3 \times 5 + 3 \times 2 \\ 4 \times 5 + 0 \times 2 \end{bmatrix} = \begin{bmatrix} 0 \\ 8 \\ 21 \\ 20 \end{bmatrix} \begin{array}{l} E1 \\ E2 \\ E3 \\ E4 \end{array}$$

Since 21 is the largest element in matrix EC, therefore the maximum value is reached at the extreme point E3 whose coordinates are (3,3).

Thus, to achieve maximum profit the manufacturer should produce 3 units each of both the products A and B.

Summary of Graphical Method

It involves:

(i) Formulating the linear programming problem, i.e. expressing the objective function and constraints in the standardised format.

(ii) Plotting the capacity constraints on the graph paper. For this purpose normally two terminal points are required. This is done by presuming simultaneously that one of the constraints is zero. When constraints concerns only one factor, then line will have only one origin point and it will run parallel to the other axis.

(iii) Identifying feasible region and coordinates of corner points. Mostly it is done by breading the graph, but a point can be identified by solving simultaneous equation relating to two lines which intersect to form a point on graph.

(iv) Testing the corner point which gives maximum profit. For this purpose the coordinates relating to the corner point should put in objectives function and the optimal point should be ascertained.

(v) For decision – making purpose, sometimes, it is required to know whether optimal point leaves some resources unutilized. For this purpose value of coordinates at the optimal point should be put with constraint to find out which constraints are not fully utilized.

Example 3: A company produces two products A and B, each of which requires processing in two machines. The first machine can be used at most for 60 hours, the second machine can be used at most for 40 hours. The product A requires 2 hours on machine one and one hour on machine two. The product B requires one hour on machine one and two hours on machine two. Express above situation using linear inequalities.

Solution: Let the company produce, x number of product A and y number of product B. As each of product A requires 2 hours in machine one and one hour in machine two, x number of product A requires 2x hours in machine one and x hours in machine two. Similarly, y number of product

B requires y hours in machine one and 2y hours in machine two. But machine one can be used for 60 hours and machine two for 40 hours. Hence 2x + y cannot exceed 60 and x + 2y cannot exceed 40. In other words,

$2x + y \leq 60$ and $x + 2y \leq 40$.

Thus, the conditions can be expressed using linear inequalities.

Example 4: A fertilizer company produces two types of fertilizers called grade I and grade II. Each of these types is processed through two critical chemical plant units. Plant A has maximum of 120 hours available in a week and plant B has maximum of 180 hours available in a week. Manufacturing one bag of grade I fertilizer requires 6 hours in plant A and 4 hours in plant B. Manufacturing one bag of grade II fertilizer requires 3 hours in plant A and 10 hours in plant B. Express this using linear inequalities.

Solution: Let us denote by x1, the number of bags of fertilizers of grade I and by x2, the number of bags of fertilizers of grade II produced in a week. We are given that grade I fertilizer requires 6 hours in plant A and grade II fertilizer requires 3 hours in plant A and plant A has maximum of 120 hours available in a week. Thus $6x_1 + 3x_2 \leq 120$.

Similarly grade I fertilizer requires 4 hours in plant B and grade II fertilizer requires 10 hours in Plant B and Plant B has maximum of 180 hours available in a week. Hence, we get the inequality $4x_1 + 10x_2 \leq 180$.

Example 5: Graph the inequalities $5x_1 + 4x_2 \geq 9$, $x_1 + x_2 \geq 3$, $x_1 \geq 0$ and $x_2 \geq 0$ and mark the common region.

Solution: We draw the straight lines $5x_1 + 4x_2 = 9$ and $x_1 + x_2 = 3$.

Table for $5x_1 + 4x_2 = 9$

X_1	0	9/5
X_2	9/4	0

Table for $x_1 + x_2 = 3$

X_1	0	3
X_2	3	0

Now, if we take the point (4, 4), we find

$5x_1 + 4x_2 \geq 9$

i.e., $5.4 + 4.4 \geq 9$

or, $36 \geq 9$ (True)

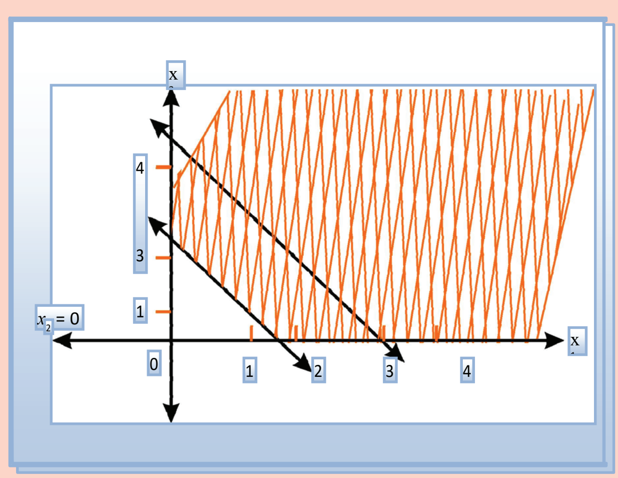
$x_1 + x_2 \geq 3$

i.e., $4 + 4 \geq 3$

$8 \geq 3$ (True)

Hence (4, 4) is in the region which satisfies the inequalities.

We mark the region being satisfied by the inequalities and note that the cross-hatched region is satisfied by all the inequalities.



Example 6: Draw the graph of the solution set of the following inequality and equality:

$x + 2y = 4.$

$x - y \leq 3.$

Mark the common region.

Solution: We draw the graph of both $x + 2y = 4$ and $x - y \leq 3$ in the same plane.

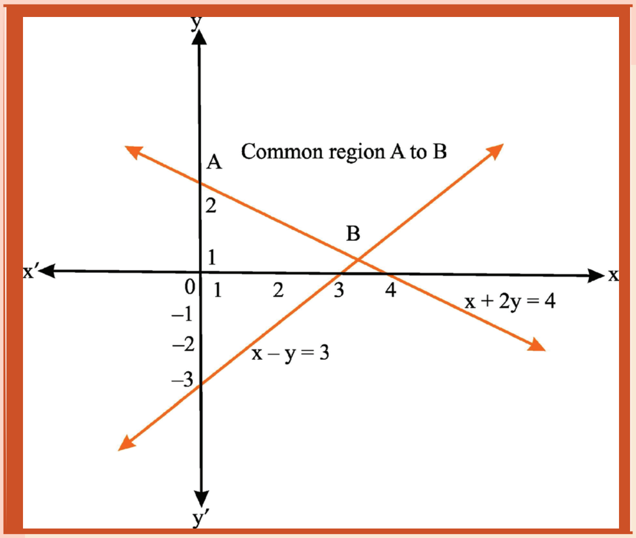
The solution set of system is that portion of the graph of $x + 2y = 4$ that lies within the half-plane representing the inequality $x - y \leq 3$.

For $x + 2y = 4$,

x	4	0
y	0	2

For $x - y = 3$,

x	3	0
y	0	-3



Example 7: Draw the graphs of the following inequalities:

$x + y \leq 4,$

$x - y \leq 4,$

$x \geq -2.$

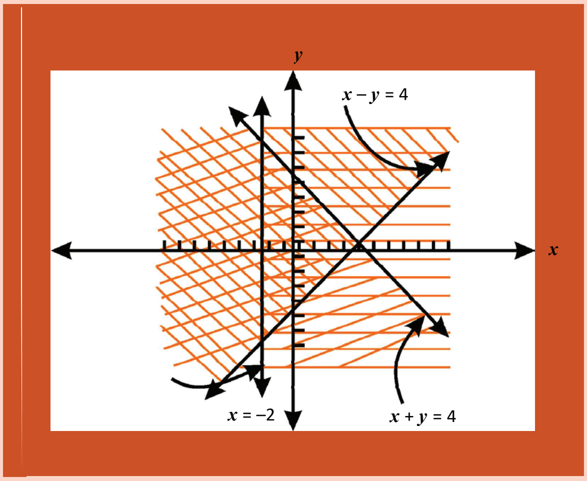
and mark the common region.

For $x - y = 4$

x	4	0
y	0	-4

For $x - y = 4$

x	0	4
y	4	0



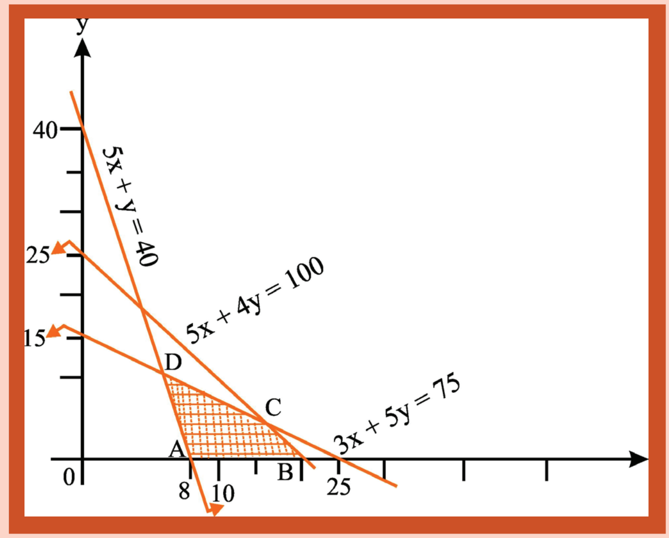
The common region is the one represented by overlapping of the shadings.

Example 8: Draw the graphs of the following linear inequalities:

$5x + 4y \leq 100$, $5x + y \geq 40$

$3x + 5y \leq 75$, $x \geq 0, y \geq 0$.

and mark the common region.



Solution:

$5x + 4y = 100$ or,

$3x + 5y = 75$ or,

$5x + y = 40$ or,

Plotting the straight lines on the graph paper we have the above diagram:

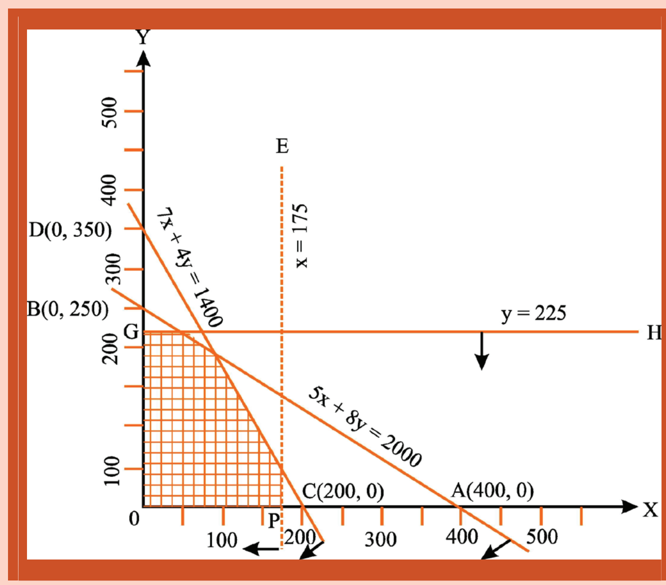
The common region of the given inequalities is shown by the shaded portion ABCD.

Example 9: Draw the graphs of the following linear inequalities:

$5x + 8y \leq 2000$, $x \leq 175$, $x \geq 0$.

$7x + 4y \leq 1400$, $y \leq 225$, $y \geq 0$.

and mark the common region:



Solution: Let us plot the line AB ($5x + 8y = 2{,}000$) by joining the points A(400, 0) and B(0, 250).

Similarly, we plot the line CD ($7x + 4y = 1400$) by joining the points C(200, 0) and D(0, 350).

x	400	0
y	0	250

x	200	0
y	0	350

Also, we draw the lines EF($x = 175$) and GH ($y = 225$).

The required graph is shown alongside in which the common region is shaded.

Example 10: Draw the graphs of the following linear inequalities

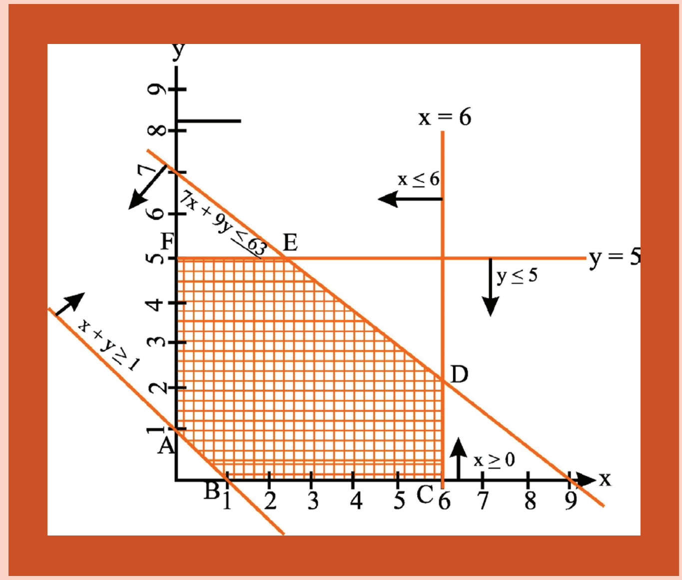
:$x + y \geq 1$, $7x + 9y \leq 63$,

$y \leq 5$, $x \leq 6$, $x \geq 0$, $y \geq 0$. and mark the common region.

Solution: $x + y = 1$; $\begin{array}{c|c|c} x & 1 & 0 \\ \hline y & 0 & 1 \end{array}$; $7x+9y=63$, $\begin{array}{c|c|c} x & 9 & 0 \\ \hline y & 0 & 7 \end{array}$

We plot the line AB ($x + y = 1$), CD ($y = 5$), EF ($x = 6$), DE ($7x + 9y = 63$).

Given inequalities are shown by arrows. Common region ABCDEF is the shaded region.



Example 11: Two machines (I and II) produce two grades of plywood, grade A and grade B. In one hour of operation machine I produces two units of grade A and one unit of grade B, while machine II, in one hour of operation produces three units of grade A and four units of grade B. The machines are required to meet a production schedule of at least fourteen units of grade A and twelve units of grade B. Express this using linear inequalities and draw the graph.

Solution: Let the number of hours required on machine I be x and that on machine II be y. Since in one hour, machine I can produce 2 units of grade A and one unit of grade B, in x hours it will produce 2x and x units of grade A and B respectively. Similarly, machine II, in one hour, can produce 3 units of grade A and 4 units of grade B. Hence, in y hours, it will produce 3y and 4y units Grade A & B respectively.

The given data can be expressed in the form of linear inequalities as follows:

$2x + 3y \geq 14$ (Requirement of grade A)

$x + 4y \geq 12$ (Requirement of grade B)

Moreover x and y cannot be negative, thus $x \geq 0$ and $y \geq 0$

Let us now draw the graphs of above inequalities. Since both x and y are positive, it is enough to draw the graph only on the positive side.

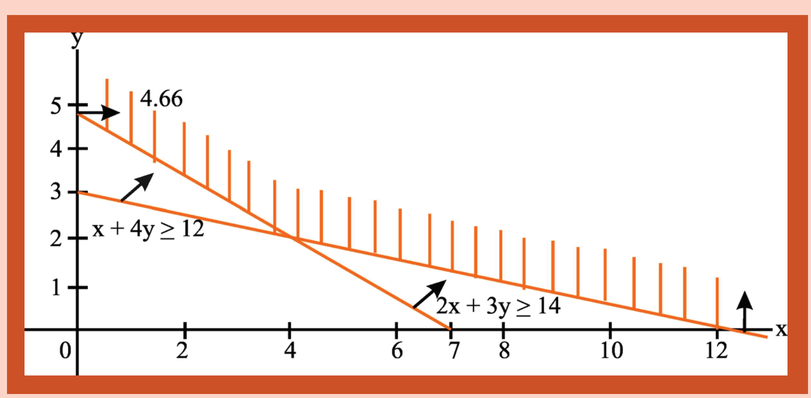
The inequalities are drawn in the following graph:

For $2x + 3y = 14$,

x	7	0
y	0	4.66

For $x + 4y = 12$,

x	0	12
y	3	0



In the above graph we find that the shaded portion is moving towards infinity on the positive side. Thus the result of these inequalities is unbounded.

Chapter 4:

Mathematics of Finance

At the foundation level with regards to Business Mathematics the topic Time Value of money is very important for students not only to acquire professional knowledge but also for examination point of view. Here in this chapter an attempt is made for solving and understanding the concepts of time value of money with the help of following questions with solutions.

Problems on Simple, Compound and Effective rate of interest

Simple interest = PTI, Where P = Principal, T = Time, I = Rate of Interest

1. A sum of money amount to Rs. 6,200 in 2 years and Rs. 7,400 in 3 years. The principal and rate of interest is

Solution: A sum of money in 2 years = P + P.2.I = Rs 6200

A sum of money in 3 years = P + P.3.I = Rs 7400

Interest in 1 year = Rs.1200; Interest in 2 years = Rs. 2400

Amount = Rs.6200, P= Principal = 6200-2400= Rs. 3800.

$2400 = 3800 \times 2 \times \frac{I}{100}$, I = rate of interest = 31.58% .

2. A sum of money doubles itself in 10 years. The number of years it would triple itself is

Solution:

$2P = P + \frac{PTR}{100}$, $P = \frac{PTR}{100}$; T = 10 then R = 10%

$3P = P + \frac{PTR}{100}$ then $2P = \frac{PTR}{100}$ and R= 10%

Time (T) = 20 years.

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

Compound rate of interest $= P(1+i)^n - P = A-P$

where P = principle i= interest n =conversion period

3. The population of a town increases every year by 2% of the population at the beginning of that year. The number of years by which the total increase of population be 40% is

Solution: $1.4P = P(1+0.02)^n$

$(1.02)^n = 1.4$, n = 17 years (app)

Depreciation $(A) = P(1-i)^n$

Where, A = Scrap Value, P = Original Cost, I =Depreciated at the rate, n = Number of years

4. A machine is depreciated at the rate of 20% on reducing balance. The original cost of the machine was Rs.1,00,000 and its ultimate scrap value was Rs. 30,000. The effective life of the machine is

Solution: Here A = Scrap Value = 30000 and P = Original Cost = Rs.1,00,000

$30000 = 100000\,(1-0.2)^n$

3/10 = 0.3 = (0.8) n , n = 5.4 years

5. The useful life of a machine is estimated to be 10 years and cost Rs. 10,000. Rate of depreciation is 10% p.a. The scrap value at the end of its life is

Solution: Here A = Scrap Value =?, P= Original cost = 10,000 n= 10 , I = 10%

$A = 100000\,(1-0.1)^{10} = 10000(0.9)^{10}$

A = 3486.78

The difference between simple and compound interest for 2 years = $P.i^2$,
where P =Principal, i= interest

6. The difference between the S.I and the C.I on Rs. 2,400 for 2 years at 5% p.a is

Solution: The difference between simple and compound interest for 2 years $=2400(0.05)^2= 6$.

The differences between simple and compound interest for 3 years = $3P.i^2 + P.i^3$, where P = Principal i= Interest

7. The difference between the S.I and the C.I on a certain money invested for 3 years at 6% p.a is Rs. 110.16 the principle is

Solution: The differences between simple and compound interest for 3 years

$= 110.16 = P(3i^2+i^3)= P(3\times0.06^2+0.06^3)$

$110.16 = P(0.0108+0.000216) = P(0.011016)$

$$P= \frac{110.16}{0.011016} = 10,000$$

9. The annual birth and death rates per 1,000 are 39.4 and 19.4 respectively. The number of years in which the population will be doubled assuming there is no immigration or emigration is

Solution: Here given, birth rate per 1,000 = 39.4 and death rates per 1,000 = 19.4

difference = 20 % per 1000 population

growth rate $= \frac{20}{1000} \times 100 = 2\%$

Future population double

P =1000, A = 2000, r = 2%

$2000= 1000(1+0.02)^n$

$(1.02)^n = 2$

Number of years = n= 35

10. What annual rate of interest compounded annually doubles an investment in 7 years? (Given that 2 1/7= 1.104090)

Solution: If the principal be P, $A_n = 2$ p

Since $A_n = P(1+ i)^n$

$2P = P(1 + i)^7$

$2^{1/7} = (1 + i)$

$1.104090 = 1 + i$

I = 0.10409, Required rate of interest = 10.41% per annum

11.Vidya deposited Rs.60000 in a bank for two years with the interest rate of 5.5% p.a. How much interest she would earn? what will be the final value of investment?

Solution: Required interest amount is given by, I = P x it = Rs.60,000 x $\frac{5.5}{100}$ x 2 = Rs.6,600

The amount value of investment is given by, A = P + I = Rs.(50,000 + 6600) = Rs. 66,600

12. Rajiv invested Rs. 75,000 in a bank at the rate of 8% p.a. simple interest rate. He received Rs. 135,000 after the end of term. Find out the period for which sum was invested by Rajiv.

Solution: $\boxed{\text{We know} \quad A = P + Pit = P(1+it)}$

i.e. $75000 = 135000 \left(1 + \frac{8}{100} \times t\right)$

$135000/75000 = \frac{100 + 8t}{100}$

$1.8 \times 100 - 100 = 8t$

$80 = 8t$

t (Time) = 10 years

13. Which is a better investment, 3.6% per year compounded monthly or 3.2% per year simple interest? Given that $(1+0.002)^{12}$ =1.0366.

$\boxed{\textbf{Effective rate of interest } E = (1 + i)^n - i}$

Solution: i=3/12 = 0.25% =0.003 , n = 12

$E = (1 + i)^n - 1$

$= (1 + 0.003)^{12} - 1 = 1.0366 - 1 = 0.0366$ or = 3.66%

Effective rate of interest (E) being more than 3.66% is more than simple interest

so the Effective rate of interest (E) is better investment.

the 3.2% per year

14. The C.I on Rs. 4,000 for 6 months at 12% p.a payable quarterly is

Solution: Here P = Rs.4000, n = 6/2= 2 , r= 0.12/4= 0.03

Compound Interest = $[P(1+i)^n-P]$ = $[4000(1+0.03)^2-4000]$ =Rs.243.60

Annuity applications

F = Future value = C.F. $(1 + i)^n$ Where C.F = Cash flow
i= rate of interest, n = time period

15: Ravi invest Rs.5000 in a two-year investment that pays you 12% per annum. Calculate the future value of the investment

Solution: We know, F= Future value = C.F.(1 + i)n, Where C.F= Cash flow =Rs.5000, i = rate of interest=0.12,n = time period=2

F= Rs. $5000(1+0.12)^2$=Rs. 5000 X 1.2544 =Rs. 6272.

Future value of the annuity=A(n,i)= $A\left[\frac{(1+i)^n - 1}{i}\right]$

16: Find the future value of an annuity of Rs.5000 is made annually for 7 years at interest rate of 14% compounded annually. [Given that (1.14)7= 2.5023]

Solution: Here annual payment A=Rs.5000, n=7, i=14% = 0.14

Future value of the annuity=A(7,0.14)=5000 $\left[\frac{(1+0.14)^7 - 1}{(0.14)}\right] = \left[\frac{5000 \times (2.5023-1)}{0.14}\right]$ = Rs. 53,652.46

17: Rs.2000 is invested at the end of each month in an account paying interest 6% per year compounded monthly. What is the future value of this annuity after 10th payment? Given that $(1.005)^{10}$ = 1.0511

Solution: Here A= Rs.2000, n= 10, i= 6% per annum= 6/12 % per month = 0.005

Future value of annuity after 10 months is given by

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

$$A(10, 0.005)=2000\left[\frac{(1 + 0.005)^{10} - 1}{0.005}\right] = 2000\left[\frac{1.0511 - 1}{0.005}\right]$$

=2000 x 10.23 = Rs. 20,460

Future value of the annuity regular or annuity due = $A\left[\frac{(1+i)^n - 1}{i}\right] \times (1 + i)$

18.Swati invests Rs. 20,000 every year starting from today for next 10 years. Suppose interest rate 8% per annum compounded annually. Calculate future value of the annuity. Given that (1 + 0.08)10 = 2.158925)

Solution: Calculate future value as though it were an ordinary annuity.

Future value of the annuity as if it were an ordinary annuity

$$=\text{Rs.}20000\left[\frac{(1+0.08)^{10}-1}{0.08}\right]$$

$$=\text{Rs.}20000 \times 14.486563 = \text{Rs.}289731.25$$

Multiply the result by $(1 + i)$ $=\text{Rs.}289731.25 \times (1+0.08)$

$= \text{Rs. } 3,12,909.75$

19. What is the present value of Rs.100 to be received after two years compounded annually at 10%.

Solution: Here A_n =Rs.100, i=10% = 0.1, n= 2

$$\text{Required present value} = \frac{A_n}{(1+i)^n} = \frac{100}{(1+0.1)^2} = \frac{100}{1.21} = 82.64$$

Thus Rs 82.64 shall grow to Rs.100 after 2 years at 10% compounded annually.

20: Find the present value of Rs.10000 to be required after 5 years if the interest rate be 9%. Given that $(1.09)^5$=1.5386.

Solution: Here i= 0.09 , n =5 , A_n=10000

$$\text{Required present value} = \frac{A_n}{(1+i)^n} = \frac{10000}{(1+0.09)^5} = \frac{10000}{1.5386} = \text{Rs.}6499.42$$

$$\boxed{\text{Present Value of Annuity regular} = A = P(n,i) = A.\left[\frac{(1+i)^n-1}{i(1+i)^n}\right], P(n,i) = \left[\frac{(1+i)^n-1}{i(1+i)^n}\right]}$$

21.Soni borrows Rs.5,00,000 to buy a car. If he pays equal instalments for 10 years and 10% interest on outstanding balance, what will be the equal annual instalment? Given [P(10,0.10) =6.14457]

Solution:

We know, $A = \frac{V}{P(n,i)}$ Here V = Rs.500000, n = 10, I = 10% p.a.= 0.10

$$\therefore \text{Annual Instalment} = \frac{V}{P(n,i)} = \text{Rs. } \frac{5,00,000}{P(10,0.10)} = \text{Rs. } \frac{5,00,000}{6.14457} = \text{Rs. } 8,1372.66$$

22. If Rs.10,000 is paid every year for ten years to pay off a loan. What is the loan amount if interest rate be 14% per annum compounded annually? Given [P(10,0.14) =5.21611]

Solution:

V = A.P(n,i) Here A =Rs.10000, n = 10, i =0.14

V = 10000 x P(10, 0.14)

= 10000 x 5.21611 = Rs.52161.10

Therefore, the loan amount is Rs. 52,161.10

23: Ram bought a Scooter costing Rs.73000 by making a down payment of Rs.3000 and agreeing to make equal annual payment for four years. How much would be each payment if the interest on unpaid amount be 14% compounded annually? Given [P(4 ,0.14) =2.91371]

Solution: In the present case we have present value of the annuity i.e. Rs.70000 (73000-3000) and we have to calculate equal annual payment over the period of four years.

We know that, V = A. P (n , i) Here n = 4 and I = 0.14

$$A = \frac{V}{P(n,i)} = \frac{70000}{P(4{,}0.14)} = \frac{70000}{2.91371}$$

Therefore, each payment = Rs .24024.35

24. Suppose your Father decides to gift you Rs. 20,000 every year starting from today for the next six years. You deposit this amount in a bank as and when you receive and get 10% per annum interest rate compounded annually. What is the present value of this annuity?

Solution: For calculating value of the annuity immediate following steps will be followed. Present value of the annuity as if it were a regular annuity for one year less i.e. for five years

= Rs.20,000 x P(5, 0.10)

= Rs.20, 000 x 3.79079 = Rs.75815.80

Add initial cash deposit to the value, Rs. (75815.80+20,000) = Rs. 95,815.80

Sinking Fund: Interest is computed at end of every period with specified interest rate.

25. How much amount is required to be invested every year so as to accumulate Rs.5,00,000 at the end of 10 years if interest is compounded annually at 10%? Given A. (10, 0.1)= 15.9374248

Solution:

Here A = 500000, n = 10, $A(n,i) = \left[\frac{(1+i)^n - 1}{i}\right] = \left[\frac{(1+0.1)^{10} - 1}{0.1}\right] = 15.9374248$

since A = P. A. (n ,i)

500000 = P A. (10, 0.1) = P x 15.9374248

$$P = \frac{500000}{15.9374248} = \text{Rs.}31372.70$$

26. ABC Ltd. wants to lease out an asset costing Rs.360000 for a five year period. It has fixed a rental of Rs.105000 per annum payable annually starting from the end of first year. Suppose rate of interest is 14% per annum compounded annually on which money can be invested by the company. Is this agreement favorable to the company?

Solution: First, we have to compute the present value of the annuity of Rs. 1,05,000 for five years at the interest rate of 14% p.a. compounded annually.

The present value V of the annuity is given by V =A.P (n ,i) = 105000 x P (5, 0.14)

= 105000 x 3.43308 = Rs. 3,60,473.40

which is greater than the initial cost of the asset and consequently leasing is favourable to the lessor.

27. A company is considering proposal of purchasing a machine either by making full payment of Rs.4000 or by leasing it for four years at an annual rate of Rs.1250. Which course of action is preferable if the company can borrow money at 14% compounded annually?

Solution: The present value V of annuity is given by

V =A.P (n, i) = 1250 x P (4,0.14)

= 1250 x 2.91371 = Rs.3642.11

which is less than the purchase price, and consequently leasing is preferable.

28.A machine can be purchased for Rs.50,000. Machine will contribute Rs. 12,000 per year for the next five years. Assume borrowing cost is 10% per annum compounded annually. Determine whether machine should be purchased or not.

Solution: The present value of annual contribution

V = A.P. (n, i)

= 12000 P(5,0.10) = 12000 x 3.79079

= Rs. 45,489.48

which is less than the initial cost of the machine. Therefore, machine must not be purchased.

29. A machine with useful life of seven years costs Rs.10000 while another machine with useful life of five years costs Rs.8000. The first machine saves labour expenses of Rs.1900 annually and the second one saves labour expenses of Rs.2200 annually. Determine the preferred course of action. Assume cost of borrowing as 10% compounded per annum.

Solution: The present value of annual cost savings for the first machine = Rs. 1900.P (7, 0.10)

= Rs.1900 x 4.86842 = Rs.9250

cost of machine being Rs.10000 it costs more by Rs.750 than it saves in terms of labour cost.

The present value of annual cost savings of the second machine = Rs. 2200.P (5,0.10) = Rs. 2200 x 3.79079

= Rs.8339.74

Cost of the second machine being Rs.8000, effective savings in labour cost is Rs.339.74. Hence the second machine is preferable.

30.An investor intends purchasing a three year Rs.1000 par value bond having nominal interest rate of 10%. At what price the bond may be purchased now if it matures at par and the investor requires a rate of return of 14%?

Solution:

$$\text{Present value of the bond} = \frac{100}{(1+0.14)^1} + \frac{100}{(1+0.14)^2} + \frac{100}{(1+0.14)^3} + \frac{100}{(1+0.14)^3}$$

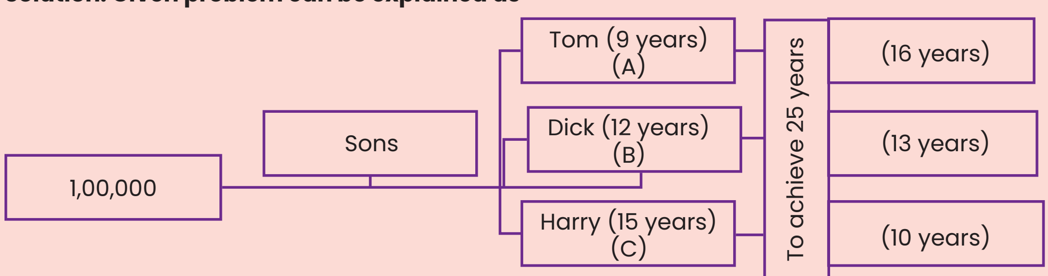
$$= 100 \times 0.87719 + 100 \times 0.769467 + 100 \times 0.674\,972 + 1000 \times 0.674972$$

$$= 87.719 + 76.947 + 67.497 + 674.972 = 907.125$$

Thus the purchase value of the bond is Rs. 907.125

31. Johnson left Rs. 1,00,000 with the direction that it should be divided in such a way that his minor sons Tom, Dick and Harry aged 9, 12 and 15 years should each receive equally after attaining the age 25 years. The rate of interest being 3.5%, how much each son receives after getting 25 years old?

Solution: Given problem can be explained as



$$A\left(1+\frac{3.5}{100}\right)^{16} = B\left(1+\frac{3.5}{100}\right)^{13} = C\left(1+\frac{3.5}{100}\right)^{10}$$

A(1.035)16= B(1.035)13= c(1.035)10

A(1.035)6= B(1.035)3= c------(I)

A : B : C = 1 : (1.035)3: (1.035)6

A + B + C= 100000

x + x (1.035)3+ x (1.035)6 = 100000

x [1+(1.035)3+ (1.035)6] = 10. 0000

x(3.337973) = 1,00,000 then x = Rs.29958.30 (A' s share)

B' s share = 29958.30 (1.035)3= Rs.33215.30

C' s share = 29958.30 (1.035)6= Rs.36826.40

32. A machine costs Rs. 5,20,000 with an estimated life of 25 years. A sinking fund is created to replace it by a new model at 25% higher cost after 25 years with a scrap value realization of Rs. 25000. what amount should be set aside every year if the sinking fund investments accumulate at 3.5% compound interest p.a.?

Solution:

Cost of new machine = $5{,}20{,}000 \times \frac{125}{100}$ = Rs. 6,50, 000, Scrap value = Rs. 25,000

For new machine = 650000-25000= Rs. 6,25, 000.

Here = Rs. 6,25, 000, n= 25 , i = 0.35

$$6{,}25{,}000 = P.\left[\frac{(1+i)^n - 1}{i}\right] = P.\left[\frac{(1+0.035)^{25} - 1}{0.035}\right]$$

$$6{,}25{,}000 = \text{then } P = \frac{625000}{38.95} = \text{Rs. } 16046.27$$

33. Appu retires at 60 years receiving a pension of Rs. 14,400 a year paid in half-yearly installments for rest of his life after reckoning his life expectation to be 13 years and that interest at 4% p.a. is payable half-yearly. What single sum is equivalent to his pension?

Solution:

Amount = 14400/2 = Rs.7200, n = 13×2 = 26 , i= 4/2 = 0.02

$$V = A.P(n,i) = A.\left[\frac{(1+i)^n - 1}{i(1+i)^n}\right] = 7200.\left[\frac{(1+0.02)^{26} - 1}{0.02(1+0.02)^{26}}\right]$$

$$= 7200 \times 20.1210357 = \text{Rs. } 1{,}44{,}871.4574$$

single sum is equivalent to his pension= Rs. 144871.46

34. A sinking fund is created for redeeming debentures worth Rs. 5 lakhs at the end of 25 years. How much provision needs to be made out of profits each year provided sinking fund investments can earn interest at 4% p.a.? [Given (1.04)25= 2.6658]

Solution: A = P.A(n,i) , where A= Amount is to be saved, P = Periodic payments, n= 25 years

$$A = P.\left[\frac{(1+i)^n - 1}{i}\right] = P\left[\frac{(1+0.04)^{25} - 1}{0.04}\right] = P.41.6459083$$

$$5{,}00{,}000 = P.416459083$$

$$P = \frac{500000}{41{,}6459083} = Rs.\ 12{,}006$$

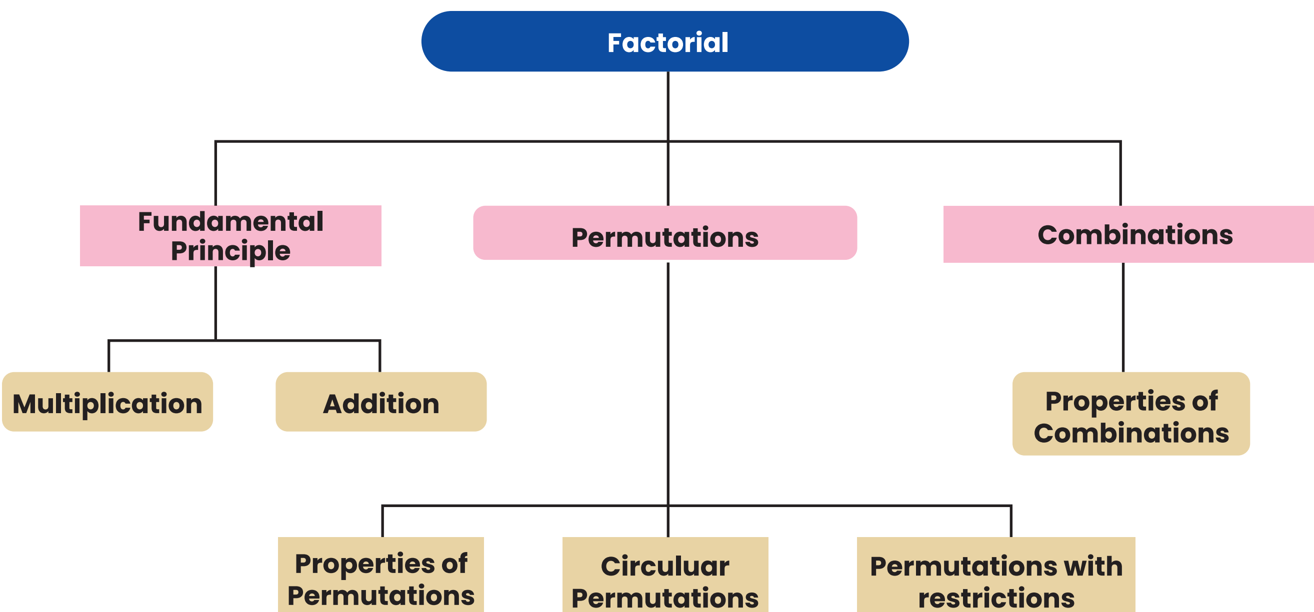
Chapter 5:

Permutations and Combinations

This chapter on **CA Foundation Course Paper 3: Quantitative Aptitude** covers the essential concepts of **"Permutations and Combinations"** for the purpose of making students understand the arrangement of different objects using concrete examples.

After reading this Chapter a student will be able to understand —

- difference between permutation and combination for the purpose of arranging different objects;
- number of permutations and combinations when r objects are chosen out of n different objects.
- meaning and computational techniques of circular permutation and permutation with restrictions.



Introduction

In this chapter we will learn problem of arranging and grouping of certain things, taking particular number of things at a time. It should be noted that (a, b) and (b, a) are two different arrangements, but they represent the same group. In case of arrangements, the sequence or order of things is also taken into account.

The manager of a large bank has a difficult task of filling two important positions from a group of five equally qualified employees. Since none of them has had actual experience, he decides to allow each of them to work for one month in each of the positions before he makes the decision. How long can the bank operate before the positions are filled by permanent appointments?

Solution to above - cited situation requires an efficient counting of the possible ways in which the desired outcomes can be obtained. A listing of all possible outcomes may be desirable, but is likely to be very tedious and subject to errors of duplication or omission. We need to devise certain techniques which will help us to cope with such problems. The techniques of permutation and combination will help in tackling problems such as above.

Fundamental Principles of Counting

(a) Multiplication Rule: If certain thing may be done in 'm' different ways and when it has been done, a second thing can be done in 'n' different ways then total number of ways of doing both things simultaneously = $m \times n$.

Eg. if one can go to school by 5 different buses and then come back by 4 different buses then total number of ways of going to and coming back from school = $5 \times 4 = 20$.

(b) Addition Rule: It there are two alternative jobs which can be done in 'm' ways and in 'n' ways respectively then either of two jobs can be done in $(m + n)$ ways.

Eg. if one wants to go school by bus where there are 5 buses or to by auto where there are 4 autos, then total number of ways of going school = $5 + 4 = 9$.

Note:- 1) **AND $\Rightarrow$ Multiply**

OR $\Rightarrow$ Add

2) The above fundamental principles may be generalised, wherever necessary.

The Factorial Definition: The factorial n, written as n! or $\lfloor\underline{n}$, represents the product of all integers from 1 to n both inclusive. To make the notation meaningful, when n = o, we define o! or $\lfloor\underline{0}$

Thus, $n! = n(n-1)(n-2)3.2.1$

Example 1: Find 5!, 4! and 6!

Solution: $5! = 5 \times 4 \times 3 \times 2 \times 1 = 120$; $4! = 4 \times 3 \times 2 \times 1 = 24$; $6! = 6 \times 5 \times 4 \times 3 \times 2 \times 1 = 720$.

Example 2: Find 9 ! / 6 ! ; 10 ! / 7 !.

Solution:

$$\frac{9!}{6!} = \frac{9 \times 8 \times 7 \times 6}{6!} = 9 \times 8 \times 7 = 504; \quad \frac{10!}{7!} = \frac{10 \times 9 \times 7!}{7!}$$

Example 3: Find x if $1/9! + 1/10! = x/11!$

Solution: $1/9!\,(1 + 1/10) = x/11 \times 10 \times 9!$ or, $11/10 = x/11 \times 10$ i.e., $x = 121$

Example 4: Find n if $\lfloor\underline{n+1} = 30\lfloor\underline{n-1}$

Solution: $\lfloor\underline{n+1} = 30\lfloor\underline{n-1} \Rightarrow (n+1).n\lfloor\underline{n-1} = 30\lfloor\underline{n-1}$

or $n^2 + n = 30$ or $n^2 + n - 30$ or, $n^2 + 6n - 5n - 30 = 0$ or, $(n + 6)(n - 5) = 0$

Permutations:

A group of persons want themselves to be photographed. They approach the photographer and request him to take as many different photographs as possible with persons standing in different positions amongst themselves. The photographer wants to calculate how many films does he need to exhaust all possibilities? How can he calculate the number?

In the situations such as above, we can use permutations to find out the exact number of films.

Definition: The ways of arranging or selecting smaller or equal number of persons or objects from a group of persons or collection of objects with due regard being paid to the order of arrangement or selection, are called permutations.

Number of Permutations when r objects are chosen out of n different objects. (Denoted by ^nP_r or nPr or P(n, r)) :

Let us consider the problem of finding the number of ways in which the first r rankings are secured by n students in a class. As any one of the n students can secure the first rank, the number of ways in which the first rank is secured is n.

Now consider the second rank. There are (n – 1) students left and the second rank can be secured by any one of them. Thus the different possibilities are (n – 1) ways. Now, applying fundamental principle, we can see that the first two ranks can be secured in n (n – 1) ways by these n students.

Theorem : The number of permutations of n things when r are chosen at a time

$^nP_r = n\,(n-1)\,(n-2)\ldots(n-r+1)$

where the product has exactly r factors.

Results

1. Number of permutations of n different things taken all n things at a time is given by nPn = n!
2. ^nP_r using factorial notation.

$$^nP_r = \frac{n!}{(n-r)!}$$

3. Justification for 0! = 1. Now applying r = n in the formula for ^nP_r, we get
$^nP_n = n!/(n-n)! = n!/0!$

Example 1: Evaluate each of $^5P_3$, $^{10}P_2$, $^{11}P_5$,

Solution: $^5P_3 = 5\times4\times(5-3+1) = 5 \times 4 \times 3 = 60$,

$^{10}P_2 = 10 \times \ldots \times (10-2+1) = 10 \times 9 = 90$,

$^{11}P_5 = 11!/(11-5)! = 11 \times 10 \times 9 \times 8 \times 7 \times 6!/6! = 11 \times 10 \times 9 \times 8 \times 7 = 55440$.

Example 2: How many three letters words can be formed using the letters of the words

(a) SQUARE and (b) HEXAGON?

(Any arrangement of letters is called a word even though it may or may not have any meaning or pronunciation).

Solution:

(a) Since the word 'SQUARE' consists of 6 different letters, the number of permutations of choosing 3 letters out of six equals $^6P_3 = 6 \times 5 \times 4 = 120$.

(b) Since the word 'HEXAGON' contains 7 different letters, the number of permutations is $^7P_3 = 7 \times 6 \times 5 = 210$.

Example 3: In how many different ways can five persons stand in a line for a group photograph?

Solution: Here we know that the order is important. Hence, this is the number of permutations of five things taken all at a time. Therefore, this equals

$^5P_5 = 5! = 5 \times 4 \times 3 \times 2 \times 1 = 120$ ways.

Example 4: First, second and third prizes are to be awarded at an engineering fair in which 13 exhibits have been entered. In how many different ways can the prizes be awarded?

Solution: Here again, order of selection is important and repetitions are not meaningful as no exhibit can receive more than one prize. Hence, the answer is the number of permutations of 13 things taken three at a time. Therefore, we find ${}^{13}P_3$ = 13!/10! = 13×12×11 = 1,716 ways.

Example 5: In how many different ways can 3 students be associated with 4 chartered accountants, assuming that each chartered accountant can take at most one student?

Solution: This equals the number of permutations of choosing 3 persons out of 4. Hence, the answer is

${}^{4}P_3 = 4×3×2 = 24$.

Example 6: If six times the number permutations of n things taken 3 at a time is equal to seven times the number of permutations of (n – 1) things taken 3 at a time, find n.

Solution: We are given that $6 \times {}^{n}P_3 = 7 \times {}^{n-1}P_3$ and we have to solve this equality to find the value of n. Therefore,

$$6\frac{\lfloor n}{\lfloor n-3} = 7\frac{\lfloor n-1}{\lfloor n-4}$$

or, 6 n (n – 1) (n – 2) = 7 (n – 1) (n – 2) (n – 3)

or, 6 n = 7 (n – 3)

or, 6 n = 7n – 21 or, n = 21

Therefore, the value of n equals 21.

Example 7: Compute the sum of 4 digit numbers which can be formed with the four digits 1, 3, 5, 7, if each digit is used only once in each arrangement.

Solution: The number of arrangements of 4 different digits taken 4 at a time is given by ${}^{4}P_4$ = 4! = 24. All the four digits will occur equal number of times at each of the positions, namely ones, tens, hundreds, thousands.

Thus, each digit will occur 24 / 4 = 6 times in each of the positions. The sum of digits in one's position will be 6 × (1 + 3 + 5 + 7) = 96. Similar is the case in ten's, hundred's and thousand's places. Therefore, the sum will be 96 + 96 × 10 + 96 × 100 + 96 × 1000 = 1,06,656.

Example 8: Find n if nP3 = 60.

Solution: $${}^{n}P_3 = \frac{n!}{(n-3)!} = 60 \text{ (given)}$$

i.e., n (n–1) (n–2) = 60 = 5 × 4 × 3

Therefore, n = 5.

Solution: We know $${}^{n}p_r = \frac{n!}{(n-r)!}$$

$$\therefore {}^{56}P_r{+6} = \frac{56!}{\{56-(r+6)\}!} = \frac{56!}{(-r)!}$$

Similarly, $${}^{54}P_{r+3} \quad \frac{54!}{\{54-(r+3)\}!} = \frac{54!}{(51-r)}$$

Thus, $$\frac{{}^{56}P_{r+6}}{{}^{54}P_{r+3}} = \frac{56!}{(50-r!)} x \frac{(51-r)!}{54!}$$

$$\frac{56 \text{ x } 55 \text{ x } 54!}{(50-r)!} x \frac{(51-r)(50-r)!}{54!} = \frac{56 \text{ x } 55 \text{ x } (51-r)}{1}$$

But we are given the ratio as 30800 : 1 ; therefore

$$\frac{56 \times 55 \times (51 - r)}{1} = \frac{30,800}{1}$$

or, $(51-r) = \frac{30,800}{56 \times 55}$ $\therefore r = 41$

Example 10: Prove the following

$(n + 1)! - n! = \Rightarrow n.n!$

Solution: By applying the simple properties of factorial, we have $(n +1)! - n! = (n+1)\,n! - n! = n!.(n+1-1) = n.n!$

Example 11: In how many different ways can a club with 10 members select a President, Secretary and Treasurer, if no member can hold two offices and each member is eligible for any office?

Solution: The answer is the number of permutations of 10 persons chosen three at a time. Therefore, it is ${}^{10}p_3 = 10\times9\times8=720$.

Example 12: When Jhon arrives in New York, he has eight shops to see, but he has time only to visit six of them. In how many different ways can he arrange his schedule in New York?

Solution: He can arrange his schedule in ${}^{8}P_6 = 8\times 7 \times 6 \times 5 \times 4 \times 3 = 20,160$ ways.

Example 13: When Dr. Ram arrives in his dispensary, he finds 12 patients waiting to see him. If he can see only one patient at a time, find the number of ways, he can schedule his patients (a) if they all want their turn, and (b) if 3 leave in disgust before Dr. Ram gets around to seeing them.

Solution: (a) There are 12 patients and all 12 wait to see the doctor. Therefore the number of ways = ${}^{12}P_{12} = 12! = 479,001,600$

(b) There are $12-3 = 9$ patients. They can be seen ${}^{12}P_9 = 79,833,600$ ways.

Circular Permutations

So for we have discussed arrangements of objects or things in a row which may be termed as linear permutation. But if we arrange the objects along a closed curve viz., a circle, the permutations are known as circular permutations.

The number of circular permutations of n different things chosen at a time is $(n-1)!$.

Example 1: In how many ways can 4 persons sit at a round table for a group discussions?

Solution: The answer can be get from the formula for circular permutations. The answer is $(4-1)! = 3! = 6$ ways.

NOTE : These arrangements are such that every person has got the same two neighbours. The only change is that right side neighbour and vice-versa.

Thus the number of ways of arranging n persons along a round table so that no person has the same two neighbours is = $\frac{1}{2}\,\underline{|n-1}$

Similarly, in forming a necklace or a garland there is no distinction between a clockwise and anti clockwise direction because we can simply turn it over so that clockwise becomes anti clockwise and vice versa. **Hence, the number of necklaces formed with n beads of different colours=** $\frac{1}{2}\,\underline{|n-1}$

Permutation with Restrictions

In many arrangements there may be number of restrictions. in such cases, we are to arrange or select the objects or persons as per the restrictions imposed. The total number of arrangements in all cases, can be found out by the application of fundamental principle.

Theorem 1. Number of permutations of n distinct objects taken r at a time when a particular object is not taken in any arrangement is $^{n-1}p_r$.

Theorem 2. Number of permutations of r objects out of n distinct objects when a particular object is always included in any arrangement is r. n⊠1 p

Example 1: How many arrangements can be made out of the letters of the word `DRAUGHT', the vowels never beings separated?

Solution: The word `DRAUGHT' consists of 7 letters of which 5 are consonants and two are vowels. In the arrangement we are to take all the 7 letters but the restriction is that the two vowels should not be separated.

We can view the two vowels as one letter. The two vowels A and U in this one letter can be arranged in 2! = 2 ways. (i) AU or (ii) UA. Further, we can arrange the six letters : 5 consonants and one letter compound letter consisting of two vowels. The total number of ways of arranging them is $^6P_6 = 6! = 720$ ways.

Hence, by the fundamental principle, the total number of arrangements of the letters of the word DRAUGHT, the vowels never being separated = 2 × 720 = 1440 ways.

Example 2: Show that the number of ways in which n books can be arranged on a shelf so that two particular books are not together. The number is (n−2).(n−1)!

Solution: We first find the total number of arrangements in which all n books can be arranged on the shelf without any restriction. The number is, $^nP_n = n!$ ..... (1)

Then we find the total number of arrangements in which the two particular books are together.

The books can be together in $^2P_2 = 2! = 2$ ways. Now we consider those two books which are kept together as one composite book and with the rest of the (n−2) books from (n−1) books which are to be arranged on the shelf; the number of arrangements = $^{n-1}P_{n-1} = (n-1)$!. Hence by the Fundamental Principle, the total number of arrangements on which the two particular books are together equals = 2 × (n−1)! ....(2)

the required number of arrangements of n books on a shelf so that two particular books are not together

$$= (1) - (2)$$
$$= n! - 2 \times (n-1)!$$
$$= n.(n-1)! - 2 \,.\, (n-1)!$$
$$= (n-1)! \,.\, (n-2)$$

Example 3: There are 6 books on Economics, 3 on Mathematics and 2 on Accountancy. In how many ways can these be placed on a shelf if the books on the same subject are to be together?

Solution: Consider one such arrangement. 6 Economics books can be arranged among themselves in 6! Ways, 3 Mathematics books can be arranged in 3! Ways and the 2 books on Accountancy can be arranged in 2! ways. Consider the books on each subject as one unit. Now there are three units. These 3 units can be arranged in 3! Ways.

Total number of arrangements $= 3! \times 6! \times 3! \times 2!$

$= 51{,}840.$

Example 4: How many different numbers can be formed by using any three out of five digits 1, 2, 3, 4, 5, no digit being repeated in any number?

How many of these will (i) begin with a specified digit? (ii) begin with a specified digit and end with another specified digit?

Solution: Here we have 5 different digits and we have to find out the number of permutations of them 3 at a time. Required number is $^3P_3 = 5.4.3 = 60$.

(i) If the numbers begin with a specified digit, then we have to find

the number of Permutations of the remaining 4 digits taken 2 at a time. Thus, desire number is = 4.3 = 12.

(ii) Here two digits are fixed; first and last; hence, we are left with the choice of finding the number of permutations of 3 things taken one at a time i.e., =3.

Example 5: How many four digit numbers can be formed out of the digits 1,2,3,5,7,8,9, if no digit is repeated in any number? How many of these will be greater than 3000?

Solution: We are given 7 different digits and a four-digit number is to be formed using any 4 of these digits. This is same as the permutations of 7 different things taken 4 at a time.

Hence, the number of four-digit numbers that can be formed = $^7P_4$ = 7 × 6 × 5 × 4 × = 840 ways.

Next, there is the restriction that the four-digit numbers so formed must be greater than 3,000. Thus, it will be so if the first digit-that in the thousand's position, is one of the five digits 3, 5, 7, 8, 9. Hence, the first digit can be chosen in 5 different ways; when this is done, the rest of the 3 digits are to be chosen from the rest of the 6 digits without any restriction and this can be done in $^6P_3$ ways.

Hence, by the Fundamental principle, we have the number of four-digit numbers greater than 3,000 that can be formed by taking 4 digits from the given 7 digits = 5 × $^6P_3$ = 5 × 6 × 5 × 4 = 5 × 120 = 600.

Example 6: Find the total number of numbers greater than 2000 that can be formed with the digits 1, 2, 3, 4, 5 no digit being repeated in any number.

Solution: All the 5 digit numbers that can be formed with the given 5 digits are greater than 2000. This can be done in

$$^5P_5 = 5! = 120 \text{ ways} \qquad (1)$$

The four digited numbers that can be formed with any four of the given 5 digits are greater than 2000 if the first digit, i.e.,the digit in the thousand's position is one of the four digits 2, 3, 4, 5. this can be done in $^4P_1$ = 4 ways. When this is done, the rest of the 3 digits are to be chosen from the rest of 5–1 = 4 digits. This can be done in $^4P_3$ = 4 × 3 × 2 = 24 ways.

Therefore, by the Fundamental principle, the number of four-digit numbers greater than 2000

= 4 × 24 = 96 .... (2)

Adding (1) and (2), we find the total number greater than 2000 to be 120 + 96 = 216.

Example 7: There are 6 students of whom 2 are Indians, 2 Americans, and the remaining 2 are Russians. They have to stand in a row for a photograph so that the two Indians are together, the two Americans are together and so also the two Russians. Find the number of ways in which they can do so.

Solution: The two Indians can stand together in $^2P_2$ = 2! = 2 ways. So is the case with the two Americans and the two Russians.

Now these 3 groups of 2 each can stand in a row in = 3 x 2 = 6 ways. Hence by the generalized fundamental principle, the total number of ways in which they can stand for a photograph under given conditions is

6 × 2 × 2 × 2 = 48

Example 8: A family of 4 brothers and three sisters is to be arranged for a photograph in one row. In how many ways can they be seated if (i) all the sisters sit together, (ii) no two sisters sit together?

Solution:

(i) Consider the sisters as one unit and each brother as one unit. 4 brothers and 3 sisters make 5 units which can be arranged in 5! ways. Again 3 sisters may be arranged amongst themselves in 3! Ways

Therefore, total number of ways in which all the sisters sit together = 5!×3! = 720 ways.

(ii) In this case, each sister must sit on each side of the brothers. There are 5 such positions as indicated below by upward arrows :

↑ B1 ↑ B2 ↑ B3 ↑ B4 ↑

4 brothers may be arranged among themselves in 4! ways. For each of these arrangements 3 sisters can sit in the 5 places in $^5P_3$ ways.

Thus the total number of ways = $^5P_3$ × 4! = 60 × 24 = 1,440

Example 9: In how many ways can 8 persons be seated at a round table? In how many cases will 2 particular persons sit together?

Solution: This is in form of circular permutation. Hence the number of ways in which eight persons can be seated at a round table is $(n-1)! = (8-1)! = 7! = 5040$ ways.

Consider the two particular persons as one person. Then the group of 8 persons becomes a group of 7 (with the restriction that the two particular persons be together) and seven persons can be arranged in a circular in 6! Ways.

Hence, by the fundamental principle, we have, the total number of cases in which 2 particular persons sit together in a circular arrangement of 8 persons $= 2! \times 6! = 2 \times 6 \times 5 \times 4 \times 3 \times 2 \times 1$

$= 1,440.$

Example 10: Six boys and five girls are to be seated for a photograph in a row such that no two girls sit together and no two boys sit together. Find the number of ways in which this can be done.

Solution: Suppose that we have 11 chairs in a row and we want the 6 boys and 5 girls to be seated such that no two girls and no two boys are together. If we number the chairs from left to right, the arrangement will be possible if and only if boys occupy the odd places and girls occupy the even places in the row. The six odd places from 1 to 11 may filled in by 6 boys in ${}^6P_6$ ways. Similarly, the five even places from 2 to 10 may be filled in by 5 girls in ${}^5P_5$ ways.

Hence, by the fundamental principle, the total number of required arrangements $= {}^6P_6 \times {}^5P_5 = 6! \times 5! = 720 \times 120 = 86,400.$

Combinations

We have studied about permutations in the earlier section. There we have said that while arranging, we should pay due regard to order. There are situations in which order is not important. For example, consider selection of 5 clerks from 20 applicants. We will not be concerned about the order in which they are selected. In this situation, how to find the number of ways of selection? The idea of combination applies here.

Definition: The number of ways in which smaller or equal number of things are arranged or selected from a collection of things where the order of selection or arrangement is not important, are called combinations.

The selection of a poker hand which is a combination of five cards selected from 52 cards is an example of combination of 5 things out of 52 things.

Number of combinations of n different things taken r at a time. (denoted by ${}^nC_r$ C(n,r), $C_{n,r}$)

Let ${}^nC_r$ denote the required number of combinations. Consider any one of those combinations. It will contain r things. Here we are not paying attention to order of selection. Had we paid attention to this, we will have permutations or r items taken r at a time. In other words, every combination of r things will have ${}^rP_r$ permutations amongst them. Therefore, ${}^nC_r$ combinations will give rise to ${}^nC_r \cdot {}^rP_r$ permutations of r things selected from n things. From the earlier section, we can say that ${}^nC_r \cdot {}^rP_r = {}^nP_r$ as ${}^nP_r$ denotes the number of permutations of r things chosen out of n things.

Since, ${}^nC_r \cdot {}^rP_r = {}^nP_r$,

${}^nC_r = {}^nP_r / {}^rP_r$ $= n!/(n-r)! ⊠ r!/(r-r)!$

$= n!/(n-r)! \times 0!/r!$

$= n!/r!(n-r)!$

$\therefore {}^nC_r = n!r!(n-r)!$

Remarks: Using the above formula, we get

(i) ${}^nC_o = n!/0!(n-0)! = n!/n! = 1.$ [As 0! = 1]

${}^nC_n = n!/n!(n-n)! = n!/n!0! = 1$ [Applying the formula for ${}^nC_r$ with r = n]

Example 1: Find the number of different poker hands in a pack of 52 playing cards.

Solution: This is the number of combinations of 52 cards taken five at a time. Now applying the formula,

$$^{52}C_5 = 52!/5!\,(52-5)! = 52!/5!\,47! = \frac{52 \times 51 \times 50 \times 49 \times 48 \times 47!}{5 \times 4 \times 3 \times 2 \times 1 \times 47!}$$

Example 2: Let S be the collection of eight points in the plane with no three points on the straight line. Find the number of triangles that have points of S as vertices.

Solution: Every choice of three points out of S determines a unique triangle. The order of the points selected is unimportant as whatever be the order, we will get the same triangle. Hence, the desired number is the number of combinations of eight things taken three at a time. Therefore, we get

$^8C_3 = 8!/3!5! = 8\times7\times6/3\times2\times1 = 56$ choices.

Example 3: A committee is to be formed of 3 persons out of 12. Find the number of ways of forming such a committee.

Solution: We want to find out the number of combinations of 12 things taken 3 at a time and this is given by

$^{12}C_3 = 12!/3!(12-3)!$ [by the definition of ^nC_r]

$= 12!/3!9! = 12\times11\times10\times9!/3!9! = 12\times11\times10/3\times2 = 220$

Example 4: A committee of 7 members is to be chosen from 6 Chartered Accountants, 4 Economists and 5 Cost Accountants. In how many ways can this be done if in the committee, there must be at least one member from each group and at least 3 Chartered Accountants?

Solution: The various methods of selecting the persons from the various groups are shown below:

Committee of 7 members			
	C.A.s	**Economists**	**Cost Accountants**
Method 1	3	2	2
Method 2	4	2	1
Method 3	4	1	2
Method 4	5	1	1
Method 5	3	3	1
Method 6	3	1	3

Number of ways of choosing the committee members by

Method 1 = $^6C_3 \times {}^4C_2 \times {}^5C_2 = \frac{6x5x4}{3x2x1} \text{ x } \frac{4x3}{2x1} \text{ x } \frac{5x4}{2x1}$

Method 2 = $^6C_4 \times {}^4C_2 \times {}^5C_1 = \frac{6x5}{2x1} \text{ x } \frac{4x3}{2x1} \text{ x } \frac{5}{1}$

Method 3 = $^6C_4 \times {}^4C_1 \times {}^5C_2 = \frac{6x5}{2x1} \text{ x } 4 \text{ x } \frac{5x4}{2x1}$

Method 4 = $^6C_5 \times {}^4C_1 \times {}^5C_1 = 6\times4\times5 = 120$.

Method 5 = $^6C_3 \times {}^4C_3 \times {}^5C_1 = \frac{6x5x4}{3x2x1} \text{ x } \frac{4x3x2}{3x2x1} \text{ x } 5 = 20\times4\times5 = 400$

Method 6 = $^6C_3 \times {}^4C_1 \times {}^5C_3 = \frac{6x5x4}{3x2x1} \text{ x } 4 \text{ x } \frac{5x4}{2x1} = 20\times4\times10 = 800$.

Therefore, total number of ways = 1,200 + 450 + 600 + 120 + 400 + 800 = 3,570

Example 5: A person has 12 friends of whom 8 are relatives. In how many ways can he invite 7 guests such that 5 of them are relatives?

Solution: Of the 12 friends, 8 are relatives and the remaining 4 are not relatives. He has to invite 5 relatives and 2 friends as his guests. 5 relatives can be chosen out of 8 in $^8C_5$ ways; 2 friends can be chosen out of 4 in $^4C_2$ ways.

Hence, by the fundamental principle, the number of ways in which he can invite 7 guests such that 5 of them are relatives and 2 are friends.

$= {}^8C_5 \times {}^4C_2$

$= \{8! / 5!\,(8-5)!\} \times \{4! / 2!\,(4-2)!\} = [(8 \times 7 \times 6 \times 5!)/5!\ X\ 3!] \times \dfrac{4x3x2x!}{2!2!} = 8x7x6$

$= 336.$

Example 6: A Company wishes to simultaneously promote two of its 6 department heads to assistant managers. In how many ways these promotions can take place?

Solution: This is a problem of combination. Hence, the promotions can be done in

$^6C_2 = 6\times5 / 2 = 15$ ways

Example 7: A building contractor needs three helpers and ten men apply. In how many ways can these selections take place?

Solution: There is no regard for order in this problem. Hence, the contractor can select in any of

10C3 ways i.e.,

$(10 \times 9 \times 8) / (3 \times 2 \times 1) = 120$ ways.

Example 8: In each case, find n:

Solution: (a) $4.\,{}^nC_2 = {}^{n+2}C_3$ (b) $^{n+2}C_n = 45.$

(a) We are given that $4.\,{}^nC_2 = {}^{n+2}C_3$. Now applying the formula,

$$4 \times \frac{n!}{2!(n-2)!} = \frac{(n+2)!}{3!(n+2-3)!}$$

$$\text{or, } \frac{4 \times n.\,(n-1)(n-2)!}{2!(n-2)!} = \frac{(n+2)(n+1).n.(n-1)!}{3!(n-1)!}$$

$4n(n-1)/2 = (n+2)(n+1)n/3!$

or, $4n(n-1) / 2 = (n+2)(n+1)n / 3\times2\times1$

or, $12(n-1)=(n+2)(n+1)$

or, $12n-12 = n2 + 3n + 2$

or, $n2 - 9n + 14 = 0.$

or, $n2 - 2n - 7n + 14 = 0.$

or, $(n-2)(n-7) = 0$

$\therefore n=2$ or 7.

(b) We are given that n+2Cn = 45. Applying the formula,

$(n+2)!/\{n!(n+2-n)!\} = 45$

or, $(n+2)(n+1)\,n! / n!\,2! = 45$

or, $(n+1)(n+2) = 45\times2! = 90$

or, $n^2+3n-88 = 0$

or, $n^2+11n-8n-88 = 0$

or, $(n+11)(n-8) = 0$

Thus, n equals either – 11 or 8. But negative value is not possible. Therefore we conclude that n=8.

Example 9: A box contains 7 red, 6 white and 4 blue balls. How many selections of three balls can be made so that (a) all three are red (b) none is red (c) one is of each colour?

Solution: (a) All three balls will be of red colour if they are taken out of 7 red balls and this can be done in

$^7C_3 = 7! / 3!(7-3)!$

$= 7! / 3!4! = 7\times6\times5\times4! / (3\times2\times4!) = 7\times6\times5 / (3\times2) = 35$ ways

Hence, 35 selections (groups) will be there such that all three balls are red.

(b) None of the three will be red if these are chosen from (6 white and 4 blue balls) 10 balls and this can be done in

$^{10}C_3 = 10!/\{3!(10-3)!\} = 10! / 3!7!$

$= 10\times9\times8\times7! / (3\times2\times1\times7!) = 10\times9\times8 /(3\times2) = 120$ ways.

Hence, the selections (or groups) of three such that none is a red ball are 120 in number.

One red ball can be chosen from 7 balls in $^7C_1 = 7$ ways. One white ball can be chosen from 6 white balls in $^6C_1$ ways. One blue ball can be chosen from 4 blue balls in $^4C_1 = 4$ ways. Hence, by generalized fundamental principle, the number of groups of three balls such that one is of each colour = $7\times6\times4 = 168$ ways.

Example 10: If $^{10}P_r = 6{,}04{,}800$ and $^{10}C_r = 120$; find the value of r,

Solution: We know that $^nC_r \cdot {}^rP_r = {}^nP_r$. We will us this equality to find r.

$^{10}P_r = {}^{10}C_r .r!$

or, $6{,}04{,}800 = 120 \times r!$

or, $r! = 6{,}04{,}800 \div 120 = 5{,}040$

But $r! = 5040 = 7\times6\times4\times3\times2\times1 = 7!$

Therefore, r=7.

Properties of nCr

1. $^nC_r = {}^nC_{n-r}$

   We have $^nC_r = n! / \{r!(n-r)!\}$ and $^nC_{n-r} = n! / [(n-r)! \{n-(n-r)\}!] = n! / \{(n-r)!(n-n+r)!\}$

   Thus $^nC_{n-r} = n! / \{(n-r)! (n-n+r)!\} = n! / \{(n-r)!r!\} = {}^nC_r$

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

   By definition,

   $^nC_{r-1} + {}^nC_r = n! / \{(r-1)! (n-r+1)!\} + n! / \{r!(n-r)!\}$

   But $r! = r\times(r-1)!$ and $(n-r+1)! = (n-r+1) \times (n-r)!$. Substituting these in above, we get

   $$^{n+1}C_{r-1} + {}^nC_r = n! \left\{ \frac{1}{r-1)!(n-r+1)(n-r)!} + \frac{1}{r(r-1)!(n-r)!} \right\}$$

   $= \{n! / (r-1)! (n-r)!\} \{(1 / n-r+1) + (1/r)\}$

   $= \{n! / (r-1)! (n-r)!\} \{(r+n-r+1) / r(n-r+1)\}$

   $= (n+1) n! / \{r . (r-1)! (n-r)! . (n-r+1)\}$

3. $^NC_o = n!/\{0! (n-0)!\} = n! / n! =1.$

4. $^nC_n = n!/\{n! (n-n)!\} = n! / n! . 0! = 1.$

   $= (n+1)! / \{r!(n+1-r)!\} = {}^{n+1}C_r$

Note

(a) ^nC_r has a meaning only when r and n are integers $0 \le r \le n$ and $^nC_{n-r}$ has a meaning only when $0 \le n - r \le n$.

(b) ^nC_r and $^nC_{n-r}$ are called complementary combinations, for if we form a group of r things out of n different things, (n−r) remaining things which are not included in this group form another group of rejected things. The number of groups of n different things, taken r at a time should be equal to the number of groups of n different things taken (n−r) at a time.

Example 11: Find r if ${}^{18}C_r = {}^{18}C_{r+2}$

Solution: As ${}^{n}C_r = {}^{n}C_{n-r}$, we have ${}^{18}C_r = {}^{18}C_{18-r}$

But it is given, ${}^{18}C_r = {}^{18}C_{r+2}$

$\therefore {}^{18}C_{18-r} = {}^{18}C_{r+2}$

or, 18 – r = r+2

Solving, we get

2r = 18 – 2 = 16 i.e., r=8.

Example 12: Prove that

${}^{n}C_r = {}^{n-2}C_{r-2} + 2\,{}^{n-2}C_{r-1} + {}^{n-2}C_r$

Solution: R.H.S $= {}^{n-2}C_{r-2} + {}^{n-2}C_{r-1} + {}^{n-2}C_{r-1} + {}^{n-2}C_r$

$= {}^{n-1}C_{r-1} + {}^{n-1}C_r$ [using Property 2 listed earlier]

$= {}^{(n-1)+1}C_r$ [using Property 2 again]

$= {}^{n}C_r$ = L.H.S.

Hence, the result

Example 13: If ${}^{28}C_{2r} : {}^{24}C_{2r-4} = 225 : 11$, find r.

Solution: We have ${}^{n}C_r = n! / \{r!(n-r)!\}$ Now, substituting for n and r, we get

${}^{28}C_{2r} = 28! / \{(2r)!(28-2r)!\}$,

${}^{24}C_{2r-4} = 24! / [(2r-4)!\,\{24-(2r-4)\}!] = 24! / \{(2r-4)!(28-2r)!\}$

We are given that ${}^{28}C_{2r} : {}^{24}C_{2r-4} = 225 : 11$. Now we calculate,

$$= \frac{{}^{28}C_{2r}}{{}^{24}C_{2r-4}} = \frac{28!}{(2r)!(28-2r)!} \times \frac{(2r-4)!(28-2r)!}{24!}$$

$$= \frac{28 \times 27 \times 26 \times 25 \times 24!}{(2r)(2r-1)(2r-2(2r-3)(2r-4)!(28-2r)!} \times \frac{(2r-4)!(28-2r)!}{24}$$

$$= \frac{28 \times 27 \times 26 \times 25}{(2r)(2r-1)(2r-2(2r-3)} \times \frac{225}{11}$$

or, $(2r)\,(2r-1)\,(2r-2)\,(2r-3) = \dfrac{11 \times 28 \times 27 \times 26 \times 25}{225}$

$= 11 \times 28 \times 3 \times 26$

$= 11 \times 7 \times 4 \times 3 \times 13 \times 2$

$= 11 \times 12 \times 13 \times 14$

$= 14 \times 13 \times 12 \times 11$

$\therefore$ 2r= 14 i.e., r = 7

Example 14: Find x if ${}^{12}C_5 + 2\,{}^{12}C_4 + {}^{12}C_3 = {}^{14}C_x$

Solution: L.H.S $= {}^{12}C_5 + 2\,{}^{12}C_4 + {}^{12}C_3$

$= {}^{12}C_5 + {}^{12}C_4 + {}^{12}C_4 + {}^{12}C_3$

$= {}^{13}C_5 + {}^{13}C_4$

$= {}^{14}C_5$

Also ${}^{n}C_r = {}^{n}C_{n-r}$. Therefore ${}^{14}C_5 = {}^{14}C_{14-5} = {}^{14}C_9$

Hence, L.H.S $= {}^{14}C_5 = {}^{14}C_9 = {}^{14}C_x$ = R.H.S by the given equality

This implies, either x = 5 or x = 9.

Example 15: Prove by reasoning that

(i) ${}^{n+1}C_r = {}^nC_r + {}^nC_{r-1}$

(ii) ${}^nP_r = {}^{n-1}P_r + r\,{}^{n-1}P_{r-1}$

Solution: (i) ${}^{n+1}C_r$ stands for the number of combinations of (n+1) things taken r at a time. As a specified thing can either be included in any combination or excluded from it, the total number of combinations which can be combinations or (n+1) things taken r at a time is the sum of :

(a) combinations of (n+1) things taken r at time in which one specified thing is always included and

(b) the number of combinations of (n+1) things taken r at time from which the specified thing is always excluded.

Now, in case (a), when a specified thing is always included, we have to find the number of ways of selecting the remaining (r–1) things out of the remaining n things which is ${}^nC_{r-1}$.

Again, in case (b), since that specified thing is always excluded, we have to find the number of ways of selecting r things out of the remaining n things, which is ${}^nC_r$.

Thus, ${}^{n+1}C_r = {}^nC_{r-1} + {}^nC_r$

(i) We divide nPr i.e., the number of permutations of n things take r at a time into two groups:

(a) those which contain a specified thing

(b) those which do not contain a specified thing.

In (a) we fix the particular thing in any one of the r places which can be done in r ways and then fill up the remaining (r–1) places out of (n–1) things which give rise to ${}^{n-1}P_{r-1}$ ways. Thus, the number of permutations in case (a) = r × ${}^{n-1}P_{r-1}$

In case (b), one thing is to be excluded; therefore, r places are to be filled out of (n–1) things. Therefore, number of permutations = ${}^{n-1}P_r$

Thus, total number of permutations = ${}^{n-1}P_r + r.\,{}^{n-1}P_{r-1}$

i.e., ${}^nP_r = {}^{n-1}P_r + r.\,{}^{n-1}P_{r-1}$

Standard Results

We now proceed to examine some standard results in permutations and combinations. These results have special application and hence are dealt with separately.

I. Permutations when some of the things are alike, taken all at a time

The number of ways p in which n things may be arranged among themselves, taking them all at a time, when n1 of the things are exactly alike of one kind, n2 of the things are exactly alike of another kind, n3 of the things are exactly alike of the third kind, and the rest all are different is given by,

$$P = \frac{n!}{n_1!n_2!n_3!}$$

II. Permutations when each thing may be repeated once, twice, ...upto r times in any arrangement.

Result: The number of permutations of n things taken r at time when each thing may be repeated r times in any arrangement is nr.

Proof: There are n different things and any of these may be chosen as the first thing. Hence, there are n ways of choosing the first thing.

When this is done, we are again left with n different things and any of these may be chosen as the second (as the same thing can be chosen again.)

Hence, by the fundamental principle, the two things can be chosen in n × n = n^2 number of ways.

Proceeding in this manner, and noting that at each stage we are to chose a thing from n different things, the total number of ways in which r things can be chosen is obviously equal to n × n ×

.........to r terms = n^r.

III. Combinations of n different things taking some or all of n things at a time

Result : The total number of ways in which it is possible to form groups by taking some or all of n things (2n −1).

In symbols, $\sum_{r=1}^{n} {}^nC_r = 2^n - 1$

Proof: Each of the n different things may be dealt with in two ways; it may either be taken or left. Hence, by the generalised fundamental principle, the total number of ways of dealing with n things:

2 × 2 × 2×   2, n times i.e., 2^n

But this include the case in which all the things are left, and therefore, rejecting this case, the total number of ways of forming a group by taking some or all of n different things is $2^n - 1$.

IV. Combinations of n things taken some or all at a time when n1 of the things are alike of one kind, n_2 of the things are alike of another kind n3 of the things are alike of a third kind. etc.

Result : The total, number of ways in which it is possible to make groups by taking some or all out of n ($=n_1 + n_2 +$ n3 +...) things, where n1 things are alike of one kind and so on, is given by

$\{(n_1 + 1)(n_2 + 1)(n_3 + 1)...\} -1$

Proof : The n_1 things all alike of one kind may be dealt with in $(n_1 + 1)$ ways. We may take 0, 1, 2,....n, of them. Similarly n_2 things all alike of a second kind may be dealt with in $(n_2 +1)$ ways and n_3 things all alike of a third kind may de dealt with in $(n_3 +1)$ ways.

Proceeding in this way and using the generalised fundamental principle, the total number of ways of dealing with all n $(= n_1 + n_2 + n_3 +...)$ things, where n_1, things are alike of one kind and so on, is given by

$(n_1 + 1)(n_2 + 1)(n_3 + 1)...$

But this includes the case in which none of the things are taken. Hence, rejecting this case, total number of ways is $\{(n_1 + 1)(n_2 + 1)(n_3 + 1)...\} -1\}$

V. The notion of Independence in Combinations

Many applications of Combinations involve the selection of subsets from two or more independent sets of objects or things. If the combination of a subset having r_1 objects form a set having n_1 objects does not affect the combination of a subset having r_2 objects from a different set having n_2 objects, we call the combinations as being independent. Whenever such combinations are independent, any subset of the first set of objects can be combined with each subset of the second set of the object to form a bigger combination. The total number of such combinations can be found by applying the generalised fundamental principle.

Result: The combinations of selecting r_1 things from a set having n_1 objects and r_2 things from a set having n_2 objects where combination of r_1 things, r_2 things are independent is given by

$n_1C_{r1} \times {}^{n}{}_2C_{r2}$

Note: This result can be extended to more than two sets of objects by a similar reasoning.

Example 1: How many different permutations are possible from the letters of the word `CALCULUS'?

Solution: The word `CALCULUS' consists of 8 letters of which 2 are C and 2 are L, 2 are U and the rest are A and S. Hence, by result (I), the number of different permutations from the letters of the word `CALCULUS' taken all at a time

$= \frac{8!}{2!2!2!1!1!}$

$= \frac{8 \times 7 \times 6 \times 5 \times 4 \times 3 \times 2}{2 \times 2 \times 2} = 7 \times 6 \times 5 \times 4 \times 3 \times 2 = 5{,}040$

Example 2: In how many ways can 17 billiard balls be arranged, if 7 of them are black, 6 red and 4 white?

Solution: We have, the required number of different arrangements:

$\frac{17!}{7!6!4!}$ = 40,84,080 7! 6! 4!

Example 3: An examination paper with 10 questions consists of 6 questions in Algebra and 4 questions in Geometry. At least one question from each section is to be attempted. In how many ways can this be done?

Solution: A student must answer atleast one question from each section and he may answer all questions from each section.

Consider Section I : Algebra. There are 6 questions and he may answer a question or may not answer it. These are the two alternatives associated with each of the six questions. Hence, by the generalised fundamental principle, he can deal with two questions in 2 × 2 ....6 factors = 2^6 number of ways. But this includes the possibility of none of the question from Algebra being attempted. This cannot be so, as he must attempt at least one question from this section. Hence, excluding this case, the number of ways in which Section I can be dealt with is $(2^6 - 1)$.

Similarly, the number of ways in which Section II can be dealt with is $(2^4 - 1)$.

Hence, by the Fundamental Principle, the examination paper can be attempted in $(2^6 - 1)(2^4 - 1)$ number of ways.

Example 4: A man has 5 friends. In how many ways can he invite one or more of his friends to dinner?

Solution: By result, (III) of this section, as he has to select one or more of his 5 friends, he can do so in 25 – 1 = 31 ways.

Note : This can also be done in the way, outlines below. He can invite his friends one by one, in twos, in threes, etc. and hence the number of ways.

$= {}^5C_1 + {}^5C_2 + {}^5C_3 + {}^5C_4 + {}^5C_5$

$= 5 + 10 + 10 + 5 + 1 = 31$ ways.

Example 5: There are 7 men and 3 ladies. Find the number of ways in which a committee of 6 can be formed of them if the committee is to include atleast two ladies?

Solution: The committee of six must include at least 2 ladies, i.e., two or more ladies. As there are only 3 ladies, the following possibilities arise:

The committee of 6 consists of (i) 4 men and 2 ladies (ii) 3 men and 3 ladies. The number of ways for (i) $= {}^7C_4 \times {}^3C_2 = 35 \times 3 = 105$;

The number of ways for (ii) $= {}^7C_3 \times {}^3C_3 = 35 \times 1 = 35$.

Hence the total number of ways of forming a committee so as to include at least two ladies = 105

+ 35 = 140.

Example 6: Find the number of ways of selecting 4 letters from the word `EXAMINATION'.

Solution: There are 11 letters in the word of which A, I, N are repeated twice. Thus we have 11 letters of 8 different kinds (A, A), (I, I), (N, N), E, X, M, T, O. The group of four selected letters may take any of the following forms:

(i) Two alike and other two alike

(ii) Two alike and other two different

(iii) All four different

In case (i), the number of ways $= {}^3C_2 = 3$.

In case (ii), the number of ways $= {}^3C_1 \times {}^7C_2 = 3 \times 21 = 63$.

In case (iii), the number of ways $= {}^8C_4 = \dfrac{8 \times 7 \times 6 \times 5}{1 \times 2 \times 3 \times 4} = 70$

Hence, the required number of ways = 3 + 63 + 70 = 136 ways

Chapter - 6:

Sequence and Series: A.P. & G.P.

(A) Arithmetic Progression (A.P.):

Exp: 1, 4, 7, 10, 13,.........

General Form: $a, a+d, a+2d, a+3d, a+4d, \ldots\ldots, a+(n-1)d$

General Term: $T_n = a + (n-1)d$, where a = first term & d = common difference

Sum of the first n terms: $S_n = \frac{n}{2}\{2a + (n-1)d\} = \frac{n}{2}(a+l)$, where l = last term (T_n)

Sum of special series:

Σn = *sum of first n natural numbers* $= 1 + 2 + 3 + \ldots\ldots + n = \frac{n(n+1)}{2}$

Σn^2 = *sum of squares of first n natural numbers* $= 1^2 + 2^2 + 3^2 + \ldots\ldots + n^2 = \frac{n(n+1)(2n+1)}{6}$

Σn^3 = *sum of cubes of first n natural numbers* $= 1^3 + 2^3 + 3^3 + \ldots\ldots + n^3 = \left\{\frac{n(n+1)}{6}\right\}^2$

$\Sigma(2n)$ = *sum of first n even natural numbers* $= 2 + 4 + 6 + \ldots\ldots + 2n = n(n+1)$

$\Sigma(2n-1)$ = *sum of first n odd natural numbers* $= 1 + 3 + 5 + \ldots\ldots + (2n-1) = n^2$

Arithmetic Mean: AM between a and b, A.M.= $\frac{a+b}{2}$

Sum of n Arithmetic Means $= A_1 + A_2 + A_3 + \ldots\ldots + A_n = nA = n\left(\frac{a+b}{2}\right)$

Assumption of terms of A.P.:

.........., a - d, a, a + d,.......(when number of terms given are odd)

........, a -3d, a - d, a + d, a + 3d,.......(when number of terms given are even)

$d = t_n - t_{n-1}$, when t_n is given $\qquad t_n = S_n - S_{n-1}$, *when* S_n *is given*

Sequence and Series: A.P. & G.P.

(B) Geometric Progression (G.P.):

Exp: 1, 2, 4, 8, 16,.........

General Form: $a, ar, ar^2, ar^3, ar^4, \ldots\ldots, ar^{n-1}$

General Term: T_n ar^{n-1}, *where* a = first term & r = common ratio

Sum of the first *n* terms: $S_n = \frac{a(1-r^n)}{(1-r)}$, when $r < 1$ $\qquad$ $S_n = \frac{a(r^n - 1)}{(r - 1)}$, when $r > 1$

Sum of infinite terms: $S_\infty = \frac{a}{(1-r)}$, when $r < 1$

Geometric Mean: G.M. between a and b, G.M.= $\sqrt{ab}$

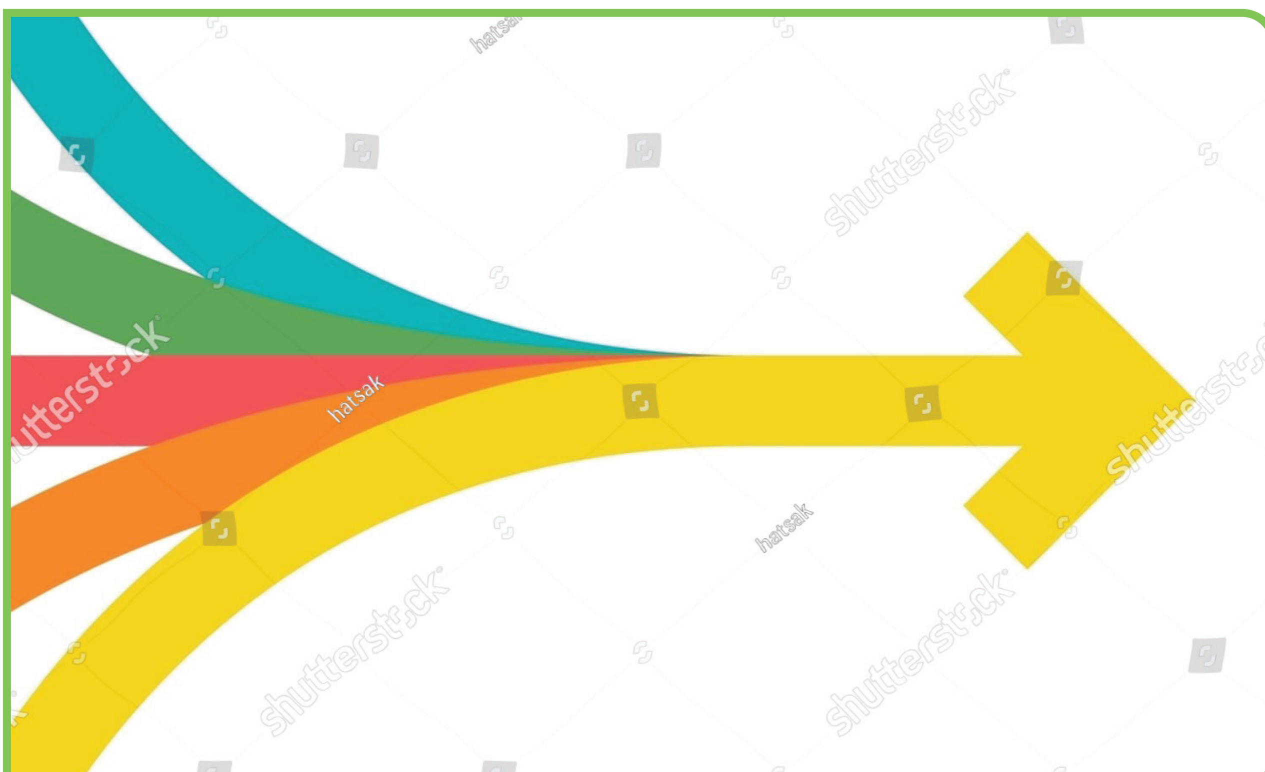
If a,b,c are in G.P., then $b^2 = ac$

Product of n Geometric Means = $G_1 G_2 G_3 \ldots\ldots .G_n = G^n = \left(\sqrt{ab}\right)^n$

Assumption of terms of G.P.:

$\ldots\ldots, \frac{a}{r}, a, ar, \ldots\ldots$ *(when number of terms given are odd)*

$\ldots\ldots, \frac{a}{r^3}, \frac{a}{r}, a, ar^3, \ldots\ldots$ *(when number of terms given are even)*



Chapter – 7:

Sets, Relations, and Functions

After reading this chapter, students will be able to understand:

- Understand the concept of set theory.
- Appreciate the basics of functions and relations.
- Understand the types of functions and relations.
- Solve problems relating to sets, functions, and relations.

In our mathematical language, everything in this universe, whether living or non-living, is called an object.

If we consider a collection of objects given in such a way that it is possible to tell beyond doubt whether a given object is in the collection under consideration or not, then such a collection of objects is called a well-defined collection of objects.

Sets

A set is defined to be a collection of well-defined distinct objects. This collection may be listed or described. Each object is called an element of the set. We usually denote sets by capital letters and their elements by small letters.

Example: A	=	{a, e, i, o, u}
B	=	{2, 4, 6, 8, 10}
C	=	{pqr, prq, qrp, rqp, qpr, rpq}
D	=	{1, 3, 5, 7, 9}
E	=	{1,2}

This form is called Roster or Braces form. In this form we make a list of the elements of the set and put it within braces { }.

Instead of listing we could describe them as follows :

A = the set of vowels in the alphabet

B = The set of even numbers between 2 and 10 both inclusive.

C = The set of all possible arrangements of the letters p, q and r

D = The set of odd digits between 1 and 9 both inclusive.

E = The set of roots of the equation $x^2 - 3x + 2 = 0$

Set B, D and E can also be described respectively as

B = $\{x : x = 2m$ and m being an integer lying in the interval $0 < m < 6\}$

D = $\{2x - 1 : 0 < x < 5$ and x is an integer$\}$

E = $\{x : x^2 - 3x + 2 = 0\}$

This form is called set-Builder or Algebraic form or Rule Method. This method of writing the set is called Property method. The symbol : or/reads 'such that'. In this method, we list the property or properties satisfied by the elements of the set.

We write, {x:x satisfies properties P}. This means, "the set of all those x such that x satisfies the properties P".

A set may contain either a finite or an infinite number of members or elements. When the number of members is very large or infinite it is obviously impractical or impossible to list them all. In such case.

we may write as :

N = The set of natural numbers = {1, 2, 3....}

W = The set of whole numbers = {0, 1, 2, 3,...) etc.

I. The members of a set are usually called elements. In A = {a, e, i, o, u}, a is an element and we write $a \in A$ i.e. a belongs to A. But 3 is not an element of B = {2, 4, 6, 8, 10} and we write $3 \notin B$. i.e. 3 does not belong to B.

II. If every element of a set P is also an element of set Q we say that P is a subset of Q. We write $P \subset Q$. Q is said to be a superset of P. For example $\{a, b\} \subset \{a, b, c\}$, $\{2, 4, 6, 8, 10\} \subset N$. If there exists even a single element in A, which is not in B then A is not a subset of B.

III. If P is a subset of Q but P is not equal to Q then P is called a proper subset of Q.

IV. "Φ" has no proper subset.

Illustration:

{3} is a proper subset of {2, 3, 5}. But {1, 2} is not a subset of {2, 3, 5}.

Thus if P = {1, 2} and Q = {1, 2 ,3} then P is a subset of Q but P is not equal to Q. So, P is a proper subset of Q.

To give completeness to the idea of a subset, we include the set itself and the empty set. The empty set is one which contains no element. The empty set is also known as **null or void** set usually denoted by { } or Greek letter Φ, to be read as phi. For example the set of prime numbers between 32 and 36 is a null set. The subsets of {1, 2, 3,} include {1, 2, 3}, {1, 2}, {1, 3}, {2, 3}, {1},{2}, {3} and { }.

A set containing n elements has 2^n subsets. Thus a set containing 3 elements has 2^3 (=8) subsets. A set containing n elements has 2^n –1 proper subsets. Thus a set containing 3 elements has 2^3 – 1 (=7) subsets. The proper subsets of {1,2,3} include {1, 2}, {1, 3}, {2, 3}, {1}, {2}, {3], { }.

Suppose we have two sets A and B. The intersection of these sets, written as

$A \cap B$ contains those elements which are in A and are also in B.

For example A = {2, 3, 6, 10, 15}, B = {3, 6, 15, 18, 21, 24} and C = { 2, 5, 7}, we have $A \cap B$ ={ 3, 6, 15}, $A \cap C$ = {2}, $B \cap C$ = Φ , where the intersection of B and C is empty set. So, we say B and C are disjoint sets since they have no common element. Otherwise sets are called overlapping or intersecting sets. The union of two sets, A and B, written as $A \cup B$ contain all these elements which are in either A or B or both.

So $A \cup B$ = {2, 3, 6, 10, 15, 18, 21, 24}

$A \cup C$ = {2, 3, 5, 6, 7, 10, 15}

A set which has at least one element is called non-empty set . Thus the set { 0 } is non-empty set. It has one element say 0.

Singleton Set:

A set containing one element is called Singleton Set. For example {1} is a singleton set, whose only member is 1.

Equal Set:

Two sets A & B are said to be equal, written as A = B if every element of A is in B and every element of B is in A.

Illustration:

If A = {2, 4, 6} and B = {2, 4, 6} then A = B.

Remarks:

(I) The elements of a set may be listed in any order.

Thus, {1, 2, 3} = {2, 1, 3} = {3, 2, 1} etc.

(II) The repetition of elements in a set is meaningless.

Example:

$\{x : x$ is a letter in the word “follow”$\} = \{f, o, l, w\}$

Example:

Show that Φ , $\{0\}$ and 0 are all different.

Solution:

Φ is a set containing no element at all; $\{0\}$ is a set containing one element, namely 0. And 0 is a number, not a set.

Hence Φ ,$\{0\}$ and 0 are all different.

The set which contains all the elements under consideration in a particular problem is called the universal set denoted by S. Suppose that P is a subset of S. Then the complement of P, written as P^c (or P’) contains all the elements in S but not in P. This can also be written as S – P or S ~ P. $S - P = \{x : x \in S, x \notin P\}$.

For example let $S = \{0, 1, 2, 3, 4, 5, 6, 7, 8, 9\}$

$P = \{0, 2, 4, 6, 8\}$

$Q = \{1, 2, 3, 4, 5)$

Then $P' = \{1, 3 ,5 ,7, 9\}$ and $Q'= \{0 , 6 , 7, 8, 9\}$

Also P “$\cup$” Q $= \{0, 1, 2, 3, 4, 5, 6, 8\}$, $(P \cup Q)^{\text{l}} = \{7, 9\}$

$P \cap Q = \{2, 4\}$

$P \cup Q' = \{0, 2, 4, 6, 7, 8, 9\}$, $(P \cap Q)' = \{0, 1, 3, 5, 6, 7, 8, 9\}$

$P' \cup Q' = \{0, 1, 3, 5, 6, 7, 8, 9\}$

$P' \cap Q' = \{7, 9\}$

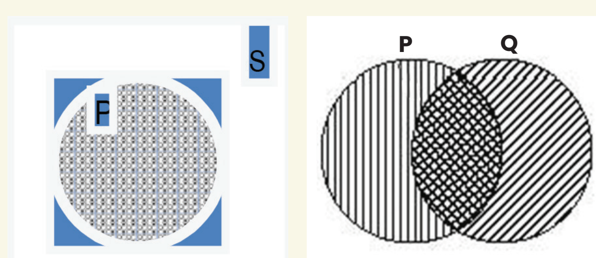
So it can be noted that $(P \cup Q)' = P' \cap Q'$ and $(P \cup Q)' = P' \cup Q'$. These are known as De Morgan’s laws.

Venn Diagrams

We may represent the above operations on sets by means of Euler - Venn diagrams.

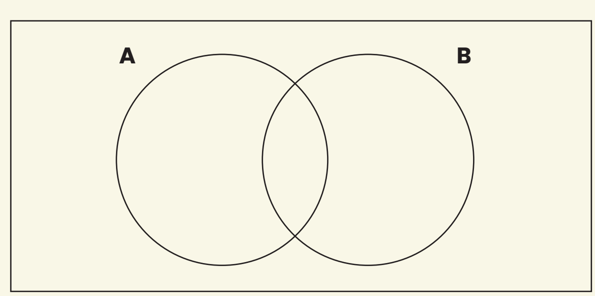
Thus Fig. 1(a) shows a universal set S represented by a rectangular region and one of its subsets P represented by a circular shaded region.

The un-shaded region inside the rectangle represents P’. Figure 1(b) shows two sets P and Q represented by two intersecting circular regions. The total shaded area represents P U Q, the cross-hatched “intersection” represents $P \cap Q$.



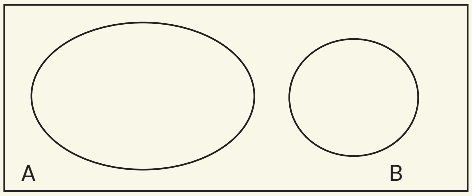
(a) Fig. 1 (b)

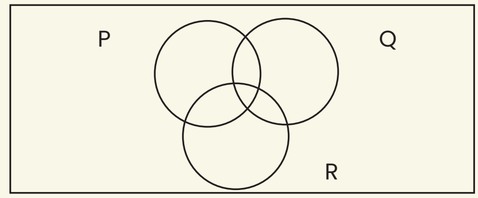
The number of distinct elements contained in a finite set A is called its **cardinal number**. It is denoted by n(A). For example, the number of elements in the set $R = \{2, 3, 5, 7\}$ is denoted by n(R). This number is called the cardinal number of the set R.



Thus n(AUB) = n(A) + n(B) – n(A ∩ B)

If A and B are disjoint sets, then n(A ∪ B) = n(A) + n(B) as n (A ∩ B) = 0





For three sets P, Q and R

n(PUQUR) = n(P) + n(Q) + n(R) – n(P ∩ Q) – n(Q ∩ R) – n(P ∩ R) + n(P ∩ Q ∩ R)

When P, Q and R are disjoint sets = n(P) + n(Q) + n(R)

Illustration:

If A = {2, 3, 5, 7} , then n(A) = 4

Equivalent Set: Two finite sets A & B are said to be equivalent if n (A) = n(B). Clearly, **equal sets are equivalent but equivalent sets need not be equal.**

Illustration:

The sets A = {1, 3, 5} and B = {2, 4, 6} are equivalent but not equal.

Here n (A) = 3 = n (B) so they are equivalent sets. But the elements of A are not in B. Hence they are not equal though they are equivalent.

Power Set : The collection of all possible subsets of a given set A is called the power set of A, to be denoted by P(A).

Illustration:

(I) If A = {1, 2, 3} then

P(A) = { {1, 2, 3}, {1, 2}, {1, 3}, {2, 3}, {1}, {2}, {3}, }

(II) If A = {1, {2}}, we may write A = {1, B} when B = {2} then

P(A) = {Q, {1}, {B }, {1, B}} = { Q, {1}, {{2}}, {1,{2}}}

RODUCT SETS

Ordered Pair : Two elements a and b, listed in a specific order, form an ordered pair, denoted by (a, b).

Cartesian Product of sets : If A and B are two non-empty sets, then the set of all ordered pairs (a, b) such that a belongs to A and b belongs to B , is called the Cartesian product of A and B, to be denoted by A × B.

Thus, A × B = {(a, b) : a ∈ A and b ∈ B}

If A = φ or B = φ, we define A × B = φ

Illustration:

Let A = {1, 2, 3}, B = {4, 5}

Then A × B = {(1, 4), (1, 5), (2, 4) (2, 5), (3, 4), (3, 5)}

Example:

If A × B = {(3, 2), (3, 4), (5, 2), (5, 4) } , find A and B.

Solution:

Clearly A is the set of all first co-ordinates of A × B, while B is the set of all second co-ordinates of elements of A × B.

Therefore A = {3, 5} and B = {2, 4}

Example:

Let P = {1, 3, 6} and Q {3, 5}

The product set P × Q = {(1, 3), (1, 5), (3, 3), (3, 5), (6, 3), (6, 5)}.

Notice that n(P×Q) = n(P) × n(Q) and ordered pairs (3, 5) and (5, 3) are not equal. and Q × P = {(3, 1), (3, 3), (3, 6), (5, 1), (5, 3), (5, 6)}

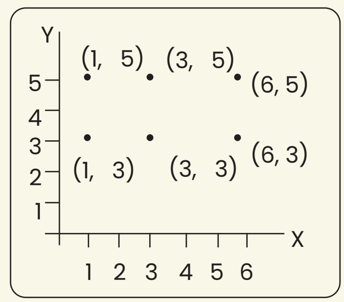
So P × Q ≠ Q × P; but n (P × Q) = n (Q × P).

Illustration:

Here n(P) = 3 and n(Q) = 2, n (P × Q) = 6. Hence n(P×Q) = n(P) × n(Q). and n(P × Q) = n(Q × P) = 6.

We can represent the product set of ordered pairs by points in the XY plane.

If X=Y= the set of all natural numbers, the product set XY represents an infinite equal lattice of points in the first quadrant of the XY plane.



RELATIONS AND FUNCTIONS

Any subset of the product set X.Y is said to define a **relation** from X to Y and any relation from X to Y in which no two different ordered pairs have the same first element is called a **function**.

Let A and B be two non-empty sets. Then, a rule or a correspondence f which associates to each element x of A, a unique element, denoted by f(x) of B, is called a function or **mapping** from A to B and we write f : A→B

The element f(x) of B is called the image of x, while x is called the pre-image of f (x).

DOMAIN & RANGE OF A FUNCTION

Let f : A→B, then A is called the domain of f, while B is called the co-domain of f. The set f(A) = {f (x) : x ∈ A} is called the range of f.

Illustration: Let A = {1, 2, 3, 4} and B = {1, 4, 9, 16, 25} We consider the rule f(x) = x^2 . Then f(1) = 1 ; f(2) =4 ; f (3) = 9 & f (4) = 16.

Then clearly each element in A has a unique image in B. So, f : A ⊠ B : f (x) = x^2 is a function from A to B.

Here domain of f = {1, 2, 3, 4} and range of f = {1, 4, 9, 16}

Example:

Let N be the set of all natural numbers. Then, the rule

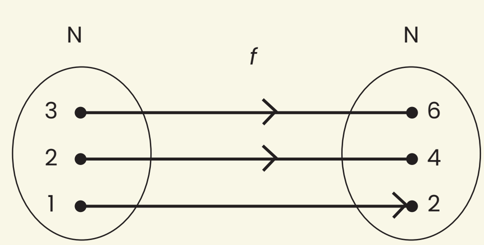
f : N → N : f(x) = 2x, for all x ∈ N

is a function from N to N, since twice a natural number is unique.

Now, f (1) = 2; f (2) = 4; f(3) = 6 and so on.

Here domain of f = N = {1, 2 ,3, 4,.......} ; range of f = { 2, 4, 6,.........}

This may be represented by the mapping diagram or arrow graph.



Various Types of Function

One-One Function : Let f : A→B. If different elements in A have different images in B, then f is said to be a one-one or an injective function or mapping.

Illustration:

(i) Let A = {1, 2, 3} and B = {2, 4, 6}

Let us consider f : A→B : f(x) = 2x.

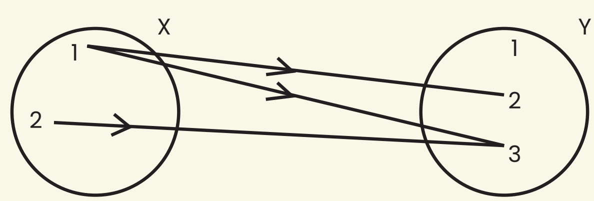
Then f(1) = 2; f(2) =4 ; f (3) = 6

Clearly, f is a function from A to B such that different elements in A have different images in B. Hence f is one-one.

Remark : Let f : A→B and let x_1 , x_2 ∈ A. Then $x_1 = x_2$ implies $f(x_1) = f(x_2)$ is always true.

But $f(x_1) = f(x_2)$ implies $x_1 = x_2$ is true only when f is one-one.

(ii) let x={1, 2, 3, 4} and y={1, 2, 3}, then the subset {(1, 2), (1, 3), (2, 3)} defines a relation on X.Y.



Notice that this particular subset contains all the ordered pair in X.Y for which the X element (x) is less than the Y element (y). So in this subset we have X<Y and the relation between the set, is "less than". This relation is not a function as it includes two different ordered pairs (1, 2), (1, 3) have same first element.

X.Y={(1, 1), (1, 2), (1, 3), (2, 1), (2, 2), (2, 3), (3, 1), (3, 2), (3, 3),(4, 1), (4, 2), (4, 3)}

The subset {(1, 1), (2, 2), (3, 3)} defines the function y= x as different ordered pairs of this subset have different first element.

Onto or Surjective Functions : Let f : A→B. If every element in B has at least one pre-image in A, then f is said to be an onto function.

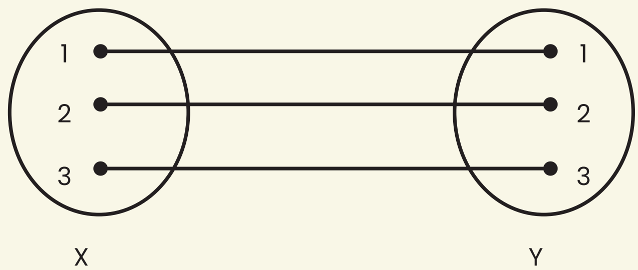
If f is onto, then corresponding to each y ∈ B, we must be able to find at least one element x ∈ A such that y = f (x)

Clearly, f is onto if and only if range of f = B

Illustration:

Let N be the set of all natural numbers and E be the set of all even natural numbers.

Then, the function f : N → E : f (x) = 2x , for all x ∈ N is onto, since each element of E is of the form 2x, where x ∈ N and the same is the f-image of x ∈ N.



Represented on a mapping diagram it is a one-one mapping of X onto Y.

Bijection Function : A one-one and onto function is said to be bijective. A bijective function is also known as a one-to-one correspondence.

Identity Function : Let A be a non-empty set. Then, the function I defined by

$I : A \to A : I(x) = x$ for all $x \in A$ is called an identity function on A.

It is a one-to-one onto function with domain A and range A.

Into Functions: Let $f : A \to B$. There exists even a single element in B having no pre-image in A, then f is said to be an into function.

Illustration:

Let $A = \{2, 3, 5, 7\}$ and $B = \{0, 1, 3, 5, 7\}$. Let us consider $f : A \to B$; $f(x) = x - 2$. Then $f(2) = 0$; $f(3) = 1$; $f(5) = 3$ & $f(7) = 5$.

It is clear that f is a function from A to B.

Here there exists an element 7 in B, having no pre-mage in A. So, f is an into function.

Constant Function: Let $f : A \to B$, defined in such a way that all the elements in A have the same image in B, then f is said to be a constant function.

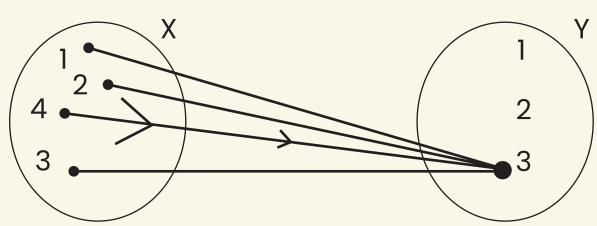
Illustration:

Let $A = \{1, 2, 3\}$ and $B = \{5, 7, 9\}$. Let $f : A \to B : f(x) = 5$ for all $x \in A$. Then, all the elements in A have the same image namely 5 in B. So, f is a constant function.

Remark: The range set of a constant function is a singleton set.

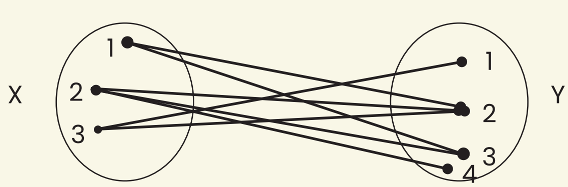
Example:

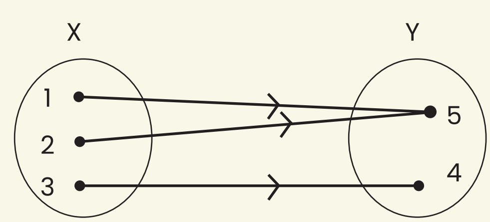
Another subset of X.Y is $\{(1, 3), (2, 3), (3, 3), (4, 3)\}$



This relation is a function (a constant function). It is represented on a mapping diagram and is a many-one mapping of X into Y.

Let us take another subset $\{(4, 1), (4, 2), (4, 3)\}$ of X.Y. This is a relation but not a function. Here different ordered pairs have same first element so it is not a function.



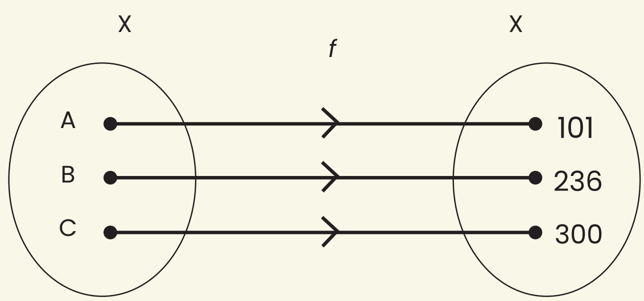


Many-one Mapping

This is an example of many-one mapping.

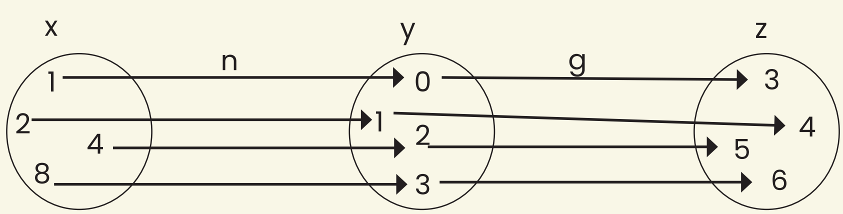
Equal Functions: Two functions f and g are said to be equal, written as f = g if they have the same domain and they satisfy the condition f(x) = g(x), for all x.

A function may simply pair people and the corresponding seat numbers in a theatre. It may simply associate weights of parcels with portal delivery charge or it may be the operation of squaring adding to doubling, taking the log of etc.



The function f here assigning a locker number to each of the persons A, B and C. Names are associated with or mapped onto, locker numbers under the function f.

We can write $f : X \to Y$ or $f(x) = y$ or $f(B) = 236$



This diagram shows the effect of two functions n and g on the sets X, Y and Z

$n : X \to Y$ and $g : Y \to Z$

If x, y, z are corresponding elements of X, Y and Z, we write n(x) = y and g(y) = z

Thus g(n(1)), n(1) = 0 and g (0) = 3, so that g (n(1)) = g(0) = 3 we can write it as g(n(1)) or g o n (1) = 3 But g(1) = 4 and n(g(1))= n(4) = 2

So gn ≠ ng (or, g o n ≠ n o g)

The function gn or ng is called a composite function. As n(8) = 3, we say that 3 is the image of 8 under the mapping (or function) n.

INVERSE FUNCTION:

Let f be a one-one onto function from A to B. Let y be an arbitrary element of B. Then f being onto, there exists an element x in A such that f (x) = y.

As f is one-one this x is unique.

Thus for each $y \in B$, there exists a unique element $x \in A$ such that f (x) = y. So, we may define a function, denoted by f^{-1} as:

$f^{-1} : B \to A : f^{-1}(y) = x$ if and only if f (x) = y. The above function f–1 is called the inverse of f.

A function is invertible if and only if f is one-one onto.

Remarks: If f is one-one onto then f^{-1} is also one-one onto.

Illustration: If $f : A \to B$ then $f^{-1} : B \to A$.

DIFFERENT TYPES OF RELATIONS

Let S = {a, b, c, ...} be any set then the relation R is a subset of the product set S×S

(i) If R contains all ordered pairs of the form (a, a) in S×S, then R is called reflexive. In a reflexive relation 'a' is related to itself .

For example, 'Is equal to' is a reflexive relation for a = a is true.

(ii) If $(a, b) \in R \Rightarrow (b, a) \in R$ for every $a, b \in S$ then R is called symmetric

For example $a=b \Rightarrow b = a$. Hence the relation 'is equal to' is a symmetric relation.

(iii) If $(a, b) \in R$ and $(b, c) \in R \Rightarrow (a, c) \Rightarrow R$ for every $a, b, c, \in S$ then R is called transistive.

For example $a = b, b = c \Rightarrow a = c$. Hence the relation 'is equal to' is a transitive relation.

A relation which is reflexive, symmetric and transitive is called an equivalence relation or simply an equivalence. 'is equal to' is an equivalence relation.

Similarly, the relation " is parallel to " on the set S of all straight lines in a plane is an equivalence relation.

Illustration:

The relation "is parallel to" on the set S is

(1) reflexive, since $a \parallel a$ for $a \in S$ (2) symmetric, since $a \parallel b \Rightarrow b \parallel a$ (3) transitive, since $a \parallel b$, $b \parallel c \Rightarrow a \parallel c$

Hence it is an equivalence relation.

DOMAIN & RANGE OF A RELATION:

If R is a relation from A to B, then the set of all first co- ordinates of elements of R is called the domain of R, while the set of all second co-ordinates of elements of R is called the range of R.

So, Dom (R) = $\{a : (a, b) \in R\}$ & Range (R) = $\{ b : (a, b) \in R\}$

Illustration:

Let A = {1, 2, 3} and B = {2, 4, 6}

Then A × B = {(1, 2), (1, 4), (1, 6), (2, 2), (2, 4), (2, 6), (3, 2), (3, 4), (3, 6)}

By definition every subset of A × B is a relation from A to B.

Thus, if we consider the relation

R = {(1, 2), (1, 4), (3, 2), (3, 4)} then Dom (R) = {1, 3} and Range (R)= {2, 4}

From the product set X. Y = {(1, 3), (2, 3), (3, 3), (4, 3), (2, 2), (3, 2), (4, 2), (1, 1), (2, 1), (3, 1), (4, 1)}, the subset {(1, 1), (2, 2), (3, 3)} defines the relation 'Is equal to', the subset {(1, 3), (2, 3), (1, 2)} defines 'Is less than', the subset {(4, 3), (3, 2), (4, 2), (2, 1), (3, 1), (4, 1)} defines 'Is greater than' and the subset {(4, 3), (3, 2), (4, 2), (2, 1), (3, 1), (4, 1), (1, 1), (2, 2) (3, 3)} defines to greater 'In greater than or equal'.

Illustration:

Let A = {1, 2, 3} and b = {2, 4, 6}

Then A × B = {(1, 2), (1, 4), (1, 6), (2, 2), (2, 4), (2, 6), (3, 2), (3, 4), (3, 6)}

If we consider the relation = {(1, 2), (1, 4), (3, 4)} then Dom (R) = {1, 3} and Range = {2, 4}. Here the relation "Is less than".

IDENTITY RELATION:

The relation I = $\{(a, a) : a \in A\}$ is called the identity relation on A.

Illustration:

Let A = $\{1, 2, 3\}$ then I = $\{(1, 1), (2, 2), (3, 3)\}$

Inverse Relation:

If R be a relation on A, then the relation R^{-1} on A, defined by $R^{-1} = \{(b, a) : (a, b) \in R\}$ is called an inverse relation on A. Clearly, Dom (R^{-1}) = Range (R) & Range (R^{-1}) = Dom (R).

Illustration:

Let A = $\{1, 2, 3\}$ and R = $\{(1, 2), (2, 2), (3, 1), (3, 2)\}$

Then R being a subset of a × a, it is a relation on A. Dom (R) = $\{1, 2, 3\}$ and Range (R) = $\{2, 1\}$

Now, $R^{-1} = \{(2, 1), (2, 2), (1, 3), (2, 3)\}$. Here, Dom $(R^{-1}) = \{2, 1\}$ = Range (R) and

Range $(R^{-1}) = \{1, 2, 3\}$ = Dom (R).

Illustration:

Let A = $\{1, 2, 3\}$, then

(i) R1 = $\{(1, 1), (2, 2), (3, 3), (1, 2)\}$

Is reflexive and transitive but not symmetric, since $(1, 2) \in$ R1 but (2, 1) does not belongs to R_1.

(ii) R2 = $\{(1, 1), (2, 2), (1, 2), (2, 1)\}$

is symmetric and transitive but not reflexive, since (3, 3) does not belong to R_2.

(iii) R3 = $\{(1, 1), (2, 2), (3, 3), (1, 2), (2, 1), (2, 3), (3, 2)\}$

is reflexive and symmetric but not transitive, since $(1, 2) \in$ R3 & $(2, 3) \in$ R3 but (1, 3) does not belong to R3.

Problems and solution using Venn Diagram

1. Out of a group of 20 teachers in a school, 10 teach Mathematics, 9 teach Physics and 7 teach Chemistry. 4 teach Mathematics and Physics but none teach both Mathematics and Chemistry. How many teach Chemistry and Physics? How many teach only Physics ?

   Let x be the no. of teachers who teach both Physics & Chemistry. $9 - 4 - x + 6 + 7 - x + 4 + x = 20$

   or $22 - x = 20$ or $x = 2$

   Hence, 2 teachers teach both Physics and Chemistry and $9 - 4 - 2 = 3$ teachers teach only Physics.

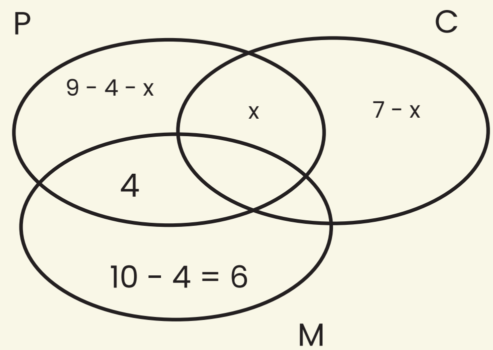
2. A survey shows that 74% of the Indians like grapes, whereas 68% like bananas.

   What percentage of the Indians like both grapes and bananas?

Solution:

Let P & Q denote the sets of Indians who like grapes and bananas respectively. Then n(P) = 74, n(Q) = 68 and $n(P \cup Q) = 100$.

We know that $n(P \cap Q) = n(P) + n(Q) - n(P \cup Q) = 74 + 68 - 100 = 42$. Hence, 42% of the Indians like both grapes and bananas.



3. In a class of 60 students, 40 students like Maths, 36 like Science, and 24 like both the subjects. Find the number of students who like

(i) Maths only

(ii) Science only

(iii) either Maths or Science

(iv) neither Maths nor Science

Solution:

Let M = students who like Maths and S = students who like Science Then $n(M) = 40$, $n(S) = 36$ and $n(M \cap S) = 24$

Hence,

(i) $n(M) - n(M \cap S) = 40 - 24 = 16$ = number of students like Maths only.

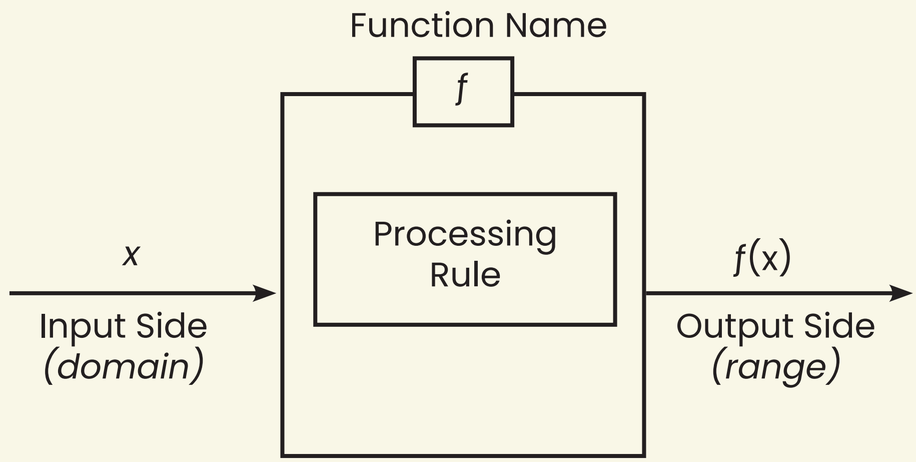
(ii) $n(S) - n(M \cap S) = 36 - 24 = 12$ = number of students like Science only.

(iii) $n(M \cup S) = n(M) + n(S) - n(M \cap S) = 40 + 36 - 24 = 52$ = number of students who like either Maths or Science.

(iv) $n(M \cup S)^c = 60 - n(M \cup S) = 60 - 52 = 8$ = number of students who like neither Maths nor Science.

Limits and Continuity

Value of the function: Value of a function is always a real number.



Exp 1:

If $f(x) = \frac{x+1}{x^2-3x-4}$, find f(0).

Solution: $f(0) = \frac{1}{-4} - \frac{1}{4}$

Note: Any function becomes undefined (i.e. mathematically cannot be evaluated) if denominator is zero.

Exp 2:

If $f(x) = \frac{x+1}{x^2-3x-4}$, find f(-1), f(4)

Solution: $f(-1) = \frac{0}{0}$ and $f(4) \frac{5}{0}$ and hence function is not defined at

$x = -1, 4$

Even Function: If a function f(x) is such that f (-x) = f (x) then it is said to be an even function of x.

Exp 3:

If $f(x) = \frac{e^x + e^{-x}}{2}$ then $f(-x) \frac{e^{-x} + e^{-(-x)}}{2} = \frac{e^{-x} + e^x}{2} = f(x)$. Hence f(x) is an even function.

Odd Function: If a function $f(x)$ is such that $f(-x) = -f(x)$ then it is said to be an odd function of *x*.

Exp 4:

If $f(x) = \frac{5^x + 1}{5^x - 1}$ then $f(-x) \frac{5^{-x} + 1}{5^{-x} - 1} = \frac{1 + 5^x}{1 - 5^x} = -f(x)$. Hence f(x) is an odd function.

Inverse Function: If y = f (x) defined in an interval (a, b) is a function such that we express x as a function of y say x = g(y) then g(y) is called the inverse of f(x).

Exp 5:

If $y = \dfrac{5x + 3}{2x + 9}$, then find its inverse.

Solution: $y = \dfrac{5x + 3}{2x + 9}$

$\Rightarrow y(2x+9)=5x+3$

$\Rightarrow 2yx+9y=5x+3$

$\Rightarrow (2y-5)x=3-9y$

$\Rightarrow x = \dfrac{3 - 9y}{2y - 5}$

Composite Function: If $y = f(u)$ and $u = g(x)$ then $y = f\{g(x)\} = fog(x)$ is called the function of a function or a composite function.

Exp 6:

If $2f(x) = x + 3$, $g(x) = x^2$, then find $gof(x)$.

Solution: $gof(x) = g\{f(x)\} = g(x + 3)=(x+3)^2$

$f(a)$ = value of the function at x = a
$\lim\limits_{x \to a} f(x)$ = limit of the function as x approaches towards a

Limit:

$$\lim_{x \to a} f(x)$$

LHL | *RHL*

$$\lim_{x \to a^-} f(x) \qquad \lim_{x \to a^+} f(x)$$

$$\lim_{h \to 0} f(a-h) \qquad \lim_{h \to 0} f(a+h)$$

$\lim\limits_{x \to a^-} f(x)$ is said to exist when both left-hand and right-hand limits exists and

they are equal. We write as $\lim\limits_{x \to a^-} f(x) = \lim\limits_{x \to a^-} f(x) = \lim\limits_{x \to a^-} f(x)$

Thus, if $\lim\limits_{h \to 0} f(a-h) = \lim\limits_{h \to 0} f(a+h)$ then $\lim\limits_{x \to a} f(x)$ exists.

Note:

$$\lim_{x\to 0^+} \frac{1}{x} = +\infty$$

$$\lim_{x\to 0^-} \frac{1}{x} = -\infty$$

$$\lim_{x\to 0^+} \frac{1}{x} \neq \lim_{x\to 0^-} \frac{1}{x}$$

Thus, $\lim_{x\to 0} \frac{1}{x}$ does not exist.

Some Important Limits:

(1) $\lim_{x\to 0} \frac{e^x - 1}{x} = 1$

(2) $\lim_{x\to 0} \frac{a^x - 1}{x} = \log_e a$

(3) $\lim_{x\to 0} \frac{\log(1+x)}{x} = 1$

(4) $\lim_{x\to 0} \frac{x^n - a^n}{x - a} = na^{n-1}$

(5) $\lim_{x\to 0} \frac{(1+x)^n - 1}{x} = n$

(6) $\lim_{x\to 0} (1+x)^{\frac{1}{x}} = \lim_{x\to\infty} \left(1+\frac{1}{x}\right)^x = e$

(7) $\lim_{x\to 0} (1+px)^{\frac{1}{x}} = \lim_{x\to\infty} \left(1+\frac{p}{x}\right)^x = e^p$

Note:

- The number *e* called exponential number is given by $e \approx 2.7183$.
- In calculus all logarithms are taken with respect to base '*e*' that is $\log x = \log_e x$.

Continuity:

A function *f* is continuous at the point x = a if the following are true:

(i) $f(a)$ is defined,

(ii) $\lim_{x\to a^-} f(x) = \lim_{x\to a^+} f(x)$ and

(iii) $\lim_{x\to a} f(x) = f(a)$

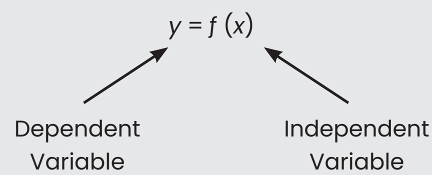
- If f and g are continuous at x = a, then $f \pm g$, fg and $\frac{f}{g}$ (g(a) ≠ 0) are continuous at x = a.
- A polynomial function $y = P(x)$ is continuous at every point x.
- A rational function $R(x)$ $\frac{P(x)}{Q(x)}$ is continuous at every point x in its domain.

Chapter – 8:

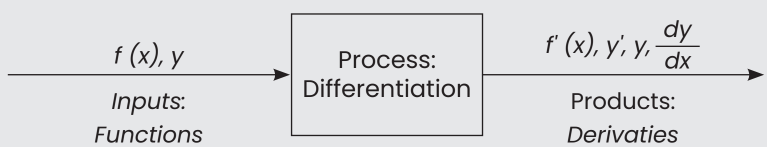
Basic Concepts of Differential and Integral Calculus

Differentiation

Differentiation is the process of finding rate of change of a dependent variable with respect to independent variable.



$\frac{dy}{dx}$ $f'(x)=y'=y_1$= derivative or differential coefficient of $f(x)$ with respect to x



The Process of Differentiation

Differentiation Formulae

$$\frac{d}{dx}(c) = 0$$

$$\frac{d}{dx}(x^n) = nx^{n-1}$$

$$\frac{d}{dx}(e^x) = e^x$$

$$\frac{d}{dx}(a^x) = a^x \log_e a$$

$$\frac{d}{dx}(\log_e x) = \frac{1}{x}$$

Also remember the following formulae:

$$\frac{d}{dx}(x) = 1$$

$$\frac{d}{dx}(\sqrt{x}) = \frac{1}{2\sqrt{x}}$$

$$\frac{d}{dx}\left(\frac{1}{x}\right) = -\frac{1}{x^2}$$

Differentiation Techniques

1. $\frac{d}{dx}(c) = 0$

Exp: $\frac{d}{dx}(2) = 0,\ \frac{d}{dx}\left(\sqrt{2}\right) = 0,\ \frac{d}{dx}(\log 2) = 0$

2. $\frac{d}{dx}[cf(x)] = c\frac{d}{dx}[f(x)]$

Exp: $\frac{d}{dx}(2x^3) = 2\frac{d}{dx}(x^3) = 2(x^3) = 6x^3$

3. $\frac{d}{dx}[c_1 f(x) \pm c_2 g(x)] = c1\frac{d}{dx}[f(x)] \pm c_2\frac{d}{dx}[g(x)]$

Exp: $\frac{d}{dx}(2x^3 + 5\log x)\ \frac{d}{dx}(2x^3) + \frac{d}{dx}(5\log x) = 2\frac{d}{dx}(x^3) + 5\frac{d}{dx}(\log x) = 6x^2 + \frac{5}{x}$

4. Product Rule:

$$\frac{d}{dx}[f(x) \text{ x } g(x)] = \frac{d}{dx}[f(x)]\,g(x) + f(x)\frac{d}{dx}[g(x)]$$

(u . v)' = u'v + uv'

In general, (u . v . w)' = u'vw+ uv'w+ uvw'

Exp:

$$\frac{d}{dx}(x^3\log x) = \frac{d}{dx}(x^3)\log x + x^3\frac{d}{dx}(\log x) = 3x^2\log x + x^3.\frac{1}{x}\ \ 3x^2\log x + x^2 = x^2(3\log x + 1)$$

5. Quotient Rule:

$$\left(\frac{u}{v}\right)' = \frac{u'v - uv'}{v^2}$$

Exp:

$$\frac{d}{dx}\left(\frac{x^3}{\log x}\right) = \frac{\frac{d}{dx}(x^3)\log x - x^3\frac{d}{dx}(\log x)}{(\log x)^2} = \frac{3x^2\log x - x^3.\frac{1}{x}}{(\log x)^2} = \frac{3x^2\log x - x^2}{(\log x)^2} = \frac{x^2(3\log x - 1)}{(\log x)^2}$$

6. Derivative of a function of function (Chain Rule):

If y = $f[h(x)]$, *then* $\frac{dy}{dx} = \frac{dy}{du} \text{ x } \frac{du}{dx}$, *where* $u = h\ x$

Exp: If y = log $(1 + x^2)$, then find $\frac{dy}{dx}$.

Solution: Let $y = \log u$ and $u = 1 + x^2$

$$\frac{dy}{dx} = \frac{dy}{du} \text{ x } \frac{du}{dx} = \frac{1}{u} \text{ x } (0 + 2x) = \frac{2x}{u} = \frac{2x}{1 + x^2}$$

7. Derivative of Implicit Functions:

A function in the form of f (x, y) = 0 e.g. 2 2 x y +3xy + y = 0 where y cannot be directly defined as a function of x is called an implicit function of x.

Exp:

If $x^3 + y^3 - 3axy = 0$, then find $\frac{dy}{dx}$.

Solution:

$x^3 + y^3 - 3axy = 0$

Differentiating with respect to x we get

$$3x^2 + 3y^2 . \frac{dy}{dx} - 3a\left(1 \cdot y + x \cdot \frac{dy}{dx}\right) = 0$$

$$(y^2 - ax) \cdot \frac{dy}{dx} = ay - x^2$$

$$\frac{dy}{dx} = \frac{ay - x^2}{y^2 - ax}$$

8. Derivative of Parametric Equation:

When both the variables are expressed in terms of a parameter (a third variable), then the involved equations are called as parametric equations.

Exp:

If $x = at^2$ and y = 2at , then find $\frac{dy}{dx}$.

Solution: $\frac{dx}{dt} = 2at; \ \frac{dy}{dt} = 2a$

$$\frac{dy}{dx} = \frac{\left(\frac{dy}{dt}\right)}{\left(\frac{dx}{dt}\right)} = \frac{2a}{2at} = \frac{1}{t}$$

9. Logarithmic Differentiation:

This procedure of finding out derivative by taking logarithm is used in the following two situations:

➤(function)$^{\text{function}}$ or (variable)$^{\text{variable}}$

➤When the function is the product of number of functions.

Exp:

If $y = x^x$, then find $\frac{dy}{dx}$.

Solution:

$y = x^x$

Taking logarithm of both the sides we get

$\log y = x \log x$

Differentiating with respect to x we get

$$\frac{1}{y}\frac{dy}{dx} = \log x + 1$$

$$\text{or } \frac{dy}{dx} = y(1 + \log x) = x^x(1 + \log x)$$

Exp:

If $y = \dfrac{(x-2)^3\,(x-3)^2}{(x-1)^4}$, then find $\dfrac{dy}{dx}$.

Solution:

$$y = \frac{(x-2)^3\,(x-3)^2}{(x-1)^4}$$

Taking logarithm of both the sides we get

$\log y = 3\log(x-2) + 2\log(x-3) + 4\log(x-1)$

Differentiating with respect to x, we get

$$\frac{1}{y}\frac{dy}{dx} = \frac{3}{(x-2)} + \frac{2}{(x-3)} - \frac{4}{(x-1)}$$

$$\text{or } \frac{dy}{dx} = y\left[\frac{3}{(x-2)} + \frac{2}{(x-3)} - \frac{4}{(x-1)}\right]$$

$$\text{or } \frac{dy}{dx}\ \frac{(x-2)^3\,(x-3)^2}{(x-1)^4}\left[\frac{3}{(x-2)} + \frac{2}{(x-3)} - \frac{4}{(x-1)}\right]$$

Higher Order Differentiation:

Second order derivative $= f''(x) = \dfrac{d^2y}{dx^2} = y'' = y_2$

Exp: If $y = x^4 + 5x^3 + 2x^2 + 9$, then find $\dfrac{d^2y}{dx^2}$.

Solution: $y = x^4 + 5x^3 + 2x^2 + 9$

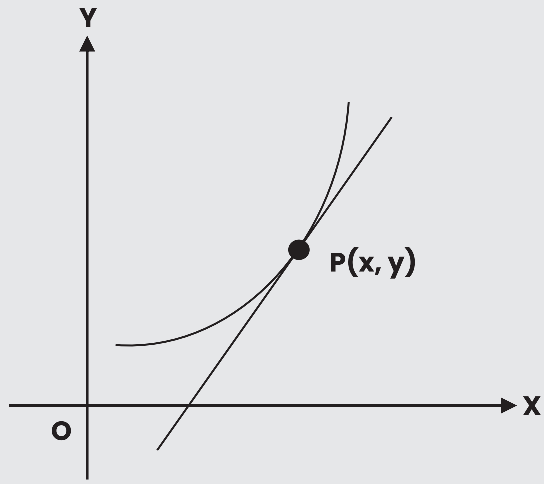
$$\frac{dy}{dx} = 4x^3 + 15x^2 + 4x$$

$$\frac{d}{dx}\left(\frac{dy}{dx}\right) = 12x^2 + 30x + 4$$

$$\frac{d^2y}{dx^2} = 12x^2 + 30x + 4$$

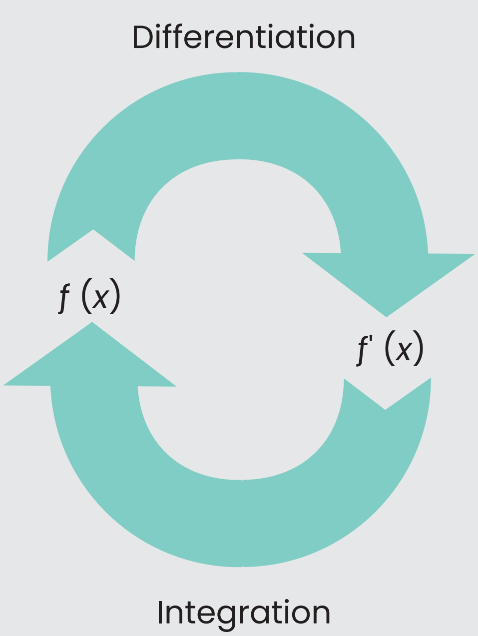
Geometric Meaning of the Derivative

The derivative of $f(x)$ at a point x represents the slope or gradient of the tangent to the curve $y = f(x)$ at the point x.



Integration

Integration is the reverse process of differentiation.



Integration Formulae

$\int x^n\, dx = \frac{x^{n+1}}{n+1} + C \ (n \neq -1)$

$\int \frac{1}{x}\, dx = \log_e x + c$

$\int e^x\, dx = e^x + c$

$\int a^x\, dx = \frac{a^x}{\log_e a} + C$

Also remember the following formulae:

$\int 1\, dx = \int dx = x + c$

$\int \frac{1}{\sqrt{x}}\, dx = 2\sqrt{x} + c$

$\int \frac{dx}{x^2 - a^2} = \frac{1}{2a} \log \frac{x-a}{x+a} + C$

$\int \frac{dx}{a^2 - x^2} = \frac{1}{2a} \log \frac{a+x}{a-x} + C$

$\int \frac{dx}{\sqrt{x^2 + a^2}} = \log\left(x + \sqrt{x^2 + a^2}\right) + C$

$\int \frac{dx}{\sqrt{x^2 - a^2}} = \log\left(x + \sqrt{x^2 - a^2}\right) + C$

$\int \sqrt{x^2 + a^2}\, dx = \frac{x}{2}\sqrt{x^2 + a^2} + \frac{a^2}{2} = \log\left(x + \sqrt{x^2 + a^2}\right) + C$

$\int \sqrt{x^2 - a^2}\, dx = \frac{x}{2}\sqrt{x^2 - a^2} - \frac{a^2}{2} = \log\left(x + \sqrt{x^2 - a^2}\right) + C$

Integration by parts

$\int u \cdot v\, dx = u \int v\, dx - \int \left\{ \frac{du}{dx} \cdot \int u \cdot v\, dx \right\} dx$

Priority of functions to be considered as a first function is as follows:

Logarithmic Functions e.g. log x

Algebric Functions e.g. x^2, $\sqrt{x}$ *etc*

Exponential Functions e.g. e^x, a^x *etc*

Definite Integration

If $\int f(x)\,dx = F(x)$, then

$$\int_a^b f(x)\,dx = [F(x)]_a^b = [F(b) - F(a)]$$

Important Property of definite Integral

$$\int_a^b f(x)\,dx = \int_a^b f(a + b - x)\,dx$$

Chapter – 9:

Number Series, Coding Decoding and Odd Man Out Series

Learning Objectives

- This Section deals with questions in which series or letters in some order, Coding and decoding
- This terms of the series or letters are follows certain pattern throughout

9.1 Series

Series Classified into Two Types, Namely

A. Number Series

B. Alphabet Series

A. Number Series

Case 1: Missing terms of the series

In this type the questions we have to identify the missing term of the series real according to a specific pattern of the series rule to form its code. The students are required to detect the missing number of the series and answer the questions accordingly.

Example 1: Find the missing term of the series 2, 7, 16,_____, 46, 67, 92

Explanation: Here the terms of the series are +5, +9, +13, +17, +21 , +25...

Thus, 2 + 5 = 7; and 7 + 9 = 16 ...

So missing term = 16 + 13 = 29

Example 2: Find the wrong terms of the series 9, 29, 65, 126, 217, 344

Explanation: 2^3+1; $3^3 + 1$; $4^3 + 1$; $5^3 + 1$; $6^3 + 1$; $7^3 + 1$

Here 29 is wrong term of series

Example 3: Find the missing term of the series 1,9, 25, 49, 81, 121, ...

Solution: The given terms of the series are consists square of consecutive odd number 12, 32, 52, 72, ...

So missing value = 132 = 169

B. ALPHABET SERIES

Alphabet series consists of letters of the alphabet placed in a specific pattern. For example, the series are in the following order of the numbers.

1	2	3	4	5	6	7	8	9	10	11	12	13	14	15	16	17	18	19	20	21	22	23	24	25	26
A	B	C	D	E	F	G	H	I	J	K	L	M	N	O	P	Q	R	S	T	U	V	W	X	Y	Z
26	25	24	23	22	21	20	19	18	17	16	15	14	13	12	11	10	9	8	7	6	5	4	3	2	1

Example 4: Find the next term of the series BKS, DJT, FIU, HHV?

Explanation: In each term, the first letter is moved two steps forward, the second letter one step backward and third letter one step forward to obtain the corresponding letter of the next term. So the missing term is JGW.

C. Letter Series:

This type of question usually consist of a series of small letters which follow a certain pattern. However, some letters are missing from the series. These missing letters are then given in a proper sequence as one of the alternatives.

Example 5: aab,____, aaa, bba,____

(a) baa (b) abb (c) bab (d) aab

1) The first blank space should be filled in by 'b' so that we have two a's by two b's.
2) The second blank place should be either `a', so that we have three a's followed by three b's.
3) The last space must be filled in by 'a'.
4) Thus we have two possible answers – 'baa'and 'bba'.
5) But only 'baa'appers in the alternatives.

So the answer (a) is correct.

9.2 Coding and Decoding

Before transmitting, the data is encoded and at receiver side encode data is decoded in order to obtain original data by determining common key in encoded data.

The Coding and Decoding is classified into the following types.

Type 1: Letter Coding

Type 2: Number Coding

Type 1: Letter Coding

In this type the real alphabets in a word are replaced by certain other alphabets according to a specific rule to form its code. The candidate is required to detect the common rule and answer the questions accordingly.

Case 1: To form the code for another word

Example 6: If in a certain language MYSTIFY is coded as NZTUJGZ, how is MENESIS coded in that language?

Explanation: Clearly, each letter in the word MYSTIFY is moved one step forward to obtain the corresponding letter of the code.

M Y S T I F Y

+1 ↓

N Z T U J G Z

So, in MENESIS, N will be coded as O, E as F, M as N and so on. Thus, the code becomes NFOFTJT.

Example 7: If TAP is coded as SZO, then how is FRIEND coded?

Explanation: Clearly each letter in the word TAP is moved one step backward to obtain the corresponding letter of the code.

S Z O

-1 ↑

T A P

Thus, in FRIEND, F will be coded as E, R as Q , I as H, E as D, N as M and D as C. So, the code becomes EQHDMC.

Example 8: In a certain code, MENTION is written as LNEITNO. How is PRESENT written in that code?

Explanation: Clearly, to obtain the code, the first letter of the word MENTION is moved one step backward and the remaining letters are.

Reversed in order, taking two at a time. So, in PRESENT, P will be coded as O, and the sequence of the remaining letter in the code would be ERESTN. Thus the code becomes OERESTN. Hence, The answer is OERESTN.

Case 2: To find the word by analysing the given code (DECODING)

Example 9: If in a certain language CARROM is coded as BZQQNL, which word will be coded as HORSE?

Explanation: each letter of the word is one step ahead of the corresponding letter of the code

B Z Q Q N L	H O R S E
| | | | | |	| | | | |
C A R R O M	I P S T F

So, H is coded as I, O as P, R as S, S as T and E as F. HORSE is coded a IPSTF.

Type 2: Number Coding

In these questions, either numerical code values are assigned to a word or alphabetical code letters are assigned to the numbers. The candidate is required to analyse the code as per the directions.

Case 1: When a numerical code values are assigned to words.

Example 10: If in a certain language A is coded as 1, B is coded as 2, and so on, how is AICCI is coded in that code?

Explanation: As given the letters are coded as

A	B	C	D	E	F	G	H	I
1	2	3	4	5	6	7	8	9

So in AICCI, A is coded as 1, I as 9,and C as 3. Thus, AICCI is coded as 19339.

Example 11: If PAINT is coded as 74128 and EXCEL is coded as 93596, then how would you encode ANCIENT ?

Explanation: Clearly, in the given code, the alphabets are coded as follows:

P	A	I	N	T	E	X	C	L
7	4	1	2	8	9	3	5	6

So, in ANCIENT, A is coded as 4, N is coded as 2, C as 5, I is coded as 1, E as 9, and T as 8. Hence, the correct code is 4251928.

Case 2: Number to letter coding.

Example 12: In a certain code, 2 is coded as P, 3 as N, 9 as Q, 5 as R, 4 as A and 6 as B. How is 423599 coded in that code?

Explanation: Clearly as given, 4 as A, 2 as P, 3 as N and 5 is coded as R, 9 as Q. So, 423599 is coded as APNRQQ.

9.3 Odd Man Out

Classification means 'to assort the items' of a given group on the basis of a certain common quality they possess and then spot the stranger or 'odd one out'.

These questions are based on words, letters and numerals. In these types of problems, we consider the defining quality of particular things. In these questions, four or five elements are given,out of which one does not belong to the group. You are required to find the 'odd one'.

Example 13: January, May, July, November

(a) January (b) May (c) July (d) November

Explanation: All the months above are 31 days , whereas , November 30 days

Answer: (d)

Example 14: 10, 14, 16, 18, 23, 24 and 26

(a) 26 (b) 17 (c) 23 (d) 9

Explanation: Each of the above series are even number, except 23.

Answer: (c)

Example 15: 6, 9, 15, 21, 24, 26, 30

(a) 9 (b) 26 (c) 24 (d) 30

Explanation: All are multiples of 3, except 26, answer (b)

Answer: (b)

Example 16: 1, 5, 14, 30, 51, 55, 91

(a) 5 (b) 55 (c) 51 (d) 91

Explanation: Each pattern is 1^2, $1^2 + 2^2$, $1^2 + 2^2 + 3^2$, $1^2 + 2^2 + 3^2 + 4^2$, $1^2 + 2^2 + 3^2 + 4^2 + 5^2$, $1^2 + 2^2 + 3^2 + 4^2 + 5^2 + 6^2$

But 51, is not of the form.

Answer: (c)

Example 17: 16, 25, 36, 62, 144, 196, 225

(a) 36 (b) 62 (c) 196 (d) 144

Explanation: Each of the number except 62, is a perfect square.

Answer: (b)

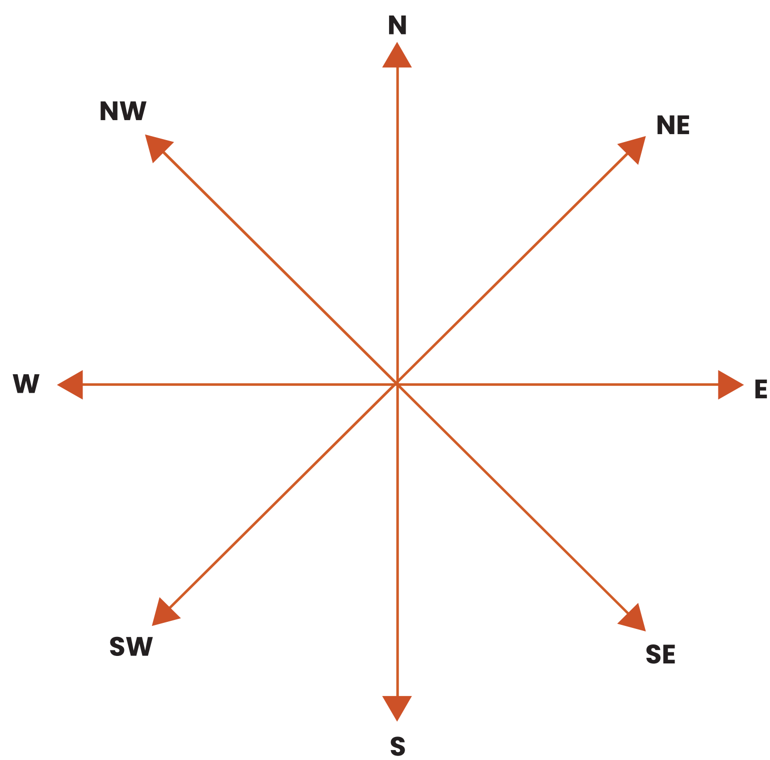
Chapter – 10:

Direction Sense Test

Introduction

After reading this chapter, students will be able to understand:

- In this test, the questions consist of a sort of direction puzzle. A successive follow-up of direction is formulated and the students is required to ascertain the final direction. The test is meant to judge then ability to trace and follow correctly and sense the direction correctly.
- The adjoining figure shows the four main directions (North N, South S, East E, and West W) and four cardinal directions (North East (NE), North West (NW), South East (SE), South West (SW) to help the students know the directions.



Always Remember:

Direction mentioned (given)	Direction indicated
Left + left	Down
Left + right	Up
Right + left	Up
Right + right	Down
Up + left	Left
Up + right	Right
Down + left	Right
Down + right	Left

Examples:

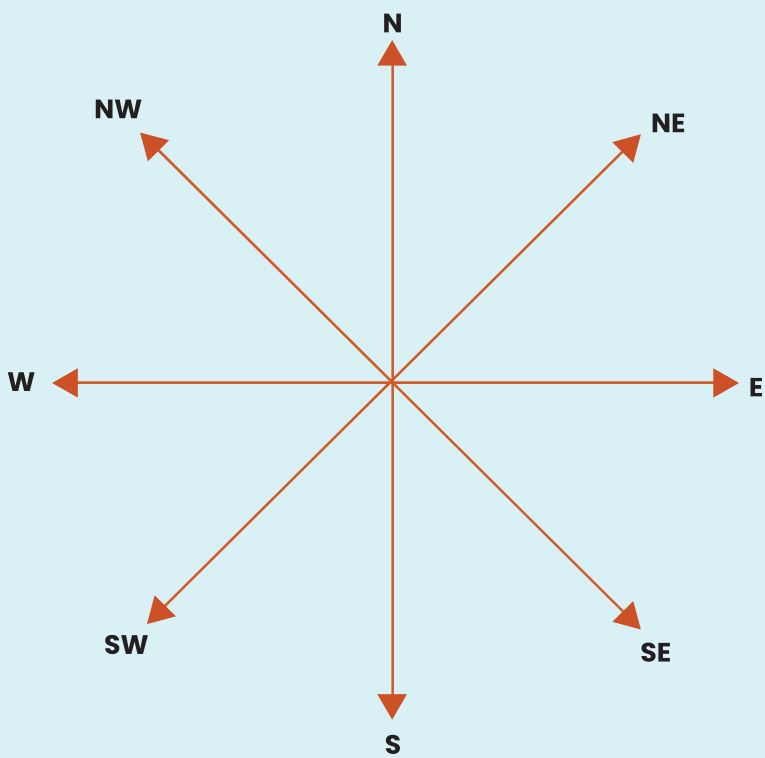
1. **A man starts from a point and walks 2 km towards North, turns towards his right and walks 2 km, turns right again and walks. What is the direction now be is facing?**

   (a) South (b) South-East

   (c) North (d) West

Explanation: (a) The diagram given below helpful solving the questions and Direction Test.

South.

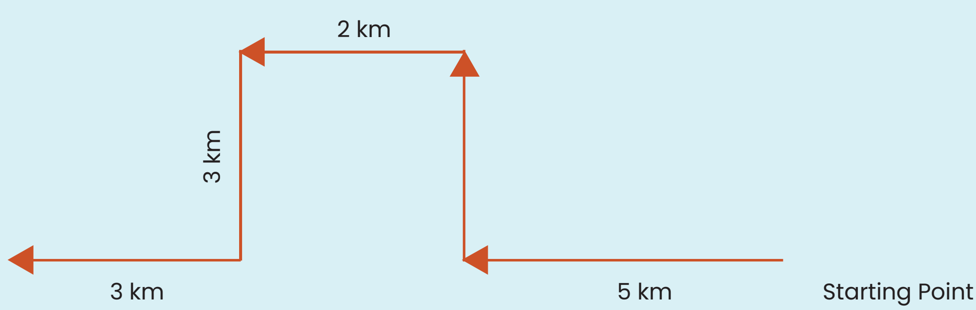


2. **Ramu walks 5 kms starting from his house towards west then turns right and walks 3 km. Thereafter she takes left turn and walks 2 km. Further, he turn left and walks 3 km. Finally, he turns right and walks 3 kms. In what direction he is now from his house?**

   (a) West (b) North

   (c) South (d) East

Explanation:



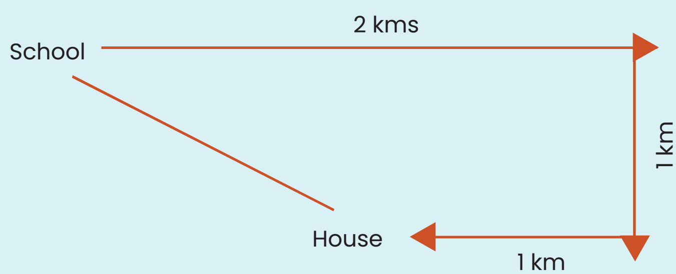
It's clear from the diagram Ramu is West of his house.

3. Gopal started walking 2 kms straight from his school. Then he turned right and walked 1 km. Again he turned right and walked 1 km to reach his house. If his house is south- east from his school, then in which direction did Gopal start walking from the school?

(a) East (b) West

(c) South (d) North

Explanation:



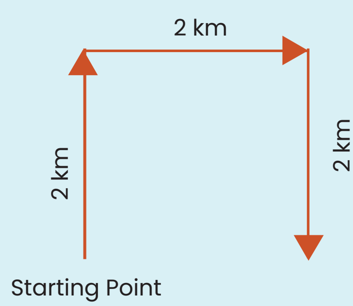
From the diagram that Gopal Started walking towards East from the school.

4. A man starts from a point, walks 2 kms towards north, turns towards his right and walks 2 kms, turns right again and walks. What is the direction now he is facing?

(a) South (b) East

(c) North (d) West

Explanation:



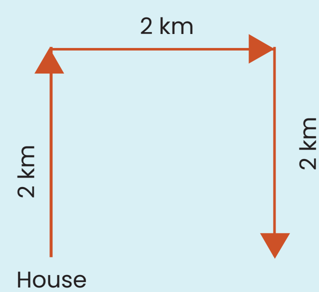
Based on the diagram the person facing towards south.

5. Janaki started from her house and walked 2 km towards North. Then she took a right turn and covered one kilometre. Then she took again a right turn and walked for 2 kms. In what direction is she going?

(a) North (b) East

(c) South (d) West

Explanation:



Based on the diagram; it's clear that Janaki is going towards South

Chapter - 11:

Seating Arrangements

Learning Objectives

- To understand the Logical statements involved in the Seating Arrangements.
- To understand the types of Seating Arrangements.

The process of making a group of people to sit as per a prefixed manner is called Seating Arrangement these questions, some conditions are given on the basis of which students are required to arrange objects, either in a row or in a circular order.

Introduction

11.1 Based on Various Pattern of Sitting Arrangements are Classified into

1) Linear Arrangements
2) Circular Arrangements
3) Polygon Arrangements

Here we are limited to our topic linear and circular arrangements only. While making arrangements, it should be noted that all the conditions given are compiled with. These type of questions generally involve five to eight individuals arranged in a certain manner or pre-conditions. They may have to be arranged in a circle or in a row accordingly.

Sometimes these questions are made more difficult by allowing an individual to a particular position with some conditions.

General instructions to Solve Seating Arrangement Questions are as follows.

1) First of all take a review on the given information. After performing this step, you would get an idea of the situation of people or objects.
2) Next, determine the usefulness of each information's and classify them accordingly into 'definite information', 'comparative information' and 'negative information'.
3) When the place of any objects or persons is definitely mentioned then we say that it is a definite information, X is sitting on the right end of the bench.
4) When the place of any object or person is not mentioned definitely but mentioned only in the comparison of another person or object, then we say that it is a comparative information.

   Example 1: A is sitting second to the right of E. This type of information can be helpful when we can get the definite information about E.
5) A part of definite information may consist of negative information. A negative information does not tell us anything definitely but it gives an idea to eliminate a possibility.

   Example 2: C is not sitting on the immediate left of A.

11.2 Type-1 Linear Arrangement

In this type of arrangement, we arrange objects or persons in a line or row. The arrangement is done only on one 'axis' and hence, the position of persons or objects assumes importance in terms of order like positions. In this type of arrangement, we take directions according to our left and right.

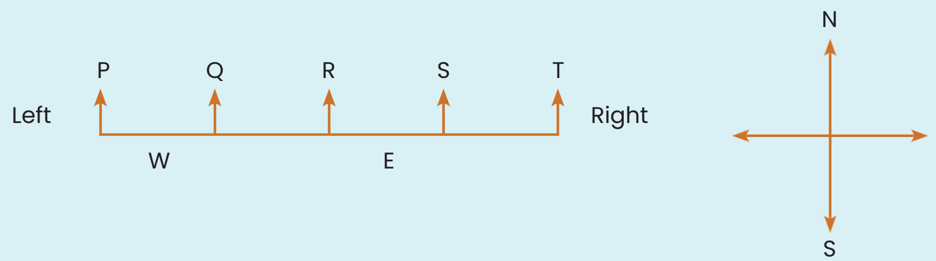
Steps to Solve the Linear Arrangements:

(a) Identify the number of objects and their names.
(b) Use pictorial method to represent the people or objects and their positions.
(c) Arrange the information with relevant facts and their positions and try to find out the solution.
(d) Answer the questions based on the arrangement having made.

There are few words which must be paid adequate attention, i.e., 'between' means sandwiched, 'immediate left' is different from 'to the left'. To understand it let us see some pictorial representation.

When direction of face is not clear, then we take One Row Sequence

(A) When direction of face is not clear, then we take based on diagram will be as follows:



From the above diagram, it is clear that

(i) Q, R, S, T are right of P but only Q is the immediate right of P.

(ii) S, R, Q, P are left of T but only S is the immediate left of T.

(iii) R, S, T are right of Q only R is the immediate right of Q.

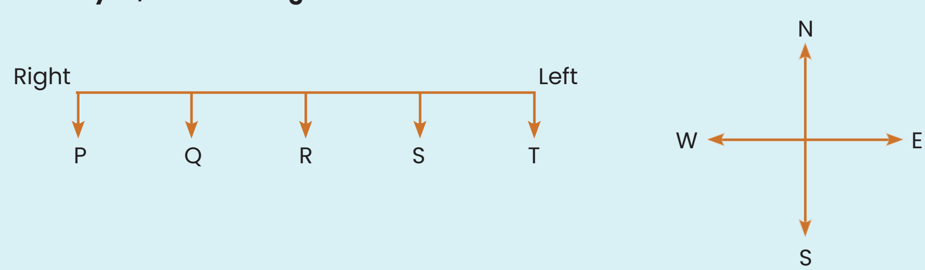
(iv) R, Q, P are left of S but only R is the immediate left of S.

(v) S and T are right of R but only S is the immediate right of R.

(vi) Q and P are left of R but only Q is the immediate left of R.

(vii) P is the immediate left of Q while T is the immediate right of S.

(B) When direction of face is towards you, then the diagram will be as follows:



From the above diagram, it is clear that

(i) Left of P = Q, R, S and T

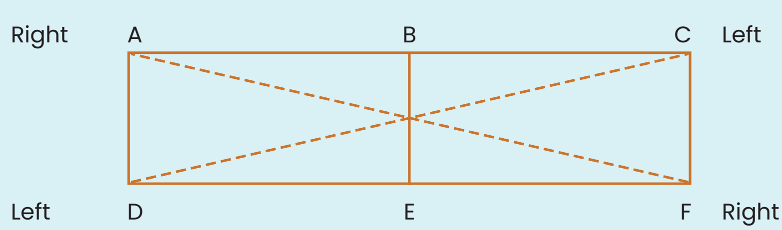
(ii) Right of T = S, R, Q and P

(iii) Q is immediate left of P; R is immediate left of Q; S is immediate left of R and T is immediate left of S.

(iv) S is immediate right of T; R is immediate right of S; Q is immediate right of R; and P is immediate right of Q.

Two Rows Sequence

Let us see 6 persons seating in two rows.



From the above diagram, it is clear that

(i) A is sitting opposite D

(ii) B is sitting opposite E

(iii) C is sitting opposite F

(iv) D and C are sitting at diagonally opposite positions

(iv) A and F are sitting at diagonally opposite positions.

Example 3: Four Children's are sitting in arrow. A is occupying seat next to B but not next to C. If C is not sitting next to D ? Who is occupying seat adjacent to D.

(a) B (b) B and A (c) Impossible to tell (d) A

Solution: (d) The arrangements as per given information is possible only if C is sitting next to B and D is sitting next to A.

Therefore, two possible arrangements are C, B, A, D, or D, A, B, C

Clearly, only A is sitting adjacent to D:

Example 4: P, Q, R, S, T, U, V and W are sitting in a row facing North.

(i) P is fourth to the right of T

(ii) W is fourth to the left of S

(iii) R and U, which are not at the ends, are neighbours of Q and T respectively.

(iv) W is next to the left of P and P is the neighbour of Q. Who are sitting at the extreme ends ?

Solution:

From information

(i) we get that there are three persons between P and TXXXP.

In the information (iv), it is given that W is next to the left of P and Q is the neighbour of P. Using the information with (i), we get TXXWPQ.

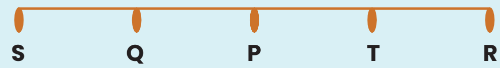
By the information (ii), TXXWPQXS By the information (iii),

So, T and S are sitting at the extreme ends.

Example 5: There are Five houses P, Q , R, S, T . P is immediate right of Q and T is immediate left of R and immediate right of P . Q is right of S. Which house in the middle.

(a) P (b) Q (c) R (d) T

Solution: According to the question the houses can be arranged as follows. Assuming all houses are facing towards North.

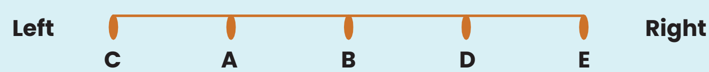


Therefore, house P is middle.

Example 6: Five friends are sitting on a bench. A is to the immediate left of B but on the immediate right of C, D is to the immediate right of B but on the immediate left of E. Who are at the extremes?

(a) A, B (b) A, D (c) C, E (d) B, D

Solution: Arrangements according to the question as follows. Assuming all students are facing towards North.

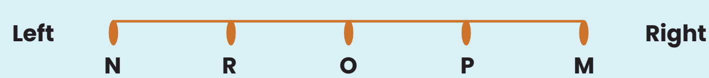


Clearly C and E are the extremes.

Example 7: In a college party, 5 girls are sitting in a row. P is to the immediate left of M and to the immediate right of O. R is sitting to the immediate right of N but to the left of O. Who is sitting in the middle?

(a) O (b) R (c) P (d) M

Solution: (a) arrangements of the question as follows.

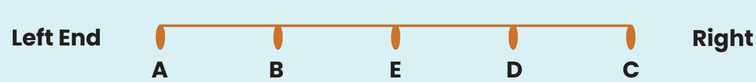


Therefore, O is sitting in the middle.

Example 8: Five boys A, B, C, D and E are standing in a row. D is on the immediate right of E, B is on the immediate left of E but on the immediate right of A. D is one the immediate left of C, who is standing on the extreme right. Who is standing in the middle?

(a) B (b) C (c) D (d) E

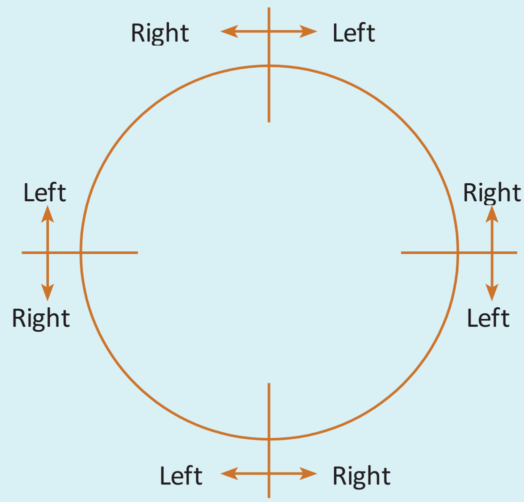
Solution: The sequence of Boys as follows. Assuming all boys are sitting towards North.



There E is standing in the middle.

Circular Arrangement:

In this arrangement, some persons are sitting around a circle and they are facing the centre.



1. Left movement is called clockwise rotation.
2. Right movement is called anti–clockwise rotation.
   (i) The above presentation is for 4 persons but for any number of persons, the direction is taken in the same manner.
   (ii) For rectangular and sequence arrangement, directions are taken as discussed in two rows sequence.

Example 9: (Q Nos. 1 to 3) **Study the following Question carefully and answer the given questions.**

Four ladies & A, B, C and D and Four Gentlemen E, F, G and H are sitting in a circle around a table facing each other .

I. No two ladies or gentlemen are sitting side by side.

II. C, who is sitting between G and E, is facing D.

III. F is between D and A and facing G.

IV. H is to the right of B.

(1) Who is sitting left of A?

(a) E (b) F (c) G (d) H

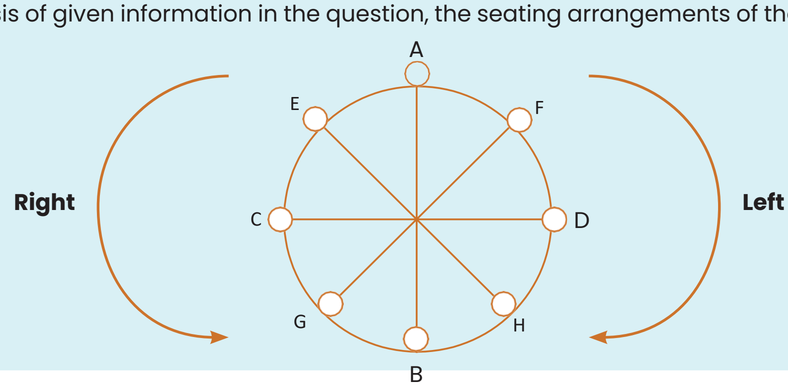
(2) E is facing whom?

(a) F (b) B (c) G (d) H

(3) Who is immediate neighbour of B?

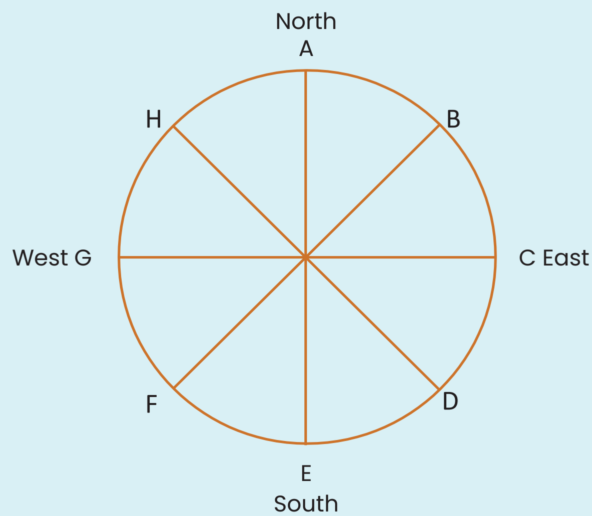
(a) G and H (b) E and F (c) E and H (d) F and H

Solution: On the basis of given information in the question, the seating arrangements of the persons are as follows.



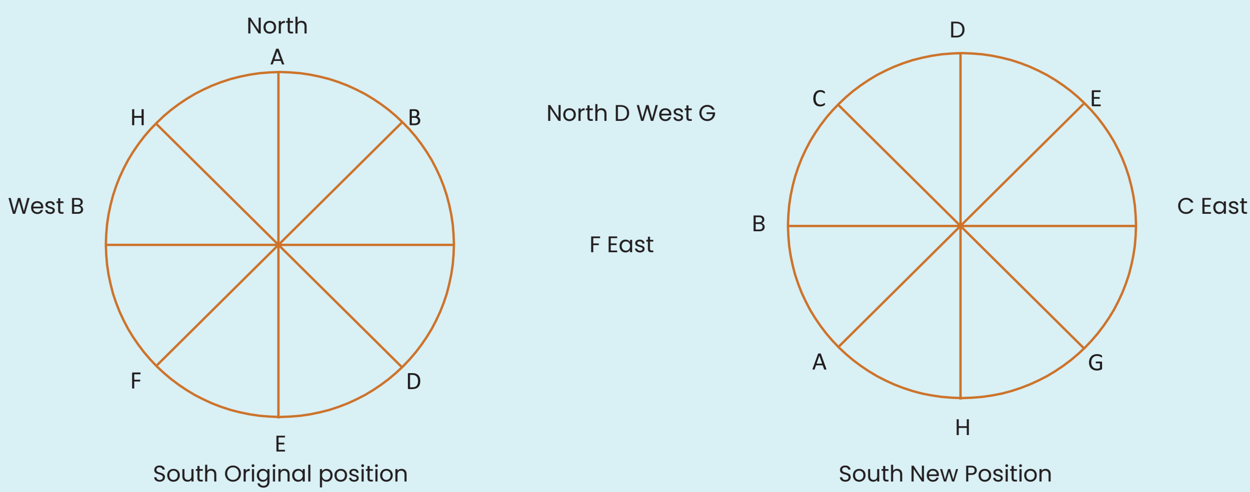
1) (b) Clearly , F is sitting left of A.
2) (d) Clearly E is facing H.
3) (a) G and H are neighbours of B.

Example 10: Eight persons A, B, C, D, E, F, G and H are sitting around the circle as given in the figure. They are facing the direction opposite to centre. If they move upto three places anti-clockwise, then.



(a) B will face West
(b) E will face East
(c) H will face North-West
(d) A will face South

Solution: Following Seating arrangement is formed from the given in formation.



Clearly B will Face west

Example 11: Five People A, B, C, D and E are seated about a round table. Every chair is spaced equidistant from adjacent chairs. (UPSC (CSAT) 2013)

I. C is seated next to A .

II. A is seated two seats from D.

III. B is not seated next to A.

Which of the following must be true?

I. D is seated next to B.

II. E is seated next to A.

Select the correct from the options given below:

(a) Only I

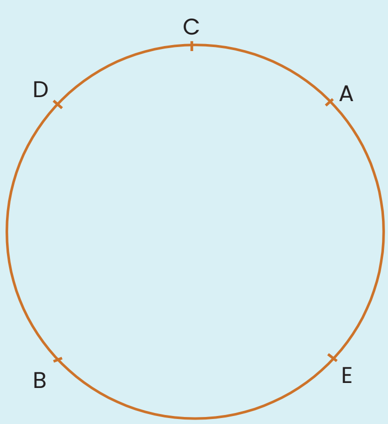
(b) Only II

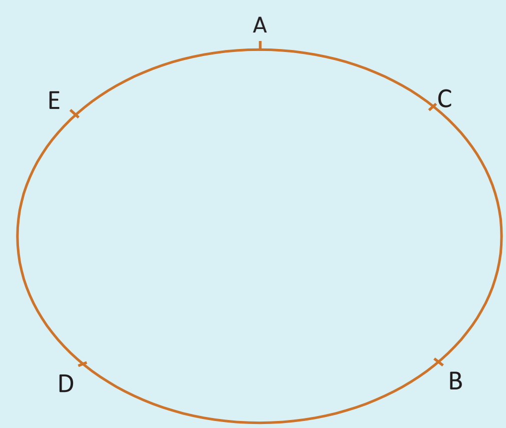
(c) Both I and II

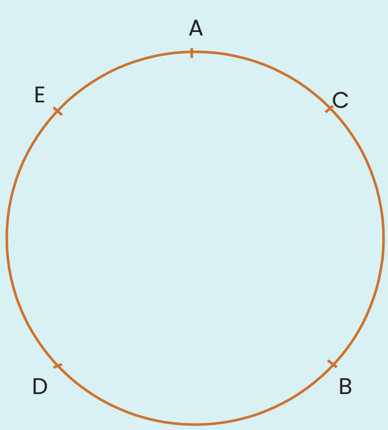
(d) Neither I nor II

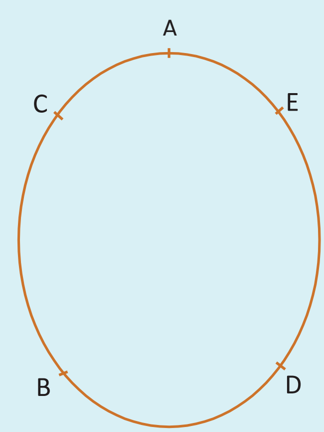
Solution:

According to the given information there are possible Seating arrangements.









From the above arrangements. It is clear that D is seated next to B . Also E is next to A.

Clearly both statements I and II are true.

Example 12: Study the following Question carefully and answer the given questions.

Eight friends A, B, C, D,E, F, G and H are sitting in a circle facing the centre, not necessarily in the same order. D sits third to the left of A. E sits to the immediate right of A. B is third to left of D. G is second to the right of B. C is neighbour of B. C is third to left of H. (GIC 2012)

1) Who amongst the following is sitting exactly between F and D?

   (a) C (b) E (c) H (d) A

2) Three of the following four are alike in a certain way based on the information given above and so form a group. Which is does not belong to that group.

   (a) DC (b) AH (c) EF (d) CB

3) Who amongst the following second to the left of H?

   (a) E (b) B (c) A (d) None of these

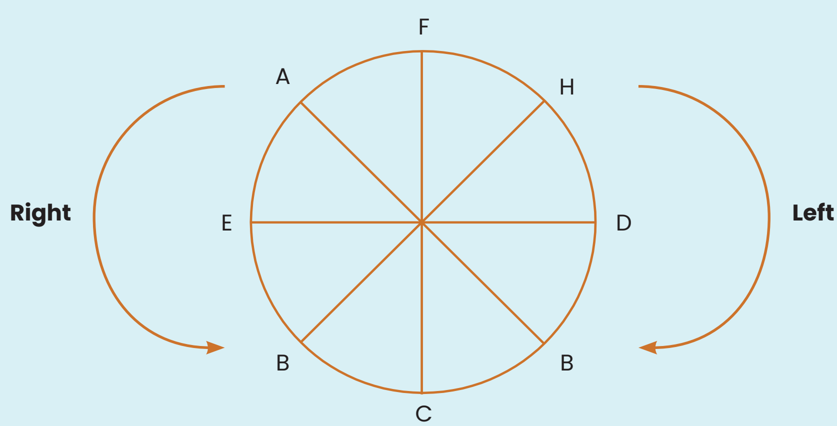
4) Who amongst the following are immediate neighbours of G?

   (a) CA (b) AF (c) DC (d) DF

5) Who amongst the following is sitting third to the right of A?

   (a) F (b) B (c) H (d) C

Solution: Arrangements according to the question is as follows.



1) (c), Clearly H is sitting exactly F and D
2) (d) DGC AFH EAF C none B

   Skipped Skipped Skipped No member is skipped in between

   So, CB does not belong to the group.
3) (d) Clearly, G is sitting second to the left of H
4) (c) Clearly D and C immediate neighbours of G
5) (d) Clearly, C is sitting third to the right of A

Chapter – 12:

Blood Relations

Learning Objectives

- Blood relations of a group of persons are given in jumbled form. In these tests, the questions which are asked in this section depend upon Relation.

12.1 Definition

A person who is related to another by birth rather than by marriage.

Prerequisites:

To remember easily the relations may be divided into two sides as given below:

(i) Relations of Paternal side:

Father's father → Grandfather

Father's mother → Grandmother

Father's brother → Uncle

Father's sister → Aunt

Children of uncle → Cousin

Wife of uncle → Aunt

Children of aunt → Cousin

Husband of aunt → Uncle

(ii) Relations of Maternal side:

Mother's father → Maternal grandfather

Mother's mother → Maternal grandmother

Mother's brother → Maternal uncle

Mother's sister → Aunt

Children of maternal uncle → Cousin

Wife of maternal uncle → Maternal aunty

Relations:

1.	Grandfather's son	Father or Uncle
2.	Grandmother's son	Father or Uncle
3.	Grandfather's only son	Father
4.	Grandmother's only son	Father
5.	Mother's or father's mother	Grandmother
6.	Son's wife	Daughter-in-Law
7.	Daughter's husband	Son-in-Law
8.	Husband's or wife's sister	Sister-in-Law
9.	Brother's son	Nephew
10.	Brother's daughter	Niece
11.	Uncle or aunt's son or daughter	Cousin
12.	Sister's husband	Brother-in-Law
13.	Brother's wife	Sister-in-Law
14.	Grandson's or granddaughter's daughter	Great grand Daughter

The efficiency in doing the problems of blood relations depends upon the knowledge of the blood relations. Some of the important relations are given below:

a) My mother's or father's son is my Brother.

b) My mother's or father's daughter is my Sister.

c) My mother's or father's father is my Grandfather.

d) My mother's or father's sister is my Aunt.

e) My mother's or father's brother is my Uncle.

f) My son's wife is my Daughter-in-law.

g) My daughter's husband is my Son-in-law.

h) My brother's son is my Nephew.

i) My brother's daughter is my Neice.

j) My sister's husband is my Brother-in-law.

k) My brother's wife is my Sister-in-law.

l) My husband's or wife's sister is my Sister-in-law.

m) My husband's or wife's brother is my Brother-in-law.

n) My uncle's or aunt's son or daughter is my Cousin.

o) My wife's father or husband's father is my Father-in-law.

p) My wife's mother or husband's mother is my Mother-in-law.

q) My father's wife is my Mother.

r) My mother's husband is my Father.

s) My son's or daughter's son is my Grandson.

t) My son's or daughter's daughter is my Grand-daughter.

Different types of questions with explanation:

(1) A is B's daughter, B is C's mother. D is C's brother. How is D related to A?

(a) Father (b) Grandfather (c) Brother (d) Son

Explanation: A is daughter of B.

B is mother of C, 'D' is Brother of 'A'

Therefore, D is Son of B. Thus 'D' is brother of 'A'

(2) P is Q's brother. R is Q's mother. S is R's father. T is S's mother. How is P related to T?

(a) Grand-daughter (b) Great grandson (c) Grandson (d) Grandmother

Explanation: P is brother of Q . Therefore, P is a male.

R is mother of P and Q and R is daughter of S. S is Son of T. S is grandfather of P. And T is great grand mother of P . Hence , P is great grand son of T .

(3) A is B's brother. C is D's father. E is B's mother. A and D are brothers. How is E related to C?

(a) Sister (b) Sister-in-law (c) Niece (d) Wife

Explanation: A is brother of B . Therefore, A is male.

C is father of D. Therefore, C is male.

E is mother of B. Therefore, E is Female. A and D are brothers.

Therefore, D is male.

Deductions:

(i) A and D are brothers of D

(ii) C is the father of A, B and D

(iii)C is the mother of A, B and D

Thus, E is wife of C

(4)A is the sister of B. B is the brother of C. C is the son of D. How is D related to A?

(a) Mother (b) Daughter (c) Son (d) Uncle

Explanation: (1) B is brother of C

C is son of D.

A is the sister of B and C. A is the daughter of 'D'.

According to the options given, we are left with no choice. But selection option (a) is correct.

(5) B is the brother of A. whose only sister is mother of C. D is maternal grandmother of C. How is A related to D?

(a) Daughter-in-law (b) Daughter (c) Aunt (d) Nephew

Explanation: Although sex of A is not mentioned clearly in the question. On the basis of information given is A is daughter of D.

(6) A and B are sisters. R and S are brothers. A's daughter is R's sister. What is B's relation to S?

(a) Mother (b) Grandmother (c) Sister (d) Aunt

Explanation: A's daughter is the sister of R and S. B is sister of A. B is aunt of S.

(7) E is the sister of B. A is the father of C. B is the son of C. How is A related to E?

(a) Grandfather (b) Grand-daughter (c) Father (d) Great-grandfather

Explanation: B' is the son of C and Grandson of A.

E is sister of B.

Therefore, A is Grandfather of E.

(8) Given that:

A is the mother of B.

C is the son of A.

D is the brother of E.

E is the daughter of B.

Who is grandmother of D?

(a) A (b) B (c) C (d) D

Explanation: E is the daughter of B and D is brother of E. Therefore, D is son B and A is son of A and C is son of A and A is mother of B.

Thus A, is Grandmother of D. Therefore, D is son of B.

(9) A is D' brother. D is B's father. B and C are sisters. How is A related to C?

(a) Son (b) Grandson (c) Father (d) Uncle

Explanation: B and C daughters of D . A is brother of D .

Therefore A is uncle of C.

(10) A is B's sister, C is B's mother, D is C's father, E is D's mother, then how A is related to D?

(a) Grandfather (b) Daughter (c) Grandmother (d) Granddaughter

Explanation: D is Father of C and C is mother of A and B

Thus, A is grandfather of D.

(11) (i) F is the brother of A.

(ii) C is the daughter of A.

(iii) K is the sister of F.

(iv) G is the brother of C.

Who is the uncle of G?

(a) A (b) C (c) K (d) F

Explanation: G is son of A and F is brother of A. Hence F is uncle of G

(12) A is father of C and D is son of B. E is brother of A. If C is sister of D how is B related to E?

(a) Sister-in-law (b) Sister (c) Brother (d) Brother-in-law

Explanation: C and D Children of A and B. B is mother of C and D.

Therefore, B is Sister-in-law of E.

(13) C is wife of B. E is the son of C, A is the brother of B and father of D. What is the relationship of E to D?

(b) Mother (b) Sister (c) Brother (d) Cousin

Explanation: E is son of B and C. A is uncle of E and Father of D.

Therefore E is cousin of D.

(14) M is the son of P. Q is the grand-daughter of O, who is the husband of P. How is M related to O?

(c) Son (b) Daughter (c) Mother (d) Father

Explanation: O is the Husband of P. M is the son of P. Therefore, M is son of O.

(15) X and Y are brothers. R is the father of Y. S is the brother of T and maternal uncle of X. What is T to R?

(d) Mother (b) Wife (c) Sister (d) Brother

Explanation: R is the Father of X and Y. S is the maternal uncle of X and Y.

Considering the option (b), T is wife of R.

Chapter 13:

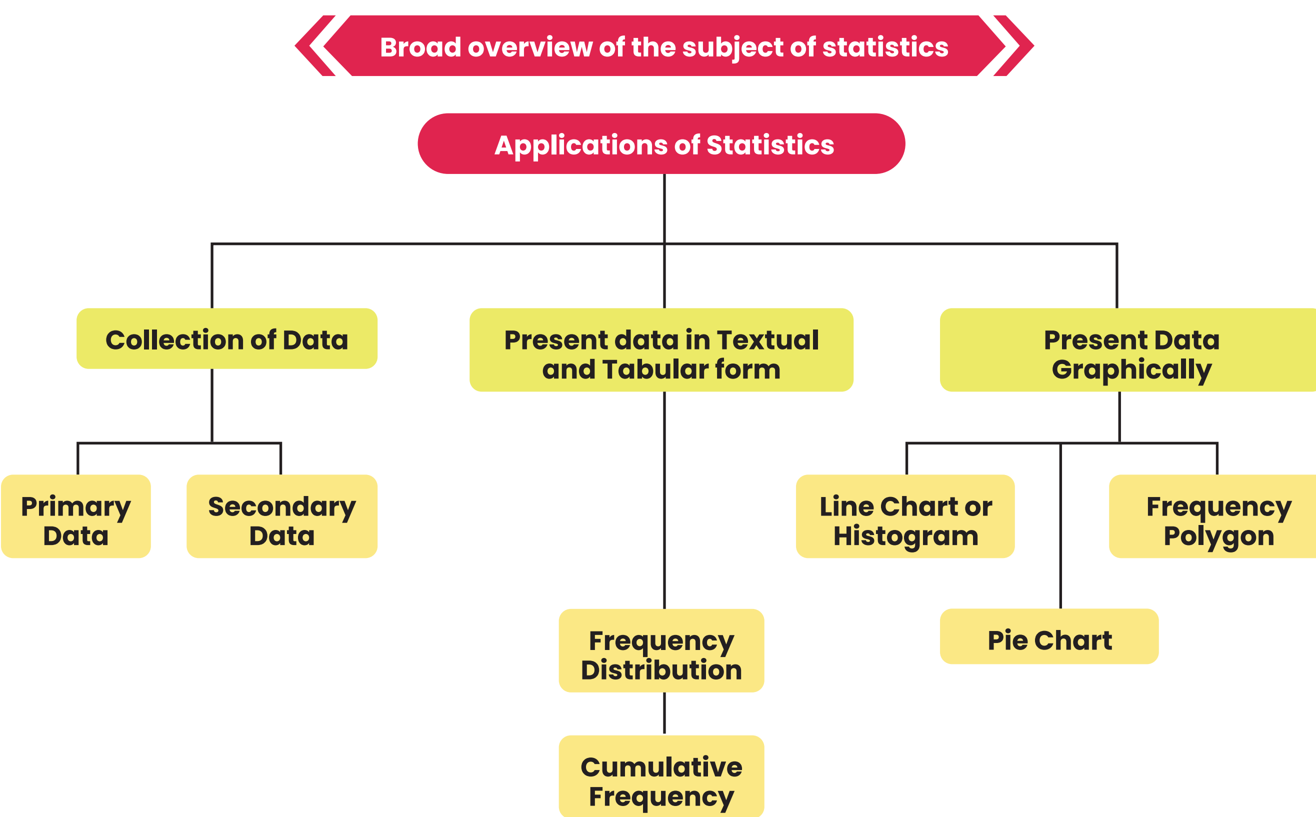
Unit-I: Statistical Descritpion of Data

Learning Objectives

After reading this chapter, students will be able to understand:

- Have a broad overview of the subject of statistics and application thereof.
- Know about data collection technique including the distinction of primary and secondary data.
- Know how to present data in textual and tabular format including the technique of creating frequency distribution and working out cumulative frequency.
- Know how to present data graphically using histogram, frequency polygon and pie chart.

Chapter Overview



Statistical Descritpion of Data

Introduction of Statistics

The modern development in the field of not only Management, Commerce, Economics, Social Sciences, Mathematics and so on but also in our life like public services, defence, banking, insurance sector, tourism and hospitality, police and military etc. are dependent on a particular subject known as statistics. Statistics does play a vital role in enriching a specific domain by collecting data in that field, analysing the data by applying various statistical techniques and finally making statistical inferences about the domain. In the present world, statistics has almost a universal application. Our Government applies statistics to make the economic planning in an effective and a pragmatic way. The businessman plan and expand their horizons of business on the basis of the analysis of the feedback data. The political parties try to impress the general public by presenting the statistics of their performances and accomplishments. Most of the research scholars of today also apply statistics to present their research papers in an authoritative manner. Thus the list of people using statistics goes on and on and on. Due to these factors, it is necessary to study the subject of statistics in an objective manner.

History of Statistics

Going through the history of ancient period and also that of medieval period, we do find the mention of statistics in many countries. However, there remains a question mark about the origin of the word 'statistics'. One view is that statistics is originated from the Latin word 'status'. According to another school of thought, it had its origin in the Italian word 'statista'. Some scholars believe that the German word 'statistik' was later changed to statistics and another suggestion is that the French word 'statistique' was made as statistics with the passage of time.

In those days, statistics was analogous to state or, to be more precise, the data that are collected and maintained for the welfare of the people belonging to the state. We are thankful to Kautilya who had kept a record of births and deaths as well as some other precious records in his famous book 'Arthashastra' during Chandragupta's reign in the fourth century B.C. During the reign of Akbar in the sixteenth century A.D. We find statistical records on agriculture in Ain-i-Akbari written by Abu Fazl. Referring to Egypt, the first census was conducted by the Pharaoh during 300 B.C. to 2000 B.C.

History of Statistics

Going through the history of ancient period and also that of medieval period, we do find the mention of statistics in many countries. However, there remains a question mark about the origin of the word 'statistics'. One view is that statistics is originated from the Latin word 'status'. According to another school of thought, it had its origin in the Italian word 'statista'. Some scholars believe that the German word 'statistik' was later changed to statistics and another suggestion is that the French word 'statistique' was made as statistics with the passage of time.

In those days, statistics was analogous to state or, to be more precise, the data that are collected and maintained for the welfare of the people belonging to the state. We are thankful to Kautilya who had kept a record of births and deaths as well as some other precious records in his famous book 'Arthashastra' during Chandragupta's reign in the fourth century B.C. During the reign of Akbar in the sixteenth century A.D. We find statistical records on agriculture in Ain-i-Akbari written by Abu Fazl. Referring to Egypt, the first census was conducted by the Pharaoh during 300 B.C. to 2000 B.C.

Definition of Statistics

We may define statistics either in a singular sense or in a plural sense Statistics, when used as a plural noun, may be defined as data qualitative as well as quantitative, that are collected, usually with a view of having statistical analysis.

However, statistics, when used as a singular noun, may be defined, as the scientific method that is employed for collecting, analysing and presenting data, leading finally to drawing statistical inferences about some important characteristics it means it is 'science of counting' or 'science of averages'.

Application of statistics

Among various applications of statistics, let us confine our discussions to the fields of Economics, Business Management and Commerce and Industry.

Economics

Modern developments in Economics have the roots in statistics. In fact, Economics and Statistics are closely associated. Time Series Analysis, Index Numbers, Demand Analysis etc. are some overlapping areas of Economics and Statistics. In this connection, we may also mention Econometrics – a branch of Economics that interact with statistics in a very positive way. Conducting socio-economic surveys and analysing the data derived from it are made with the help of different statistical methods. Regression analysis, one of the numerous applications of statistics, plays a key role in Economics for making future projection of demand of goods, sales, prices, quantities etc. which are all ingredients of Economic planning.

Business Management

Gone are the days when the managers used to make decisions on the basis of hunches, intuition or trials and errors. Now a days, because of the never-ending complexity in the business and industry environment, most of the decision making processes rely upon different quantitative techniques which could be described as a combination of statistical methods and operations research techniques. So far as statistics is concerned, inferences about the universe from the knowledge of a part of it, known as sample, plays an important role in the development of certain criteria. Statistical decision theory is another component of statistics that focuses on the analysis of complicated business strategies with a list of alternatives – their merits as well as demerits.

Statistics in Commerce and Industry

In this age of cut-throat competition, like the modern managers, the industrialists and the businessmen are expanding their horizons of industries and businesses with the help of statistical procedures. Data on previous sales, raw materials, wages and salaries, products of identical nature of other factories etc are collected, analysed and experts are consulted in order to maximise profits. Measures of central tendency and dispersion, correlation and regression analysis, time series analysis, index numbers, sampling, statistical quality control are some of the statistical methods employed in commerce and industry.

Limitations of Statistics

Before applying statistical methods, we must be aware of the following limitations:

I Statistics deals with the aggregates. An individual, to a statistician has no significance except the fact that it is a part of the aggregate.

II Statistics is concerned with quantitative data. However, qualitative data also can be converted to quantitative data by providing a numerical description to the corresponding qualitative data.

III Future projections of sales, production, price and quantity etc. are possible under a specific set of conditions. If any of these conditions is violated, projections are likely to be inaccurate.

IV The theory of statistical inferences is built upon random sampling. If the rules for random sampling are not strictly adhered to, the conclusion drawn on the basis of these unrepresentative samples would be erroneous. In other words, the experts should be consulted before deciding the sampling scheme.

Statistical Descritpion of Data

Collection of Data

We may define 'data' as quantitative information about some particular characteristic(s) under consideration. Although a distinction can be made between a qualitative characteristic and a quantitative characteristic but so far as the statistical analysis of the characteristic is concerned, we need to convert qualitative information to quantitative information by providing a numeric descriptions to the given characteristic. In this connection, we may note that a quantitative characteristic is known as a variable or in other words, a variable is a measurable quantity. Again, a variable may be either discrete or continuous. When a variable assumes a finite or a countably infinite number of isolated values, it is known as a discrete variable. Examples of discrete variables may be found in the number of petals in a flower, the number of misprints a book contains, the number of road accidents in a particular locality and so on. A variable, on the other hand, is known to be continuous if it can assume any value from a given interval. Examples of continuous variables may be provided by height, weight, sale, profit and so on. Finally, a qualitative characteristic is known as an attribute. The gender of a baby, the nationality of a person, the colour of a flower etc. are examples of attributes.

We can broadly classify data as

(a) Primary;

(b) Secondary.

Collection of data plays the very important role for any statistical analysis. The data which are collected for the first time by an investigator or agency are known as primary data whereas the data are known to be secondary if the data, as being already collected, are used by a different person or agency. Thus, if Prof. Das collects the data on the height of every student in his class, then these would be primary data for him. If, however, another person, say, Professor Bhargava uses the data, as collected by Prof. Das, for finding the average height of the students belonging to that class, then the data would be secondary for Prof. Bhargava.

Collection of Primary Data

The following methods are employed for the collection of primary data:

(i) Interview method;

(ii) Mailed questionnaire method;

(iii) Observation method;

(iv) Questionnaries filled and sent by enumerators.

Interview method again could be divided into (a) Personal Interview method, (b) Indirect Interview method and (c) Telephone Interview method.

In personal interview method, the investigator meets the respondents directly and collects the required information then and there from them. In case of a natural calamity like a super cyclone or an earthquake or an epidemic like plague, we may collect the necessary data much more quickly and accurately by applying this method.

If there are some practical problems in reaching the respondents directly, as in the case of a rail accident, then we may take recourse for conducting Indirect Interview where the investigator collects the necessary information from the persons associated with the problems.

Telephone interview method is a quick and rather non-expensive way to collect the primary data where the relevant information can be gathered by the researcher himself by contacting the interviewee over the phone. The first two methods, though more accurate, are inapplicable for covering a large area whereas the telephone interview, though less consistent, has a wide coverage.

The amount of non-responses is maximum for this third method of data collection.

Mailed questionnaire method comprises of framing a well-drafted and soundly-sequenced questionnaire covering all the important aspects of the problem under consideration and sending them to the respondents with pre-paid stamp after providing all the necessary guidelines for filling up the questionnaire. Although a wide area can be covered using the mailed questionnaire method, the amount of non-responses is likely to be maximum in this method.

In observation method, data are collected, as in the case of obtaining the data on the height and weight of a group of students, by direct observation or using instrument. Although this is likely to be the best method for data collection, it is time consuming, laborious and covers only a small area. Questionnaire form of data collection is used for larger enquiries from the persons who are surveyed. Enumerators collects information directly by interviewing the persons having information : Question are explained and hence data is collected.

Sources of Secondary Data

There are many sources of getting secondary data. Some important sources are listed below:

(a) International sources like WHO, ILO, IMF, World Bank etc.

(b) Government sources like Statistical Abstract by CSO, Indian Agricultural Statistics by the Ministry of Food and Agriculture and so on.

(c) Private and quasi-government sources like ISI, ICAR, NCERT etc.

(d) Unpublished sources of various research institutes, researchers etc.

Scrutiny of Data

Since the statistical analyses are made only on the basis of data, it is necessary to check whether the data under consideration are accurate as well as consistence. No hard and fast rules can be recommended for the scrutiny of data. One must apply his intelligence, patience and experience while scrutinising the given information.

Errors in data may creep in while writing or copying the answer on the part of the enumerator. A keen observer can easily detect that type of error. Again, there may be two or more series of figures which are in some way or other related to each other. If the data for all the series are provided, they may be checked for internal consistency. As an example, if the data for population, area and density for some places are given, then we may verify whether they are internally consistent by examining whether the relation

$$\text{Density} = \frac{\text{Area}}{\text{Population}} \text{ holds}$$

A good statistician can also detect whether the returns submitted by some enumerators are exactly of the same type thereby implying the lack of seriousness on the part of the enumerators. The bias of the enumerator also may be reflected by the returns submitted by him. This type of error can be rectified by asking the enumerator(s) to collect the data for the disputed cases once again.

Presentation Of Data

Once the data are collected and verified for their homogeneity and consistency, we need to present them in a neat and condensed form highlighting the essential features of the data. Any statistical analysis is dependent on a proper presentation of the data under consideration.

Classification or Organisation of Data

It may be defined as the process of arranging data on the basis of the characteristic under consideration into a number of groups or classes according to the similarities of the observations. Following are the objectives of classification of data:

(a) It puts the data in a neat, precise and condensed form so that it is easily understood and interpreted.

(b) It makes comparison possible between various characteristics, if necessary, and thereby finding the association or the lack of it between them.

(c) Statistical analysis is possible only for the classified data.

(d) It eliminates unnecessary details and makes data more readily understandable. Data may be classified as -

   (i) Chronological or Temporal or Time Series Data;

   (ii) Geographical or Spatial Series Data;

   (iii) Qualitative or Ordinal Data;

   (iv) Quantitative or Cardinal Data.

When the data are classified in respect of successive time points or intervals, they are known as time series data. The number of students appeared for CA final for the last twenty years, the production of a factory per month from 2000 to 2015 etc. are examples of time series data.

Data arranged region wise are known as geographical data. If we arrange the students appeared for CA final in the year 2015 in accordance with different states, then we come across Geographical Data.

Data classified in respect of an attribute are referred to as qualitative data. Data on nationality, gender, smoking habit of a group of individuals are examples of qualitative data. Lastly, when the data are classified in respect of a variable, say height, weight, profits, salaries etc., they are known as quantitative data.

Data may be further classified as frequency data and non-frequency data. The qualitative as well as quantitative data belong to the frequency group whereas time series data and geographical data belong to the non-frequency group.

Mode of Presentation of Data

Next, we consider the following mode of presentation of data:

(a) Textual presentation;

(b) Tabular presentation or Tabulation;

(c) Diagrammatic representation.

(a) Textual presentation

This method comprises presenting data with the help of a paragraph or a number of paragraphs. The official report of an enquiry commission is usually made by textual presentation. Following is an example of textual presentation.

'In 2009, out of a total of five thousand workers of Roy Enamel Factory, four thousand and two hundred were members of a Trade Union. The number of female workers was twenty per cent of the total workers out of which thirty per cent were members of the Trade Union.

In 2010, the number of workers belonging to the trade union was increased by twenty per cent as compared to 2009 of which four thousand and two hundred were male. The number of workers not belonging to trade union was nine hundred and fifty of which four hundred and fifty were females.'

The merit of this mode of presentation lies in its simplicity and even a layman can present data by this method. The observations with exact magnitude can be presented with the help of textual presentation. Furthermore, this type of presentation can be taken as the first step towards the other methods of presentation.

Textual presentation, however, is not preferred by a statistician simply because, it is dull, monotonous and comparison between different observations is not possible in this method. For manifold classification, this method cannot be recommended.

(b) Tabular presentation or Tabulation

Tabulation may be defined as systematic presentation of data with the help of a statistical table having a number of rows and columns and complete with reference number, title, description of rows as well as columns and foot notes, if any.

We may consider the following guidelines for tabulation :

I A statistical table should be allotted a serial number along with a self-explanatory title.

II The table under consideration should be divided into caption, Box-head, Stub and Body. Caption is the upper part of the table, describing the columns and sub-columns, if any. The Box-head is the entire upper part of the table which includes columns and sub-column numbers, unit(s) of measurement along with caption. Stub is the left part of the table providing the description of the rows. The body is the main part of the table that contains the numerical figures.

III The table should be well-balanced in length and breadth.

IV The data must be arranged in a table in such a way that comparison(s) between different figures are made possible without much labour and time. Also the row totals, column totals, the units of measurement must be shown.

V The data should be arranged intelligently in a well-balanced sequence and the presentation of data in the table should be appealing to the eyes as far as practicable.

VI Notes describing the source of the data and bringing clarity and, if necessary, about any rows or columns known as footnotes, should be shown at the bottom part of the table.

The textual presentation of data, relating to the workers of Roy Enamel Factory is shown in the following table.

Table 13.1

Status of the workers of Roy Enamel factory on the basis of their trade union membership for 2009 and 2010.

Status	Member of TU			Non-member			Total		
Year	M (1)	F (2)	T (3)=(1)+(2)	M (4)	F (5)	T (6)=(4)+ (5)	M (7)	F (8)	T (9)=(7)+(8)
2009 2010	3900 4200	300 840	4200 5040	300 500	500 450	800 950	4200 4700	800 1290	5000 5990

Source:

Footnote: TU, M, F and T stand for trade union, male, female and total respectively.

The tabulation method is usually preferred to textual presentation as

(i) It facilitates comparison between rows and columns.

(ii) Complicated data can also be represented using tabulation.

(iii) It is a must for diagrammatic representation.

(iv) Without tabulation, statistical analysis of data is not possible.

(c) Diagrammatic representation of data

Another alternative and attractive representation of statistical data is provided by charts, diagrams and pictures. Unlike the first two methods of representation of data, diagrammatic representation can be used for both the educated section and uneducated section of the society. Furthermore, any hidden trend present in the given data can be noticed only in this mode of representation. However, compared to tabulation, this is less accurate. So if there is a priority for accuracy, we have to recommend tabulation.

We are going to consider the following types of diagrams :

I Line diagram or Historiagram;

II Bar diagram;

III Pie chart.

I Line diagram or Historiagram

When the data vary over time, we take recourse to line diagram. In a simple line diagram, we plot each pair of values of (t, yt), yt representing the time series at the time point t in the t–yt plane. The plotted points are then joined successively by line segments and the resulting chart is known as line-diagram.

When the time series exhibit a wide range of fluctuations, we may think of logarithmic or ratio chart where Log yt and not yt is plotted against t. We use Multiple line chart for representing two or more related time series data expressed in the same unit and multiple – axis chart in somewhat similar situations if the variables are expressed in different units.

II Bar diagram

There are two types of bar diagrams namely, Horizontal Bar diagram and Vertical Bar diagram. While horizontal bar diagram is used for qualitative data or data varying over space, the vertical bar diagram is associated with quantitative data or time series data. Bars

i.e. rectangles of equal width and usually of varying lengths are drawn either horizontally or vertically. We consider Multiple or Grouped Bar diagrams to compare related series. Component or sub-divided Bar diagrams are applied for representing data divided into a number of components. Finally, we use Divided Bar charts or Percentage Bar diagrams for comparing different components of a variable and also the relating of the components to the whole. For this situation, we may also use Pie chart or Pie diagram or circle diagram.

Illustrations

Example 13.1:
The profits in lakhs of Rupees of an industrial house for 2009, 2010, 2011, 2012, 2013, 2014, and 2015 are 5, 8, 9, 6, 12, 15 and 24 respectively. Represent these data using a suitable diagram.

Solution:
We can represent the profits for 7 consecutive years by drawing either a line chart or a vertical bar chart. Fig. 13.1 shows a line chart and figure 13.2 shows the corresponding vertical bar chart.

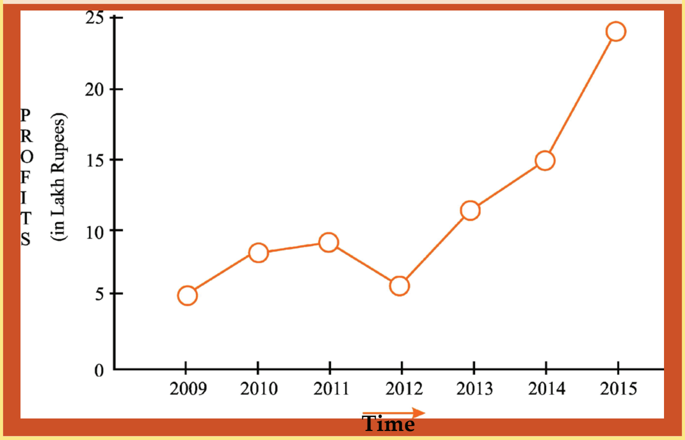


Figure 13.1

Showing line chart for the Profit of an Industrial House during 2009 to 2015.

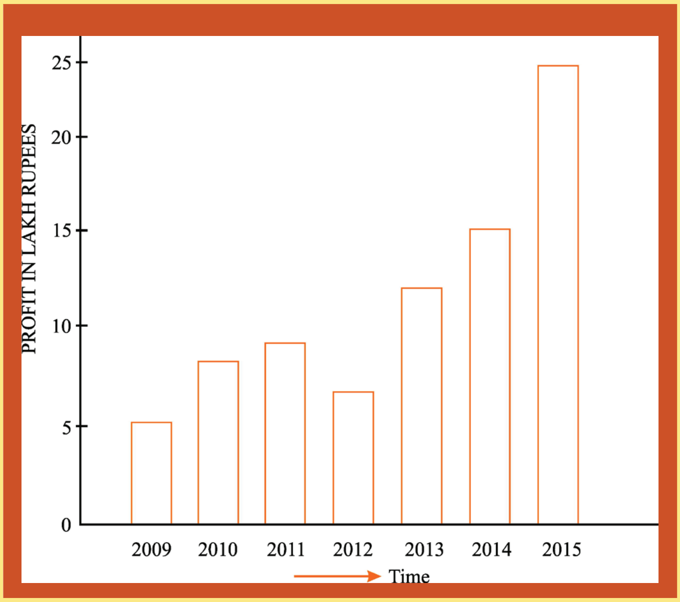


Figure 13.2

Showing vertical bar diagram for the Profit of an Industrial house from 2009 to 2015.

Example 13.2:

The production of wheat and rice of a region are given below :

Year	Production in metric tones	
	Wheat	Rice
2012	12	25
2013	15	30
2014	18	32
2015	19	36

Represent this information using a suitable diagram.

Solution:

We can represent this information by drawing a multiple line chart. Alternately, a multiple bar diagram may be considered. These are depicted in figure 13.3 and 13.4 respectively.

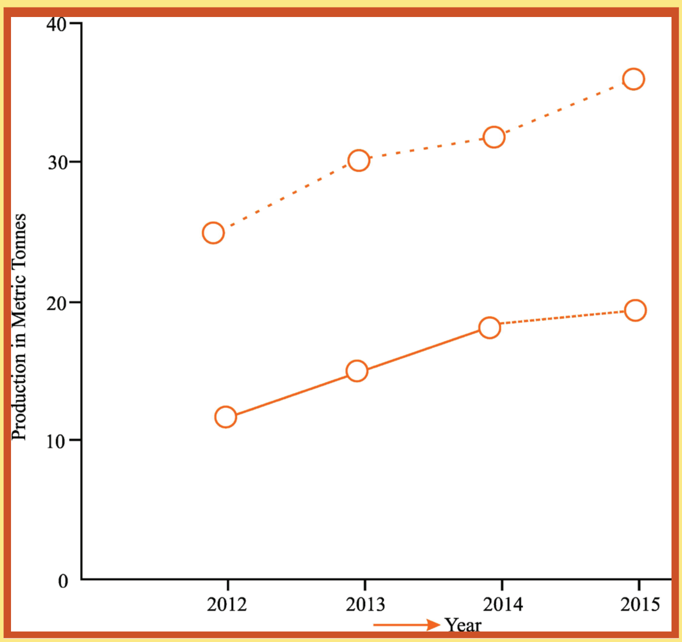


Figure 13.3

Multiple line chart showing production of wheat and rice of a region during 2012–2015. (Dotted line represent production of rice and continuous line that of wheat).

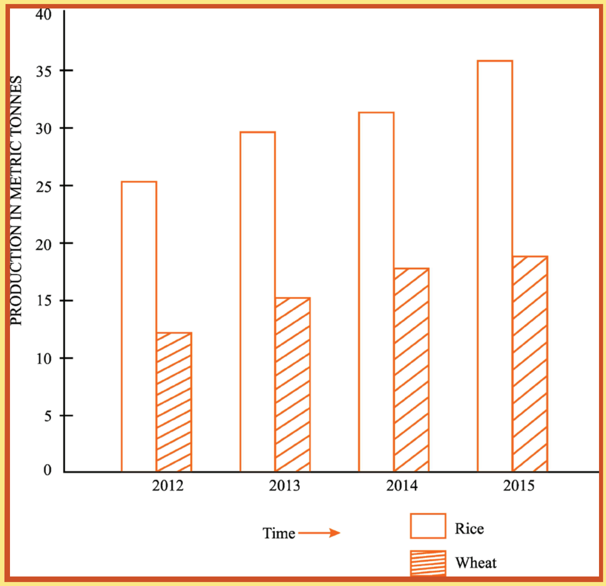


Figure 13.4

Multiple bar chart representing production of rice and wheat from 2012 to 2015.

Example 13.3:

Draw an appropriate diagram with a view to represent the following data :

Source	Revenue in millions of (`)
Customs	80
Excise	190
Income Tax	160
Corporate Tax	75
Miscellaneous	35

Solution:

Pie chart or divided bar chart would be the ideal diagram to represent this data. We consider Pie chart.

Table 13.2

Computation for drawing Pie chart

Source (1)	Revenue in Million rupees (2)	Central angle = $\frac{(2)}{\text{Total of } (2)} \times 360^\circ$
Customs	80	$\frac{80}{540} \times 360^\circ = 53^\circ$ (approx.)
Excise	190	$\frac{190}{540} \times 360^\circ = 127^\circ$
Income Tax	160	$\frac{160}{540} \times 360^\circ = 107^\circ$
Corporate Tax	75	$\frac{75}{540} \times 360^\circ = 50^\circ$
Miscellaneous	35	$\frac{35}{540} \times 360^\circ = 23^\circ$
Total	540	360°

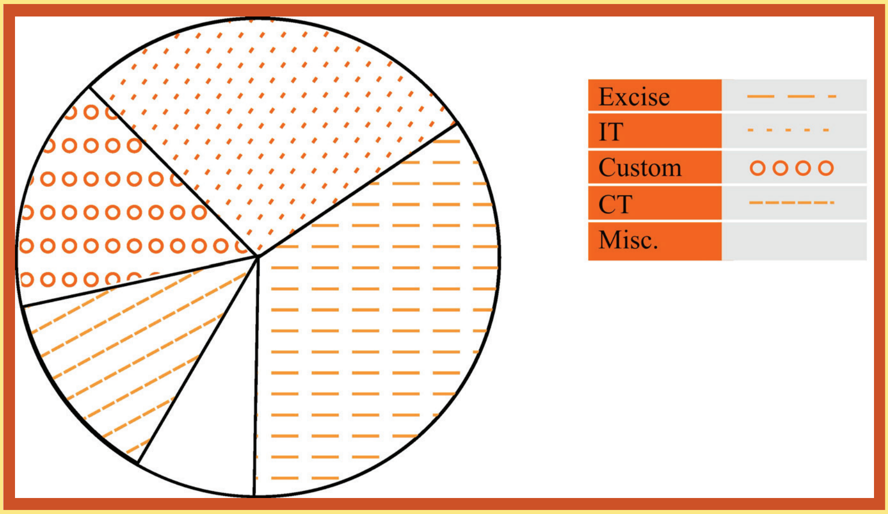


Figure 13.5

Pie chart showing the distribution of Revenue

13.4 Frequency Distribution

As discussed in the previous section, frequency data occur when we classify statistical data in respect of either a variable or an attribute. A frequency distribution may be defined as a tabular representation of statistical data, usually in an ascending order, relating to a measurable characteristic according to individual value or a group of values of the characteristic under study.

In case, the characteristic under consideration is an attribute, say nationality, then the tabulation is made by allotting numerical figures to the different classes the attribute may belong like, in this illustration, counting the number of Indian, British, French, German and so on. The qualitative characteristic is divided into a number of categories or classes which are mutually exclusive and exhaustive and the figures against all these classes are recorded. The figure corresponding to a particular class, signifying the number of times or how frequently a particular class occurs is known as the frequency of that class. Thus, the number of Indians, as found from the given data, signifies the frequency of the Indians. So frequency distribution is a statistical table that distributes the total frequency to a number of classes.

When tabulation is done in respect of a discrete random variable, it is known as Discrete or Ungrouped or simple Frequency Distribution and in case the characteristic under consideration is a continuous variable, such a classification is termed as Grouped Frequency Distribution. In case of a grouped frequency distribution, tabulation is done not against a single value as in the case of an attribute or a discrete random variable but against a group of values. The distribution of the number of car accidents in Delhi during 12 months of the year 2005 is an example of a ungrouped frequency distribution and the distribution of heights of the students of St. Xavier's College for the year 2004 is an example of a grouped frequency distribution.

Example 13.4:

Following are the records of babies born in a nursing home in Bangalore during a week (B denoting Boy and G for Girl) :

B	G	G	B	G	G	B	B	G	G
G	G	B	B	B	G	B	B	G	B
B	B	G	B	B	B	G	G	B	G

Construct a frequency distribution according to gender.

Solution:

In order to construct a frequency distribution of babies in accordance with their gender, we count the number of male births and that of female births and present this information in the following table.

Table 13.3

Frequency distribution of babies according to Gender

Category	Number of births
Boy (B)	16
Girl (G)	14
Total	30

Frequency Distribution of a Variable

For the construction of a frequency distribution of a variable, we need to go through the following steps :

I Find the largest and smallest observations and obtain the difference between them, known as Range, in case of a continuous variable.

II Form a number of classes depending on the number of isolated values assumed by a discrete variable. In case of a continuous variable, find the number of class intervals using the relation, No. of class Interval × class length ≅ Range.

III Present the class or class interval in a table known as frequency distribution table.

IV Apply 'tally mark' i.e. a stroke against the occurrence of a particulars value in a class or class interval.

V Count the tally marks and present these numbers in the next column, known as frequency column, and finally check whether the total of all these class frequencies tally with the total number of observations.

Example 13.5:

A review of the first 30 pages of a statistics book reveals the following printing mistakes:

0	1	3	3	2	5	6	0	1	0
4	1	1	0	2	3	2	5	0	4
2	3	2	2	3	3	4	6	1	4

Make a frequency distribution of printing mistakes.

Solution:

Since x, the printing mistakes, is a discrete variable, x can assume seven values 0, 1, 2, 3, 4, 5 and 6. Thus we have 7 classes, each class comprising a single value.

Table 13.4

Frequency Distribution of the number of printing mistakes of the first 30 pages of a book

Printing Mistake	Tally marks	Frequency (No. of Pages)
0	𝍸	5
1	𝍸	5
2	𝍸 I	6
3	𝍸 I	6
4	IIII	4
5	II	2
6	II	2
Total	–	30

Example 13.6:

Following are the weights in kgs. of 36 BBA students of St. Xavier's College.

70	73	49	61	61	47	57	50	59
59	68	45	55	65	68	56	68	55
70	70	57	44	69	73	64	49	63
65	70	65	62	64	73	67	60	50

Construct a frequency distribution of weights, taking class length as 5.

Solution:

We have, Range = Maximum weight – Minimum weight

= 73 kgs. – 44 kgs.

= 29 kgs.

No. of class interval × class lengths $\cong$ Range

$\Rightarrow$ No. of class interval × 5 $\cong$ 29

$\Rightarrow$ No. of class interval = $\frac{29}{5} \cong 6.$

(We always take the next integer as the number of class intervals so as to include both the minimum and maximum values).

Table 13.5

Frequency Distribution of weights of 36 BBA Students

Weight in kg (Class Interval)	Tally marks	No. of Students (Frequency)
44-48	III	3
49-53	IIII	4
54-58	𝍸	5
59-63	𝍸 II	7
64-68	𝍸 IIII	9
69-73	𝍸 III	8
Total	–	36

Some important terms associated with a frequency distribution

Class Limit (CL)

Corresponding to a class interval, the class limits may be defined as the minimum value and the maximum value the class interval may contain. The minimum value is known as the lower class limit (LCL) and the maximum value is known as the upper class limit (UCL). For the frequency distribution of weights of BBA Students, the LCL and UCL of the first class interval are 44 kgs. and 48 kgs. respectively.

Class Boundary (CB)

Class boundaries may be defined as the actual class limit of a class interval. For overlapping classification or mutually exclusive classification that excludes the upper class limits like 10–20, 20–30, 30–40, ... etc. the class boundaries coincide with the class limits. This is usually done for a continuous variable. However, for non-overlapping or mutually inclusive classification that includes both the class limits like 0–9, 10–19, 20–29,...... which is usually applicable for a discrete variable, we have

$$\text{LCB} = \text{LCL} - \frac{D}{2}$$

and $\text{UCB} = \text{UCL} + \frac{D}{2}$

where D is the difference between the LCL of the next class interval and the UCL of the given class interval. For the data presented in table 10.5, LCB of the first class interval

$= 44 \text{ kgs.} - \frac{(49 \times 48)}{2} \text{ kgs.}$

$= 43.50$ kgs.

and the corresponding UCB

$= 48 \text{ kgs.} + \frac{(49 \times 48)}{2} \text{ kgs.}$

$= 48.50$ kgs.

Mid-point or Mid-value or class mark

Corresponding to a class interval, this may be defined as the total of the two class limits or class boundaries to be divided by 2. Thus, we have

$$\text{mid-point} = \frac{\text{LCL} + \text{UCL}}{2}$$

$$= \frac{\text{LCB} + \text{UCB}}{2}$$

Referring to the distribution of weight of BBA students, the mid-points for the first two class intervals are

$\frac{44 \text{ kgs.} + 48 \text{ kgs.}}{2}$ and $\frac{49 \text{ kgs.} + 53 \text{ kgs.}}{2}$

i.e. 46 kgs. and 51 kgs. respectively.

Width or size of a class interval

The width of a class interval may be defined as the difference between the UCB and the LCB of that class interval. For the distribution of weights of BBA students, C, the class length or width is 48.50 kgs. – 43.50 kgs. = 5 kgs. for the first class interval. For the other class intervals also, C remains same.

Cumulative Frequency

The cumulative frequency corresponding to a value for a discrete variable and corresponding to a class boundary for a continuous variable may be defined as the number of observations less than the value or less than or equal to the class boundary. This definition refers to the less than cumulative frequency. We can define more than cumulative frequency in a similar manner. Both types of cumulative frequencies are shown in the following table.

Table 13.6

Cumulative Frequency Distribution of weights of 36 BBA students

Weight in kg (CB)	Cumulative Frequency	
	Less than	More than
43.50	0	33 + 3 or 36
48.50	0 + 3 or 3	29 + 4 or 33
53.50	3 + 4 or 7	24 + 5 or 29
58.50	7 + 5 or 12	17 + 7 or 24
63.50	12 + 7 or 19	8 + 9 or 17
68.50	19 + 9 or 28	0 + 8 or 8
73.50	28 + 8 or 36	0

Frequency density of a class interval

It may be defined as the ratio of the frequency of that class interval to the corresponding class length. The frequency densities for the first two class intervals of the frequency distribution of weights of BBA students are 3/5 and 4/5 i.e. 0.60 and 0.80 respectively.

Relative frequency and percentage frequency of a class interval

Relative frequency of a class interval may be defined as the ratio of the class frequency to the total frequency. Percentage frequency of a class interval may be defined as the ratio of class frequency to the total frequency, expressed as a percentage. For the last example, the relative frequencies for the first two class intervals are 3/36 and 4/36 respectively and the percentage frequencies are 300/36 and 400/36 respectively. It is quite obvious that whereas the relative frequencies add up to unity, the percentage frequencies add up to one hundred.

13.5 Graphical Representation of A Frequency Distribution

We consider the following types of graphical representation of frequency distribution :

(i) Histogram or Area diagram;

(ii) Frequency Polygon;

(iii)Ogives or cumulative Frequency graphs.

(i) Histogram or Area diagram

This is a very convenient way to represent a frequency distribution. Histogram helps us to get an idea of the frequency curve of the variable under study. Some statistical measure can be obtained using a histogram. A comparison among the frequencies for different class intervals is possible in this mode of diagrammatic representation.

In order to draw a histogram, the class limits are first converted to the corresponding class boundaries and a series of adjacent rectangles, one against each class interval, with the class

interval as base or breadth and the frequency or frequency density usually when the class intervals are not uniform as length or altitude, is erected. The histogram for the distribution of weight of 36 BBA students is shown below. The mode of the weights has also been determined using the histogram.

i.e. Mode = 66.50 kgs.

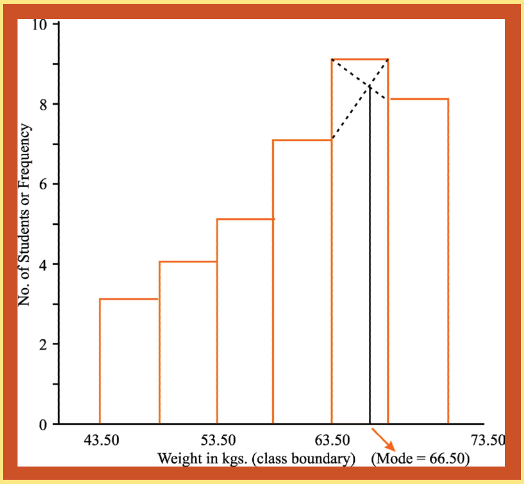


Figure 13.6

Showing histogram for the distribution of weight of 36 BBA students

(ii) Frequency Polygon

Usually frequency polygon is meant for single frequency distribution. However, we also apply it for grouped frequency distribution provided the width of the class intervals remains the same. A frequency curve can be regarded as a limiting form of frequency polygon. In order to draw a frequency polygon, we plot (x_i, f_i) for i = 1, 2, 3, n with x_i denoting the mid-point of the its class interval and f_i, the corresponding frequency, n being the number of class intervals. The plotted points are joined successively by line segments and the figure, so drawn, is given the shape of a polygon, a closed figure, by joining the two extreme ends of the drawn figure to two additional points (x_0,0) and (x_{n+1},0).

The frequency polygon for the distribution of weights of BBA students is shown in Figure 13.7. We can also obtain a frequency polygon starting with a histogram by adding the mid- points of the upper sides of the rectangles successively and then completing the figure by joining the two ends as before.

Mid-points	No. of Students(Frequency)
46	3
51	4
56	5
61	7
66	9
71	8

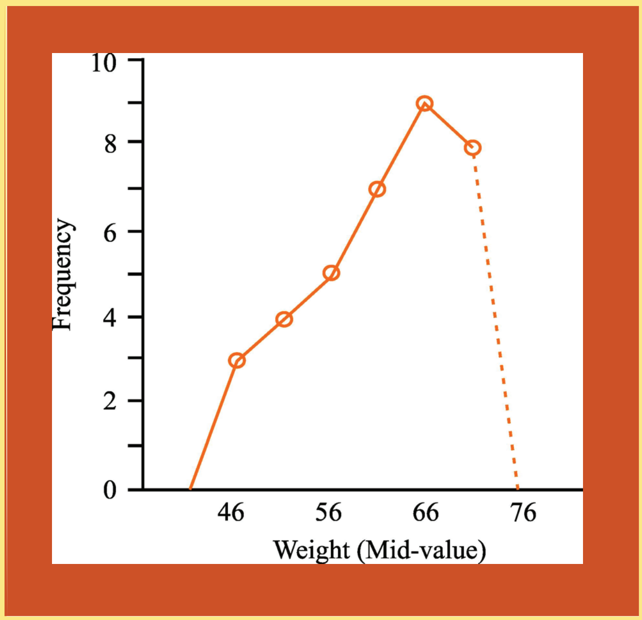


Figure 13.7

Showing frequency polygon for the distribution of height of 36 BBA students

(iii) Ogives or Cumulative Frequency Graph

By plotting cumulative frequency against the respective class boundary, we get ogives. As such there are two ogives – less than type ogives, obtained by taking less than cumulative frequency on the vertical axis and more than type ogives by plotting more than type cumulative frequency on the vertical axis and thereafter joining the plotted points successively by line segments. Ogives may be considered for obtaining quartiles graphically. If a perpendicular is drawn from the point of intersection of the two ogives on the horizontal axis, then the x-value of this point gives us the value of median, the second or middle quartile. Ogives further can be put into use for making short term projections.

Figure 13.8 depicts the ogives and the determination of the quartiles. This figure give us the following information.

1st quartile or lower quartile (Q_1) = 55 kgs.

2nd quartile or median (Q_2 or Me) = 62.50 kgs.

3rd quartile or upper quartile (Q_3) = 68 kgs.

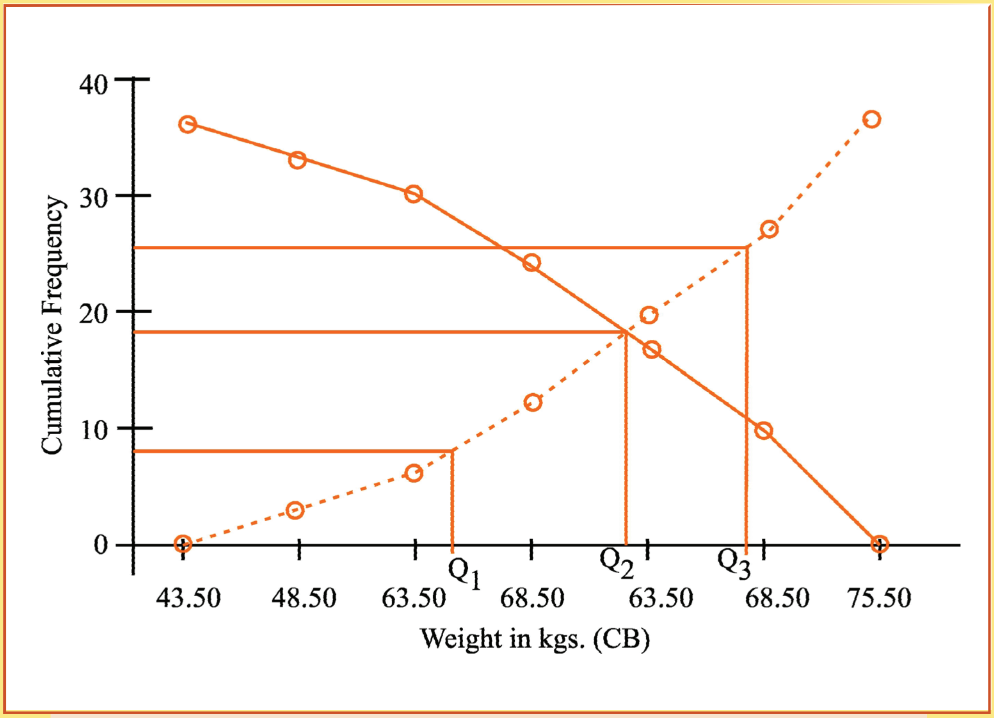


Figure 13.8

Showing the ogives for the distribution of weights of 36 BBA students

We find Q_1 = 55 kgs.

Q_2 = Me = 62.50 kgs.

Q_3 = 68 kgs.

Frequency Curve

A frequency curve is a smooth curve for which the total area is taken to be unity. It is a limiting form of a histogram or frequency polygon. The frequency curve for a distribution can be obtained by drawing a smooth and free hand curve through the mid-points of the upper sides of the rectangles forming the histogram.

There exist four types of frequency curves namely

(a) Bell-shaped curve;

(b) U-shaped curve;

(c) J-shaped curve;

(d) Mixed curve.

Most of the commonly used distributions provide bell-shaped curve, which, as suggested by the name, looks almost like a bell. The distribution of height, weight, mark, profit etc. usually belong to this category. On a bell-shaped curve, the frequency, starting from a rather low value, gradually reaches the maximum value, somewhere near the central part and then gradually decreases to reach its lowest value at the other extremity.

For a U-shaped curve, the frequency is minimum near the central part and the frequency slowly but steadily reaches its maximum at the two extremities. The distribution of Kolkata bound commuters belongs to this type of curve as there are maximum number of commuters during the peak hours in the morning and in the evening.

The J-shaped curve starts with a minimum frequency and then gradually reaches its maximum frequency at the other extremity. The distribution of commuters coming to Kolkata from the early morning hour to peak morning hour follows such a distribution. Sometimes, we may also come across an inverted J-shaped frequency curve.

Lastly, we may have a combination of these frequency curves, known as mixed curve. These are exhibited in the following figures.

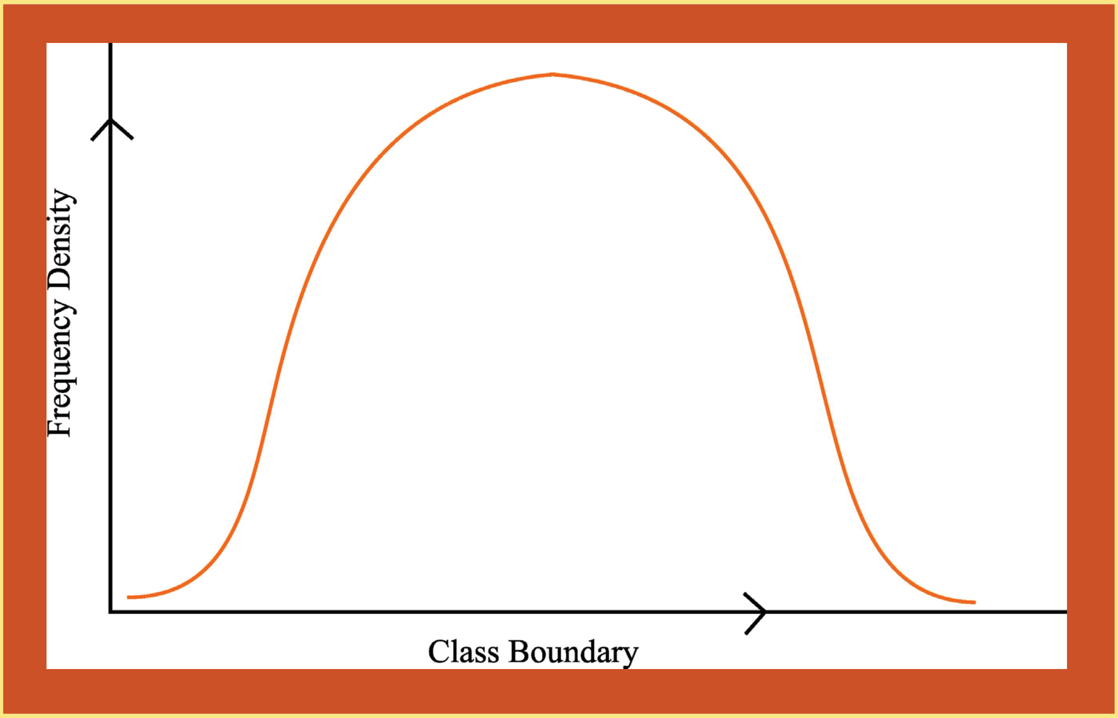


Figure 13.9

Bell-shaped curve

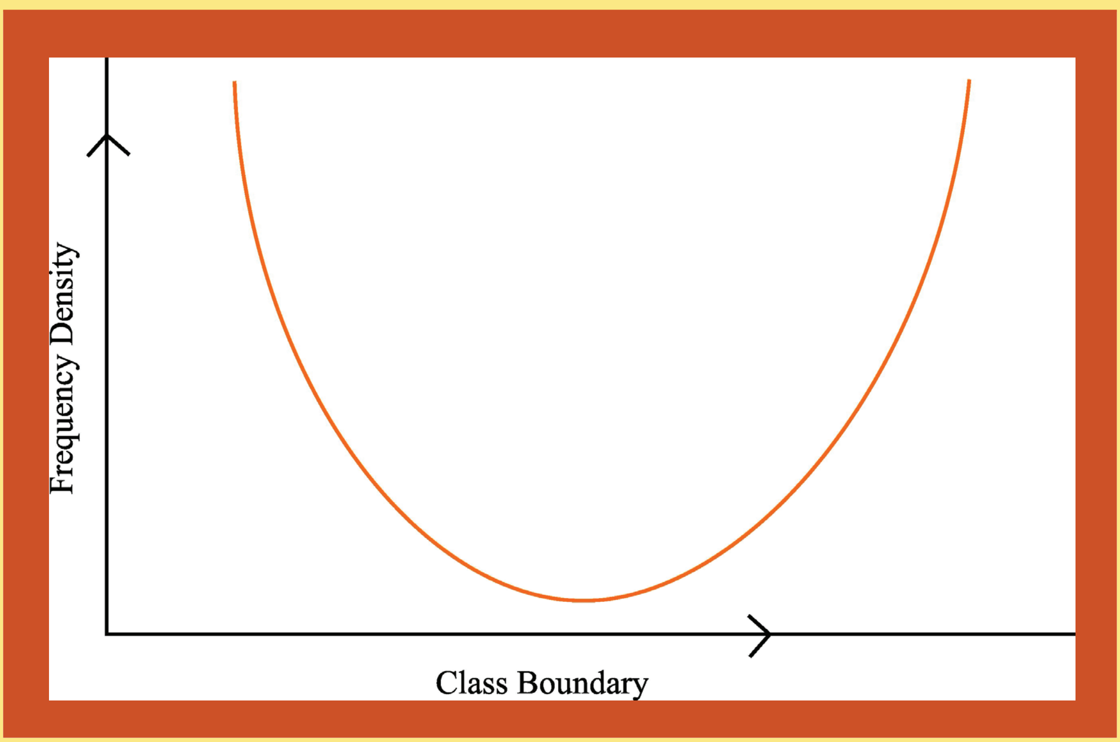


Figure 13.10

U-shaped curve

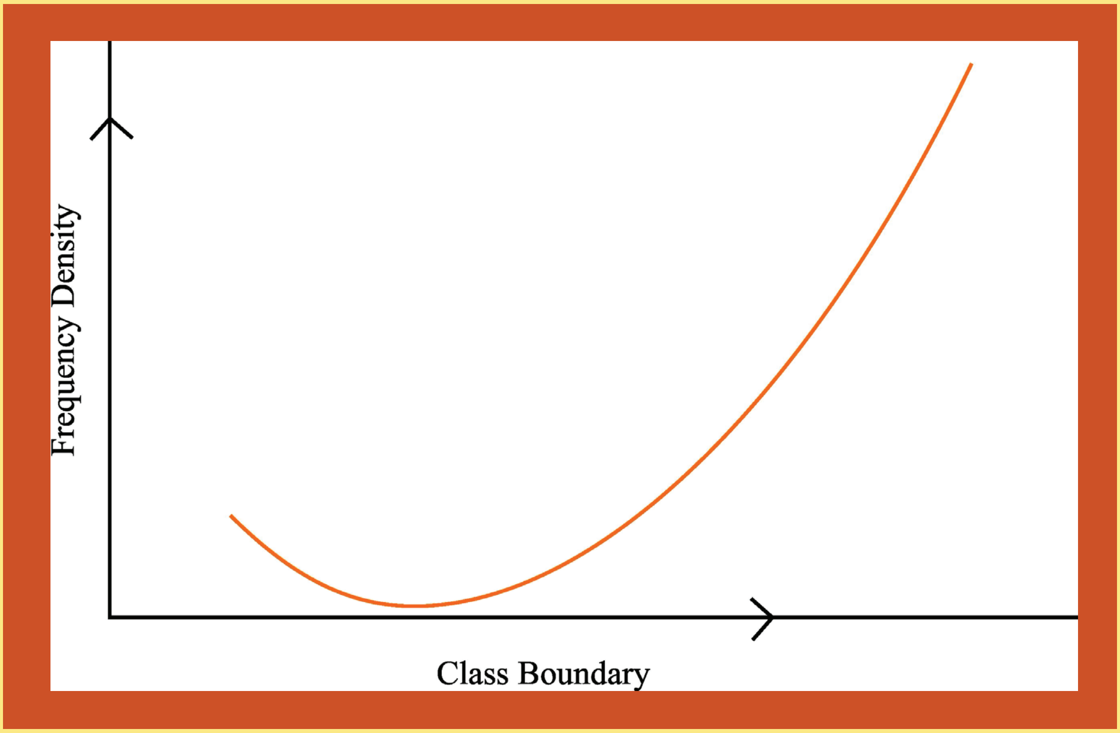


Figure 13.11
J-shaped curve

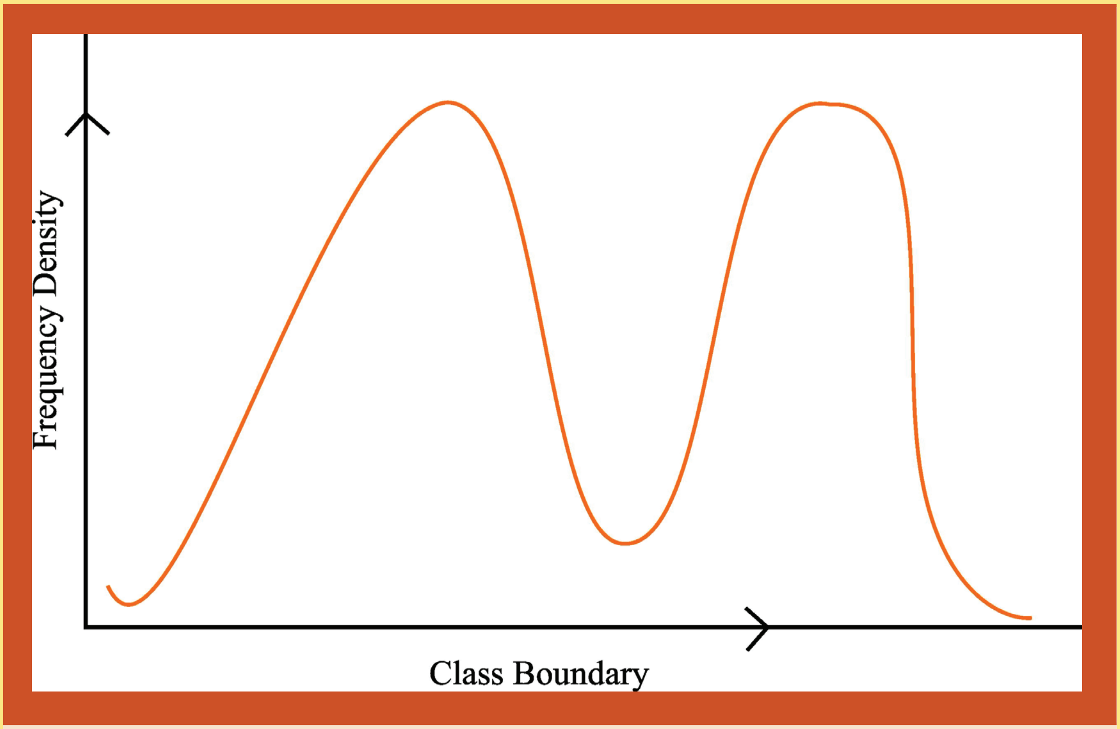


Figure 13.12
Mixed curve

Chapter 13:

Unit-II: Sampling

In this chapter the student will learn-

- Different procedure of sampling which will be the best representative of the population;

13.2.1 Introduction

There are situations when we would like to know about a vast, infinite universe or population. But some important factors like time, cost, efficiency, vastness of the population make it almost impossible to go for a complete enumeration of all the units constituting the population. Instead, we take recourse to selecting a representative part of the population and infer about the unknown universe on the basis of our knowledge from the known sample. A somewhat clear picture would emerge out if we consider the following cases.

In the first example let us share the problem faced by Mr. Basu. Mr. Basu would like to put a big order for electrical lamps produced by Mr. Ahuja's company "General Electricals". But before putting the order, he must know whether the claim made by Mr. Ahuja that the lamps of General Electricals last for at least 1500 hours is justified.

Miss Manju Bedi is a well-known social activist. Of late, she has noticed that the incidence of a particular disease in her area is on the rise. She claims that twenty per cent of the people in her town have been suffering from the disease.

In both the situations, we are faced with three different types of problems. The first problem is how to draw a representative sample from the population of electrical lamps in the first case and from the population of human beings in her town in the second case. The second problem is to estimate the population parameters i.e., the average life of all the bulbs produced by General Electricals and the proportion of people suffering form the disease in the first and second examples respectively on the basis of sample observations. The third problem relates to decision making i.e., is there enough evidence, once again on the basis of sample observations, to suggest that the claims made by Mr. Ahuja or Miss Bedi are justifiable so that Mr. Basu can take a decision about buying the lamps from General Electricals in the first case and some effective steps can be taken in the second example with a view to reducing the outbreak of the disease. We consider tests of significance or tests of hypothesis before decision making.

13.2.2 Basic Principles Of Sample Survey

Sample Survey is the study of the unknown population on the basis of a proper representative sample drawn from it. How can a part of the universe reveal the characteristics of the unknown universe? The answer to this question lies in the basic principles of sample survey comprising the following components:

(a) Law of Statistical regularity

(b) Principle of Inertia

(c) Principle of Optimization

(d) Principle of Validity

(a) According to the law of statistical regularity, if a sample of fairly large size is drawn from the population under discussion at random, then on an average the sample would posses the characteristics of that population.

Thus the sample, to be taken from the population, should be moderately large. In fact larger the sample size, the better in revealing the identity of the population. The reliability of a statistic in estimating a population characteristics varies as the square root of the sample size. However, it is not always possible to increase the sample size as it would put an extra burden on the available resource. We make a compromise on the sample size in accordance with some factors like cost, time, efficiency etc.

Apart from the sample size, the sample should be drawn at random from the population which means that each and every unit of the population should have a pre-assigned probability to belong to the sample.

(b) The results derived from a sample, according to the principle of inertia of large numbers, are likely to be more reliable, accurate and precise as the sample size increases, provided other factors are kept constant. This is a direct consequence of the first principle.

(c) The principle of optimization ensures that an optimum level of efficiency at a minimum cost or the maximum efficiency at a given level of cost can be achieved with the selection of an appropriate sampling design.

(d) The principle of validity states that a sampling design is valid only if it is possible to obtain valid estimates and valid tests about population parameters. Only a probability sampling ensures this validity.

13.2.3 Comparison Between Sample Survey and Complete Enumeration

When complete information is collected for all the units belonging to a population, it is defined as complete enumeration or census. In most cases, we prefer sample survey to complete enumeration due to the following factors:

(a) Speed: As compared to census, a sample survey could be conducted, usually, much more quickly simply because in sample survey, only a part of the vast population is enumerated.

(b) Cost: The cost of collection of data on each unit in case of sample survey is likely to be more as compared to census because better trained personnel are employed for conducting a sample survey. But when it comes to total cost, sample survey is likely to be less expensive as only some selected units are considered in a sample survey.

(c) Reliability: The data collected in a sample survey are likely to be more reliable than that in a complete enumeration because of trained enumerators better supervision and application of modern technique.

(d) Accuracy: Every sampling is subjected to what is known as sampling fluctuation which is termed as sampling error. It is obvious that complete enumeration is totally free from this sampling error. However, errors due to recording observations, biases on the part of the enumerators, wrong and faulty interpretation of data etc. are prevalent in both sampling and census and this type of error is termed as non-sampling errors. It may be noted that in sample survey, the sampling error can be reduced to a great extent by taking several steps like increasing the sample size, adhering to a probability sampling design strictly and so on. The non-sampling errors also can be contained to a desirable degree by a proper planning which is not possible or feasible in case of complete enumeration.

(e) Necessity: Sometimes, sampling becomes necessity. When it comes to destructive sampling where the items get exhausted like testing the length of life of electrical bulbs or sampling from a hypothetical population like coin tossing, there is no alternative to sample survey.

However, when it is necessary to get detailed information about each and every item constituting the population, we go for complete enumeration. If the population size is not large, there is hardly any merit to take recourse to sampling. If the occurrence of just one defect may lead to a complete destruction of the process as in an aircraft, we must go for complete enumeration.

13.2.4 Errors in Sample Survey

Errors or biases in a survey may be defined as the deviation between the value of population parameter as obtained from a sample and its observed value. Errors are of two types.

I. Sampling Errors

II. Non-Sampling Errors

Sampling Errors: Since only a part of the population is investigated in a sampling, every sampling design is subjected to this type of errors. The factors contributing to sampling errors are listed below:

(a) Errors arising out due to defective sampling design: Selection of a proper sampling design plays a crucial role in sampling. If a non- probabilistic sampling design is followed, the bias or prejudice of the sampler affects the sampling technique thereby resulting some kind of error.

(b) Errors arising out due to substitution: A very common practice among the enumerators is to replace a sampling unit by a suitable unit in accordance with their convenience when difficulty arises in getting information from the originally selected unit. Since the sampling design is not strictly adhered to, this results in some type of bias.

(c) Errors owing to faulty demarcation of units: It has its origin in faulty demarcation of sampling units. In case of an agricultural survey, the sampler has, usually, a tendency to underestimate or overestimate the character under consideration.

(d) Errors owing to wrong choice of statistic: One must be careful in selecting the proper statistic while estimating a population characteristic.

(e) Variability in the population: Errors may occur due to variability among population units beyond a degree. This could be reduced by following somewhat complicated sampling design like stratified sampling, Multistage sampling etc.

Non-sampling Errors

As discussed earlier, this type of errors happen both in sampling and complete enumeration. Some factors responsible for this particular kind of biases are lapse of memory, preference for certain digits, ignorance, psychological factors like vanity, non- responses on the part of the interviewees wrong measurements of the sampling units, communication gap between the interviewers and the interviewees, incomplete coverage etc. on the part of the enumerators also lead to non-sampling errors.

13.2.5 Some Important Terms Associated with Sampling

Population or Universe

It may be defined as the aggregate of all the units under consideration. All the lamps produced by "General Electricals" in our first example in the past, present and future constitute the population. In the second example, all the people living in the town of Miss Manju form the population. The number of units belonging to a population is known as population size. If there are one lakh people in her town then the population size, to be denoted by N, is 1 lakh.

A population may be finite or infinite. If a population comprises only a finite number of units, then it is known as a finite population. The population in the second example is obviously, finite. If the population contains an infinite or uncountable number of units, then it is known as an infinite population. The population of electrical lamps of General Electricals is infinite. Similarly, the population of stars, the population of mosquitoes in Kolkata, the population of flowers in Mumbai, the population of insects in Delhi etc. are infinite population.

Population may also be regarded as existent or hypothetical. A population consisting of real objects is known as an existent population. The population of the lamps produced by General Electricals and the population of Miss Manju's town are example of existent populations. A population that exists just hypothetically like the population of heads when a coin is tossed infinitely is known as a hypothetical or an imaginary population.

Sample

A sample may be defined as a part of a population so selected with a view to representing the population in all its characteristics selection of a proper representative sample is pretty important because statistical inferences about the population are drawn only on the basis of the sample observations. If a sample contains n units, then n is known as sample size. If a sample of 500 electrical lamps is taken from the production process of General Electricals, then n = 500. The units forming the sample are known as "Sampling Units". In the first example, the sampling unit is electrical lamp and in the second example, it is a human. A detailed and complete list of all the sampling units is known as a "Sampling Frame". Before drawing sample, it is a must to have a updated sampling frame complete in all respects before the samples are actually drawn.

Parameter

A parameter may be defined as a characteristic of a population based on all the units of the population. Statistical inferences are drawn about population parameters based on the sample observations drawn from that population. In the first example, we are interested about the parameter "Population Mean". If x□□ denotes the □ th member of the population, then population mean □, which represents the average length of life of all the lamps produced by General Electricals is given by

$$\mu = \frac{\sum_{\alpha=1}^{n} X_\alpha}{N} \quad ....(13.2.1)$$

Where N denotes the population size i.e. the total number of lamps produced by the company. In the second example, we are concerned about the population proportion P, representing the ratio of the people suffering from the disease to the total number of people in the town. Thus if there are X people possessing this attribute i.e. suffering from the disease, then we have

$$P = \frac{X}{N} ....(15.2)$$

Another important parameter namely the population variance, to be denoted by □2 is given by

$$\sigma^2 = \frac{\Sigma\,(x_\alpha - \mu)^2}{N} \quad ....(13.2.3)$$

Also we have SD = $\sigma = \sqrt{\frac{\Sigma\,(x_\alpha - \mu)^2}{N}}$ ....(13.2.4)

Statistics

A statistic may be defined as a statistical measure of sample observation and as such it is a function of sample observations. If the sample observations are denoted by $x_1, x_2, x_3, ...\ x_n$, then a statistic T may be expressed as $T = f(x_1, x_2, x_3, ...\ x_n)$

A statistic is used to estimate a particular population parameter. The estimates of population mean, variance and population proportion are given by

$$\bar{x} = \hat{\mu} = \frac{\Sigma\, x_i}{n} \quad ....(13.2.5)$$

$$S_2 = \hat{\sigma}^2 = \frac{\Sigma \left(x_i - \bar{x}\right)^2}{n} \quad ....(13.2.6)$$

$$\text{and } p = \hat{P} = \frac{x}{n} \quad ....(13.2.7)$$

Where x, in the last case, denotes the number of units in the sample in possession of the attribute under discussion.

Sampling Distribution and Standard Error of a Statistic

Starting with a population of N units, we can draw many a sample of a fixed size n. In case of sampling with replacement, the total number of samples that can be drawn is and when it comes to sampling without replacement of the sampling units, the total number of samples that can be drawn is Ncn.

If we compute the value of a statistic, say mean, it is quite natural that the value of the sample mean may vary from sample to sample as the sampling units of one sample may be different from that of another sample. The variation in the values of a statistic is termed as "Sampling Fluctuations".

If it is possible to obtain the values of a statistic (T) from all the possible samples of a fixed sample size along with the corresponding probabilities, then we can arrange the values of the statistic, which is to be treated as a random variable, in the form of a probability distribution. Such a probability distribution is known as the sampling distribution of the statistic. The sampling distribution, just like a theoretical probability distribution possesses different characteristics. The mean of the statistic, as obtained from its sampling distribution, is known as "Expectation" and the standard deviation of the statistic T is known as the "Standard Error (SE)" of T. SE can be regarded as a measure of precision achieved by sampling. SE is inversely proportional to the square root of sample size. It can be shown that

$SE(\bar{x}) = \frac{\sigma}{\sqrt{n}}$ for SRS WR

$= \frac{\sigma}{\sqrt{n}} \cdot \sqrt{\frac{N-n}{N-1}}$ for SRS WOR .....(13.2.8)

Standard Error for Proporation

$SE(p) = \sqrt{\frac{Pq}{n}}$ for SRS WR

$\sqrt{\frac{Pq}{n}} \cdot \sqrt{\frac{N-n}{N-1}}$ for SRS WOR .... (13.2.9)

SRSWR and SRSWOR stand for simple random sampling with replacement and simple random sampling without replacement.

The factor $\sqrt{\frac{N-n}{N-1}}$ is known as finite population correction (fpc) or finite population multiplier and may be ignored as it tends to 1 if the sample size (n) is very large or the population under consideration is infinite when the parameters are unknown, they may be replaced by the corresponding statistic.

Illustrations

Example 13.2.1: A population comprises the following units: a, b, c, d, e. Draw all possible samples of size three without replacement.

Solution: Since in this case, sample size (n) = 3 and population size (N) = 5. the total number of possible samples without replacement = $^5c_3 = 10$

These are abc, abd, abe, acd, ace, ade, bcd, bce, bde, cde.

Example 13.2.2: A population comprises 3 member 1, 5, 3. Draw all possible samples of size two

(i) with replacement

(ii) without replacement

Find the sampling distribution of sample mean in both cases.

Solution:

(i) With replacement :- Since n = 2 and N = 3, the total number of possible samples of size 2 with replacement = 32 = 9.

These are exhibited along with the corresponding sample mean in table 15.1. Table 15.2 shows the sampling distribution of sample mean i.e., the probability distribution of $\bar{x}$.

Table 13.2.1

All possible samples of size 2 from a population comprising 3 units under WR scheme

Serial No.	Sample of size 2 with replacement	Sample mean ($\overline{X}$)
1	1, 1	1
2	1, 5	3
3	1, 3	2
4	5, 1	3
5	5, 5	5
6	5, 3	4
7	3, 1	2
8	3, 5	4
9	3, 3	3

Table 13.2.2

Sampling distribution of sample mean

$\overline{X}$	1	2	3	4	5	Total
P	1 / 9	2 / 9	3 / 9	2 / 9	1 / 9	1

(ii) without replacement:

As N = 3 and n = 2, the total number of possible samples without replacement = NC2 = 3C2 = 3.

Table 13.2.3

Possible samples of size 2 from a population of 3 units under WOR scheme

Serial No	Sample of size 2 without replacement	Sample mean ($\overline{X}$)
1	1 , 3	2
2	1 , 5	3
3	3 , 5	4

Table 13.2.4

Sampling distribution of mean

:	2	3	4	Total
P:	1 / 3	1/3	1/3	1

Example 13.2.3: Compute the standard deviation of sample mean for the last problem. Obtain the SE of sample mean applying 15.8 and show that they are equal.

Solution:

We consider the following cases:

(i) with replacement :

(i) with replacement :

Let U = $\bar{x}$ The sampling distribution of U is given by

U:	1	2	3	4	5
P:	1/9	2/9	3/9	2/9	1/9

$E(U) = \Sigma P_i U_i$

$= 1/9 \times 1 + 2/9 \times 2 + 3/9 \times 3 + 2/9 \times 4 + 1/9 \times 5$

$= 3$

$E(U^2) = \Sigma Pi\ Ui^2$

$= 1/9 \times 12 + 2/9 \times 22 + 3/9 \times 32 + 2/9 \times 42 + 1/9 \times 52$

$= 31/3$

$\therefore v(\bar{x}) = v(U) = E(U^2) - [E(U)]^2$

$= 31/3 - 3^2$

$= 4/3$

Hence SE= $\frac{2}{\sqrt{3}}$ ....(1)

Since the population comprises 3 units, namely 1, 5, and 3 we may take $X_1 = 1$, $X_2 = 5$, $X_3 = 3$

The population mean (⊠) is given by

$$\mu = \frac{\Sigma x_\alpha}{N}$$

$$= \frac{1 + 5 + 3}{3} = 3$$

and the population variance $\sigma^2 = \frac{\Sigma (x_\alpha - \mu)^2}{N}$

$$= \frac{(1-3)^2 + (5-3)^2 + (3-3)^2}{3} = 8/3$$

Applying 15.8 we have, $SE_{\bar{x}} = \frac{\sigma}{\sqrt{n}} = \frac{8}{\sqrt{3}} \times \frac{1}{\sqrt{2}} = \frac{2}{\sqrt{3}}$ ...(2)

Thus comparing (1) and (2), we are able to verify the validity of the formula.

(ii) without replacement :

In this case, the sampling distribution of V = is given by

V:	2	3	4
P:	1/3	1/3	1/3

$E(\bar{x}) = E(V) = 1/3 \times 2 + 1/3 \times 3 + 1/3 \times 4$

$= 3$

$V(\bar{x}) = Var(V) = E(v2) - [E(v)]2$

$= 1/3 \times 22 + 1/3 \times 32 + 1/3 \times 42 - 32$

$= 29/3 - 9$

$= 2/3$

$\therefore SE_{\bar{x}} = \frac{2}{\sqrt{3}}$

Applying 13.2.8, we have

$$\therefore SE_{\bar{x}} = \frac{\sigma}{\sqrt{n}} \cdot \sqrt{\frac{N-n}{N-1}}$$

$$= \frac{8}{\sqrt{3}} x \frac{1}{\sqrt{2}} x \frac{8}{\sqrt{3}} x \frac{1}{\sqrt{2}} x \sqrt{\frac{3-2}{3-1}} = \frac{2}{\sqrt{3}}$$

and thereby, we make the same conclusion as in the previous case.

13.2.6 Types of Sampling

There are three different types of sampling which are

I. Probability Sampling

II. Non – Probability Sampling

III. Mixed Sampling

In the first type of sampling there is always a fixed, pre assigned probability for each member of the population to be a part of the sample taken from that population . When each member of the population has an equal chance to belong to the sample, the sampling scheme is known as Simple Random Sampling. Some important probability sampling other than simple random sampling are stratified sampling, Multi Stage sampling, Multi Phase Sampling, Cluster Sampling and so on. In non- probability sampling , no probability attached to the member of the population and as such it is based entirely on the judgement of the sampler. Non-probability sampling is also known as Purposive or Judgement Sampling. Mixed sampling is based partly on some probabilistic law and partly on some pre decided rule. Systematic sampling belongs to this category. Some important and commonly used sampling process are described now.

Simple Random Sampling (SRS)

When the units are selected independent of each other in such a way that each unit belonging to the population has an equal chance of being a part of the sample, the sampling is known as Simple random sampling or just random sampling. If the units are drawn one by one and each unit after selection is returned to the population before the next unit is being drawn so that the composition of the original population remains unchanged at any stage of the sampling then the sampling procedure is known as Simple Random Sampling with replacement. If, however, once the units selected from the population one by one are never returned to the population before the next drawing is made, then the sampling is known as sampling without replacement. The two sampling methods become almost identical if the population is infinite i.e. vary large or a very large sample is taken from the population. The best method of drawing simple random sample is to use random sampling numbers.

Simple random sampling is a very simple and effective method of drawing samples provided (i) the population is not very large (ii) the sample size is not very small and (iii) the population under consideration is not heterogeneous i.e. there is not much variability among the members forming the population. Simple random sampling is completely free from Sampler's biases. All the tests of significance are based on the concept of simple random sampling.

Stratified Sampling

If the population is large and heterogeneous, then we consider a somewhat, complicated sampling design known as stratified sampling which comprises dividing the population into a number of strata or sub-populations in such a way that there should be very little variations among the units comprising a stratum and maximum variation should occur among the different strata. The stratified sample consists of a number of sub samples, one from each stratum. Different sampling scheme may be applied to different strata and , in particular, if simple random sampling is applied for drawing units from all the strata, the sampling procedure is known as stratified random sampling. The purpose of stratified sampling are (i) to make representation of all the sub populations (ii) to provide an estimate of parameter not only for all the strata but also and overall estimate (iii) reduction of variability and thereby an increase in precision.

There are two types of allocation of sample size. When there is prior information that there is not much variation between the strata variances. We consider "Proportional allocation" or "Bowely's allocation where the sample sizes for different strata are taken as proportional to the population sizes. When the strata-variances differ significantly among themselves, we take recourse to "Neyman's allocation" where sample size vary jointly with population size and population standard deviation i.e. $n_i \propto N_iS_i$. Here n_i denotes the sample size for the ith stratum, N_i and S_i being the corresponding population size and population standard deviation. In case of Bowley's allocation, we have $n_i \propto N_i$.

Stratified sampling is not advisable if (i) the population is not large (ii) some prior information is not available and (iii) there is not much heterogeneity among the units of population.

Multi Stage Sampling

In this type of complicated sampling , the population is supposed to compose of first stage sampling units, each of which in its turn is supposed to compose of second stage sampling units, each of which again in its turn is supposed to compose of third stage sampling units and so on till we reach the ultimate sampling unit.

Sampling also, in this type of sampling design, is carried out through stages. Firstly, only a number of first stage units is selected. For each of the selected first stage sampling units, a number of second stage sampling units is selected. The process is carried out until we select the ultimate sampling units. As an example of multi stage sampling, in order to find the extent of unemployment in India, we may take state, district, police station and household as the first stage, second stage, third stage and ultimate sampling units respectively.

The coverage in case of multistage sampling is quite large. It also saves computational labour and is cost-effective. It adds flexibility into the sampling process which is lacking in other sampling schemes. However, compared to stratified sampling, multistage sampling is likely to be less accurate.

Systematic Sampling

It refers to a sampling scheme where the units constituting the sample are selected at regular interval after selecting the very first unit at random i.e., with equal probability. Systematic sampling is partly probability sampling in the sense that the first unit of the systematic sample is selected probabilistically and partly non- probability sampling in the sense that the remaining units of the sample are selected according to a fixed rule which is non-probabilistic in nature.

If the population size N is a multiple of the sample size n i.e. $N = nk$, for a positive integer k which must be less than n, then the systematic sampling comprises selecting one of the first k units at random, usually by using random sampling number and thereby selecting every kth unit till the complete, adequate and updated sampling frame comprising all the members of the population is exhausted. This type of systematic sampling is known as "linear systematic sampling ". K is known as "sample interval".

However, if N is not a multiple of n, then we may write $N = nk + p$, $p < k$ and as before, we select the first unit from 1 to k by using random sampling number and thereafter selecting every kth unit in a cyclic order till we get the sample of the required size n. This type of systematic sampling is known as "circular systematic sampling."

Systematic sampling is a very convenient method of sampling when a complete and updated sampling frame is available. It is less time consuming, less expensive and simple as compared to the other methods of sampling. However, systematic sampling has a severe drawback. If there is an unknown and undetected periodicity in the sampling frame and the sampling interval is a multiple of that period, then we are going to get a most biased sample, which, by no stretch of imagination, can represent the population under investigation. Furthermore, since it is not a probability sampling, no statistical inference can be drawn about population parameter.

Purposive or Judgement sampling

This type of sampling is dependent solely on the discretion of the sampler and he applies his own judgement based on his belief, prejudice, whims and interest to select the sample. Since this type of sampling is non-probabilistic, it is purely subjective and, as such, varies from person to person. No statistical hypothesis can be tested on the basis of a purposive sampling.

Chapter 14:

Unit-I: Measures of Central Tendency

At the foundation level with regards to Paper 3 Statistics part of the topic Measures of Central Tendency is very important for students not only to acquire professional knowledge but also for examination point of view. Here in this chapter an attempt is made for solving and understanding the concepts of Central Tendency with the help of following questions with solutions.

Definition of Central Tendency: Central tendency defined as the tendency of a given set of observations to cluster around a single central or middle value and the single value that represents the given set of observations is described as a measure of central tendency or, location, or average.

Following are the different measures of central tendency:

- Arithmetic Mean (AM)
- Median (Me)
- Mode (Mo)
- Geometric Mean (GM)
- Harmonic Mean (HM)

Criteria for an Ideal Measure of Central Tendency

- It should be properly and unambiguously defined.
- It should be easy to comprehend.
- It should be simple to compute.
- It should be based on all the observations.
- It should have certain desirable mathematical properties.
- It should be least affected by the presence of extreme observations.

Arithmetic Mean: defined as the sum of all the observations divided by the number of observations. Thus, if a variable x assumes n values $x_1, x_2, x_3, \dots\dots\dots x_n$, then the AM of x, to

be denoted by , is given by, $\frac{\sum_{i=1}^{n} x_i}{n}$ $\bar{x} = \frac{X_1 + X_2 + X_3 + \dots\dots\dots + X_n}{n}$

In case of a simple frequency distribution relating to an attribute, we have $= \frac{\sum_{i=1}^{n} f_i . x_i}{\sum_{i=1}^{n} f_i}$

$\bar{X} = \frac{\sum_{i=1}^{n} f_i . x_i}{N}$ assuming the observation xi occurs fi times, i=1,2,3,........n and $N = \sum f_i$.

In case of grouped frequency distribution also we may use formula with x_i as the mid value of the i-th class interval, on the assumption that all the values belonging to the i-th class interval are equal to x_i.

If classification is uniform, we consider the following formula for the computation of AM from grouped

frequency the distribution: $\bar{X} = A + \frac{\sum f_i d_i}{N} \times C$

Where, $d_i = \frac{x_i - A}{C}$ A = Assumed Mean C = Class Length

- **Properties of AM**
- **If all the observations assumed by a variable are constants, say k, then the AM is also k.**
- The algebraic sum of deviations of a set of observations from their AM is zero
- i.e. for unclassified data, $\sum (\bar{x}_i - \bar{x}) = 0$ and for grouped frequency distribution, $\sum (f_i (\bar{x}_i - \bar{x}) = 0$
- AM is affected due to a change of origin and/or scale which implies that if the original variable x is changed to another variable y by effecting a change of origin, say a, and scale say b, of x i.e. y=a+bx, then the AM of y is given by $\bar{y} = a + b\bar{x}$
- If there are two groups containing n1 and n2 observations and and as the respective arithmetic means, then

  combined AM is given by $\bar{X} = \frac{n_1\bar{x}_1 + n_2\bar{x}_2}{n_1 + n_2}$

Question 1: Following are the daily wages in rupees of a sample of 10 workers: 58, 62, 48, 53, 70, 52, 60, 84, 75, 100. Compute the mean wage.

Solution : Let x denote the daily wage in rupees.

Applying the mean wage is given by, $\overline{X} = \frac{\sum_{i=1}^{10} x_i}{10} = \frac{(58 + 62 + 48 + 53 + 70 + 52 + 60 + 84 + 75 + 100}{10} =$

$\frac{₹\ 662}{10} = ₹\ 66.2$

Question 2: Compute the mean weight of a group of B. Com students of Sri Ram College from the following data:

Weight in kgs	44-48	49-53	54-58	59-63	64-68	69-73
No. of students	3	4	5	7	9	8

Solution : Computation of mean weight of 40 B. Com students Applying formula, we get the average weight as

$\overline{X} = \frac{\Sigma f_i x_i}{N} \frac{2495}{40}$ kgs.=62.38 kgs.

Weight in kgs. (1)	No. of Student (f_i) (2)	Mid-Value (x_i) (3)	$f_i x_i$ (4) = (2) x (3)
44 – 48	3	46	138
49 – 53	4	51	204
54 – 58	5	56	280
59 – 63	7	61	427
64 – 68	9	66	594
69 – 73	12	71	852
Total	40	–	2495

Question 3: Find the AM for the following distribution:

Class Interval	5-14	15-24	25-34	35-44	45-54	55-64
Frequency	10	18	32	26	14	10

Solution : Any mid value can be taken as A. However, usually A is taken as the middle most mid-value for an odd number of class intervals and any one of the two middle most mid-values for an even number of class intervals. The class length is taken as C.

The required AM is given by $\overline{X} = A + \frac{\Sigma f_i d_i}{N} \text{ x } C = 39.5 + \frac{-64}{110} \text{ X } 10 = 39.5 - 5.82 = 33.68$

Table: Computation of AM

Class Interval	Frequency(f_i)	Mid-Value (x_i)	$d_i = \frac{X_i - A}{C}$	$f_i d_i$
(1)	(2)	(3)	(4)	(5) = (2) x (4)
5-14	10	9.5	– 3	– 30
15-24	18	19,5	– 2	– 36
25-34	32	29.5	– 1	– 32
35-44	26	39.5(A)	0	0
45-54	14	49.5	1	14
55-64	10	59.5	2	20
Total	110	-	-	-64

Question 4: Given that the mean height of a group of students is 67.45 inches. Find the missing frequencies for the following incomplete distribution of height of 100 students.

Height in inches	60-62	63-65	66-68	69-71	72-74
No. of students	5	18	-	-	8

Solution: Let x denote the height and f3 and f4 as the two missing frequencies

Table: Estimation of missing frequencies

Class Interval	Frequency	Mid-Value (x_i)	$d_i = \frac{X_i - A}{C} = \frac{X_i - 67}{3}$	$F_i d_i$ (f_i)
(1)	(2)	(3)	(4)	(5) = (2) x (4)
60-62	5	61	-2	-10
63 – 65	18	64	– 1	– 18
66 – 68	f_3	67 (A)	0	0
69 – 71	f_4	70	1	f_4
72 – 74	8	73	2	16
Total	$31+ f_3+ f_4$	–	–	$– 12+f_4$

As given, we have

$31+ f_3+ f_4= 100, f_3+ f_4= 69$.................(1)

$\overline{X} = 67.45$

and $A + \frac{\Sigma f_i d_i}{N} \times C = 67.45 = 67 + \frac{(12 + f_4)}{100} X\ 3 = 67.45$

$(-12 + f_4) \times 3 = (67.45- 67) \times 100$

$-12 + f_4 = 15, \ f_4 = 27$

On substituting 27 for f_4 in (1), we get, $f_3 + 27 = 69$, $f_3 = 42$, Thus, the missing frequencies would be 42 and 27.

Question 5: The mean salary for a group of 40 female workers is ₹ 5,200 per month and that for a group of 60 male workers is ₹ 6800 per month. What is the combined mean salary?

Solution: As given $n_1 = 40$, $n_2 = 60$, = ₹ 5,200 and = ₹ 6,800 hence, the combined mean salary per month is

$$\overline{X} = \frac{n_1\overline{x}_1 + n_2\overline{x}_2}{n_1 + n_2} = \frac{40 \times ₹5{,}200 + 60 \times ₹6{,}800}{40 + 60} = ₹\ 6{,}160.$$

Question 6: The mean weight of 150 students (boys and girls) in a class is 60 kg. The mean weight of boy student is 70 kg and that of girl student is 55 kg . Find number of boys and girls in that class.

Solution: Let the number of boy students be n1 and girl students be n_2, as given $n_1 + n_2 = 150$,

Then n2 = 150−n1, also $\overline{X} = 60$, $\overline{X}_1 = 70$, $\overline{X}_2 = 55$

$$\overline{X} = \frac{n_1\overline{x}_1 + n_2\overline{x}_2}{n_1 + n_2},\ 60 = \frac{n_1 \times 70 + (150 - n_1) \times 55}{150}$$

$$60 = \frac{70n_1 + 8250 - 55n_1}{150} = \frac{15n_1 + 8250}{150}$$

$9000 = 15n_1 + 8250$

$15n_1 = 750$, $n_1 = 50$

$n_2 = 150 - n_1 = 150 - 50 = 100$

Question 7: The average salary of a group of unskilled workers is Rs. 10,000 and that of a group of skilled workers is Rs. 15000. If the combined salary is Rs.12000, then what is the percentage of skilled workers

Solution: Let x be unskilled and y be skilled

10000x+15000y= 12000(x+y) = 12000x+12000y

2000x=3000y then 2x=3y

skilled workers is 2x/3

total workers x+2x/3=5x/3

percentage of skilled=2x/3 divided by 5x/3=40%

Question 8: The average age of a group of 10 students was 20 years. The average age increased by two years when the two new students joined in the group. What is the average age of two new students joined who joined in the group?

Solution: Average age of 10 students = 20 years, then sum of ages of 10 students = 200 years

If the two boys are included, then total number of students = 10+2 = 12

And average increased by two years = 20 + 2 =22

The average age of 12 students = 22, then sum of ages of 12 students = 22×12= 264

The Sum of ages of two boys = 264−200= 64

Average age of boys = 64/2 = 32

Median

- Partioned Values
- As compared to AM, median is a positional average which means that the value of the median is dependent upon the position of the given set of observations for which the median is wanted. Median, for a given set of observations, may be defined as the middle-most value when the observations are arranged either in an ascending order or a descending order of magnitude.

Question 9: The median of the data 13, 8, ,11, 6, 4, 15, 2, 18 is

Solution: Arranging the data in a ascending order, we get 2, 4, 6, 8, 11, 13, 15,18

Median = $\left(\frac{n+1}{2}\right)$th item = $\left(\frac{8+1}{2}\right)$th item = 4.5 th item

= 4^{th} item +0.5(5^{th} item -4^{th} item) = 8 +0.5(11-8) = 8+1.5= 9.5

Question 10: What is the median for the observations 5,8,6,9,11 and 4

Solution: We write in ascending order 4, 5, 6, 8, 9 and 11

Here n = 6, So, Median = Average of 3rd and 4th term =Median = $\frac{6+8}{2}$ = 7

In case of a grouped frequency distribution, we find median from the cumulative frequency distribution of the variable under consideration. $M = l_1 + \left(\frac{\frac{N}{2} - N_l}{N_u - N_l}\right) \times C$

Where,
l_1 = lower class boundary of the median class i.e. the class containing median.
N = total frequency.
N_l = less than cumulative frequency corresponding to l_1. (Pre median class)
N_u = less than cumulative frequency corresponding to l_2. (Post median class)
l_2 being the upper class boundary of the median class.
$C = l_2 - l_1$ = length of the median class.

Question 11: Compute the median for the distribution as given in

Solution: First, we find the cumulative frequency distribution which is exhibited in **Table Computation of Median**

Class boundary	Less than cumulative frequency
349.50	0
369.50	23
389.50	61
409.50 (l_1)	119 (N_l)
429.50 (l_2)	201(N_u)
449.50	266
469.50	297
489.50	308

We find, from the **Table ;** $\frac{N}{2} = \frac{308}{2}$ = 154 lies between the two cumulative frequencies 119 and 201 i.e. 119 < 154 < 201 . Thus, we have N_l = 119, N_u = 201, l_1 = 409.50 and l_2 = 429.50.

Hence C = 429.50 – 409.50 =20.

Substituting these values in formula , we get,

$M = 409.50 + \frac{154 - 119}{201 - 119} \times 20$ = 409.50+8.54 = 418.04.

Question 12: Find the missing frequency from the following data, given that the median mark is 23.

Marks	0-10	10-20	20-30	30-40	40-50
No.of Students	5	8	?	6	3

Solution: Let us denote the missing frequency by f_3. Following table shows the relevant computation.

Table (Estimation of Missing frequency)	
Marks	**Less than cumulative frequency**
0	0
10	5
20 (l_1)	13(N_l)
30(l_2)	13+f_3(N_u)
40	19+f_3
50	22+f_3

Going through the mark column, we find that 20<23<30. Hence l_1=20, l_2 =30 and accordingly N_l=13, N_u=13+f_3. Also the total frequency i.e. N is 22+f_3. Thus,

$$M=l_1 + \left(\frac{\frac{N}{2} - N_l}{N_u - N_l}\right) x\ C$$

$$23 = 20 + \frac{\left(\frac{22 + f_3}{2}\right) - 13}{(13 + f_3) - 13} x\ 10$$

$$3 = + \frac{22 + f_3 - 26}{f_3} x\ 5,\ \ 3f_3 = 5f_3 - 20, f_3 = 20$$

f_3 = 10, So, the missing frequency is 10.

Properties of median: We cannot treat median mathematically; the way we can do with arithmetic mean. We consider below two important features of median.

(i) If x and y are two variables, to be related by y=a+bx for any two constants a and b, then the median of y is given by y_{me} = a + bx_{me}

For example, if the relationship between x and y is given by 2x – 5y = 10 and if x_{me} i.e. the median of x is known to be 16.

Then 2x – 5y = 10

$\Rightarrow$ y = –2 + 0.40x

$\Rightarrow$ y_{me} = –2 + 0.40 x_{me}

$\Rightarrow$ y_{me}= –2 + 0.40×16

$\Rightarrow$ y_{me}= 4.40.

(ii) For a set of observations, the sum of absolute deviations is minimum when the deviations are taken from the median. This property states that $\Sigma|x_i - A|$ is minimum if we choose A as the median.

Partition Values or Quartiles

- These may be defined as values dividing a given set of observations into a number of equal parts. When we want to divide the given set of observations into two equal parts, we consider median. Similarly, quartiles are values dividing a given set of observations into four equal parts. So there are three quartiles – first quartile or lower quartile denoted by Q1, second quartile or median to be denoted by Q2 or Me and third quartile or upper quartile denoted by Q3. First quartile is the value for which one fourth of the observations are less than or equal to Q1 and the remaining three – fourths observations are more than or equal to Q1. In a similar manner, we may define Q2 and Q3.

Deciles

- Deciles are the values dividing a given set of observation into ten equal parts. Thus, there are nine deciles to be denoted by D1, D2, D3,.....D9. D1 is the value for which one-tenth of the given observations are less than or equal to D1 and the remaining nine-tenth observations are greater than or equal to D1 when the observations are arranged in an ascending order of magnitude.

Percentiles or Centiles

- that divide a given set of observations into 100 equal parts. The points of sub-divisions being P1, P2,..........P99. P1 is the value for which one hundredth of the observations are less than or equal to P1 and the remaining ninety-nine hundredths observations are greater than or equal to P1 once the observations are arranged in an ascending order of magnitude.

For unclassified data, the p^{th} quartile is given by the $(n+1)p^{th}$ value, where n denotes the total number of observations. $p = 1/4, 2/4, 3/4$ for Q_1, Q_2 and Q_3 respectively. p=1/10, 2/10,............9/10.

For D_1, D_2,......,D_9 respectively and l astly p=1/100, 2/100,.....,99/100 for P_1, P_2, P_3....P_{99} respectively.

In case of a grouped frequency distribution, we consider the following formula for the computation of quartiles.

$$Q = l_1 + \left(\frac{N_p - N_l}{N_u - N_l}\right) x\, C$$

The symbols, except p, have their usual interpretation which we have already discussed while computing median and just like the unclassified data, we assign different values to p depending on the quartile.

Another way to find quartiles for a grouped frequency distribution is to draw the ogive (less than type) for the given distribution. In order to find a particular quartile, we draw a line parallel to the horizontal axis through the point Np. We draw perpendicular from the point of intersection of this parallel line and the ogive. The x-value of this perpendicular line gives us the value of the quartile.

Question 13: Following are the wages of the labourers: Rs.82, Rs.56, Rs.90, Rs.50, Rs.120, Rs.75, Rs.75, Rs.80, Rs.130, Rs.65. Find Q_1, D_6 and P_{82}.

Solution: Arranging the wages in an ascending order, we get Rs.50, Rs.56, Rs.65, Rs.75, Rs.75, Rs.80, Rs.82, Rs.90, Rs.120, Rs.130.

Hence, we have

$$Q1 = \frac{(n+1)}{4} \text{ th value} = \frac{(10+1)}{4} \text{ th value} = 2.75^{th} \text{ value}$$

$= 2^{nd}$ value + 0.75 × difference between the third and the 2^{nd} values.

= Rs.[56 + 0.75 × (65 – 56)] = Rs. 62.75

$$D_6 = (10+1) \times \frac{6}{10} \text{ th value} = 6.60^{th} \text{ value}$$

$= 6^{th}$ value + 0.60 × difference between the 7^{th} and the 6^{th} values.

= Rs.(80 + 0.60 × 2) = Rs.81.20

$$P_{82} = (10+1) \times \frac{82}{100} \text{ th value} = 9.02^{th} \text{ value}$$

$= 9^{th}$ value + 0.02 × difference between the 10^{th} and the 9^{th} values

= Rs.(120 + 0.02 ×10) = Rs.120.20

Question 14: Following distribution relates to the distribution of monthly wages of 100 workers.

	Less Than					**More than**
Wages	500	500-699	700-899	900-1099	1100-1499	1500
No.of Workers	5	23	29	27	10	6

ompute Q_3, D_7 and P_{23}

Solution: This is a typical example of an open end unequal classification as we find the lower class limit of the first class interval and the upper class limit of the last class interval are not stated, and theoretically, they can assume any value between 0 and 500 and 1500 to any number respectively. The ideal measure of the central tendency in such a situation is median as the median or second quartile is based on the fifty percent central values. Denoting the first LCB and the last UCB by the L and U respectively, we construct the following cumulative frequency distribution:

Computation of Quartiles	
Wages in (Rs) CB	**No.of Workers (Less than Cumulative frequency)**
L	0
499.50	5
699.50	28
899.50	57
1099.50	84
1049.50	94
U	100

For Q3, $\frac{3N}{4} = \frac{3 \times 100}{4} = 75$

since, 57<75 <84, we take N_l = 57, N_u=84, l_1=899.50, l_2=1099.50, c = $l_2 - l_1$ = 200

in the formula (11.8) for computing Q_3.

Therefore, Q3 = ₹ $\left[899.50 + \frac{75 - 57}{84 - 57} \times 200\right]$ = ₹ 1032.83

Similarly, for D7, = ₹ $\frac{7N}{10} = \frac{7 \times 100}{10} = 70$ which also lies between 57 and 84.

Thus, = ₹ $\left[899.50 + \frac{70 - 57}{84 - 57} \times 200\right]$ ₹ 995.80

Lastly for P23, ₹ $\frac{23N}{100} = \frac{23}{100} \times 100 = 23$ and as 5 < 23 < 28, we have

P23 = ₹[499.50 ₹ $\frac{23 - 5}{28 - 5} \times 200$]

= ₹656.02

Mode: Mode may be defined as the value that occurs the maximum number of times. Thus, mode is that value which has the maximum concentration of the observations around it. This can also be described as the most common value with which, even, a layman may be familiar with.

Thus, if the observations are 5, 3, 8, 9, 5 and 6, then Mode (Mo) = 5 as it occurs twice and all the other observations occur just once. The definition for mode also leaves scope for more than one mode. Thus sometimes we may come across a distribution having more than one mode. Such a distribution is known as a multi-modal distribution. Bi-modal distribution is one having two modes.

Furthermore, it also appears from the definition that mode is not always defined. As an example, if the marks of 5 students are 50, 60, 35, 40, 56, there is no modal mark as all the observations occur once i.e. the same number of times.

We may consider the following formula for computing mode from a grouped frequency distribution:

Mode $= l_1 + \left(\frac{f_0 - f_{-1}}{2f_0 - f_{-1} - f_1}\right) \times C$ where,

L_1 = LCB of the modal class.i.e. the class containing mode.

f_0 = frequency of the modal class, f_{-1} = frequency of the pre–modal class

f_1 = frequency of the post modal class, C = class length of the modal class

Question 15: Compute mode for the distribution.

Solution: The frequency distribution is shown below

Class Interval	Frequency
350-369	23
370-389	38
390-409	58(f_{-1})
410-429	82(f_0)
430-449	65 (f_1)
450-469	31
470-489	11

Going through the frequency column, we note that the highest frequency i.e. f_0 is 82. Hence, f_{-1} = 58 and f_1 = 65. Also the modal class i.e. the class against the highest frequency is 410 – 429. Thus = LCB=409.50 and c=429.50 – 409.50 = 20 Hence, applying formulas , we get

Properties of Mode:

$$Mo = 409.5 + \frac{82 - 58}{2 \times 82 - 58 - 65} \times 20 = 421.21$$ which belongs to the modal class. (410 – 429)

When it is difficult to compute mode from a grouped frequency distribution, we may consider the following empirical relationship between mean, median and mode:

Mean – Mode = 3(Mean – Median) or Mode = 3 Median – 2 Mean

holds for a moderately skewed distribution. We also note that if y = a+bx, then $y_{mo}=a+bx_{mo}$

Question 16: For a moderately skewed distribution of marks in statistics for a group of 200 students, the mean mark and median mark were found to be 55.60 and 52.40. What is the modal mark?

Solution: Since in this case, mean = 55.60 and median = 52.40, applying, we get the modal mark as, Mode = 3 × Median – 2 × Mean = 3 × 52.40 – 2 × 55.60 = 46.

Question 17: If x and y related by x-y-10=0 and mode of x is known to be 23, then the mode of y is :

Solution: Mode of x = 23 , x-y-10= 0 then y = x-10 , Mode of y = 23-10 = 13

Geometric Mean: For a given set of n positive observations, the geometric mean is defined as the n-th root of the product of the observations. Thus if a variable x assumes n values $x_1, x_2, x_3, \ldots\ldots, x_n$, all the values being positive, then the GM of x is given by $G= (x_1 \times x_2 \times x_3 \ldots\ldots \times x_n)^{1/n}$

For a grouped frequency distribution, the GM is given by

$G= (x_1^{f1} \times x_2^{f2} \times x_3^{f3} \ldots\ldots \times x_n^{fn})^{1/N}$, Where $N = \Sigma f_i$

In connection with GM, we may note the following properties:

(i) Logarithm of G for a set of observations is the AM of the logarithm of the observations

(ii) if all the observations assumed by a variable are constants, say K > 0, then the GM of the observations is also K.

(iii) GM of the product of two variables is the product of their GM's i.e. if z = xy, then GM of z = (GM of x) × (GM of y)

(iv) GM of the ratio of two variables is the ratio of the GM's of the two variables i.e. if z = x/y then

$$\text{GM of } z = \frac{\text{GM of } x}{\text{GM of } y}$$

Question 18: Find the GM of 3, 6 and 12.

Solution: As given x_1=3, x_2=6, x_3=12 and n=3.

Applying, we have $G= (3\times6\times12)^{1/3} = (6^3)^{1/3}=6$.

Question 19: Find the GM for the following distribution:

x	2	4	8	16
f	2	3	3	2

Solution: According to the GM is given by $G = (x_1^{f1} \times x_2^{f2} \times x_3^{f3} \times x_4^{f4})^{1/N}$

Harmonic Mean: For a given set of non-zero observations, harmonic mean is defined as the reciprocal of the AM of the reciprocals of the observation. So, if a variable x assumes n non-zero values $x_1, x_2, x_3, \ldots\ldots\ldots, x_n$, then the HM of x is given by

$$H = \frac{n}{\Sigma (1/x_i)}$$

For a grouped frequency distribution, we have $H = \dfrac{N}{\Sigma \left[\dfrac{f_i}{X_i}\right]}$

Properties of HM

(i) If all the observations taken by a variable are constants, say k, then the HM of the observations is also k.

(ii) If there are two groups with n1 and n2 observations and H1 and H2 as respective HM's than the combined HM is

given by $= \dfrac{n_1 + n_2}{\dfrac{n_1}{H_1} + \dfrac{n_2}{H_2}}$

Question 20: A man travels at a speed of 20 km/ hr and then returns at a speed of 30 km/hr. His average of the whole journey is

Solution: Harmonic Mean is the method which is preferred for the computation of average speed

$$HM = \frac{2ab}{a+b} = \frac{2 \text{ x } 20 \text{ x } 30}{20 + 30} = 24\text{km/hr}$$

Question 21: Find the HM for 4, 6 and 10.

Solution: Applying formula, we have $H = \dfrac{3}{\dfrac{1}{4} + \dfrac{1}{6} + \dfrac{1}{10}} = \dfrac{3}{0.25 + 0.17 + 0.10} = 5.77$

Question 22: An aeroplane flies from A to B at the rate of 500 km/ hr and comes back B to A at the rate of 700 km/ hr . The average speed of the aeroplane is ;

Solution: Required average speed of the aeroplane $= \dfrac{2}{\left(\dfrac{1}{500} + \dfrac{1}{700}\right)} = \dfrac{2 \text{ x } 3500}{7 + 5} = 583.33$ km/hr

Question 23: Find the HM for the following data:

x	2	4	8	16
f	2	3	3	2

Solution: Using formula, we get $H = \dfrac{10}{\frac{2}{2}+\frac{3}{4}+\frac{3}{8}+\frac{2}{16}} = 4.44$

Question 24: compute AM, GM, and HM for the numbers 6, 8, 12, 36.

Solution: In accordance with the definition, we have

$$AM = \frac{6+8+12+36}{4} = 15.50$$

$$GM = (6 \times 8 \times 12 \times 36)^{1/4}$$
$$= (2^8 \times 3^4)^{1/4} = 12$$

$$HM = \frac{4}{\frac{1}{6}+\frac{1}{8}+\frac{1}{12}+\frac{1}{36}} = 9.93$$

The computed values of AM, GM, and HM establish AM ≥ GM ≥ HM

Question 25: If there are two groups with 75 and 65 as harmonic means and containing 15 and 13 observations, then the combined HM is given by

Solution: Combined HM is given by $= \dfrac{n_1+n_2}{\left(\frac{n_1}{H_1}+\frac{n_2}{H_2}\right)} = \dfrac{15+13}{\left(\frac{15}{75}+\frac{13}{65}\right)} = 70$

Weighted average

When the observations under consideration have a hierarchical order of importance, we take recourse to computing weighted average, which could be either weighted AM or weighted GM or weighted HM.

Weighted AM $= \dfrac{\sum w_i x_i}{\sum w_i}$

Weighted GM = Ante log $\left(\frac{\Sigma w_i \log x_i}{\Sigma w_i}\right)$

Weighted HM = $\frac{\Sigma w_i}{\Sigma\left(\frac{w_i}{x_i}\right)}$

Question 26: Given two positive numbers a and b, prove that AH=G2. Does the result hold for any set of observations?

Solution: For two positive numbers a and b, we have,

$A = \frac{a+b}{2}, G = \sqrt{ab}$ And $H = \frac{2}{\frac{1}{a}+\frac{1}{b}} = \frac{2ab}{a+b}$

Thus, $AH = \frac{a+b}{2} \times \frac{2ab}{a+b} = ab = G^2$

This result holds for only two positive observations and not for any set of observations.

Question 27: The AM and GM for two observations are 5 and 4 respectively. Find the two observations.

Solution: If a and b are two positive observations then as given

$\frac{a+b}{2} = 5, a+b = 10$

and $\sqrt{ab} = 4, ab = 16$

$\therefore (a-b)^2 = (a+b)^2 - 4ab = 10^2 - 4 \times 16 = 36$

$a - b = 6$ (ignoring the negative sign)

$2a = 16$ then $a = 8$

By substituting, we get $b = 10 - a = 2$, Thus, the two observations are 8 and 2.

Question 28: Find the mean and median from the following data:

Marks	Less than 10	Less than 20	Less than 30	Less than 40	Less than 50
No.of Students	5	13	23	27	30

Also compute the mode using the approximate relationship between mean, median and mode.

Solution: What we are given in this problem is less than cumulative frequency distribution. We need to convert this cumulative frequency distribution to the corresponding frequency distribution and thereby compute the mean and median.

Table : Computation of Mean Marks for 30 students

Marks (Class Interval)	Number of students	Mid-Vlaue(xi)	fixi
0-10	5	5	25
10-20	13-5=8	15	120
20-30	23-13=10	25	250
30-40	27-23=4	35	140
40-50	30-27=3	45	135
Total	30		670

Hence the mean mark is given by = $\bar{x} = \frac{\Sigma f_i x_i}{N} = \frac{670}{30}$ = **22.33**

Table: Computation of Median Marks

Marks	Less than 10	Less than 20	Less than 30	Less than 40	Less than 50
No.of Students	5	13	23	27	30

Since $= \frac{N}{2} = \frac{30}{2} = 15$ **lies between 13 and 23, we have l_1 = 20, N_l = 13, N_u = 23** and $C = l_2 - l_1 = 30 - 20 = 10$

Thus, Median = $20 + \frac{15 - 13}{23 - 13} \times 16 = 22$

Since Mode = 3 Median – 2 Mean, approximately, we find that Mode = 3x22 – 2x22.33 = 21.34

A General review of the different Measures of Central Tendency:

1. The best measure of central tendency, usually, is the AM. It is rigidly defined, based on all the observations,
2. Easy to comprehend, simple to calculate and amenable to mathematical properties. However, AM has one drawback in the sense that it is very much affected by sampling fluctuations.
3. In case of frequency distribution, mean cannot be advocated for open-end classification.
4. Like AM, median is also rigidly defined and easy to comprehend and compute.
5. But median is not based on all the observation and does not allow itself to mathematical treatment.
6. Median is not much affected by sampling fluctuation and it is the most appropriate measure of central tendency for an open-end classification.
7. Although mode is the most popular measure of central tendency, there are cases when mode remains undefined. Unlike mean, it has no mathematical property. Mode is also affected by sampling fluctuations.
8. GM and HM, like AM, possess some mathematical properties. They are rigidly defined and based on all the observations. But they are difficult to comprehend and compute and, as such, have limited applications for the computation of average rates and ratios and such like things.

Chapter 14:

Unit-II: Measures of Dispersion

At the foundation level with regards to Paper 3 Statistics part of the topic Measures of Dispersion is very important for students not only to acquire professional knowledge but also for examination point of view. Here in this Chapter an attempt is made for solving and understanding the concepts of Measures of Dispersion with the help of following questions with solutions

The second important characteristic of a distribution is given by dispersion. Two distributions may be identical in respect of its first important characteristic i.e. central tendency and yet they may differ on account of scatterness. The following figure shows a number of distributions having identical measure of central tendency and yet varying measure of scatterness. Obviously, distribution is having the maximum amount of dispersion.

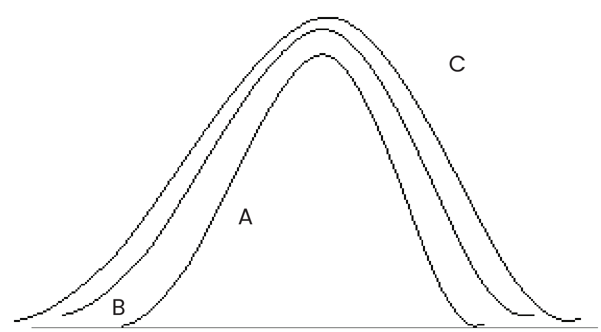


Figure: Showing distributions with identical measure of central tendency and varying amount of dispersion

Dispersion

- for a given set of observations may be defined as the amount of deviation of the observations, usually, from an appropriate measure of central tendency. Measures of dispersion may be broadly classified into

Absolute measures of dispersion.

(i) Range
(ii) Mean Deviation
(iii) Standard Deviation
(iv) Quartile Deviation

Relative measures of dispersion.

(i) Coefficient of range.
(ii) Coefficient of Mean Deviation
(iii) Coefficient of Variation
(iv) Coefficient of Quartile Deviation

Measures of Dispersion

Distinction between the absolute and relative measures of dispersion:

- Absolute measures are dependent on the unit of the variable under consideration whereas the relative measures of dispersion are unit free.
- For comparing two or more distributions, relative measures and not absolute measures of dispersion are considered.
- Compared to absolute measures of dispersion, relative measures of dispersion are difficult to compute and comprehend.

Characteristics for an ideal measure of dispersion

- An ideal measure of dispersion should be properly defined, easy to comprehend, simple to compute, based on all the observations, unaffected by sampling fluctuations and amenable to some desirable mathematical treatment

Range :

- For a given set of observations, range may be defined as the difference between the largest and smallest of observations.
- Thus if L and S denote the largest and smallest observations respectively thenRange = L – S

Coefficient of range = $\frac{L - S}{L + S} \times 100$

For a grouped frequency distribution: Range is defined as the difference between the two extreme class boundaries. The corresponding relative measure of dispersion is given by the ratio of the difference between the two extreme class boundaries to the total of these class boundaries, expressed as a percentage.

Range remains unaffected due to a change of origin but affected in the same ratio due to a change in scale i.e., if for any two constants a and b, two variables x and y are related by y = a + bx,

Then the range of y is given by $R_y = |b| \times R_x$

Example 1:

Following are the wages of 8 workers expressed in rupees: 80, 65, 90, 60, 75, 70, 72, 85. Find the range and find its coefficient.

Solution:

The largest and the smallest wages are L = ₹96 and S= ₹50

Thus range = ₹90 – ₹60 = ₹30

Coefficient of range = $\frac{90 - 60}{90 + 60}$ x 100 = 20

Example 2:

What is the range and its coefficient for the following distribution of weights?

Weight in kgs	10-19	20-29	30-39	40-49	50-59
No. of Students	11	25	16	7	3

Solution:

The lowest class boundary is 9.50 kgs. and the highest-class boundary is 59.50 kgs. Thus, we have

Range = 59.50 kgs. – 9.50 kgs. = 50 kgs.

Also, coefficient of range = $\frac{59.50 - 9.50}{59.50 + 9.50}$ x 100 = $\frac{50}{69}$ x 100 = 72.46

Example 3:

If the relationship between x and y is given by 2x+3y=10 and the range of x is ` 15, what would be the range of y?

Solution:

Since 2x+3y=10

Therefore, y = $\frac{10}{3} - \frac{2}{3}$ x, Applying the range of y is given by

R_y | b | x R_x = 2/3 × ₹ 15 = ₹10.

Mean Deviation : Since range is based on only two observations, it is not regarded as an ideal measure of dispersion. A better measure of dispersion is provided by mean deviation which, unlike range, is based on all the observations. For a given set of observation, mean deviation is defined as the arithmetic mean of the absolute deviations of the observations from an appropriate measure of central tendency. Hence if a variable x assumes n values $x_1, x_2, x_3 \ldots x_n$, then the mean deviation of x about an average A is given by

$MD_A = \frac{1}{n}\Sigma | x_i - A |$ For a grouped frequency distribution, mean deviation about A is given by MDA =

$\frac{1}{n}\Sigma | x_i - A | . f_i$

Where x_i and f_i denote the mid value and frequency of the i^{th} class interval and $N = \Sigma f_i$

In most cases we take A as mean or median and accordingly, we get mean deviation about mean or mean deviation about median.

A relative measure of dispersion applying mean deviation is given by

Coefficient of Mean Deviation = $\frac{\text{Mean deviation about A}}{\text{A}} \times 100$

Mean deviation takes its minimum value when the deviations are taken from the median. Also mean deviation remains unchanged due to a change of origin but changes in the same ratio due to a change in scale i.e. if y = a + bx, a and b being constants,

Example 4:

What is the mean deviation about mean for the following numbers?
50,60,50,50,60,60,60,50,50,50,60,60,60,50.

Solution:

The mean is given by

$$\bar{X} = \frac{50 + 60 + 50 + 50 + 60 + 60 + 60 + 50 + 50 + 50 + 60 + 60 + 60 + 50}{14} = \frac{770}{14} = 55$$

x_i	50	60	50	50	60	60	60	50	50	50	60	60	60	50	Total
$\lvert x_i - \bar{x} \rvert$	5	5	5	5	5	5	5	5	5	5	5	5	5	5	70

Thus, mean deviation about mean is given by $\frac{\Sigma \lvert x_i - \bar{x} \rvert}{n} = \frac{70}{14} = 5$

Example 5:

The coefficient of Mean Deviation about the first 9 natural numbers ?

Solution:

The Mean of first 9 natural numbers = $\frac{n+1}{2} = \frac{9+1}{2} = 5$

coefficient of Mean Deviation about the first 9 natural numbers = $\frac{\text{Mean deviation about A}}{\text{A}} \times 100 = \frac{400}{9}$

Example 6:

The mean deviation about the mode for the following observations 4/11, 6/11, 8/11, 9/11, ,12/11, 8/11 is

Answer:

For the 4/11, 6/11, 8/11, 9/11, ,12/11, 8/11 Mode is 8/11

Mean deviation from Mode = $\dfrac{\Sigma \mid x_i - \text{Mode} \mid}{n}$

$$= \frac{\left|\frac{4}{11} - \frac{8}{11}\right| + \left|\frac{6}{11} - \frac{8}{11}\right| + \left|\frac{8}{11} - \frac{8}{11}\right| + \left|\frac{9}{11} - \frac{8}{11}\right| + \left|\frac{12}{11} - \frac{8}{11}\right| + \left|\frac{8}{11} - \frac{8}{11}\right|}{6}$$

$$= \frac{\frac{4}{11} + \frac{2}{11} + 0 + \frac{1}{11} + \frac{4}{11} + 0}{6} = \frac{\frac{1}{11}}{6} = \frac{6}{11}$$

Example 7:

Find mean deviations about median and the corresponding coefficient for the following profits ('000 ₹) of a firm during a week. 82, 56, 75, 70, 52, 80, 68.

Solution:

The profits in thousand rupees is denoted by x. Arranging the values of x in an ascending order, we get 52, 56, 68, 70, 75, 80, 82.

Therefore, Me = 70. Thus, Median profit = ₹ 70,000.

Computation of Mean deviation about median

x_i	52	56	68	70	75	80	82	Total
$\lvert x_i - Me \rvert$	18	14	2	0	5	10	12	61

Thus mean deviation about median = $\dfrac{\Sigma \mid x_i - \text{Median} \mid}{n}$ = ₹ $\dfrac{61}{7}$ = 1000 = ₹ 8714.28

Coefficient of mean deviation = $\dfrac{\text{MD about median}}{\text{Median}} \times 100 = \dfrac{8714.28}{70000} \times 100 = 12.45$

Example 8:

Compute the mean deviation about the arithmetic mean for the following data:

Variable (X)	5	10	15	20	25	30
Frequency (f)	3	4	6	5	3	2

Solution:

We are to apply formula as these data refer to a grouped frequency distribution the AM is given by

$$\bar{x} = \frac{\Sigma f_i x_i}{N} = \frac{5 \times 3 + 10 \times 4 + 15 \times 6 + 20 \times 5 + 25 \times 3 + 30 \times 2}{3 + 4 + 6 + 5 + 3 + 2} = 16.52$$

$$\text{Mean deviation from Mean} = \frac{\Sigma fi.\,|x_i - \bar{x}|}{n} = \frac{139.56}{23} = 6.07$$

$$\text{Coefficient of MD about its AM} = \frac{\text{MD about AM}}{\text{AM}} \times 100 = \frac{6.07}{16.52} \times 100 = 44.33$$

Example 9:

Compute the coefficient of mean deviation about median for the following distribution:

Weight in kgs.	60-62	63-65	66-68	69-71	72-74
No. of persons	5	22	28	17	3

Solution:

We need to compute the median weight in the first stage

$$\text{Hence, } M = l_1 + \left(\frac{\frac{N}{2} - Nl}{N_u - Nl}\right) \times C = \left[60 + \frac{25 - 20}{40 - 20} \times 10\right] \text{kg.} = 62.50\text{kg.}$$

Weight (kgs)(1)	Mid-value (xi) kgs (2)	No.of persons (3)	kgs (4)	(kgs) (5) = (3)×(4)
60-62	45	5	17.50	140
63-65	55	22	7.50	90
66-68	65	28	2.50	50
69-71	75	17	12.50	125
72-74		3		
Total				405

$$\text{Mean deviation about median} = \frac{\Sigma f_i\,|x_i - \text{Median}|}{N} = \frac{405}{50} \text{kg.} = 8.10 \text{ kg.}$$

Coefficient of mean deviation about median= $\frac{\text{Mean deviation about median}}{\text{Median}}$ x 100 = $\frac{8.10}{62.50}$ x 100 = 12.96

Example 10:

If x and y are related as 4x+3y+11 = 0 and mean deviation of x is 5.40, what is the mean deviation of y?

Solution:

Since 4x + 3y + 11 = 0

Therefore, y = $\left(\frac{-11}{3}\right) + \left(\frac{-4}{3}\right)x$

Hence MD of y= |b| φ MD of x = $\frac{4}{3}$ x 5.40 = 7.20

Standard Deviation: Although mean deviation is an improvement over range so far as a measure of dispersion is concerned, mean deviation is difficult to compute and further more, it cannot be treated mathematically. The best measure of dispersion is, usually, standard deviation which does not possess the demerits of range and mean deviation.

Standard deviation for a given set of observations is defined as the root mean square deviation when the deviations are taken from the AM of the observations. If a variable x assumes n values x_1, x_2, x_3 ....x_n then its standard deviation(s) is given by

$$S = \sqrt{\frac{\Sigma(x_i - \bar{x})^2}{n}}$$

For a grouped frequency distribution, the standard deviation is given by

$$S = \sqrt{\frac{\Sigma f_i(x_i - \bar{x})^2}{n}}$$

can be simplified to the following forms for unclassified data $S = \sqrt{\frac{\Sigma x_i^2}{n} - \bar{x}^2}$

$= \sqrt{\frac{\Sigma f_i x_i^2}{N} - \bar{x}^2}$ for a grouped frequency distribution.

Variance: The square of standard deviation, known as variance

Variance = $S^2 = \sqrt{\frac{\Sigma x_i - \bar{x}^2}{n}}$ for unclassified data

$= \sqrt{\frac{\Sigma f_i(x_i - \bar{x}^2}{N}}$ for a grouped frequency distribution

Coefficient of variation (CV) = $\frac{SD}{AM}$ x 100

(A relative measure of dispersion using standard deviation is given by Coefficient of Variation (CV) which is defined as the ratio of standard deviation to the corresponding arithmetic mean, expressed as a percentage.)

Example 11:

Find the standard deviation and the coefficient of variation for the following numbers: 5, 8, 9, 2, 6

Solution:

We present the computation in the following table: **Computation of standard deviation**

xi	**5**	**8**	**9**	**2**	**6**	**Total**
x_i^2	25	64	81	4	36	210

Applying, we get the standard deviation as

$$= \sqrt{\frac{\Sigma x_i^2}{n} - \bar{x}^2} = \sqrt{\frac{210}{5} - \left(\frac{30}{5}\right)^2} \qquad \left(\text{since } \bar{x} = \frac{\Sigma x_i}{n}\right)$$

$$= \sqrt{42 - 36} = \sqrt{6} = 2.45$$

The coefficient of variation is $CV = 100 \frac{SD}{AM} = 100 \frac{2.45}{6} = 40.83$

We consider the following formula for computing standard deviation from grouped frequency distribution with a view to saving time and computational labour:

$$S = \sqrt{\frac{\Sigma f_i d_i^2}{N} - \left(\frac{\Sigma f_i d_i}{N}\right)^2}, \text{ Where } d_i = \frac{x_i - A}{C}$$

Properties of standard deviation

1. If all the observations assumed by a variable are constant i.e. equal, then the SD is zero. This means that if all the values taken by a variable x is k, say , then s = 0. This result applies to range as well as mean deviation.
2. SD remains unaffected due to a change of origin but is affected in the same ratio due to a change of scale i.e., if there are two variables x and y related as y = a+bx for any two constants a and b, then SD of y is given by $s_y = |b| s_x$
3. If there are two groups containing n_1 and n_2 observations, $_1$ and $_2$ as respective AM's, s_1 and s_2 as respective SD's, then the combined SD is given by

   $$S = \sqrt{\frac{n^1 s_1^2 + n_2 s_2^2 + n_1 d_1^2 + n_2 d_2^2}{n_1 + n_2}} \text{ where, } d_1 = \bar{X}1 - \bar{X},\ d2 = \bar{X}2 - \bar{X} \text{ and } \quad \bar{X} = \frac{n_1\bar{X}_1 + n_2\bar{X}_2}{n_1 + n_2} = \text{combined AM}$$

   This result can be extended to more than 2 groups, we have

   $$S = \sqrt{\frac{\Sigma n_i s_i^2 + \Sigma n_i d_i^2}{\Sigma n_i}} \quad \text{With } d = xi - \bar{x} \text{ and } \bar{x} = \frac{\Sigma n_i \bar{x}_i}{\Sigma n_i}$$

   Where $\bar{X}_1 = \bar{X}_2$ is reduced to $s = \sqrt{\frac{n_1 s_1^2 + n_2 s_2^2}{n_1 + n_2}}$
4. For any two numbers a and b, standard deviation is given by $\frac{|a - b|}{2}$
5. SD of first n natural numbers is $SD = \frac{\sqrt{n^2 - 1}}{12}$

Example 12:

If the S.D. of x is 3, what is the variance of (5 – 2x)?

Solution:

If y = a + bx, then σ_y |b| σ_x

Let y = 5 – 2x

$\therefore \sigma_y = |-2|\, \sigma x$

= 2 × 3 = 6

$\therefore$ Variance (5 – 2x) = $(2)^2 \times 9 = 36$

Example 13:

The coefficient of variation of the following numbers.53, 52, 61, 60 64, is

Ans :

$$\bar{x} = \frac{(53 + 52 + 61 + 60 + 64)}{5} = 58$$

$$\therefore \sigma = \sqrt{\frac{\Sigma\,(x - \bar{x})^2}{n}}$$

$$\therefore \sigma = \sqrt{\frac{(-5)^2 + (-6)^2 + 3^2 + 2^2 + 6^2}{5}} = 4.69.$$

$$\text{Coefficient of variation} = \frac{\text{S.D.}}{\text{A.M.}} \times 100$$

$$= \frac{4.6904}{58} \times 100 = 8.09.$$

Example 14:

What is the standard deviation of 5,5,9,9,9,10,5,10,10?

Solution:

$$\text{Mean} = \frac{3 \times 5 + 3 \times 9 + 3 \times 10}{9} = \frac{72}{9} = 8.$$

$$\text{Standard Deviation} = \sqrt{\frac{\Sigma f(x_i - \bar{x})^2}{n}} = \sqrt{\frac{3(9) + 3(1) + 3(4)}{9}} = \sqrt{42 / 9} = 2.16.$$

Example 15:

If x and y are related by 2x + 3y + 4 = 0 and S.D. of x is 6, then S.D. of y is :

Solution: y = (-2/3)x – 2 ;

$\sigma_y = | -2 | 3 | \sigma_x = (2 / 3)\ 6 = 4.$

Example 16:

If x and y are related by y = 2x + 5 and the S.D. and A.M. of x are known to be 5 and 10 respectively, then the coefficient of variation of y is:

Ans Y = 2x+5

$\sigma_y = | -2 | \sigma_x = 2 \times 5 = 10.$

Also $\bar{y} = 2\bar{x} + 5 = 20 + 5 = 25$

Coefficient of variation of y $\frac{\sigma_y}{y} \times 100 = \frac{10}{25} \times 100 = 40.$

Example 17:

If the mean and S.D. of x are a and b respectively, then the S.D. of $\frac{x - a}{b}$ is

Solution: Let $y = \frac{(x - a)}{b} = \frac{1}{b} . x - \frac{a}{b}$

$\sigma_y = \left| \frac{1}{b} \right| \sigma_x$

$\frac{1}{b} b = 1.$

Example 18:

If x and y all related by 3y = 7x – 9 and the S.D. of y is 7, then what is he variance of x?

Solution: 3y = 7x – 9

$X = \frac{3}{7}\ y + 9$

Also $\sigma_x \left| \frac{3}{7} \right| \sigma_y = \frac{3}{7} \times 7$

$\therefore$ Variance : $\sigma_x^2 = 3^2 = 9.$

Example 19:

Which of the following companies A and B is more consistent so far as the payment of dividend is concerned ?

Dividend paid by A	5	9	6	12	15	10	8	10
Dividend paid by B:	4	8	7	15	18	9	6	6

Solution: Here $\Sigma x_A = 75$

$\therefore \bar{x}_A = 75 / 8 = 9.375$

$\Sigma x_A^2 = 775$

$$\sigma_A^2 = \frac{\Sigma x_A^2}{N} - \left(\frac{\Sigma x_A}{N}\right)^2$$

$$= \frac{775}{8} - \left(\frac{75}{8}\right)^2 = 9$$

$\sigma_A = 3.$

$$C.V_A = \frac{\sigma_A}{\bar{x}} x\,100 = \frac{30}{9.375} x\,100 = 32.$$

Also $\Sigma x_B = 73$, $\therefore x_B = 73.18 = 9.125$

$\Sigma xB^2 = 831$,

$$\sigma_B^2 = \frac{831}{8} - \left(\frac{73}{8}\right)^2 = 20.61$$

$$\therefore \sigma_B = \sqrt{20.61} = 4.54$$

$$C.V_B = \frac{4.54}{9.125} x\,100 = 49.75$$

$C.V_A < C.V_B$

Company A is more consistent

Example 20:

Find the SD of the following distribution:

Weight(kgs)	50-52	52-54	54-56	56-58	58-60
No.of Students	17	35	28	15	5

Solution:

Weight	No .of Students	Mid – Value	$d_i = x_i - 55$	$f_i d_i$	$f_i d_i^2$
50-52	17	51	-2	-34	68
52-54	35	53	-1	-35	35
54-56	28	55	0	0	0
56-58	15	57	1	15	15
58-60	5	59	2	10	20
Total	100			-44	138

Applying, we get the SD of weight as

$$= \sqrt{\frac{\Sigma f_i d_i^2}{N} - \left(\frac{\Sigma f_i d_i}{N}\right)^2} \times C = \sqrt{\frac{138}{100} - \frac{(-44)^2}{100}} \text{ x 2 kgs.} = \sqrt{1.38 - 0.1936} \text{ x 2 kgs.} = 2.18 \text{ kgs}$$

Example 21:

The mean and variance of the 10 observations are found to be 17 and 33 respectively. Later it is found that one observation (i.e.26) is inaccurate and is removed. What is mean and standard deviation of remaining?

Solution:

Mean of 10 observations = 17 then Total of the observations = 17 ×10 = 170

Total of the 9 observations = 170-26 = 144

Changed Mean = 144/9 = 16

Variance (σ^2) = 33

$$\frac{\Sigma x^2}{n} - (17)^2 = 33 \Rightarrow \frac{\Sigma x^2}{10} = 33 + 289 = 322$$

$$\frac{\Sigma x^2}{10} = 322$$

$\Sigma x^2 = 3220 - (26)^2 = 3220 - 676 = 2544$

$$\text{Chnged Variance} = \frac{\text{Changed } \Sigma x^2}{n} - (\text{Changed } \bar{x})^2$$

$$= \frac{2544}{9} - (16)^2 = 26.67$$

SD of remaining observations = $\sqrt{26.67} = 5.16$

Example 22:

If AM and coefficient of variation of x are 10 and 40 respectively, what is the variance of (15−2x)?

Solution: let y = 15 − 2x

Then applying (11.34), we get,

$s_y = 2 \times s_x$

As given cv_x = coefficient of variation of x = 40 and= 10

Thus $cv_x = \frac{S_x}{X} \times 100$ $\quad 40 = 10\frac{S_x}{X} \times 100$

$\Rightarrow s_x = 4$

Then, $S_y = 2 \times 4 = 8$ Therefore, variance of (15-2x) = $S_y^2 = 64$

Example 23:

Compute the SD of 9, 5, 8, 6, 2. Without any more computation, obtain the SD of

Sample I	-1	-5	-2	-4	-8
Sample II	90	50	80	60	20
Sample III	23	15	21	17	9

Solution:

x_i	9	5	8	6	2	30
x_i^2	81	25	64	36	4	210

The SD of the original set of observations is given by

$$s = \sqrt{\frac{210}{5} - \left(\frac{30}{5}\right)^2} = \sqrt{\frac{\Sigma x_i^2}{n} - \left(\frac{\Sigma x_i}{n}\right)^2} = \sqrt{42 - 36} = \sqrt{6} = 2.45$$

If we denote the original observations by x and the observations of sample I by y, then we have

$y = -10 + x$

$y = (-10) + (1)x$

$\therefore S_y = |1| \times S_x = 1 \times 2.45 = 2.45$

In case of sample II, x and y are related as

$Y = 10x = 0 + (10)x$

$\therefore s_y = |10| \times s_x$

$= 10 \times 2.45 = 24.50$

And lastly, $y = (5)+(2)x$

$\Rightarrow s_y = 2 \times 2.45 = 4.90$

Example 24:

For a group of 60 boy students, the mean and SD of stats. marks are 45 and 2 respectively. The same figures for a group of 40 girl students are 55 and 3 respectively. What is the mean and SD of marks if the two groups are pooled together?

Solution:

As given $n_1 = 60, \bar{x}_1 = 45, s_1 = 2\ n_2 = 40,\ \bar{x}_1 = \bar{x}_1 = 55, s_2 = 3$

Thus the combined mean is given by

$$\bar{x}_1 = \frac{n_1\bar{x}_1 + n_2\bar{x}_2}{n_1 + n_2} = \frac{60 \times 45 + 40 \times 55}{60 + 40} = 49$$

Thus $d_1 = \bar{x}_1 = \bar{x} = 45 - 49 = -4$

$d_2 = \bar{x}_2 = \bar{x} = 55 - 49 = 6$

Applying (11.35), we get the combined SD as

$$s = \sqrt{\frac{n_1s_1^2 + n_2s_2^2 + n_1d_1^2 + n_2d_2^2}{n_1 + n_2}}$$

$$s = \sqrt{\frac{60 \times 2^2 + 40 \times 3^2 + 60 \times (-4)^2 + 40 \times 6^2}{60 + 40}}$$

$$= \sqrt{30}$$

$$= 5.48$$

Example 25:

The mean and standard deviation of the salaries of the two factories are provided below:

Factory	No. of Employees	Mean Salary	SD of Salary
A	30	₹4800	₹10
B	20	₹5000	₹12

(i) Find the combined mean salary and standard deviation of salary.

(ii) Examine which factory has more consistent structure so far as satisfying its employees are concerned.

Solution: Here we are given

$n_1 = 30,\ \bar{x}_1 = ₹4800, s_1 = ₹10,$

$n_2 = 20, \bar{x}_2 = ₹5000, s_2 = ₹12$

(i) $= \dfrac{30 \times ₹4800 + 20 \times 5000}{30 + 20}$ ₹4800

$d_1 = x_1 = \bar{x} = ₹4,800 - ₹4880 = -₹80$

$d_2 = \bar{x}_2 - \bar{x} = ₹5,000 - ₹4880 = ₹120$

hence, the combined SD in rupees is given by $s = \sqrt{\dfrac{30 \times 10^2 + 20 \times 12^2 + 30 \times (-80)^2 + 20 \times 120^2}{30 + 20}} = \sqrt{9717.60} = 98.58$

thus the combined mean salary and the combined standard deviation of salary are `4880 and `98.58 respectively.

(ii) In order to find the more consistent structure, we compare the coefficients of variation of the two factories.

Letting $CV_A = 100 \times \frac{S_A}{X_A}$ and $CV_B = 100 \times \frac{S_B}{X_B}$

We would say factory A is more consistent

if $CV_A < CV_B$. Otherwise factory B would be more consistent.

Now $CA_A = 100 \times \frac{S_A}{X_A} \times 100 \times \frac{S_1}{X_1} = \frac{100 \times 10}{4800} = 0.21$

and $CA_B = 100 \times \frac{S_B}{X_B} \times 100 \times \frac{S_2}{X_2} = \frac{100 \times 12}{5000} = 0.24$

Thus we conclude that factory A has more consistent structure.

Quartile Deviation: Another measure of dispersion is provided by quartile deviation or semi - inter –quartile range which is given by

$$Q_d = \frac{Q_3 - Q_1}{2}$$

A relative measure of dispersion using quartiles is given by coefficient of quartile deviation which is

$$\text{Coefficient of quartile deviation} = \frac{Q_3 - Q_1}{Q_3 + Q_1} \times 100$$

Merits

1) Quartile deviation provides the best measure of dispersion for open-end classification.
2) It is also less affected due to sampling fluctuations.
3) Like other measures of dispersion, quartile deviation remains unaffected due to a change of origin but is affected in the same ratio due to change in scale.

Example 26:

The quartiles of a variable are 45,52, and 65 respectively. Its quartile deviation is:

Solution:

$$\text{Quartile Deviation} = \frac{Q_3 - Q_1}{2} = \frac{65 - 45}{2} = 10$$

Example 27:

If x and y are related as $3x + 4y = 20$ and the quartile deviation of x is 12, then the quartile deviation of y is

Solution: If $y = ax + b$

Q.D of y = a × (Q.D. of x)

$3x + 4y = 20$

$= y = \frac{3}{4}x + 5$

Q.D. of y = (3/4) (Q.D. of x)

= (3/4) 12 = 9.

Example 28:

Following are the marks of the 10 students : 56, 48, 65, 35, 42, 75, 82, 60, 55, 50. Find quartile deviation and also its coefficient.

Solution:

After arranging the marks in an ascending order of magnitude, we get 35, 42, 48, 50, 55, 56, 60, 65, 75, 82

First quartile observation $(Q1 = \frac{(n+1)}{4}$ th observation$= \frac{(10+1)}{4}$ th observation

= 2.75^{th} observation

= 2^{nd} observation + 0.75 × difference between the third and the 2^{nd} observation.

= 42 + 0.75 × (48 – 42)

= 46.50

Third quartile (Q3) = $\frac{3(n+1)}{4}$ th observation

= 8.25^{th} observation

= 65 + 0.25 × 10

= 67.50

Thus applying, we get the quartile deviation as

$$\frac{Q_3 - Q_1}{2} = \frac{67.50 - 46.50}{2} = 10.50$$

Also, using the coefficient of quartile deviation$= \frac{Q_3 - Q_1}{Q_3 + Q_1} \times 100 = \frac{67.50 - 46.50}{67.50 - 46.50} = 18.42$

Example 29:

If the quartile deviation of x is 6 and 3x + 6y = 20, what is the quartile deviation of y?

Solution: 3x + 6y = 20

$$y\left(\frac{20}{6}\right)+\left(\frac{-3}{6}\right)x$$

Therefore, quartile deviation of y = $\frac{|-3|}{6}$ X quartile deviation of X

$$= \frac{1}{2} \times 6 = 3.$$

Example 30:

Find an appropriate measures of dispersion from the following data:

Daily wages (`)	upto 20	20-40	40-60	60-80	80-100
No. of workers	5	11	14	7	3

Solution: Since this is an open-end classification, the appropriate measure of dispersion would be quartile deviation as quartile deviation does not taken into account the first twenty five percent and the last twenty five per cent of the observations.

Table Computation of Quartile

Daily wages in ` (Class boundary)	No. of workers (less than cumulative frequency)
a	0
20	5
40	16
60	30
80	37
100	40

Here a denotes the first Class Boundary

$$Q1 = \left[20 + \frac{10-5}{16-5} \times 20\right] = ` 29.09$$

Q3 = 60

Thus quartile deviation of wages is given by = $\frac{Q_3 - Q_1}{2} = \frac{₹60 - ₹29.09}{2} = ₹15.46$

Example 31:

The mean and variance of 5 observations are 4.80 and 6.16 respectively. If three of the observations are 2,3 and 6, what are the remaining observations?

Solution: Let the remaining two observations be a and b, then as given

$$\frac{2 + 3 + 6 + a + b}{5}$$

$\Rightarrow$ 11+a+b =24 $\Rightarrow$ a+b =13 .... (1)

and $\dfrac{2^2 + a^2 + b^2 + 3^2 + 6^2}{5} - (4.80)^2 \Rightarrow \dfrac{49^2 + a^2 + b^2}{5} - 23.04 = 6.16$

$\Rightarrow 49 + a^2 + b^2 = 146$

$\Rightarrow a^2 + b^2 = 97$ ....(2)

From (1), we get a = 13 – b .... (3)

Eliminating a from (2) and (3), we get

$(13 - b)^2 + b^2 = 97 \Rightarrow$ $169 - 26b + 2b^2 = 97$

$\Rightarrow b^2 - 13b + 36 = 0$

$\Rightarrow (b-4)(b-9) = 0$

$\Rightarrow$ b = 4 or 9

From (3), a= 9 or 4

Thus the remaining observations are 4 and 9.

Comparison between different measures of dispersion

We may now have a review of the different measures of dispersion on the basis of their relative merits and demerits.

1. Standard deviation, like AM, is the best measure of dispersion. It is rigidly defined, based on all the observations, not too difficult to compute, not much affected by sampling fluctuations and moreover it has some desirable mathematical properties. All these merits of standard deviation make SD as the most widely and commonly used measure of dispersion.

2. Range is the quickest to compute and as such, has its application in statistical quality control. However, range is based on only two observations and affected too much by the presence of extreme observation(s).

3. Mean deviation is rigidly defined, based on all the observations, and not much affected by sampling fluctuations. However, mean deviation is difficult to comprehend and its computation is also time consuming and laborious. Furthermore, unlike SD, mean deviation does not possess mathematical properties.

4. Quartile deviation is also rigidly defined, easy to compute and not much affected by sampling fluctuations. The presence of extreme observations has no impact on quartile deviation since quartile deviation is based on the central fifty-percent of the observations. However, quartile deviation is not based on all the observations and it has no desirable mathematical properties. Nevertheless, quartile deviation is the best measure of dispersion for open-end classifications.

Chapter – 15:

Probability

At the foundation level the concept of Probability is used in accounting and finance to understand the likelihood of occurrence or non- occurrence of a variable. It helps in developing financial forecasting in which you need to develop expertise at an advanced stage of chartered accountancy course. Here in this chapter an attempt is made for solving and understanding the concepts of with the help of following questions with solutions.

The terms 'Probably' 'in all likelihood', 'chance', 'odds in favour', 'odds against' are too familiar nowadays and they have their origin in a branch of Mathematics, known as Probability. In recent time, probability has developed itself into a full-fledged subject and become an integral part of statistics.

Random Experiment: An experiment is defined to be random if the results of the experiment depend on chance only. For example if a coin is tossed, then we get two outcomes—Head (H) and Tail (T). It is impossible to say in advance whether a Head or a Tail would turn up when we toss the coin once. Thus, tossing a coin is an example of a random experiment. Similarly, rolling a dice (or any number of dice), drawing items from a box containing both defective and non—defective items, drawing cards from a pack of well shuffled fifty—two cards etc. are all random experiments.

Events: The results or outcomes of a random experiment are known as events. Sometimes events may be combination of outcomes. The evyents are of two types:

(i) Simple or Elementary,

(ii) Composite or Compound.

An event is known to be simple if it cannot be decomposed into further events. Tossing a coin once provides us two simple events namely Head and Tail. On the other hand, a composite event is one that can be decomposed into two or more events. Getting a head when a coin is tossed twice is an example of composite event as it can be split into the events HT and TH which are both elementary events.

Mutually Exclusive Events or Incompatible Events:

A set of events A_1, A_2, A_3, ... is known to be mutually exclusive if not more than one of them can occur simultaneously. Thus, occurrence of one such event implies the non-occurrence of the other events of the set. Once a coin is tossed, we get two mutually exclusive events Head and Tail.

Exhaustive Events:

The events A_1, A_2, A_3, ... are known to form an exhaustive set if one of these events must necessarily occur. As an example, the two events Head and Tail, when a coin is tossed once, are exhaustive as no other event except these two can occur.

Equally Likely Events or Mutually Symmetric Events or Equi-Probable Events:

The events of a random experiment are known to be equally likely when all necessary evidence are taken into account, no event is expected to occur more frequently as compared to the other events of the set of events. The two events Head and Tail when a coin is tossed is an example of a pair of equally likely events because there is no reason to assume that Head (or Tail) would occur more frequently as compared to Tail (or Head).

Classical Definition of Probability or A Prior Definition

Let us consider a random experiment that result in n finite elementary events, which are assumed to be equally likely. We next assume that out of these n events, nA (n) events are favourable to an event A. Then the probability of occurrence of the event A is defined as the ratio of the number of events favourable to A to the total number of events. Denoting this by P(A), we have

$$P(A) = \frac{n_A}{n} = \frac{\textit{Number} \text{ of equally likely events favourable to A}}{\text{Total number of equally likely events}}$$

However, if instead of considering all elementary events, we focus our attention to only those composite events, which are mutually exclusive, exhaustive and equally likely and if m(n) denotes such events and is furthermore mA(nA) denotes the no. of mutually exclusive, exhaustive and equally likely events favourable to A, then we have

$$P(A) = \frac{m_A}{m} = \frac{\text{Number of mutually exclusive, exhaustive and equally likely events favourable to A}}{\text{Total Number of mutually exclusive, exhaustive and equally likely events}}$$

Probability and Expected Value by Mathematical Expectation

For this definition of probability, we are indebted to Bernoulli and Laplace. This definition is also termed as a priori definition because probability of the event A is defined based on prior knowledge.

This classical definition of probability has the following demerits or limitations:

(i) It is applicable only when the total no. of events is finite.

(ii) It can be used only when the events are equally likely or equi-probable. This assumption is made well before the experiment is performed.

(iii) This definition has only a limited field of application like coin tossing, dice throwing, drawing cards etc. where the possible events are known well in advance. In the field of uncertainty or where no prior knowledge is provided, this definition is inapplicable.

In connection with classical definition of probability, we may note the following points:

(a) The probability of an event lies between 0 and 1, both inclusive.

i.e. $0 \le P(A) \le 1$

When P(A) = 0, A is known to be an impossible event and when P(A) = 1, A is known to be a sure event.

(b) Non-occurrence of event A is denoted by A′ or AC or and it is known as complimentary event of A. The event A along with its complimentary A′ forms a set of mutually exclusive and exhaustive events.

i.e. $P(A) + P(A') = 1 \quad P(A') = 1 - P(A) = 1 - \frac{m_A}{m} = \frac{m - mA}{m}$

(c) The ratio of no. of favourable events to the no. of unfavourable events is known as odds in favour of the event A and its inverse ratio is known as odds against the event A.

i.e. odds in favour of A $= m_A : (m - m_A)$

and odds against A $= (m - m_A) : m_A$

Statistical definition of Probability :

Owing to the limitations of the classical definition of probability, there are cases when we consider the statistical definition of probability based on the concept of relative frequency. This definition of probability was first developed by the British mathematicians in connection with the survival probability of a group of people.

Let us consider a random experiment repeated a very good number of times, say n, under an identical set of conditions. We next assume that an event A occurs fA times. Then the limiting value of the ratio of fA to n as n tends to infinity is defined as the probability of A.

ie. $P(A) \ \frac{\lim}{n \to \infty} \ \frac{F_A}{n}$

This statistical definition is applicable if the above limit exists and tends to a finite value.

Two events A and B are mutually exclusive if $P(A \cap B) = 0$ or more precisely

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

Similarly, three events A, B and C are mutually exclusive if.

$P(A \cup B \cup C) = P(A) + P(B) + P(C)$

Two events A and B are exhaustive if.

$P(A \cup B) = 1$

Similarly, three events A, B and C are exhaustive if.

$P(A \cup B \cup C) = 1$

Three events A, B and C are equally likely if

$P(A) = P(B) = P(C)$

Axiomatic or modern definition of probability: Then a real valued function p defined on s is known as a probability measure and p(a) is defined as the probability of A if P satisfies the following axioms:

(i) $P(A) \leq 0$ for every $A \subseteq S$ (subset)

(ii) $P(S) = 1$

(iii) For any sequence of mutually exclusive events $A_1, A_2, A_3,..$

$P(A_1 U A_2 U A_3 U....) = P(A_1) + P(A_2) + P(A_3)$

Addition theorems or theorems on total probability: For any two mutually exclusive events a and b, the probability that either a or b occurs is given by the sum of individual probabilities of A and B. i.e. P $(A \cup B)$ or $P(A + B) = P(A) + P(B)$ or P(A or B) whenever A and B are mutually exclusive

For any three events A, B and C, the probability that at least one of the events occurs is given by $P(A \cup B \cup C) = P(A) + P(B) + P(C) - P(A \cap B) - P(A \cap C) - P(B \cap C) + P(A \cap B \cap C)$

(d) For any two events A and B, the probability that either A or B occurs is given by the sum of individual probabilities of A and B less the probability of simultaneous occurrence of the events A and B.

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

For any three events A, B and C, the probability that at least one of the events occurs is given by

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

(e) Two events A and B are mutually exclusive if

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

Similarly, three events A, B and C are mutually exclusive if

$(A \cup B \cup C) = P(A) + P(B) + P(C)$

(f) $P(A-B) = P(A \cap B') = P(A) - P(A \cap B)$

And $P(B - A) = P(B \cap A') = P(B) - P(A \cap B)$

Some important Results

1. **If A and b are two independent events, then the probability of occurrence of both is given by $P(A \cap B) = P(A).P(B)$**
2. **If A, B and C are three events, then. $P(A \cap B \cap C) = P(A).P(B/A).P(C/A \cap B)$**
3. **If A and B are two mutually exclusive events of a random experiment, then.**

   $A \cap B = \phi$, $P(AUB) = P(A) + P(B)$
4. **If A and B are associated with a random experiment, then.**

   $P(AUB) = P(A) + P(B) - P(A \cap B)$
5. **If A, B and C are three events connected with random experiment, then**

   $P(AUBUC) = P(A) + P(B) + P(C) - P(A \cap B) - P(B \cap C) - P(C \cap A) + P(A \cap B \cap C)$

(g) Compound Probability or Joint Probability

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

(h) For any three events A, B and C, the probability that they occur jointly is given by

$P(A \cap B \cap C) = P(A) \quad P(B/A) \quad P(C/(A \cap B))$ Provided $P(A \cap B) > 0$

(i) $P(A'/B) = \dfrac{P(A' \cap B)}{P(B)} = \dfrac{P(B) - P(A \cap B)}{P(B)}$

(j) $P(A/B') = = \dfrac{P(A \cap B')}{P(B')} = \dfrac{P(A) - P(A \cap B)}{1 - P(B)}$

(k) P(A′/B′) = P(A/B′)= = $\frac{P('A \cap B')}{P(B')}$

= $\frac{P(A \cap B)'}{P(B')}$ [by De-Morgan's Law A′∩B′ = (AUB)′]

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

(7) A random variable or stochastic variable is a function defined on a sample space associated with a random experiment assuming any value from R and assigning a real number to each and every sample point of the random experiment.

(8) Expected value or Mathematical Expectation or Expectation of a random variable may be defined as the sum of products of the different values taken by the random variable and the corresponding probabilities.

When x is a discrete random variable with probability mass function f(x), then its expected value is given by

$= \mu = \sum_x xf(x)$

and its variance is

$\sigma^2 = E(x^2) - \mu 2$

Where $E(x^2) = \sum_x{}^2(x)f(x)f(x)$

For a continuous random variable x defined in [−∞, ∞], its expected value (i.e. mean) and variance are given by

$E(x) = \int_{\infty}^{\infty} xf(x)dx$

and $\sigma^2 = E(x^2) - \mu^2$

where $E(x2) = \int_{\infty}^{\infty} x^2 f(x)dx$

Properties of Expected Values

1. **Expectation of a constant k is k**

   i.e. E(k) = k for any constant k

2. **Expectation of sum of two random variables is the sum of their expectations.**

   i.e. E(x + y) = E(x) + E(y) for any two random variables x and y.

3. **Expectation of the product of a constant and a random variable is the product of the constant and the expectation of the random variable.**

   i.e. E(k x) = k.E(x) for any constant k

4. **Expectation of the product of two random variables is the product of the expectation of the two random variables, provided the two variables are independent.**

   i.e. E(xy) = E(x) E(y)

   Whenever x and y are independent.

1. A speaks truth in 60% and B in 75% of the cases. In what percentage of cases are they likely to contradict each other in stating the same fact?

Solution:

The Probability that A speaks the truth and B a lie = $\frac{60}{100} \times \frac{(100-75)}{100} = \frac{60}{100} \times \frac{25}{100} = \frac{3}{20}$

The Probability that B speaks the truth and A a lie = $\frac{75}{100} \times \frac{(100-60)}{100} = \frac{75}{100} \times \frac{40}{100} = \frac{3}{20}$

$\therefore$ Total Probability = $\frac{3}{20} + \frac{3}{10} = \frac{9}{20}$

Hence, the percentage of cases in which they contradict each other = (9/20) × 100 or 45%

2. A Committee of 4 persons is to be appointed from 7 men and 3 women. The probability that the committee contains (i) exactly two women, and (ii) at least one woman is

Solution:

Total number of persons = 7+3 = 10. Since 4 our out them can be formed in 10 C_4 ways the exhaustive number of cases is 10 C_4 or 210 ways.

(i) P (exactly 2 women in a committee) of four = 7 C3 / 210 = 63/210 = 3/10.

(ii) P (at least one women in committee)

= 1 – p (no women) = 1 – 1 /6 = 5/6.

3. If A and B are two events, such that P(A) = ¼, P(B) = 1/3 and P(A B) = ½; then P(B/A) is equal to

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

$\frac{1}{2} = \frac{1}{4} + 1/3 - P(A \cap B)$

Or $P(A \cap B) = \frac{1}{4} + 1/3 - 1/2 = 1/12$

Hence, $P(B/A) = \frac{P(A \cap B)}{P(A)} = \frac{1/12}{1/4} = \frac{1}{3}$

4. A person applies for a job in two firms, say X and Y. the probability of his being selected in firm X is 0.7 and being rejected in firm Y is 0.5. The probability of at least one of his applications being rejected is 0.6. What is the probability that he will be selected in one of the two firms?

Solution:

Events S : Failed in Statistics, and

Event M : Failed in Mathematics,

P (S) = 0.20, P (M) = 0.30 and $P(M \cap S) = 0.1$

Hence, $= P(M \cup S) = P(M) + P(S) - P(M \cap S)$

= 0.2+0.3−0.1= 0.4

5. A person is known to hit a target in 5 out of 8 shots, whereas another person is known to hit in 3 out of 5 shots. Find the probability that the target is hit at all when they both try.

Solution: Event A = First person hits the target and

Event B = Another person hits the target.

P (A) = 5/8 and P(B) = 3/5

$P(A^c)$ = 1-5/8 = 3/8 and P(BC) = 1-3/5= 2/5

Event X = target is hit when they both try i.e.,

When at least one of them hit the target.

P (X^c) = P (the target is not hit at all)

= $P(A^c \cap B^c)$ = P $(A^c$ x $P(B^c)$ - 3/20

Hence P(X) = 1-P(X^c) = 1-3/20 = 17/20

6. The probability that a man will be alive in 25 years is 3/5, and the probability that his wife will be alive in 25 years in 2/3. Find the probability that :
(i) Both will be alive (ii) at least one of the will be alive

Solution:

P (M) = 3/5 and P(W) = 2/3

$P(M^c)$ = 1 – 3/5 and $P(W^c)$ = 1-2/3 = 1/3.

The probability that both will be alive

= P(M) × P(W) = 3/5 × 2/3 = 2/5.

Probability that at least one of them will be alive is given by

$P(M \cup W)$ P(M) + P(W) - $P(M \cap W)$

= 3/5+2/3-6/15 = 13/15.

7. Given the data in Previous Problem find the probability that (i) only wife will be alive, (ii) only man will be alive.

Solution:

Probability that only wife will be alive.

= Probability that wife will be alive nut not man

= P(W) × $P(M^c)$ = 2/3×2/5 = 4/15

Probability that only man will be alive

= Probability that man will be alive nut not wife

= P(M) × $P(W^c)$ = 3/5×1/3 = 1/5.

8. A random variable X has the following probability distribution:

Value of X	**0**	**1**	**2**	**3**
P [X =x]	**1/3**	**1/2**	**0**	**1/6**

Find E ({X – E (X)}2]

Solution : E (X) = 0×1/3+1×1/2+2×0+3×1/6= 1,

$E(X^2)$ = 0×1/3+1×1/2+4×0+9×1/6= 2

E $[X – E(X)]^2$ = E (X^2)- $[E (X)]^2$ = 2-1 = 1.

9. Given the data in previous Problem, Find Var (Y), where Y = 2X – 1.

Solution:

$E(Y) = E(2X-1) = 2E(X) - 1 = 1$

$E(Y^2) = E(2x-1)^2 = 2E(X^2) - 4E(X)+1 = 1$

$Var.(Y) = E(Y^2) - [E(Y)]^2 = 1.1 = 0$

10. Daily demand for transistors is having the following probability distribution:

Demand	1	2	3	4	5	6
Probability	0.10	0.15	0.20	0.25	0.18	0.12

Solution:

$E(X) = 1\times0.10+2\times0.15+3\times0.20+4\times0.25+5\times0.18+6\times0.12 = 3.62$

Given the data in previous problem obtain the variance of the demand.

$E(X^2) = 1\times0.10+4\times0.15+9\times0.20+16\times0.25+25\times0.18+36\times0.12= 15.32$

$Var.(X) = 15.32 - (3.62)^2 = 2.22$

Chapter – 16:

Theoretical Distributions

This chapter will enable the students to understand and apply the techniques of developing discrete and continuous probability distributions.

In this chapter we will discuss the probability theory by considering a concept and analogous to the idea of frequency distribution. In frequency distribution where to the total frequency to different class intervals, the total probability (i.e. one) is distributed to different mass points is known as theoretical probability distributions.

- Discrete Random variable.
- Continues Random variable.

Importance of theoretical probability distribution.☞:

(a) An observed frequency distribution, in many a case, may be regarded as a sample i.e. a representative part of a large, unknown, boundless universe or population and we may be interested to know the form of such a distribution.

For Example: By fitting a theoretical probability distribution.

- Length of life of the lamps produced by manufacturer up to a reasonable degree of accuracy.
- The effect of a particular type of missiles, it may be possible for our scientist to suggest the number of such missiles necessary to destroy an army position.
- By knowing the distribution of smokers, a social activist may warn the people of a locality about the nuisance of active and passive smoking and so on.

(b) Theoretical probability distribution may be profitably employed to make short term projections for the future.

(c) Statistical analysis is possible only on the basis of theoretical probability distribution.

A probability distribution also possesses all the characteristics of an observed distribution. We define population mean, population median, population mode, population standard deviation etc. exactly same way we have done earlier. These characteristics are known as population parameters.

A probability distribution may be either a discrete probability distribution or a Continuous probability distribution depending on the random variable under study.

Two important discrete probability distributions

(a) Binomial Distribution
(b) Poisson distribution.

Important continuous probability distribution

- Normal Distribution

Binomial Distribution

One of the most important and frequently used discrete probability distribution is Binomial Distribution. It is derived from a particular type of random experiment known as Bernoulli process named after the famous mathematician Bernoulli. Noting that a 'trial' is an attempt to produce a particular outcome which is neither certain nor impossible, the characteristics of Bernoulli trials are stated below:

(i) Each trial is associated with two mutually exclusive and exhaustive outcomes, the occurrence of one of which is known as a 'success' and as such its non occurrence as a 'failure'. As an example, when a coin is tossed, usually occurrence of a head is known as a success and its non–occurrence i.e. occurrence of a tail is known as a failure.

(ii) The trials are independent.

(iii) The probability of a success, usually denoted by p, and hence that of a failure, usually denoted by q = 1–p, remain unchanged throughout the process.

(iv) The number of trials is a finite positive integer.

A discrete random variable x is defined to follow binomial distribution with parameters n and p, to be denoted by $x \sim B(n, p)$, if the probability mass function of x is given by

$f(x) = p(X = x) =$ for x = 0, 1, 2, ., n

= 0, otherwise

We may note the following important points in connection with binomial distribution:

(a) As $n > 0$, $p, q \geq 0$, it follows that $f(x) \geq 0$ for every x

Also $\sum_x f(x) = f(0) + f(1) + f(2) + \ldots + f(n) = 1$

(b) Binomial distribution is known as biparametric distribution as it is characterised by two parameters n and p. This means that if the values of n and p are known, then the distribution is known completely.

(c) The mean of the binomial distribution is given by $\mu = np$

(d) Depending on the values of the two parameters, binomial distribution may be unimodal or bi- modal., the mode of binomial distribution, is given by μ_0 = the largest integer contained in $(n+1)p$ if $(n+1)p$ is a non-integer $(n+1)p$ and $(n+1)p - 1$ if $(n+1)p$ is an integer

(e) The variance of the binomial distribution is given by

$\sigma^2 = npq$

Since p and q are numerically less than or equal to 1, $npq < np$

variance of a binomial variable is always less than its mean.

Also variance of X attains its maximum value at p = q = 0.5 and this maximum value is n/4.

(f) Additive property of binomial distribution.

If X and Y are two independent variables such that

$X \sim \beta(n_1, P)$

and $Y \sim \beta(n_2, P)$

Then $(X+Y) \sim (n_1 + n_2, P)$

Applications of Binomial Distribution

Binomial distribution is applicable when the trials are independent and each trial has just two outcomes success and failure. It is applied in coin tossing experiments, sampling inspection plan, genetic experiments and so on.

Poisson Distribution

Poisson distribution is a theoretical discrete probability distribution which can describe many processes. Simon Denis Poisson of France introduced this distribution way back in the year 1837.

Poisson Model

Let us think of a random experiment under the following conditions:

I. The probability of finding success in a very small time interval (t, t + dt) is kt, where k (>0) is a constant.

II. The probability of having more than one success in this time interval is very low.

III. The probability of having success in this time interval is independent of t as well as earlier successes.

The above model is known as Poisson Model. The probability of getting x successes in a relatively long time interval T containing m small time intervals t i.e. T = mt. is given by

$\frac{e^{-kt}.(kt)^x}{x!}$ for x = 0, 1, 2, ...

Taking kT = m, the above form is reduced to

$\frac{e^{-m}.m^x}{x!}$ for x = 0, 1, 2, ...

Definition of Poisson Distribution

A random variable X is defined to follow Poisson distribution with parameter λ, to be denoted by X ~ P (m) if the probability mass function of x is given by

$$f(x) = P(X = x) = \frac{e^{-m}.m^x}{x!} \quad \text{for } x = 0, 1, 2, ...$$

$$= 0 \quad \text{otherwise}$$

Here e is a transcendental quantity with an approximate value as 2.71828.

Important points in connection with Poisson distribution:

(i) Since $e^{-m} = 1/e^m > 0$, whatever may be the value of m, m > 0, it follows that $f(x) \geq 0$ for every x.

Also it can be established that $\sum_x f(x) = 1$ i.e. $f(0) + f(1) + f(2) + \ldots\ldots = 1$

(ii) Poisson distribution is known as a uniparametric distribution as it is characterised by only one parameter m.

(iii) The mean of Poisson distribution is given by m i,e $\mu = m$

(iv) The variance of Poisson distribution is given by $\sigma^2 = m$

(v) Like binomial distribution, Poisson distribution could be also unimodal or bimodal depending upon the value of the parameter m.

We have μ_0 = The largest integer contained in m if m is a non-integer

= m and m−1 if m is an integer

(vi) Poisson approximation to Binomial distribution

If n, the number of independent trials of a binomial distribution, tends to infinity and p, the probability of a success, tends to zero, so that m = np remains finite, then a binomial distribution with parameters n and p can be approximated by a Poisson distribution with parameter m (= np).

In other words when n is rather large and p is rather small so that m = np is moderate then $\beta(n, p) \cong P(m)$

(vii) Additive property of Poisson distribution

If X and y are two independent variables following Poisson distribution with parameters m_1 and m_2 respectively, then Z = X + Y also follows Poisson distribution with parameter $(m_1 + m_2)$.

i.e. if $X \sim P(m_1)$

and $Y \sim P(m_2)$

and X and Y are independent, then

$Z = X + Y \sim P(m_1 + m_2)$

Application of Poisson distribution

Poisson distribution is applied when the total number of events is pretty large but the probability of occurrence is very small. Thus we can apply Poisson distribution, rather profitably, for the following cases:

a) The distribution of the no. of printing mistakes per page of a large book.

b) The distribution of the no. of road accidents on a busy road per minute.

c) The distribution of the no. of radio-active elements per minute in a fusion process.

d) The distribution of the no. of demands per minute for health centre and so on.

Normal or Gaussian distribution

The two distributions discussed so far, namely binomial and Poisson, are applicable when the random variable is discrete. In case of a continuous random variable like height or weight, it is impossible to distribute the total probability among different mass points because between any two unequal values, there remains an infinite number of values. Thus a continuous random variable is defined in term of its probability density function f (x), provided, of course, such a function really exists, f (x) satisfies the following condition:

$$f(x) \geq 0 \text{ for } x\ (-\infty,\infty) \text{ and } = \int_{-\infty}^{+\infty} f(x) = 1$$

The most important and universally accepted continuous probability distribution is known as normal distribution. Though many mathematicians like De-Moivre, Laplace etc. contributed towards the development of normal distribution, Karl Gauss was instrumental for deriving normal distribution and as such normal distribution is also referred to as Gaussian Distribution.

A continuous random variable x is defined to follow normal distribution with parameters μ and σ^2, to be denoted by

$X \sim N(\mu, \sigma^2)$

If the probability density function of the random variable x is given by

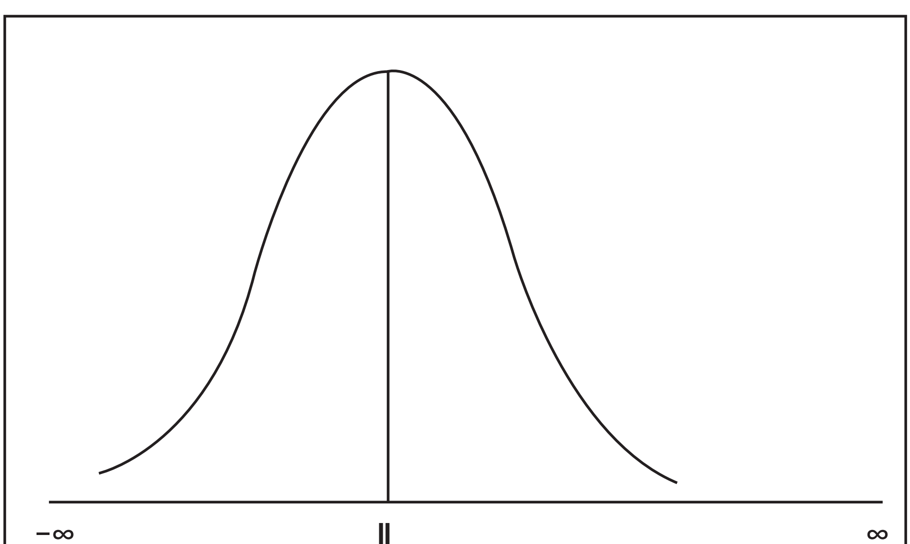
$$f(x) = \frac{1}{\sigma\sqrt{2\pi}} . e^{-(\bar{x}-\mu)^2/2\sigma^2} \qquad \text{for } -\infty < x < \infty$$

where μ and σ are constants, and > 0

Some important points relating to normal distribution are listed below:

(a) The name Normal Distribution has its origin some two hundred years back as the then mathematician were in search for a normal model that can describe the probability distribution of most of the continuous random variables.

(b) If we plot the probability function y = f (x), then the curve, known as probability curve, takes the following shape:



Showing Normal Probability Curve. A quick look at figure reveals that the normal curve is bell shaped and has one peak, which implies that the normal distribution has one unique mode. The line drawn through x = μ has divided the normal curve into two parts which are equal in all respect. Such a curve is known as symmetrical curve and the corresponding distribution is known as symmetrical distribution. Thus, we find that the normal distribution is symmetrical about x = μ. It may also be noted that the binomial distribution is also symmetrical about p = 0.5. We next note that the two tails of the normal curve extend indefinitely on both sides of the curve and both the left and right tails never touch the horizontal axis. The total area of the normal curve or for that any probability curve is taken to be unity i.e. one. Since the vertical line drawn through x = μ divides the curve into two equal halves, it automatically follows that,

The area between – ∞ to μ= the area between μ to ∞ = 0.5

When the mean is zero, we have

the area between – ∞ to 0 = the area between 0 to ∞ = 0.5

(c) If we take μ = 0 and = 1 , we have

$$f(x) = \frac{1}{Ó\sqrt{2\pi}} e^{-2^2/2} \qquad \text{for } -\infty<z<\infty$$

The random variable z is known as standard normal variate (or variable) or standard normal deviate. The probability that a standard normal variate X would take a value less than or equal to a particular value say X = x is given by
$\varphi(x) = p\,(X \le x)$
(x) is known as the cumulative distribution function.
We also have (0) = P (X ≤ 0) = Area of the standard normal curve between −∞ and 0 = 0.5

(d) The normal distribution is known as biparametric distribution as it is characterised by two parameters μ and σ^2. Once the two parameters are known, the normal distribution is completely specified.

Properties of Normal Distribution

1. Since π = 22/7 , $e^{-\theta} = 1 / e^{\theta} > 0$, whatever θ may be,

   it follows that f (x) 0 for every x.

   It can be shown that $\int_{-\infty}^{\infty} f(x)dx=1$

2. The mean of the normal distribution is given by μ. Further, since the distribution is symmetrical about x = μ, it follows that the mean, median and mode of a normal distribution coincide, all being equal to μ.

3. The standard deviation of the normal distribution is given by

   Mean deviation of normal distribution is $\sigma \frac{\sqrt{2}}{\pi}$

   $\sigma \frac{\sqrt{2}}{\pi} \cong 0.8\sigma$

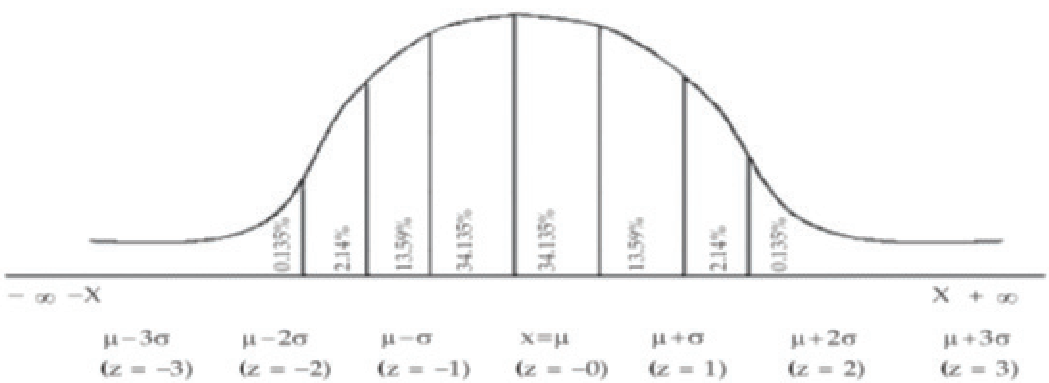
   The first and third quartiles are Q_1 = μ– 0.675 σ and Q_3 = μ+ 0.675 σ

   so that, quartile deviation = 0.675 σ

4. The normal distribution is symmetrical about x =μ . As such, its skewness is zero i.e. the normal curve is neither inclined move towards the right (negatively skewed) nor towards the left (positively skewed).

5. The normal curve $y = f(x)$ has two points of inflexion to be given by $x = \mu - \sigma$ and $x = \mu + \sigma$ i.e. at these two points, the normal curve changes its curvature from concave to convex and from convex to concave.
6. If $x \sim N(\mu, \sigma^2)$ then $z = x - \mu/\sigma \sim N(0, 1)$, z is known as standardised normal variate or normal deviate.

   We also have $P(z \leq k) = \phi(k)$ The values of $\phi(k)$ for different k are given in a table known as "Biometrika."
7. Area under the normal curve is shown in the following figure:



Area Under Normal Curve

From the figure, we have

$P(\mu - \sigma < x < \mu + \sigma) = 0.6828$

$\Rightarrow P(-1 < z < 1) = 0.6828$

$P(\mu - 2\sigma < x < \mu + 2\sigma) = 0.9546$

$\Rightarrow P(-2 < z < 2) = 0.9546$

and $P(\mu - 3\sigma < x < \mu + 3\sigma) = 0.9973$

$\Rightarrow P(-3 < z < 3) = 0.9973$.

We note that 99.73 per cent of the values of a normal variable lies between $(\mu - 3\sigma)$ and $(\mu + 3\sigma)$. Thus the probability that a value of x lies outside that limit is as low as 0.0027.

8. If x and y are independent normal variables with means and standard deviations as μ_1 and μ_2 and σ_1 and σ_2 respectively, then $z = x + y$ also follows normal distribution with
   mean $(\mu_1 + \mu_2)$ and SD = $\sqrt{\sigma_1^2 + \sigma_2^2}$ respectively.

   i.e. If $x \sim N(\mu_1, \sigma_1^2)$ and $y \sim N(\mu_2, \sigma_2^2)$ and x and y are independent,
   then $z = x + y \sim N(\mu_1 + \mu_2, \sigma_1^2 + \sigma_2^2)$

Applications of Normal Distribution

Most of the continuous variables like height, weight, wage, profit etc. follow normal distribution. If the variable under study does not follow normal distribution, a simple transformation of the variable, in many a case, would lead to the normal distribution of the changed variable.

When n, the number of trials of a binomial distribution, is large and p, the probability of a success, is moderate i.e. neither too large nor too small then the binomial distribution, also, tends to normal distribution. Poisson distribution, also for large value of m approaches normal distribution. Such transformations become necessary as it is easier to compute probabilities under the assumption of a normal distribution.

1. What is the probability that out of 10 missiles fired, atleast 2 will hit the target

Solution: Probability of atleast 2 will hit the target is given as,

$P(X \geq 2) = 1 - P(X < 2)$

The probability of a missile hitting a target is 1 /8

$P(X \geq 2) = 1 - [P(X = 0) + P(X = 1)]$

$= 1 - [{}^{10}C_0 \times (1/8)^0 \times (7/8)^{10} + {}^{10}C_1 \times (1/8)^1 \times (7/8)^9]$

$= 1 - [7^{10}/8^{10} + (10 \times 7^9)/8^{10}]$

$= 1 - [7^{10} + 10 \times 7^9] / 8^{10}$

$= 1 - (17 \times 7^9) / 8^{10}$

$P(X \geq 2) \approx 0.3611 = 36.11\%$

Therefore, the probability that out of 10 missiles fired, atleast 2 will hit the target is 0.3611.

2. Given X is a binomial variable such that 2 P(X = 2) = P(X = 3) and mean of X is known to be 10/3. What would be the probability that X assumes at most the value 2

Solution:

mean = 10/3 , Mean = np = 10/3

$P(X) = {}^nC_x p^x (1-p)^{n-x}$

$\Rightarrow P(2) = {}^nC_2 p^2 (1-p)^{n-2}$

$P(3) = {}^nC_3 p^3 (1-p)^{n-3}$

$P(3) = 2\,P(2)$

$\Rightarrow {}^nC_3 p^3 (1-p)^{n-3} = 2\,{}^nC_2 p^2 (1-p)^{n-2}$

$\Rightarrow p/3!(n-3)! = (2/2!(n-2)!)(1 - p)$

$\Rightarrow p/6 = (1-p)/(n - 2)$

$\Rightarrow np - 2p = 6 - 6p$

$\Rightarrow 4p = 6 - np$

$\Rightarrow 4p = 6 - 10/3$

$\Rightarrow 4p = 8/3$

$\Rightarrow p = 2/3$

$np = 10/3 \Rightarrow n = 5$

probability that X assumes at most the value 2 = P(0) + P(1) + P(2)

$= {}^5C_0(2/3)^0(1/3)_5 + {}^5C_1(2/3)^1(1/3)^4 + {}^5C_2(2/3)^2(1/3)^3$

$= 1/243 + 10/243 + 40/243 = 51/243 = 17/81$

3. Assuming that one-third of the population is tea drinkers and each of 1000 enumerators takes a sample of 8 individuals to find out whether they are tea drinkers or not, how many enumerators are expected to report that five or more people are tea drinkers?

Solution : Assume that being a tea drinker is like taking a flip of coin.(i.e. either the person drinks tea or not) with the probability of heads being 1/3 and tails being 2/3

now P(X≥5) =P(X=5)+P(X=6)+P(X=7)+P(X=8) = (56*8+28*4+8*2+1)/6561 = 577/6561 =0.087943 so for 1000 trials the number of trials which report greater than or equal to 5 is 1000*P(X>=5) = 87.94 =88

4. If a random variable x follows binomial distribution with mean as 5 and satisfying the condition 10. P (x= 0) = P(x= 1) , what is the value of P(X≥1/X>1) ?

Solution:
Here Mean np = 5
10.P(x=0) = P(x=1)
$10.{}^{n}C_{0}.p^{0}.(1-p)^{n} = {}^{n}C_{1}.p.(1-p)^{n-1}$
$10.(1-p)^{n} = np(1-p)^{n-1}$
10(1-p)=5
1-p=1/2
p=1/2, n=10

$$p\left(x \geq 1 \big/ x \geq 0\right) = \frac{p(x \geq 1)}{p(x \geq 0)}$$

= 1-p(x=0)

$$= 1 - \frac{1}{2^{10}}$$

= 0.99

5. Out of 128 families with four children each, how many are expected to have atleast have one boy and one girl ?

Solution : 4 children in a family can be in

2 * 2 * 2 * 2 = 16 ways

at least one boy and one girl = Total cases - all boys - all girs

All boys = 1 case, All girls = 1 case

=> at least one boy and one girl = 16 - 1 - 1 = 14

Probability of at least one boy and one girl = 14/16

out of 128 families expected to have = 128 * 14/16 = 8 * 14 = 112

112 Families expected to have at least one boy and one girl

6. In 10 independent rollings of a biased die, the probability that an even number will appear 5 times is twice the probability that an even number will appear 4 times. What is the probability that an even number will appear twice when the die is rolled 8 times?

Solution : Probability of even number p = p

Then probability of odd number (or not even number) = q = 1 - p

Probability Appearing 5 times

${}^{10}C_5 * p^5 * q^{10-5}$

$= {}^{10}C_5 * p^5 * q^5$

Probability Appearing 5 times

${}^{10}C_4 * p^4 * q^{10-4}$

$= {}^{10}C_4 * p^4 * q^6$

${}^{10}C_5 * p^5 * q^5 = 2 * {}^{10}C_4 * p^4 * q^6$

=>p * 10!/5!5! = 2q * 10!/6!4!

=> p * 6 = 2q * 5

=> 3p = 5q => 3p = 5(1 - p)

=> 8p = 5

=> p = 5/8& q = 1 - 5/8 = 3/8

probability that an even number will appear twice when the die is rolled 8 times

$= {}^{8}C_2 * p^2q^{8-2}$

$= 28 * (5/8)^2 (3/8)^6$

$= 28 * 25 * 3^6 / 8^8$

$= 700 * 3^6 / 8^8$

= 5,10,300/1,67,77,216 = 0.0304

8. Suppose that weather records show that on an average 5 out of 31 days in October are rainy days. Assuming a binomial distribution which each day of October as an independent trail, then the probability that the next October will have at most three rainy days is:

Answer : p = Probability of a rainy day in October :p = 5/31, q = probability of a non-rainy day in October q = 1- p = 1-5/31 = 26/31

n = 31 (number of days in October)

$$P(x) = 31Cr \left(\frac{5}{31}\right)^r \cdot \left(\frac{26}{31}\right)^{31-r}$$

Required Probability = P (0) + P (1) +P(2) +P(3)

$$= 31C_0 \cdot \left(\frac{5}{31}\right)^0 \cdot \left(\frac{26}{31}\right)^{31} + 31C_1 \cdot \left(\frac{5}{31}\right)^1 \cdot \left(\frac{26}{31}\right)^{31-1} + 31C_2 \cdot \left(\frac{5}{31}\right)^2 \cdot \left(\frac{26}{31}\right)^{31-2} + 31C_3 \cdot \left(\frac{5}{31}\right)^3 \cdot \left(\frac{26}{31}\right)^{31-3}$$

= 0.2403

9. If 5 days are selected at random, then the probability of getting two Sundays is:

Solution: Let P = Probability of getting a Sunday in a week (P) = 1/7

Therefore P = 1/7 and q= 1-p = 1-1/7 = 6/7

$$\text{Required probability} = 15C_2 \cdot \left(\frac{1}{7}\right)^r \cdot \left(\frac{6}{7}\right)^{15-2} = 0.288 = 0.29$$

10. An experiment of succeeds twice as often as it falls. What is the probability that in next five trials there will at least three successes?

Answer: According to the given statement p = 2q

We know that p = 2/3 q= 1/3

Required probability P(X≥3) = P(3) +P(4) +P(5)

$$= 5C_3 \cdot \left(\frac{2}{3}\right)^2 \cdot \left(\frac{6}{7}\right)^{15-2} + 5C_4 \cdot \left(\frac{2}{3}\right)^2 \cdot \left(\frac{6}{7}\right)^{15-2} + 5C_5 \cdot \left(\frac{2}{3}\right)^3 \cdot \left(\frac{1}{3}\right)^2$$

11. What is the probability of getting 3 head if 6 unbiased coins are tossed simultaneously?

Answer : if x denotes the number of heads, then x follows binomial distribution with parameters n = 6 and p = probability of success = ½

q= probability of failure = 1-1/2 = ½ , being given the coins are unbiased

The probability mass function of x is given by f(x) = 6C3 . $\left(\frac{1}{2}\right)^3 . \left(\frac{1}{2}\right)^3$ = 20 x $\left(\frac{1}{2}\right)^3$ = 0.3125

12. In Binomial Distribution n = 9 and p = 1/3 , What is the value of variance.

Solution: In Binomial Distribution variance = npq , here n= 9 , p = 1/3 and q= 2/3

Therefore variance = 9*1/3*2/3 = 2

13. For a Binomial Distribution E(x) = 2, V(X)= 4/3 . Find the value of n

Solution: Here E(X) = np = 2

V(x)= npq = 4/3 then substituting the value of np

2×q = 4/3

2q = 4/3

q= 2/3 then p = 1-2/3 = 1/3

np = 2, n×1/3 = 2, n= 6

14. The node of Binomial Distribution for which the mean is 4 and variance 3 is equal to?

Solution: In Binomial Distribution Mean = np = 4 and Variance = npq = 3

Then 4q = 3 q = ¾ , p = 1-q = 1-3/4 = ¼

n×1/4 = 4 therefore n = 16

(n+1) p = (16+1) ×1/4 = 4.25 which is no longer integer.

So mode = 4

15. In a Binomial Distribution with 5 independent trials, probability of 2 and 3 successes are 0.4362 and 0.2181 respectively. Parameter 'p 'of the Binomial Distribution is

Answer : Given n = 5 , P(x=2) = 0.4362

P(x=3) =0.2181

P(x=3) = $5C_3 . (p)^3 . (q)^{5-3}$ = $10\ (p)^3 . (q)^2$

0.2181 = $10.P^{3.q2}$

And P(X=2) =$5C_2 . (p)^3 . (q)^{5-2}$ = $10\ (p)^2 . (q)^3$

0.4362 = $10\ (p)^2 . (q)^3$

By dividing

0.2181/0.4362 = $\frac{10.p^3.q^2}{10.p^2.q^3}$

½= q/p

q= 2p 2p+p = 1

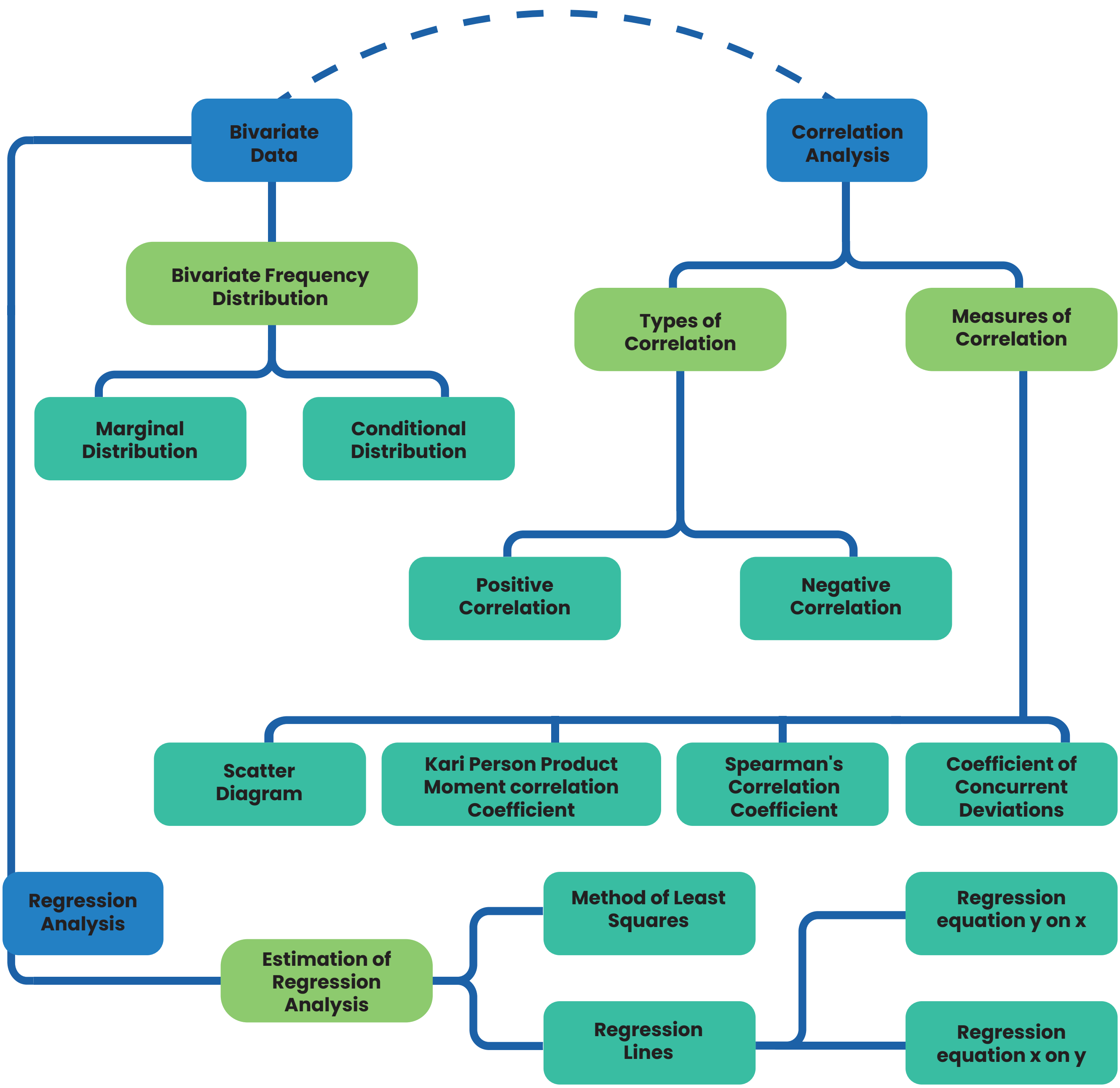
3p = 1 then p = 1/3

Chapter – 17:

Correlation And Regression

At the foundation level with regards to Paper 3 Statistics part of the topic Correlation and Regression is very important for students not only to acquire professional knowledge but also for examination point of view. Here in this chapter an attempt is made for solving and understanding the concepts of Correlation and Regression with the help of following questions with solutions

Chapter Overview



Univariate Distribution: Statistical measure relating to Univariate distribution i.e. distribution of one variable like height, weight, mark, profit, wage and so on. However, there are situations that demand study of more than one variable simultaneously. A businessman may be keen to know what amount of investment would yield a desired level of profit or a student may want to know whether performing better in the selection test would enhance his or her chance of doing well in the final examination. With a view to answering this series of questions, we need to study more than one variable at the same time.

Bivariate Data: When data are collected on two variables simultaneously, they are known as bivariate data and the corresponding frequency distribution, derived from it, is known as Bivariate Frequency Distribution. If x and y denote marks in Maths and Stats for a group of 30 students, then the corresponding bivariate data would be (x_i, y_i) for i = 1, 2, ... 30 where (x_1, y_1) denotes the marks in Mathematics and Statistics for the student with serial number or Roll Number 1, (x_2, y_2), that for the student with Roll Number 2 and so on and lastly (x_{30}, y_{30}) denotes the pair of marks for the student bearing Roll Number 30.

Correlation Analysis and Regression Analysis are the two analyses that are made from a multivariate distribution i.e. a distribution of more than one variable. In particular when there are two variables, say x and y, we study bivariate distribution. We restrict our discussion to bivariate distribution only.

Correlation analysis helps us to find an association or the lack of it between the two variables x and y. Thus, if x and y stand for profit and investment of a firm or the marks in Statistics and Mathematics for a group of students, then we may be interested to know whether x and y are associated or independent of each other. The extent or amount of correlation between x and y is provided by different measures of Correlation namely Product Moment Correlation Coefficient or Rank Correlation Coefficient or Coefficient of Concurrent Deviations. In Correlation analysis, we must be careful about a cause-and-effect relation between the variables under consideration because there may be situations where x and y are related due to the influence of a third variable although no causal relationship exists between the two variables.

Regression analysis, on the other hand, is concerned with predicting the value of the dependent variable corresponding to a known value of the independent variable on the assumption of a mathematical relationship between the two variables and also an average relationship between them.

As in the case of a Univariate Distribution, we need to construct the frequency distribution for bivariate data. Such a distribution takes into account the classification in respect of both the variables simultaneously. Usually, we make horizontal classification in respect of x and vertical classification in respect of the other variable y. Such a distribution is known as Bivariate Frequency Distribution or Joint Frequency Distribution or Two way classification of the two variables x and y. Frequency Distribution, we can obtain two types of univariate distributions which are known as:

(a) Marginal distribution.: Marginal distributions always divide the column or row totals by the table total

(b) Conditional distribution. To calculate a conditional distribution, you must first establish a condition. For instance, we could ask what the distribution of gender is among students who watched the last football game. So, the condition here would be that the student watched the game. In particular, if there are m classifications for x and n classifications for y, then there would be altogether (m + n) conditional distribution.

Correlation Analysis: While studying two variables at the same time, if it is found that the change in one variable is reciprocated by a corresponding change in the other variable either directly or inversely, then the two variables are known to be associated or correlated. Otherwise, the two variables are known to be dissociated or uncorrelated or independent. There are two types of correlation.

Positive correlation

If two variables move in the same direction i.e. an increase (or decrease) on the part of one variable introduces an increase (or decrease) on the part of the other variable, then the two variables are known to be positively correlated.

As for example, height and weight yield and rainfall, profit and investment etc. are positively correlated.

Negative correlation

On the other hand, if the two variables move in the opposite directions i.e. an increase (or a decrease) on the part of one variable results a decrease (or an increase) on the part of the other variable, then the two variables are known to have a negative correlation.

The price and demand of an item, the profits of Insurance Company and the number of claims it has to meet etc. are examples of variables having a negative correlation.

The two variables are known to be uncorrelated if the movement on the part of one variable does not produce any movement of the other variable in a particular direction. As for example, Shoe-size and intelligence are uncorrelated

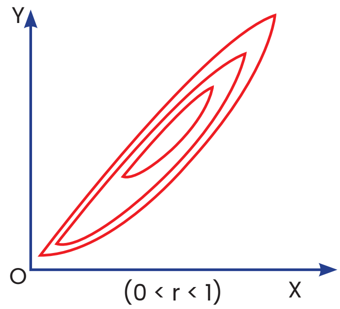
Measures of correlation

- Scatter diagram
- Karl Pearson's Product moment correlation coefficient
- Spearman's rank correlation co-efficient
- Co-efficient of concurrent deviations

(a) SCATTER DIAGRAM: This is a simple diagrammatic method to establish correlation between a pair of variables. Unlike product moment correlation co-efficient, which can measure correlation only when the variables are having a linear relationship, scatter diagram can be applied for any type of correlation – linear as well as non-linear i.e. curvilinear. Scatter diagram can distinguish between different types of correlation although it fails to measure the extent of relationship between the variables.

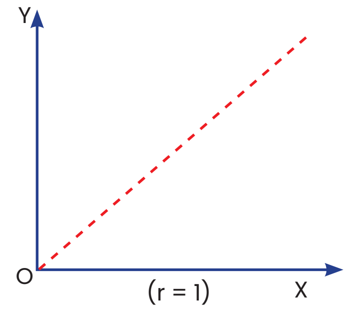
Each data point, which in this case a pair of values (xi, yi) is represented by a point in the rectangular axes of cordinates. The totality of all the plotted points forms the scatter diagram. The pattern of the plotted points reveals the nature of correlation. In case of a positive correlation, the plotted points lie from lower left corner to upper right corner, in case of a negative correlation the plotted points concentrate from upper left to lower right and in case of zero correlation, the plotted points would be equally distributed without depicting any particular pattern. The following figures show different types of correlation and the one-to-one correspondence between scatter diagram and product moment correlation coefficient.

Showing Positive Correlation



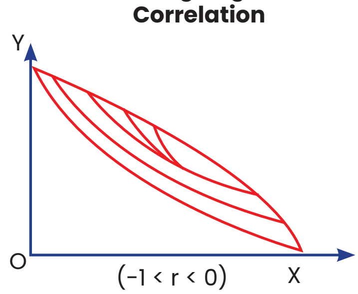
(0 < r < 1)
FIGURE 17.3

Showing Perfect Correlation

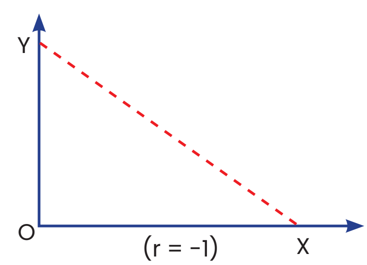


(x=1)
FIGURE 17.4

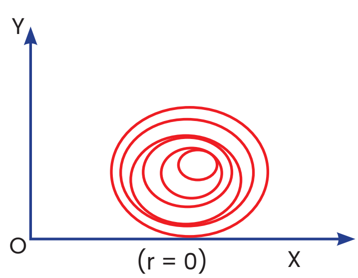
Showing Negative Correlation



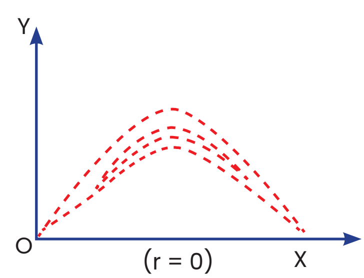
Showing Perfect Negative Correlation



Showing No Correlation



Showing Curvilinear Correlation



(b) KARL PEARSON'S PRODUCT MOMENT CORRELATION COEFFICIENT

This is by for the best method for finding correlation between two variables provided the relationship between the two variables is linear. Pearson's correlation coefficient may be defined as the ratio of covariance between the two variables to the product of the standard deviations of the two variables. If the two variables are denoted by x and y and if the corresponding bivariate data are (xi, yi) for i = 1, 2, 3, ..., n, then the coefficient of correlation between x and y, due to Karl Pearson, in given by

$$r = r_{xy} = \frac{Cov(x,y)}{S_x \times S_y}$$

$$\text{where, } cov\,(x, y) = \frac{\Sigma(x_i - x)\,(y_i - \bar{y})}{n} = \frac{\Sigma x_i\, y_i}{n} - \bar{x}\,\bar{y}$$

$$S_x = \sqrt{\frac{\Sigma(x_i - \bar{x})^2}{n}} = \sqrt{\frac{\Sigma x_i^2}{n} - \bar{x}^2} \quad \text{and} \quad S_y = \sqrt{\frac{\Sigma(y_i - \bar{y})^2}{n}} = \sqrt{\frac{\Sigma y_i^2}{n} - \bar{y}^2}$$

A single formula for computing correlation coefficient is given by

$$r = \frac{n\Sigma x_i y_i - \Sigma x_i \times \Sigma y_i}{\sqrt{n\Sigma x_i^2 - (\Sigma x_i)^2}\ \sqrt{n\Sigma_i^2 - (\Sigma y_i)^2}}$$

In case of a bivariate frequency distribution, we have

$$Cov(x,y) = \frac{\sum_{i,j} x_i y_i f_{ij}}{n} - \bar{x} \times \bar{y}$$

$$Sx = \sqrt{\frac{\sum_i f_{io} x_i^2}{N} - \bar{x}^2} \qquad and.Sy = \sqrt{\frac{\sum_i f_{oj} y_j^2}{N} - \bar{y}^2}$$

where	x_i	=	Mid-value of the ith class interval of x. ,y_j	= Mid-value of the j[th] class interval of y
	f_{io}	=	Marginal frequency of x f_{oj}	= Marginal frequency of y
	f_{ij}	=	frequency of the (i, j)[th] cell , N	= Total frequency

PROPERTIES OF CORRELATION COEFFICIENT

(i) The Coefficient of Correlation is a unit-free measure: This means that if x denotes height of a group of students expressed in cm and y denotes their weight expressed in kg, then the correlation coefficient between height and weight would be free from any unit.

(ii) The coefficient of correlation remains invariant under a change of origin and/or scale of the variables under consideration depending on the sign of scale factors.

This property states that if the original pair of variables x and y is changed to a new pair of variables u and v by effecting a change of origin and scale for both x and y i.e.

$$u = \frac{x - a}{b} \text{ and } v = \frac{y - c}{d}$$

where a and c are the origins of x and y and b and d are the respective scales and then we have

$$r_{xy} = \frac{bd}{|b||d|}\, r_{uv}$$

r_{xy} and r_{uv} being the coefficient of correlation between x and y and u and v respectively, (12.10) established, numerically, the two correlation coefficients remain equal and they would have opposite signs only when b and d, the two scales, differ in sign.

(iii) The coefficient of correlation always lies between –1 and 1, including both the limiting values i.e.

$-1 \le r \le 1$

(c) SPEARMAN'S RANK CORRELATION COEFFICIENT:

When we need finding correlation between two qualitative characteristics, say, beauty and intelligence, we take recourse to using rank correlation coefficient. Rank correlation can also be applied to find the level of agreement (or disagreement) between two judges so far as assessing a qualitative characteristic is concerned. As compared to product moment correlation coefficient, rank correlation coefficient is easier to compute, it can also be advocated to get a first hand impression about the correlation between a pair of variables.

Spearman's rank correlation coefficient is given by

$$r_R = 1 - \frac{6 \sum d_i^2}{n(n^2 - 1)}$$

where r_R denotes rank correlation coefficient and it lies between –1 and 1 inclusive of these two values.

$d_i = x_i - y_i$ represents the difference in ranks for the i-th individual and n denotes the number of individuals.

In case u individuals receive the same rank, we describe it as a tied rank of length u. In case of a tied rank, formula is changed to

$$r_R = 1 - \frac{6\left[\sum_i d_i + \sum_j \frac{(tj^3 - t_j)}{12}\right]}{n(n^2 - 1)}$$

In this formula, t_j represents the j^{th} tie length and the summation extends over the lengths of all the ties for both the series.

(d) COEFFICIENT OF CONCURRENT DEVIATIONS

A very simple and casual method of finding correlation when we are not serious about the magnitude of the two variables is the application of concurrent deviations. This method involves in attaching a positive sign for a x-value (except the first) if this value is more than the previous value and assigning a negative value if this value is less than the previous value. This is done for the y-series as well. The deviation in the x-value and the corresponding y-value is known to be concurrent if both the deviations have the same sign.

Denoting the number of concurrent deviation by c and total number of deviations as m (which must be one less than the number of pairs of x and y values), the coefficient of concurrent deviation is given by

$$r_C = \pm\sqrt{\pm\frac{(2c - m)}{m}}$$

If (2c–m) >0, then we take the positive sign both inside and outside the radical sign and if (2c–m) <0, we are to consider the negative sign both inside and outside the radical sign.

Like Pearson's correlation coefficient and Spearman's rank correlation coefficient, the coefficient of concurrent deviations also lies between –1 and 1, both inclusive.

spurious correlation: There are some cases when we may find a correlation between two variables although the two variables are not causally related. This is due to the existence of a third variable which is related to both the variables under consideration. Such a correlation is known as spurious correlation or non-sense correlation. As an example, there could be a positive correlation between production of rice and that of iron in India for the last twenty years due to the effect of a third variable time on both these variables. It is necessary to eliminate the influence of the third variable before computing correlation between the two original variables.

Correlation Coefficient: Correlation coefficient measuring a linear relationship between the two variables indicates the amount of variation of one variable accounted for by the other variable. A better measure for this purpose is provided by the square of the correlation coefficient, Known as 'coefficient of determination'. This can be interpreted as the ratio between the explained variance to total variance i.e.

$$r^2 = \frac{\text{Explained variance}}{\text{Total variance}}$$

Thus, a value of 0.6 for r indicates that $(0.6)^2 \times 100\%$ or 36 per cent of the variation has been accounted for by the factor under consideration and the remaining 64 per cent variation is due to other factors.

Coefficient of non-determination: The 'coefficient of non-determination' is given by $(1-r^2)$ and can be interpreted as the ratio of unexplained variance to the total variance.

Coefficient of non-determination = $(1-r^2)$

Regression Lines:

(i) The two lines of regression coincide i.e. become identical when r = –1 or 1 or in other words, there is a perfect negative or positive correlation between the two variables.

(ii) If r = 0 Regression lines are perpendicular to each other.

Regression Analysis: In regression analysis, we are concerned with the estimation of one variable for a given value of another variable (or for a given set of values of a number of variables) on the basis of an average mathematical relationship between the two variables (or a number of variables). Regression analysis plays a very important role in the field of every human activity. A businessman may be keen to know what would be his estimated profit for a given level of investment on the basis of the past records. Similarly, an outgoing student may like to know her chance of getting a first class in the final University Examination on the basis of her performance in the college selection test.

When there are two variables x and y and if y is influenced by x i.e. if y depends on x, then we get a simple linear regression or simple regression. y is known as dependent variable or regression or explained variable and x is known as independent variable or predictor or explanator. In the previous examples since profit depends on investment or performance in the University Examination is dependent on the performance in the college selection test, profit or performance in the University Examination is the dependent variable and investment or performance in the selection test is the In-dependent variable.

In case of a simple regression model if y depends on x, then the regression line of y on x in given by

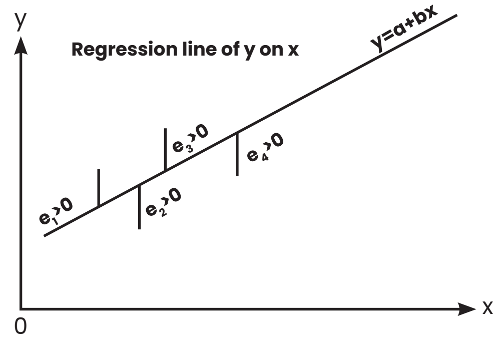
$$y = a + bx$$

Here a and b are two constants and they are also known as regression parameters. Furthermore, b is also known as the regression coefficient of y on x and is also denoted by byx. We may define the regression line of y on x as the line of best fit obtained by the method of least squares and used for estimating the value of the dependent variable y for a known value of the independent variable x.

The method of least squares involves in minimizing

$$\Sigma e_i^2 = \Sigma\,(y_i^2 - \hat{y}_i)^2 = \Sigma(y_i - a - bxi)^2$$

where yi demotes the actual or observed value and y^i = a + bxi, the estimated value of yi for a given value of xi, ei is the difference between the observed value and the estimated value and ei is technically known as error or residue. This summation intends over n pairs of observations of (xi, yi). The line of regression of y or x and the errors of estimation are shown in the following figure.



SHOWING REGRESSION LINE OF y on x AND ERRORS OF ESTIMATION

Minimisation of the equation yields the following equations known as 'Normal Equations'

$$\Sigma y_i = na + b\Sigma x_i$$
$$\Sigma x_i y_i = a\Sigma x_i + b\Sigma x_i^2$$

Solving there two equations for b and a, we have the "least squares" estimates of b and a as

$$b = \frac{cov(x_i y_i)}{S_x^2} = \frac{r.S_x.S_y}{S_x^2}$$

$$= \frac{r.S_x.S_y}{S_x^2}$$

After estimating b, estimate of a is given by

$$a = \overline{y} - b\overline{x}$$

Substituting the estimates of b and a in equation, we get

$$\frac{(y - \overline{y})}{S_y} = \frac{r(x - \overline{x})}{S_x}$$

There may be cases when the variable x depends on y and we may take the regression line of x on y as

$x = a^{\wedge} + b^{\wedge}y$

Unlike the minimization of vertical distances in the scatter diagram as shown in figure (12.7) for obtaining the estimates of a and b, in this case we minimize the horizontal distances and get the following normal equation in a^ and b^, the two regression parameters:

$$\Sigma x_i = na^{\wedge} + b^{\wedge}\Sigma y_i$$
$$\Sigma x_i y_i = a^{\wedge}\Sigma y_i + b^{\wedge}\Sigma y_i^2$$

or solving these equations, we get

$$b^{\wedge} = b_{xy} = \frac{cov(x, y)}{S_y^2} = \frac{r.S_x}{S_y}$$

$$b^{\wedge} = \overline{x} - b^{\wedge}\,\overline{y}$$

A single formula for estimating b is given by

$$b\hat{} = b_{yx} = \frac{n \Sigma x_i y_i - \Sigma x_i . \Sigma y_i}{n \Sigma y_i^2 - (\Sigma y_i)^2}$$

Similarly, $b\hat{} = b_{yx} = \frac{n \Sigma x_i y_i - \Sigma x_i . \Sigma y_i}{n \Sigma y_i^2 - (\Sigma y_i)^2}$

The standardized form of the regression equation of x on y is given by

$$\frac{x - \bar{x}}{S_x} = r \frac{(y - \bar{y})}{S_y}$$

PROPERTIES of Regression lines:
We consider the following important properties of regression lines:

(i) The regression coefficients remain unchanged due to a shift of origin but change due to a shift of scale.

This property states that if the original pair of variables is (x, y) and if they are changed to the pair (u, v) where

$$u = \frac{x - a}{p} \text{ and } v = \frac{(y - c)}{q}$$

$$b_{yx} = \frac{p}{q} \times b_{vu} \text{ and bxy} = \frac{q}{p} \times b_{uv}$$

(ii) The two lines of regression intersect at the point, where x and y are the variables under consideration.
According to this property, the point of intersection of the regression line of y on x and the regression line of x on y is i.e. the solution of the simultaneous equations in x and y.

(iii) The coefficient of correlation between two variables x and y in the simple geometric mean of the two regression coefficients. The sign of the correlation coefficient would be the common sign of the two regression coefficients.

This property says that if the two regression coefficients are denoted by byx (=b) and bxy (=b') then the coefficient of correlation is given by

$$r = \pm\sqrt{b_{yx} \times b_{xy}}$$

If both the regression coefficients are negative, r would be negative and if both are positive, rwould assume a positive value.

Example :1 Compute the correlation coefficient between x and y from the following data n = 10, Σxy = 220, Σx^2 = 200, Σy^2 = 262

Σx = 40 and Σy = 50

Solution From the given data, we have by applying

$$r = \frac{n\Sigma xy - \Sigma x \times \Sigma y}{\sqrt{n\Sigma x^2 - (\Sigma x)^2}\ \sqrt{n\Sigma y^2 - (\Sigma y)^2}}$$

$$= \frac{10 \times 220 - 40 \times 50}{\sqrt{10 \times 200 - (40)^2} \times \sqrt{10 \times 262 - (50)^2}}$$

$$= \frac{2200 - 2000}{\sqrt{2000 - 1600} \times \sqrt{2620 - 2500}}$$

$$= \frac{200}{20 \times 10.9545} = 0.91$$

Thus, there is a good amount of positive correlation between the two variables x and y.

Alternately

As given, $\bar{x} = \frac{\Sigma x}{n} = \frac{40}{10} = 4 \quad \bar{y} = \frac{\Sigma y}{n} = \frac{50}{10} = 5$

$$\text{Cov}(x, y) = \frac{\Sigma xy}{n} - \bar{x}.\bar{y} = \frac{220}{10} - 4.5 = 2$$

$$S_x = \sqrt{\frac{\Sigma x^2}{n} - (\bar{x})^2} = \sqrt{\frac{200}{10} - 4^2} = 2$$

$$S_y = \sqrt{\frac{\Sigma y_i^2}{n} - \bar{y}^2} = \sqrt{\frac{262}{10} - 5^2} = \sqrt{26.20 - 25} = 1.0954$$

Thus, applying formula, we get

$$r = \frac{\text{cov}(x,y)}{S_x.S_y} = \frac{2}{2 \times 1.0954} = 0.91$$

As before, we draw the same conclusion.

Example: 2 For a group of 8 students, the sum of squares of differences in ranks for Mathematics and Statistics marks was found to be 50 what is the value of rank correlation coefficient?

Solution: As given n = 8 and = 50. Hence the rank correlation coefficient between marks in Mathematics and Statistics is given by

$$r_R = 1 - \frac{6\Sigma d_i^2}{n(n^2 - 1)} = 1 - \frac{6 \times 50}{8(8^2 - 1)} = 0.40$$

Example 3 For a number of towns, the coefficient of rank correlation between the people living below the poverty line and increase of population is 0.50. If the sum of squares of the differences in ranks awarded to these factors is 82.50, find the number of towns.

Solution: As given $r_R = 0.50$, $\Sigma d_i^2 = 82.50$.

Thus $r_R = 1 - \dfrac{6\Sigma d_i^2}{n(n^2 - 1)}$

$0.50 = 1 - \dfrac{6 \times 82.50}{n(n^2 - 1)}$

$= n(n^2 - 1) = 990$

$= n(n^2 - 1) = 10(10^2 - 1)$

Therefore n = 10 as n must be a positive integer.

Example: 4 While computing rank correlation coefficient between profits and investment for 10 years of a firm, the difference in rank for a year was taken as 7 instead of 5 by mistake and the value of rank correlation coefficient was computed as 0.80. What would be the correct value of rank correlation coefficient after rectifying the mistake?

Solution:

We are given that n = 10,

$r_R = 0.80$ and the wrong $d_i = 7$ should be replaced by 5.

$r_R = 1 - \dfrac{6\Sigma d_i^2}{n(n^2 - 1)}$

$0.80 = 1 - \dfrac{6\Sigma d_i^2}{10(10^2 - 1)}$

$\Sigma d_i^2 = 33$

Corrected $\Sigma d_i^2 = 33 - 7^2 + 5^2 = 9$

Hence rectified value of rank correlation coefficient = $1 - \dfrac{6 \times 9}{10(10^2 - 1)} = 0.95$

Example: 5 Find product moment correlation coefficient from the following information:

X	:	2	3	5	5	6	8
Y	:	9	8	8	6	5	3

Solution: In order to find the covariance and the two-standard deviation, we prepare the following table:

Table: Computation of Correlation Coefficient

x_i	y_i	x_iy_i	x_i^2	y_i^2
(1)	**(2)**	**(3)= (1) x (2)**	**(4)= (1)²**	**(5)= (2)²**
2	9	18	4	81
3	8	24	9	64
5	8	40	25	64
5	6	30	25	36
6	5	30	36	25
8	3	24	64	9
29	39	166	163	279

We have

$$\bar{x} = \frac{29}{6} = 4.8333 \quad \bar{y} = \frac{39}{6}$$

$$\text{cov}(x, y) = \frac{\Sigma x_i y_i}{n} - \bar{x}\,\bar{y}$$

$$= 166/6 - 4.8333 \times 6.50 = -3.7498$$

$$= \sqrt{\frac{\Sigma x_i^2}{n} - (\bar{x})^2}$$

$$= \sqrt{\frac{163}{6}\ (4.8333)^2}$$

$$= \sqrt{27.1667 - 23.3608} = 1.95$$

$$Sy = \sqrt{\frac{\Sigma y_i^2}{n} - (\bar{y})^2}$$

$$= \sqrt{\frac{279}{6} - (6.50)^2}$$

$$= \sqrt{46.50 - 42.25} = 2.0616$$

Thus the correlation coefficient between x and y in given by

$$r = \frac{cov(x,y)}{S_x.S_y} = \frac{-3.7498}{1.9509 \times 2.0616} = -0.93$$

We find a high degree of negative correlation between x and y. Also, we could have applied formula as we have done for the first problem of computing correlation coefficient.

Sometimes, a change of origin reduces the computational labor to a great extent. This we are going to do in the next problem.

Example: 6 Given that the correlation coefficient between x and y is 0.8, write down the correlation coefficient between u and v where

(i) $2u + 3x + 4 = 0$ and $4v + 16x + 11 = 0$

(ii) $2u - 3x + 4 = 0$ and $4v + 16x + 11 = 0$

(iii) $2u - 3x + 4 = 0$ and $4v - 16x + 11 = 0$

(iv) $2u + 3x + 4 = 0$ and $4v - 16x + 11 = 0$

Solution

Using formula , we find that

$$r_{xy} = \frac{bd}{|b||d|} r_{uv}$$

i.e. $r_{xy} = r_{uv}$ if b and d are of same sign and $r_{uv} = -r_{xy}$ when b and d are of opposite signs, b and d being the scales of x and y respectively. In (i), $u = (-2) + (-3/2)\ x$ and $v = (-11/4) + (-4)y$.

Since $b = -3/2$ and $d = -4$ are of same sign, the correlation coefficient between u and v would be the same as that between x and y i.e. $r_{xy} = 0.8 = r_{uv}$

In (ii), $u = (-2) + (3/2)x$ and $v = (-11/4) + (-4)y$ Hence $b = 3/2$ and $d = -4$ are of opposite signs and we have $r_{uv} = -r_{xy} = -0.8$

Proceeding in a similar manner, we have $r_{uv} = 0.8$ and -0.8 in (iii) and (iv).

Example 7: For the variables x and y, the regression equations are given as $7x - 3y - 18 = 0$ and $4x - y - 11 = 0$

(i) Find the arithmetic means of x and y.

(ii) Identify the regression equation of y on x.

(iii) Compute the correlation coefficient between x and y.

(iv) Given the variance of x is 9, find the SD of y.

Solution

(i) Since the two lines of regression intersect at the point, ( ) replacing x and y by and respectively in the given regression equations, we get

$$7\bar{x} - 3\bar{y} - 18 = 0$$

and $$4\bar{x} - \bar{y} - 11 = 0$$

Solving these two equations, we get $\bar{x} = 3$ and $\bar{y} = 1$

Thus the arithmetic means of x and y are given by 3 and 1 respectively.

(ii) Let us assume that $7x - 3y - 18 = 0$ represents the regression line of y on x and $4x - y - 11 = 0$ represents the regression line of x on y.

Now $7x - 3y - 18 = 0$

$$\Rightarrow \; y = (-6) + \frac{(7)}{3}x$$

$$\Rightarrow \; b_{yx} = \frac{7}{3}$$

Again $4x - y - 11 = 0$

$$\Rightarrow \; x = \frac{(11)}{4} + \frac{(11)}{4}y \qquad \therefore b_{xy} = \frac{1}{4}$$

Thus $r^2 = b_{yx} \times b_{xy}$

$$= \frac{7}{3} \times \frac{1}{4}$$

$$= \frac{7}{12} < 1$$

Since $|r| \leq 1 \Rightarrow r^2 \leq 1$, our assumptions are correct. Thus, $7x - 3y - 18 = 0$ truly represents the regression line of y on x.

(iii) Since $r^2 = \frac{7}{12}$

$\therefore \; r = \sqrt{\frac{7}{12}}$ (We take the sign of r as positive since both the regression coefficients are positive)

$= 0.7638$

(iv) $b_{yx} = r \text{ X } \frac{S_y}{S_x}$

$$\Rightarrow \frac{7}{3} = 0.7638 \text{ X } \frac{S_y}{3} \qquad (\therefore S_x^{\,2} = 9 \text{ as given})$$

$$\Rightarrow S_y = \frac{7}{0.7638} = 9.1647$$

Example 8: The following data relate to the test scores obtained by eight salesmen in an aptitude test and their daily sales in thousands of rupees:

Salesman :	1	2	3	4	5	6	7	8
scores :	60	55	62	56	62	64	70	54
Sales :	31	28	26	24	30	35	28	24

Solution

Let the scores and sales be denoted by x and y respectively. We take a, origin of x as the average of the two extreme values i.e. 54 and 70. Hence a = 62 similarly, the origin of y is taken

as $b = \frac{24 + 35}{2} \cong 30$

Table 12.4

Computation of Correlation Coefficient Between Test Scores and Sales.

Scores	**Sales in**	**ui**	**vi**	**uivi**	**ui2**	**vi2**
(xi)	**Rs. 1000**	**= xi – 62**	**= yi – 30**			
(1)	**(yi)**					
	(2)	**(3)**	**(4)**	**(5)=(3)x(4)**	**(6)=(3) 2**	**(7)=(4) 2**
60	31	–2	1	–2	4	1
55	28	–7	–2	14	49	4
62	26	0	–4	0	0	16
56	24	–6	–6	36	36	36
62	30	0	0	0	0	0
64	35	2	5	10	4	25
70	28	8	–2	–16	64	4
54	24	–8	–6	48	64	36
Total	—	–13	–14	90	221	122

Since correlation coefficient remains unchanged due to change of origin, we have

$$r = r_{xy} = r_{uv} = \frac{n\Sigma u_i v_i - \Sigma u_i \times \Sigma v_i}{\sqrt{n\Sigma u_i^2 - (\Sigma u_i)^2}\ \sqrt{n\Sigma v_i^2 - (\Sigma v_i)^2}}$$

$$= \frac{8 \times 90 - (-13) \times (-14)}{\sqrt{8 \times 221 - (-13)^2} \times \sqrt{8 \times 122 - (-14)^2}} = \frac{538}{\sqrt{1768 - 169} \times \sqrt{976 - 196}}$$

$$= 0.48$$

In some cases, there may be some confusion about selecting the pair of variables for which correlation is wanted. This is explained in the following problem.

Example 9 Examine whether there is any correlation between age and blindness on the basis of the following data:

Age in years :	0-10	10-20	20-30	30-40	40-50	50-60	60-70	70-80
No. of Persons (in thousands) :	90	120	140	100	80	60	40	20
No. of blind Persons :	10	15	18	20	15	12	10	06

Solution

Let us denote the mid-value of age in years as x and the number of blind persons per lakh as y. Then as before, we compute correlation coefficient between x and y.

Table : Computation of correlation between age and blindness

Age in years (1)	Mid-value x (2)	No. of Persons ('000) P (3)	No. of blind B (4)	No. of blind per lakh y=B/P × 1 lakh (5)	xy (2)×(5) (6)	x^2 $(2)^2$ (7)	y^2 $(5)^2$ (8)
0-10	5	90	10	11	55	25	121
10-20	15	120	15	12	180	225	144
20-30	25	140	18	13	325	625	169
30-40	35	100	20	20	700	1225	400
40-50	45	80	15	19	855	2025	361
50-60	55	60	12	20	1100	3025	400
60-70	65	40	10	25	1625	4225	625
70-80	75	20	6	30	2250	5625	900
Total	320	—	—	150	7090	17000	3120

The correlation coefficient between age and blindness is given by

$$r = \frac{n\Sigma xy - \Sigma x.\Sigma y}{\sqrt{n\Sigma x^2-(\Sigma x)^2}\ \sqrt{n\Sigma y^2-(\Sigma y)^2}}$$

$$= \frac{8.7090 - 320.150}{\sqrt{8.17000-(320)^2} \times \sqrt{8.3120 - (150)^2}}$$

$$= \frac{8720}{183.3030.49.5984} = 0.96$$

which exhibits a very high degree of positive correlation between age and blindness.

Example-10: Coefficient of correlation between x and y for 20 items is 0.4. The AM's and SD's of x and y are known to be 12 and 15 and 3 and 4 respectively. Later on, it was found that the pair (20, 15) was wrongly taken as (15, 20). Find the correct value of the correlation coefficient.

Solution

We are given that n = 20 and the original r = 0.4, $\bar{x}$ = 12, $\bar{y}$ = 15, S_x = 3 and S_y = 4

$$r = \frac{cov(x,y)}{S_x.S_y} = 0.4 = \frac{cov(x,y)}{3 \times 4}$$

$$= cov(x,y)=4.8$$

$$= \frac{\Sigma xy}{n} - \bar{x}\,\bar{y}=4.8$$

$$= \frac{\Sigma xy}{20} - 12 \times 15=4.8$$

$$= \Sigma xy = 3696$$

Hence, corrected Σxy= 3696 – 20 × 15 + 15 × 20 = 3696

Also, $S_x^2 = 9$

$= (\Sigma x^2/ 20) - 12^2 = 9$

$\Sigma x^2 = 3060$

Similarly, $S_y^2 = 16$

$$S_y^2 = \frac{\Sigma y^2}{20} - 15^2=16$$

$\Sigma y^2 = 4820$

Thus corrected Σx = n – wrong x value + correct x value.

= 20 × 12 – 15 + 20

= 245

Similarly corrected Σy = 20 × 15 – 20 + 15 = 295

Corrected $\Sigma x^2 = 3060 - 15^2 + 20^2 = 3235$

Corrected $\Sigma y^2 = 4820 - 20^2 + 15^2 = 4645$

Thus corrected value of the correlation coefficient by applying formula (12.5)

$$= \frac{20.3696 - 245.295}{\sqrt{20.3235-(245)^2} \times \sqrt{20.4645 - (295)^2}}$$

$$= \frac{73920 - 72275}{68.3740 \times 76.6480}$$

$$= 0.31$$

Chapter 18:

Index Numbers

Introduction

Index numbers are convenient devices for measuring relative changes of differences from time to time or from place to place. Just as the arithmetic mean is used to represent a set of values, an index number is used to represent a set of values over two or more different periods or localities.

The basic device used in all methods of index number construction is to average the relative change in either quantities or prices since relatives are comparable and can be added even though the data from which they were derived cannot themselves be added. For example, if wheat production has gone up to 110% of the previous year's production and cotton production has gone up to 105%, it is possible to average the two percentages as they have gone up by 107.5%. This assumes that both have equal weight; but if wheat production is twice as important as cotton, percentage should be weighted 2 and 1. The average relatives obtained through this process are called the index numbers.

- **Relatives:** One of the simplest examples of an index number is a price relative, which is the ratio of the price of single commodity in a given period to its price in another period called the base period or the reference period. It can be indicated as follows:

Price relative = $\frac{P_n}{P_0} \times 100$

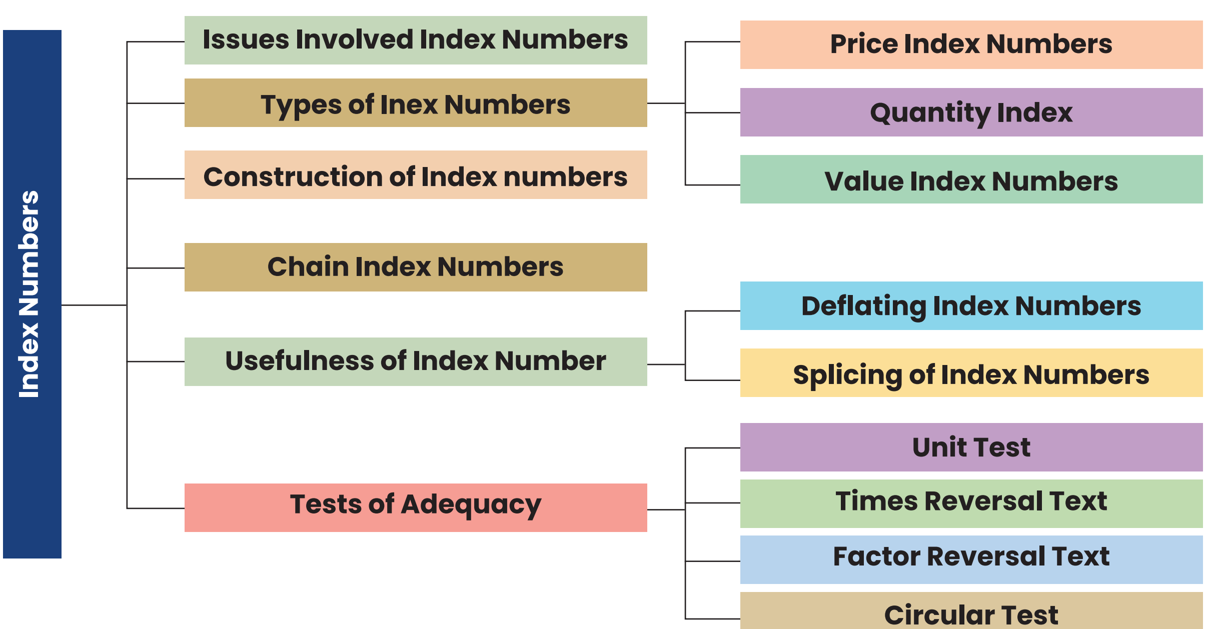
There can be other relatives such as of quantities, volume of consumption, exports, etc. The relatives in that case will be:

Quantity relative = $\frac{Q_n}{Q_0} \times 100$

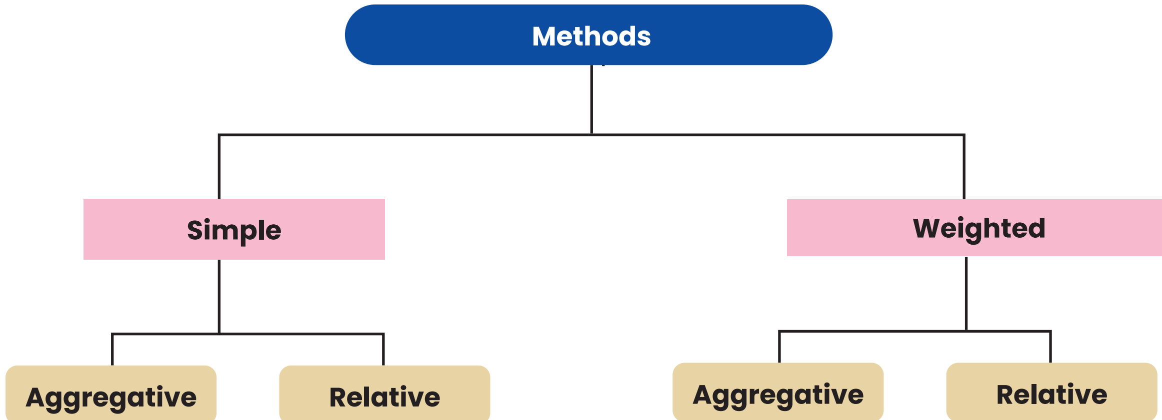
Similarly, there are value relatives:

Value relative = $\frac{V_n}{V_0} = \frac{P_n Q_n}{P_0 Q_0} = \left(\frac{P_n}{P_0} \times \frac{Q_n}{Q_0}\right) \times 100$

Index number Overview



Methods of Index numbers:



Price Index numbers

(a) Simple aggregative price index $= \frac{\Sigma P_n}{\Sigma P_0} \text{ x } 100$

(b) Laspeyres' Index: In this Index base year quantities are used as weights:

Laspeyres Index $= \frac{\Sigma P_n Q_0}{\Sigma P_0 Q_0} \text{ x } 100$

(c) Paasche's Index: In this Index current year quantities are used as weights:

Passche's Index $= \frac{\Sigma P_n Q_n}{\Sigma P_0 Q_n} \text{ x } 100$

(d) The Marshall-Edgeworth index uses this method by taking the average of the base year and the current year

Marshall-Edgeworth Index $= \frac{\Sigma P_n (Q_0 + Q_n)}{\Sigma P_0 (Q_0 + Q_n)} \text{ x } 100$

(e) Fisher's ideal Price Index: This index is the geometric mean of Laspeyres' and Paasche's.

Fisher's Index $= \sqrt{\frac{P_n Q_0}{P_0 Q_0} \text{ x } \frac{P_n Q_n}{P_0 Q_n}} \text{ x } 100$

(g) Weighted Average of Relative Method: $\frac{\Sigma \left(\frac{P_n}{P_0}\right) \text{x} (P_0 Q_0)}{\Sigma P_0 Q_0} \text{ x } 100 \quad \frac{\Sigma P_n Q_0}{\Sigma P_0 Q_0} \text{ x } 100$

(h) Chain Index $= \frac{\text{Link relative of current year x Chain Index of the previous year}}{100}$

Quantity Index Numbers

- Simple aggregate of quantities: $\frac{\Sigma Q_n}{\Sigma Q_0} \times 100$

- The simple average of quantity relatives: $\frac{\frac{\Sigma Q_n}{\Sigma Q_0}}{N} \times 100$

- Weighted aggregate quantity indices:

  (i) With base year price as weight : $\frac{\Sigma Q_n P_0}{\Sigma Q_0 P_0} \times 100$ (Laspeyre's index)

  (ii) With current year price as weight : $\frac{\Sigma Q_n P_n}{\Sigma Q_0 P_n} \times 100$ (Paasche's index)

  (iii) Geometric mean of (i) and (ii) : $\sqrt{\frac{\Sigma Q_n P_0}{\Sigma Q_0 P_0} \times \frac{\Sigma Q_n P_n}{\Sigma Q_0 P_n}} \times 100$ (Fisher's Ideal)

- Base-year weighted average of prices as relatives in Marshall-Edgeworth quantity index number.

Weighted Relative method formula $\frac{\Sigma\left(\frac{Q_n}{Q_0} P_0 Q_0\right)}{\Sigma P_0 Q_0} \times 100$

- Value Indices $\frac{V_n}{V_0} = \frac{\Sigma P_n Q_n}{\Sigma P_0 Q_0} \times 100$

The Chain Index Numbers

So far we concentrated on a fixed base but it does not suit when conditions change quite fast. In such a case the changing base for example, 2018 for 2019, and 2019 for 2020, and so on, may be more suitable. If, however, it is desired to associate these relatives to a common base the results may be chained. Thus, under this method the relatives of each year are first related to the preceding year called the link relatives and then they are chained together by successive multiplication to form a chain index.

The formula is:

$$\text{Chain Index} = \frac{\text{Link relative of current year x Chain Index of the previous year}}{100}$$

Illustrations

Commodities	2018	2019	2020
Cheese (per 100 gms)	120	150	160
Egg (per piece)	30	36	40
Potato (per kg)	50	60	60
Aggregrate	200	246	260
Index	100	123	130

Simple Aggregative Index for 2019 over 2018 = $\frac{\sum P_n}{\sum P_0} = \frac{246}{200} \times 100 = 123$

and for 2020 over 2018 = $\frac{\sum P_n}{\sum P_0} \times 100 = \frac{246}{200} \times 100 = 130$

The above method is easy to understand but it has a serious defect. It shows that the first commodity exerts greater influence than the other two because the price of the first commodity is higher than that of the other two. Further, if units are changed then the Index numbers will also change. Students should independently calculate the Index number taking the price of eggs per dozen i.e., ₹ 36, ₹ 43.20, ₹ 39.60 for the three years respectively. This is the major flaw in using absolute quantities and not the relatives. Such price quotations become the concealed weights which have no logical significance.

Limitations of Index Numbers

So far we have studied various types of index numbers. However, they have certain limitations. They are:

1. As the indices are constructed mostly from deliberate samples, chances of errors creeping in cannot be always avoided.
2. Since index numbers are based on some selected items, they simply depict the broad trend and not the real picture.
3. Since many methods are employed for constructing index numbers, the result gives different values and this at times create confusion.

Usefulness of Index Numbers

In spite of its limitations, index numbers are useful in the following areas:

1. Framing suitable policies in economics and business. They provide guidelines to make decisions in measuring intelligence quotients, research etc.
2. They reveal trends and tendencies in making important conclusions in cyclical forces, irregular forces, etc.
3. They are important in forecasting future economic activity. They are used in time series analysis to study long-term trend, seasonal variations and cyclical developments.
4. Index numbers are very useful in deflating, i.e., they are used to adjust the original data for price changes and thus transform nominal wages into real wages.
5. Cost of living index numbers measure changes in the cost of living over a given period.

Deflating Time Series Using Index Numbers

Sometimes a price index is used to measure the real values in economic time series data expressed in monetary units. For example, GNP initially is calculated in current price so that the effect of price changes over a period of time gets reflected in the data collected. Thereafter, to determine how much the physical goods and services have grown over time, the effect of changes in price over different values of GNP is excluded. The real economic growth in terms of constant prices of the base year therefore is determined by deflating GNP values using price index.

Deflating Time Series Using Index Numbers

Sometimes a price index is used to measure the real values in economic time series data expressed in monetary units. For example, GNP initially is calculated in current price so that the effect of price changes over a period of time gets reflected in the data collected. Thereafter, to determine how much the physical goods and services have grown over time, the effect of changes in price over different values of GNP is excluded. The real economic growth in terms of constant prices of the base year therefore is determined by deflating GNP values using price index.

Year	Wholesale Price Index	GNP at Current Prices	Real GNP
2019	113.1	7499	6630
2020	116.3	7935	6823
2021	121.2	8657	7143
2022	127.7	9323	7301

The formula for conversion can be stated as

$$\text{Deflated Value} = \frac{\text{Current Value}}{\text{Price Index of the current year}} \times 100$$

$$\text{Or Current Value} \times \frac{\text{Base Price } (P_0)}{\text{Current Price } (P_n)}$$

Shifting and Splicing of Index Numbers

These refer to two technical points: (i) how the base period of the index may be shifted, (ii) how two index covering different bases may be combined into single series by splicing.

Shifted Price Index

Year	Original Price Index	Shifted Price Index to base 2021
2011	100	71.4
2012	104	74.3
2013	106	75.7
2014	107	76.4
2015	110	78.6
2016	112	80.0
2017	115	82.1
2018	117	83.6
2019	125	89.3
2020	131	93.6
2021	140	100.0
2022	147	105.0

$$\text{The formula used is, Shifted Price Index} = \frac{\text{Original Price Index}}{\text{Price Index of the year on which it has to be shifted}} \times 100$$

Splicing two sets of price index numbers covering different periods of time is usually required when there is a major change in quantity weights. It may also be necessary on account of a new method of calculation or the inclusion of new commodity in the index.

Splicing Two Index Number Series

Year	Old Price Index [2010 = 100]	Revised Price Index [2015=100]	Spliced Price Index [2015 = 100]
2010	100.0		87.6
2011	102.3		89.6
2012	105.3		92.2
2013	107.6		94.2
2014	111.9		98.0
2015	114.2	100.0	100.0
2016		102.5	102.5
2017		106.4	106.4
2018		108.3	108.3
2019		111.7	111.7
2020		117.8	117.8

You will notice that the old series 2010 has to be converted shifting to the base. 2015 i.e, 114.2 to have a continuous series, even when the two parts have different weights

Test of Adequacy

There are four tests:

1. Unit Test: This test requires that the formula should be independent of the unit in which or for which prices and quantities are quoted. Except for the simple (unweighted) aggregative index all other formulae satisfy this test.

(ii) Time Reversal Test: It is a test to determine whether a given method will work both ways in time, forward and backward. The test provides that the formula for calculating the index number should be such that two ratios, the current on the base and the base on the current should multiply into unity. In other words, the two indices should be reciprocals of each other. Symbolically,

$P_{01} \times P_{10} = 1$

where P_{01} is the index for time 1 on 0 and P_{10} is the index for time 0 on 1. You will notice that Laspeyres' method and Paasche's method do not satisfy this test, but Fisher's Ideal Formula does.

While selecting an appropriate index formula, the Time Reversal Test and the Factor Reversal test are considered necessary in testing the consistency.

Laspeyres: $P_{01} = \frac{\Sigma P_1 Q_0}{\Sigma P_0 Q_0}$ $\qquad$ $P_{10} = \frac{\Sigma P_0 Q_1}{\Sigma P_1 Q_1}$

$$P_{01} \times P_{10} = \frac{\Sigma P_1 Q_0}{\Sigma P_0 Q_0} \times \frac{\Sigma P_0 Q_1}{\Sigma P_1 Q_1} \neq 1$$

Paasche's: $P_{01} = \frac{\Sigma P_1 Q_1}{\Sigma P_0 Q_1}$ $\quad P_{10} = \frac{\Sigma P_0 Q_0}{\Sigma P_1 Q_0}$

$$P_{01} \times P_{10} = \frac{\Sigma P_1 Q_1}{\Sigma P_0 Q_1} \times \frac{\Sigma P_0 Q_0}{\Sigma P_1 Q_0} \neq 1$$

Fisher's: $P_{01} = \sqrt{\frac{\Sigma P_1 Q_0}{\Sigma P_0 Q_0} \times \frac{\Sigma P_1 Q_1}{\Sigma P_0 Q_1}}$ $\quad P_{10} = \sqrt{\frac{\Sigma P_0 Q_1}{\Sigma P_1 Q_1} \times \frac{\Sigma P_0 Q_0}{\Sigma P_1 Q_0}}$

$$P_{01} \times P_{10} = \sqrt{\frac{\Sigma P_1 Q_0}{\Sigma P_0 Q_0} \times \frac{\Sigma P_1 Q_1}{\Sigma P_0 Q_1} \times \frac{\Sigma P_0 Q_1}{\Sigma P_1 Q_1} \times \frac{\Sigma P_0 Q_0}{\Sigma P_1 Q_0}} = 1$$

(iii) Factor Reversal Test: This holds when the product of price index and the quantity index should be equal to the corresponding value index, i.e.,

Symbolically: $P_{01} \times Q_{01} = V_{01}$

Fishers' $P_{01} = \sqrt{\frac{\Sigma P_1 Q_0}{\Sigma P_0 Q_0} \times \frac{\Sigma P_1 Q_1}{\Sigma P_0 Q_1}}$ $\quad Q_{01} = \sqrt{\frac{\Sigma Q_1 P_0}{\Sigma Q_0 P_0} \times \frac{\Sigma Q_1 P_1}{\Sigma Q_0 P_1}}$

$$P_{01} \times Q_{01} = \sqrt{\frac{\Sigma P_1 Q_0}{\Sigma P_0 Q_0} \times \frac{\Sigma P_1 Q_1}{\Sigma P_0 Q_1} \times \frac{\Sigma Q_1 P_0}{\Sigma Q_0 P_0} \times \frac{\Sigma Q_1 P_1}{\Sigma Q_0 P_1}} = \sqrt{\frac{\Sigma P_1 Q_1}{\Sigma P_0 Q_0} \times \frac{\Sigma P_1 Q_1}{\Sigma P_0 Q_0}} = \frac{\Sigma P_1 Q_1}{\Sigma P_0 Q_0} = V_{01}$$

Thus Fisher's Index satisfies Factor Reversal test. Because Fisher's Index number satisfies both the tests in (ii) and (iii), it is called an Ideal Index Number.

(iv) Circular Test:
It is concerned with the measurement of price changes over a period of years, when it is desirable to shift the base. For example, if the 1970 index with base 1965 is 200 and 1965 index with base 1960 is 150, the index 1970 on base 1960 will be 300. This property therefore enables us to adjust the index values from period to period without referring each time to the original base. The test of this shiftability of base is called the circular test.

- **This test is not met by Laspeyres, or Paasche's or the Fisher's ideal index. The simple geometric mean of price relatives and the weighted aggregative with fixed weights meet this test.**

Example 1:

Compute Fisher's Ideal Index from the following data:

Commodities	Base Year		Current Year	
	Price	Quantity	Price	Quantity
A	4	3	6	2
B	5	4	6	4
C	7	2	9	2
D	2	3	1	5

Show how it satisfies the time and factor reversal tests.

Solution:

Commodities	P_0	Q_0	P_1	Q_1	P_0Q_0	P_1Q_0	P_0Q_1	P_1Q_1
A	4	3	6	2	12	18	8	12
B	5	4	6	4	20	24	20	24
C	7	2	9	2	14	18	14	18
D	2	3	1	5	6	3	10	5
					52	63	52	59

Fisher's Ideal Index: $P_{01} = \sqrt{\frac{\Sigma P_1Q_0}{\Sigma P_0Q_0} \times \frac{\Sigma P_1Q_1}{\Sigma P_0Q_1}} \text{ X } 100 = \sqrt{\frac{63}{52} \text{ X } \frac{59}{52}} \text{ X } 100$

$= \sqrt{1.375} \text{ X } 100 = 1.172 \text{ X } 100 = 117$

Time Reversal Test: $P_{01} \text{ X } P_{10} = \sqrt{\frac{63}{52} \text{X} \frac{59}{52} \text{X} \frac{52}{59} \text{X} \frac{52}{63}} = \sqrt{1} = 1$

∴ Time Reversal Test is satisfied.

Factor Reversal Test:

$P_{01} \text{ X } P_{01} = \sqrt{\frac{63}{52} \text{X} \frac{59}{52} \text{X} \frac{52}{59} \text{X} \frac{52}{63}} = \sqrt{\frac{59}{52} \text{X} \frac{59}{52}} = \frac{59}{52}$

Since, $\frac{\Sigma P_1Q_1}{\Sigma P_0Q_0}$ is also equal to $\frac{59}{52}$, the Factor Reversal Test is satisfied.

Example 2:

If the ratio between Laspeyre's index number and Paasche's Index number is 28 : 27. Then the missing figure in the following table P is :

Commodity	Base Year		Current Year	
	Price	Quantity	Price	Quantity
X	L	10	2	5
Y	L	5	P	2

Solution:

So Laspeyre's index number = ∑ PxQo / ∑ PoQo
= 2 x 10 + P x 5 / L x 10 + L x 5
= 20 + 5P / 15 L
= 5 (4 + P) / 15 L
= 4 + P / 3L

Now for the Paasche's index number we have, ∑ Px Qx / ∑ Po Qx
= 2 x 5 + P x 2 / L x 5 + L x 2
= 2P + 10 / 7L
= 2 (P + 5) / 7L

Given Ratio = L : P = 28:27
So 4 + p / 3L / 2 (P + 5) / 7L = 28/27
or 7 (4 + P) / 6(P + 5) = 28/27
or 9(4 + P) = 8(P + 5)
36 + 9P = 8P + 40
or P = 40 – 36
or P = 4

Example 3:

The consumer price index for 2006 on the basis for 2006 on the basis of 2005 from the following data is:

Commodities in 2005	Quantities consumed in 2005	Prices in 2005	Prices in 2006
A	6	5.75	6.00
B	6	5.00	8.00
C	1	6.00	9.00
D	6	8.00	10.00
E	4	2.00	1.50
F	1	20.00	15.00

Solution:

Commodities in 2005	Quantities consumed in 2005(q_0)	Price in 2005 (P_0)	Prices in 2006 (P_1)	P_1q_0	P_0q_0
A	6	5.75	6.00	36	34.50
B	6	5.00	8.00	48	30.00
C	1	6.00	9.00	9	6.00
D	6	8.00	10.00	60	48.00
E	4	2.00	1.50	6	8.00
F	1	20.00	15.00	15	20.00
				$\Sigma P_1q_0=174$	$\Sigma P_0q_0=146.5$

Consumer Price Index = $\frac{\Sigma P_1Q_0}{\Sigma P_0Q_0}$ X 100 = $\frac{174}{146.5}$ X 100 = 118.77

Example 4:

Net monthly salary of an employee was R30,000 in 2000. The consumer price index 2015 is 250 with 2000 as base year, if he has to rightly compensated, then dearness allowance to be paid to the employee is:

Solution:

The consumer price index number in 2015 is 250 with 2000 as base year.

if in 2000 = 100 then in 2015 = 250

if in 2000 = 1 then in 2015 = 250/100 = 2.5

if in 2000 = 30000 then 250/100* 30000 = 75,000

additional dearness allowance to be paid to the employee is = 75000 – 30000 = R45000

additional dearness allowance to be paid to the employee = ₹45,000

Example 5:

Consumer price index number goes up from 110 to 200 and the Salary of a worker is also raised from R330 to R500. Therefore, in real terms, to maintain his previous standard of living, he should get an additional amount of:

Solution:

Cost of Living Index in base year = 110,

Cost of Index in current year = 200

Salary of worker in base year = 330

Salary of worker in current year = 500

Real wages (for base year) = $\frac{\text{Money wages in base year}}{\text{Cost of living index in base year}}$ x 100

= $\frac{330}{110}$ x 100 = 300

$$\text{Real wages (for current year)} = \frac{\text{Money wages in current year}}{\text{Cost of living index number in current year}} \times 100$$

$$= \frac{300}{200} \times 100 = 250$$

Thus, we can say that even though money wage of the worker had increased from ₹330 to ₹500, his real wage has fallen from ₹300 to ₹250. This implies a loss of ₹50 in real terms.

Example 6:

If the price of a commodity in a place have decreased by 30% over the base period prices, then the index number of that place is:

Solution:

Base price of any commodity = 10 decreased price = 30% of 100 = 30

Index number of that place now = 100-30 =70

Example 7:

If with an increase of 10% in prices, the rise in wages is 20% then the real wages has increased by

Solution:

$$\text{Real wages} = \frac{\text{Real wage of current year}}{\text{Real wage of base year}} \times 100$$

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

$$= 120-100 = 20\%$$

Example 8:

In the year 2010, the monthly salary of clerk was ₹24,000. The consumer price index was 140 in the year 2010, which rises to 224 in 2016. If he has to be rightly compensated, what dditional monthly salary to be paid to him?

Solution:

Year	CPI	Salary
2010	140	24000
2016	224	X

$= 140/224 = 24000/x$

$$x = \frac{24000 \times 224}{140} = 38,400$$

DA = 38,400 - 24,000

= 14,400

Important Points:

1. Circular Test is an extension of time reversal test.
2. Time Reversal Test and Factor Reversal Test is satisfied by: Fisher's Ideal Index.
3. The ratio of price of single commodity in a given period to its price in the preceding year price is called the relative price.
4. Fisher's Ideal Formula does not satisfy Circular test.
5. The best average for constructing an index number is: Geometric Mean.
6. The time reversal test is satisfied by Fisher's index number.
7. The factor reversal test is satisfied by : Fisher's index.

6. The circular test is satisfied by Simple GM price relative.
7. Fisher's index number is based on Geometric mean of Laspeyre's and Paasche's index numbers.
8. Paasche index is based on: Current year quantities.
9. Purchasing Power of Money is the Reciprocal of price index number and inverse relationship between Purchasing power of money and price index number.