FOUNDATION : PAPER - 4

SYLLABUS - 2016

FUNDAMENTALS OF BUSINESS MATHEMATICS AND STATISTICS

FOUNDATION

STUDY NOTES

The Institute of Cost Accountants of India
CMA Bhawan, 12, Sudder Street, Kolkata - 700 016

First Edition : August 2016
Reprint : April 2017
Reprint : January 2018

Published by :
Directorate of Studies
The Institute of Cost Accountants of India (ICAI)
CMA Bhawan, 12, Sudder Street, Kolkata - 700 016
www.icmai.in

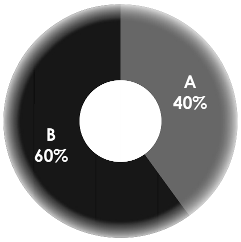
Printed at :
Mega Calibre Enterprises Pvt. Ltd.
06/315 Action Area 3,
New Town, Rajarhat,
Kolkata 700 160.

Syllabus - 2016

PAPER 4: FUNDAMENTALS OF BUSINESS MATHEMATICS AND STATISTICS (FBMS)

Syllabus Structure

A	Fundamentals of Business Mathematics	40%
B	Fundamentals of Business Statistics	60%



ASSESSMENT STRATEGY

There will be written examination paper of three hours.

OBJECTIVES

To gain understanding on the fundamental concepts of mathematics and statistics and its application in business decision-making

Learning Aims

The syllabus aims to test the student's ability to:

- Understand the basic concepts of basic mathematics and statistics
- Identify reasonableness in the calculation
- Apply the basic concepts as an effective quantitative tool
- Explain and apply mathematical techniques
- Demonstrate to explain the relevance and use of statistical tools for analysis and forecasting

Skill sets required

Level A: Requiring the skill levels of knowledge and comprehension

CONTENTS	
Section A: Fundamentals of Business Mathematics	
1. Arithmetic	20%
2. Algebra	20%
Section B: Fundamentals of Business Statistics	
3. Statistical representation of Data	10%
4. Measures of Central Tendency and Dispersion	30%
5. Correlation and Regression	10%
6. Probability	10%

SECTION A: FUNDAMENTALS OF BUSINESS MATHEMATICS [40 MARKS]

1. Arithmetic

(a) Ratios, Variations and Proportions

(b) Simple and Compound interest

(c) Arithmetic Progression and Geometric Progression

2. Algebra

(a) Set Theory

(b) Indices and Logarithms (basic concepts)

(c) Permutation and Combinations (basic concepts)

(d) Quadratic Equations (basic concepts)

SECTION B: FUNDAMENTALS OF BUSINESS STATISTICS [60 MARKS]

3. Statistical Representation of Data

(a) Diagrammatic representation of data

(b) Frequency distribution

(c) Graphical representation of Frequency Distribution – Histogram, Frequency Polygon Curve, Ogive, Pie-chart

4. Measures of Central Tendency and Dispersion

(a) Mean, Median, Mode, Mean Deviation

(b) Range, Quartiles and Quartile Deviation

(c) Standard Deviation

(d) Co-efficient of Variation

(e) Karl Pearson and Bowley's Coefficient of Skewness

5. Correlation and Regression

(a) Scatter diagram

(b) Karl Pearson's Coefficient of Correlation

(c) Regression lines, Regression equations, Regression coefficients

6. Probability

(a) Independent and dependent events; Mutually exclusive events

(b) Total and Compound Probability; Baye's theorem; Mathematical Expectation

Contents

SECTION - A

BUSINESS MATHEMATICS

Study Note 1 : Arithmetic

Study Note 2 : Algebra

SECTION - B

STATISTICS

Study Note 3 : Statistical Representation of Data

Study Note 4 : Measures of Central Tendency and Dispersion

Study Note 5 : Correlation and Regression

Study Note 6 : Probability

Section A
Fundamentals of Business Mathematics
(Syllabus - 2016)

Study Note - 1

ARITHMETIC

This Study Note includes

1.1 RATIOS, VARIATIONS AND PROPORTIONS

Ratios:

Ratio is the comparative relation between two quantities of same kind expressed in the same units.

Example: In a class test A secured 80 marks and B secured 40 marks out of 100 then we can compare that A secures double that of B.

i.e. Ratio is 80/40 = 2, is a pure natural number, and no unit is associated with it.

Note:

1. In the ratio a:b and a and b are called the terms of the ratio. Here a is called antecedent and B is called consequent.

   Properties of Ratio:

   The value of ratio remains unchanged when the terms of the ratio are multiplied and divided by the same number.

   Ex: 2:4, Multiplied by 2, 4:8   9:27, divided by 9, 1:3

2. Two or more ratios can be compared by reducing them to the same denominator.

   Ex: In the ratios 3:4 and 8:12 which is greater

   $\frac{3}{4}, \frac{8}{12}$

   $\frac{9}{12} > \frac{8}{12}$

Ratio of equality and in-equality:

1. If a = b then the ratio a:b is called equal ratio.

   Ex: 3:3, 4:4 etc are equal ratios.

2. If a>b then the ratio a:b is called greater inequality.

   Ex: 4:3, 9:7 etc are greater inequalities.

3. If a < b then the ratio a:b is called lesser inequality.

   Ex: 3:4, 7:9 etc

Inverse ratio or reciprocal ratio:

The inverse ratio of a:b is b:a.

Different kinds of ratios:

1. **Compound ratio:**

   If two or more ratios are multiplied together then the ratio is called compounded ratio.

   Ex: For the ratios a:b, c:d, e:f the compounded ratio is ace : bdf

Note:

The compounded ratio of two reciprocal ratios is unity. i.e a:b is the reciprocal ratio of b:a then the compounded ratio is ab:ba..

$\rightarrow \frac{ab}{ba} = 1$

2. **Duplicate ratio:**

   If two equal ratios are compounded together then the resulting ratio is called duplicate ratio i.e the duplicate ratio of two equal ratios a:b and a:b is a^2:b^2.

3. **Triplicate ratio:**

   If three equal ratios are multiplied together then the resulting ratio is called triplicate ratio. i.e. the triplicate ratio of a:b,a:b,a:b is a^3:b^3.

4. **Sub duplicate ratio and sub triplicate ratios:**

   $\sqrt{a} : \sqrt{b}$ is the sub duplicate ratio of a:b and $\sqrt[3]{a} : \sqrt[3]{b}$ is the sub triplicate ratio of a:b.

   Ex: The sub-duplicate ratio of 4:9 is $\sqrt{4} : \sqrt{9}$ is 2:3. The sub triplicate ratio of 8:64 is $\sqrt[3]{8} : \sqrt[3]{64}$ is 2:4= 1:2

5. **Continued ratio:**

   The continued ratio is the relation between the magnitudes of two or more ratios and is denoted by a:b:c.

   Ex: The continued ratio of 2:3 and 4:10. 8:12:30, 4:6:15

Points to be remember:-

1. Reduce the quantities to same units.

   Ex: if A= ₹ 2 and B = 50 p

   Then a:b = 200 : 50

   = 4 : 1 = 4

2. When the quantity is increased by given ratio multiply the quantity by greater ratio.
3. When the quantity is decreased by given ratio multiply the quantity by lesser ratio.
4. When both increasing and decreasing of quantities are present in a problem multiply the quantity by greater ratio in increase and multiply the result by lesser ratio to obtain the final result.

Proportions (OR) proportional:

If two rations are equal then we say that the two ratios are in proportion. In other words the four quantities a,b,c and d are said to be in proportion if a:b = c:d

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

Here the first and last quantities i.e a and d are extremes and the two middle terms b and c are called means.

Property:

The four quantities a, b, c, d are in proportion of $\Leftrightarrow \frac{a}{b} = \frac{c}{d} \Leftrightarrow ad = bc$ converse is also true.

Continued proportion:

If 3 quantities a,b, and c such that a:b = b:c. then we say that those 3 qualities are in continued proportion. If 3 quantities are in continued proportion then we get $b^2=ac$.

Ex: 3,6,12 are in continued proportion.

$= 6^2 = 3\times12 = 36$

Ex: 2, -4, 8 are in continued proportion

$\Leftrightarrow (-4)^2 = 8\times2$

$\Leftrightarrow 16 = 16$

Basic rules of proportions:

1. **Invertendo:**

   If a:b = c:d which implies b:a = d:c then we say that the proportion is invertendo.

2. **Alternendo:**

   If a:b = c:d

   $\Rightarrow$ a:c = b:d, then we say that the proportion in alternendo.

3. **Componendo:**

   If a:b = c:d

   $\Rightarrow$ a + b:b = c + d:d then we say that proportion is componendo.

4. **Dividendo:**

   If a:b = c:d

   $\Rightarrow$ a − b:b = c − d:d then we say that the proportion is in dividendo

5. **Componendo and dividendo:**

   If a:b = c:d

   $\Rightarrow \frac{a+b}{a-b} = \frac{c+d}{c-d}$ then we say that the proportion is in componendo and dividendo.

6. **Important theorem:**

   If $\frac{a}{b} = \frac{c}{d} = \frac{e}{f} = \cdots$ ... ... then each ratio $= \left\{\frac{pa^n + qc^n + re^n + \cdots\ldots\ldots}{pb^n + qd^n + rf^n + \cdots\ldots\ldots}\right\}^{1/n}$

   Where p, q, r…… are quantities

   Proof:

   Let

   $\frac{a}{b} = \frac{c}{d} = \frac{e}{f}$ ... ... k (say)

   $\therefore \frac{a}{b} = k \Rightarrow a = bk.$

   $\frac{c}{d} = k \Rightarrow c = dk.$

   $\frac{e}{f} = k = e = fk.$

R.H.s:

$$= \left[\frac{pa^n + qc^n + re^n + \cdots \ldots \ldots}{pb^n + qd^n + rf^n + \cdots \ldots \ldots} \right]^{1/n}$$

$$= \left[\frac{pb^n k^n + qc^n k^n + rf^n k^n + \cdots \ldots \ldots}{pb^n + qd^n + rf^n + \cdots \ldots \ldots} \right]^{1/n}$$

$= (k^n)^{1/n} = k$

Note:

1. Put n = 1

   $\therefore \dfrac{pa + qc + re + \cdots \ldots \ldots}{pb + qd + rf + \cdots \ldots \ldots}$ = each ratio = k

2. Put p = q = r =……… = 1

   $\dfrac{a^n + c^n + e^n + \cdots \ldots \ldots}{b^n + d^n + f^n + \cdots \ldots \ldots}$ = each ratio

3. The continued ratio

   x:y:z = a:b:c

   Can be written as $\dfrac{x}{a} = \dfrac{y}{b} = \dfrac{z}{c}$

4. If x:y = a:b it does not mean that x = a and y = b

   But x = ka, y = kb, k is a constant

Illustrations:

1. If $\dfrac{4x-3z}{4c} = \dfrac{4z-3y}{3b} = \dfrac{4y-3x}{2a}$, show that each ratio is equal to $\dfrac{x+y+z}{2a+3b+4c}$

   sol: Each of the given ratio $= \dfrac{4x-3z+4z-3y+4y-3x}{4c+3b+2a} = \dfrac{x+y+z}{2a+3b+4c}$

2. If $\dfrac{a}{b} = \dfrac{c}{d} = \dfrac{e}{f} = \dfrac{g}{h}$ show $\dfrac{aceg}{bdfh} = \dfrac{a^4+c^4+e^4+g^4}{b^4+d^4+f^4+h^4}$

   sol: $\dfrac{a}{b} = \dfrac{c}{d} = \dfrac{e}{f} = \dfrac{g}{h} = k$ (say), so that a = bk, c = dk, e = fk, g = hk

   L.H.S. $= \dfrac{bk.dk.fk.hk}{bdfh} = k^4$

   R.H.S $= \dfrac{b^4k^4 + d^4k^4 + f^4k^4 + h^4k^4}{b^4+d^4+f^4+h^4} = \dfrac{k^4\left(b^4+d^4+f^4+h^4\right)}{b^4+d^4+f^4+h^4} = k^4$. Hence the result.

3. If $a_1, a_2, \ldots\ldots a_n$, be continued proportion, show that $\dfrac{a_1}{a_n} = \left(\dfrac{a_1}{a_2}\right)^{n-1}$

   Sol: $\dfrac{a_1}{a_2} = \dfrac{a_2}{a_3} = \dfrac{a_3}{a_4} = \ldots\ldots \dfrac{a_{n-1}}{a_n} = k(\text{say})$ $\qquad k^{n-1} = \dfrac{a_1}{a_2} \times \dfrac{a_2}{a_3} \times \dfrac{a_3}{a_4} \times \ldots \times \dfrac{a_{n-1}}{a_n} = \dfrac{a_1}{a_n}$

   again, $k^{n-1} = \left(\dfrac{a_1}{a_2}\right)^{n-1}$

$$\therefore \frac{a_1}{a_n} = \left(\frac{a_1}{a_2}\right)^{n-1}$$

4. $\frac{x}{a} = \frac{y}{b} = \frac{z}{c}$ prove that $\frac{x^2 - yz}{a^2 - bc} = \frac{y^2 - zx}{b^2 - ca} = \frac{z^2 - xy}{c^2 - ab}$ (ICWA. (F) June 2007)

Sol: $\frac{x}{a} = \frac{y}{b} = \frac{z}{c}$ = k (say) x = ak, y = bk, z = ck

$\frac{x^2 - yz}{a^2 - bc} = \frac{k^2(a^2 - bc)}{(a^2 - bc)}$ = k^2, similarly $\frac{y^2 - zx}{b^2 - ca} = k^2 = \frac{z^2 - xy}{c^2 - ab}$ (to show in detail). Hence the result

5. If $\frac{p}{b-c} = \frac{q}{c-a} = \frac{r}{a-b}$ prove that p + q + r = 0 = pa + qb + rc (ICWA. (F) Dec. 2007)

Sol: $\frac{p}{b-c} = \frac{q}{c-a} = \frac{r}{a-b}$ = k (say), p = k (b – c), q = k (c – a), r = k (a – b)

Now p + q + r = k (b - c + c - a + a - b) = k × 0 = 0

And pa + qb + rc = ka (b-c) + kb(c-a) + kc(a-b) = k(ab - ac + bc - ba + ca - cb) = k×0 = 0. Hence the result.

6. If $\frac{x}{b+c} = \frac{y}{c+a} = \frac{z}{a+b}$ prove that $\frac{x(y-z)}{b^2-c^2} = \frac{y(z-x)}{c^2-a^2} = \frac{z(x-y)}{a^2-b^2}$ (ICWA(F) Dec 2005)

Sol : $\frac{x}{b+c} = \frac{y}{c+a} = \frac{z}{a+b}$ = k, x = k (b + c), y = k (c + a), z = k (a + b)

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

similarly, $\frac{y(Z-X)}{c^2-a^2} = -k^2 = \frac{z(x-y)}{a^2-b^2}$ (To show in detail) Hence the result.

7. The marks obtained by four examinees are as follows: A:B = 2:3, B:C = 4:5, C:D = 7:9, find the continued ratio.

A:B = 2:3

Sol: B:C = 4:5 = $4 \times \frac{3}{5} : 5 \times \frac{3}{4} = 3 : \frac{15}{4}$ (for getting same number in B, we are to multiply by $\frac{3}{4}$)

C:D = 7:9 = $7 \times \frac{15}{28} = \frac{15}{4} : \frac{135}{28}$ (to same term of C, multiply by $\frac{15}{28}$)

$\therefore$ A : B : C : D = 2 : 3 : $\frac{15}{4}$: $\frac{135}{28}$ 56 : 84 : 105 : 135

8. Two numbers are in the ratio of 3:5 and if 10 be subtracted from each of them, the remainders are in the ratio of 1:5, find the numbers.

Sol: Let the numbers be x and y, so that $\frac{x}{y} = \frac{3}{5}$ or, 5x = 3y ... ... (1)

Again $\frac{x-10}{y-10} = \frac{1}{5}$ or, 5x – y = 40 ....(ii), solving (I) & (ii), x = 12, y = 20

$\therefore$ regd. numbers are 12 and 20

9. The ratio of annual incomes of A an B is 4:3 and their annual expenditure is 3:2. If each of them saves ₹ 1000 a year, find their annual income.

   Sol: Let the incomes be 4x and 3x (in ₹)

   Now $\frac{4x-1000}{3x-1000}=\frac{3}{2}$ or, x = 1000 (on reduction)

   $\therefore$ Income of A = ₹ 4000, that of B = ₹ 3000

10. The prime cost of an article was three times the value of material used. The cost of raw materials was increased in the ratio 3:4 and the productive wage was increased in the ratio 4:5. Find the present price cost of an article, which could formerly be made for ₹ 180. (ICWA. (F) June 2007)

    Sol: prime cost = x+y, where x = productive wage, y = material used

    Now prime cost = 180 = 3y or, y = 60, again x+y =180, x = 180 – y = 180 – 60 = 120

    Present material cost = $\frac{4y}{3}$, present wage $\frac{5x}{4}$

    $\therefore$ present prime cost = $\frac{4\times 60}{3}+\frac{5\times 120}{4}$ = 80 + 150 = ₹ 230.

Practice Problems

1. The ratio of the present age of a father to that of his son is 5:3. Ten years hence the ratio would be 3:2. Find their present ages. (ICWA (F) 84) (Ans. 50,30)
2. The monthly salaries of two persons are in the ratio of 3:5. If each receives an increase of ₹ 20 in salary, the ratio is altered to 13:21. Find the respective salaries. (Ans. ₹ 240, ₹ 400)
3. What must be subtracted from each of the numbers 17, 25, 31, 47 so that the remainders may be in proportion. (Ans. 3)
4. If $\frac{x}{b+c}=\frac{y}{c+a}=\frac{z}{a+b}$ show that (b – c) x + (c – a) y + (a – b) z = 0
5. If $\frac{4x-3z}{4c}=\frac{4z-3y}{3b}=\frac{4y-3x}{2a}$ show that each ratio $=\frac{x+y+z}{2a+3b+4c}$
6. If $\frac{x}{y+z}=\frac{y}{z+x}=\frac{z}{x+y}=k$ prove that $k=\frac{1}{2}$, if $(x + y + z) \neq 0$
7. If $\frac{\sqrt{a}-\sqrt{b}}{\sqrt{a}+\sqrt{b}}=\frac{1}{2}$, prove that $\frac{a^2+ab+b^2}{a^2-ab+b^2}=\frac{91}{73}$ (ICWA (F) june 2001)
8. If = $\frac{a}{4}=\frac{b}{5}=\frac{c}{9}$, prove that $\frac{a+b+c}{9}=2$ (ICWA Dec. 2000)

   $\left(\text{Hint}: \frac{a}{4}=\frac{b}{5}=\frac{c}{9}=K; a=4k, c=9k \ \& \text{ etc}\right)$
9. (i) if $\frac{b+c}{a}=\frac{c+a}{b}=\frac{a+b}{c}$ and a+b+c ≠ 0 then show that each of these ratios is equal to 2.
   Also prove that $a^2 + b^2 + c^2 = ab + bc + ca$. (ICWA (prel.) Dec. 90)

   (ii) if a:b = c:d show that xa + yb: : $a\alpha - b\beta = xc + yd : c\alpha - d\beta$
10. Given $\frac{\alpha}{q-r}=\frac{\beta}{r-p}=\frac{\gamma}{p-q}$, prove that $\alpha+\beta+\gamma=0,\ p\alpha+q\beta+r\gamma=0$ (ICWA, July, 62)

11. Monthly incomes of two persons Ram and Rahim are in the ratio 5:7 and their monthly expenditures are in the ratio 7:11. If each of them saves ₹ 60 per month. Find their monthly income.

(ICWA (F) 2003) (Ans. 200, ₹ 280)

12. There has been increment in the wages of labourers in a factory in the ratio of 22:25, but there has also been a reduction in the number of labourers in the ratio of 15:11. Find out in what ratio the total wage bill of the factory would be increased or decreased. (Ans. 6:5 decrease)

OBJECTIVE PROBLEMS

I. One (or) two steps problems:

1. Find the value of x when x is an mean proportional between: (i)x-2 and x+6 (ii) 2 and 32

Ans: (i) 3 (ii) ±8 (ICWA (F))

2. If the mean proportional between x and 2 is 4, find x (ICWA (F) June 2007)(Ans. 8)

3. If the two numbers 20 and x+2 are in the ratio of 2:3; find x (ICWA(F) Dec. 2006) (Ans. 28)

4. If $\frac{\sqrt{a} \quad \sqrt{b}}{\sqrt{a} \quad \sqrt{b}} \quad \frac{1}{}$ find $\frac{a}{b}$ (ICWA (F) Dec. 2006) (Ans. 9/1)

5. If $\frac{a+b}{a-b} = 2$ find the value of $\frac{a^2-ab+b^2}{a^2+ab+b^2}$ (Ans. $\frac{7}{13}$) (ICWA (F) Dec. 2005)

6. If $\frac{4x-3y}{4c} = \frac{4z-3y}{3b} = \frac{4y-3x}{2a}$ show that each ratio is equal to $\frac{x+y+z}{2a+bb+4c}$ (ICWA (F) June, 2005)

(Hints: each ratio $\frac{4x-3z+4z-3y+4y-3x}{4c+3b+2a}$ & etc.

7. The ratio of the present age of mother to her daughter is 5:3. Ten years hence the ratio would be 3:2, Find their present ages (ICWA (F) Dec. 2004) (Ans. 50;30 years)

8. If 15 men working 10 days earn ₹ 500. How much will 12 men earn working 14 days ? (Ans. ₹ 560)

9. Fill up the gaps: $\frac{a}{b} = \frac{a^2}{-} = \frac{1}{-} = \frac{-}{1} = \frac{b^2}{-}$ (Ans. ab, b/a, a/b, b^3/a, in order)

10. The ratio of work done by (x-1) men in (x+1) days to that of (x+2) men in (x-1) days is 9:10, find the value of x (Ans. 8) (hints: $\frac{(x-1)(x+1)}{(x+2)(x-1)} = \frac{9}{10}$ & etc.)

II. Multiple choice questions

1. If A:B = 2:3, B:C = 4:5, then A:C =

(a) 6:7 (b) 7:6 (c) 8:15 (d) 15:8

2. The inverse ration of $1\frac{3}{5} : 2\frac{1}{4}$ is

(a) 32:45 (b) 45:32 (c) 18:5 (d) 5:18

3. The ratio of 10 meters to ₹ 15

(a) The ratio can not be determined (b) 2:3 (c) 3:2 (d) 5:10

4. If twice of money of A = 5 times of money of B, then the ratio of money of A to that of B

   (a) 2:5 (b)15:25 (c)12:30 (d) 5:2

5. The ratio $\frac{5}{3}:2\frac{1}{4}$ is

   (a) ration of lesser in equality (b) ratio of greater inequality (c) 20:9 (d) 5:27

6. The ratio of present age of jadu to that of madhu is 4:5. If the present age of madhu is 30 years, then the present age of jadu is:

   (a) 20 years (b) 25 years (c) 24 years (d) 35 year

7. The ratio of 5 kg 55 gm to 35 kg 50 gm:

   (a) 5:7 (b) 1,011:7,010 (c) 111:710 (d) none of these

8. The ration 1 year 6 month : 2 years : 2 year 6 months

   (a) 3:4:5 (b) 2:3:5 (c) 2:4:5 (d) none of these

9. If $\frac{1}{2}$ of money of A = $\frac{1}{3}$rd money of B = $\frac{1}{4}$ of money of C, then the continued ratio of money of A, B and C

   (a) 2:3:4 (b) 6:4:3 (c) 4:3:2 (d) 3:2:1

10. Some money is distributed between A and B in the ratio 2:3. If A receives ₹ 72, then B receives

    (a) ₹ 90 (b) ₹ 144 (c) ₹ 108 (d) none of these

III. Fill in the blanks

11. ₹2530 is distributed between Ram and Hari such that Ram gets $\frac{11}{12}$ part that Hari gets. Then hari gets_____
12. Some amount of money is distributed amount Ram, Mitra and shipra such that twice the money that Rama gets = thrice the amount of money that Mitra gets = four times the amount of money that shipra gets. Then the continued ratio of their money is ______________
13. In a map 2cm denotes a distance of 3 km., then the seale in the map is_________
14. The ratio of two numbers is 2:3. If 6 is subtracted from the second number then the number which is subtracted from the first number so that the new ratio becomes the same as that of the previous, is ______
15. The sub-duplicate ratio of 49:81 is ________
16. $\left(\frac{1}{2}+\frac{1}{3}\right):\left(\frac{1}{2}\times\frac{1}{3}\right):$ ______
17. The compound ratio of 1.2:2.5, 2.1:3.2 and 5:3 is : ________
18. If A:B = 3:4, B:C = 2:5, then A:B:C: __________
19. Two numbers are in the ratio is 5:8 and if 6 be subtracted from each of them then the reminders are in the ratio 1:2, then the numbers are______
20. If3,x,27 are in continued proportion then x=_________

IV. State whether the following statements are true or false

21. If 3x+4y:5x-3y=5:3 then x:y = 27.16 ( )
22. The ratio of two numbers is 12:5. If antecedent is 45 then the consequent is 108 ( )

23. If the ration of two positive numbers is 4:5 and their L.C.M is 140 then the number are 35,45 ( )

24. The compound ratio of sub- duplicate ratio and sub-triplicate ratio of 729:64 is 81:8 ( )

25. The ratio of two numbers is 11:15. The sum of 3 times the first number and twice the second numbers is 630. The H.C.F of the numbers is 10 ( )

26. The mean proportional of 4x and $16x^3$ is $12x^2$ ( )

27. The third proportional of 1 hr 20 minutes, 1hr 40 minutes is 2 hrs. ( )

28. The fourth proportional of ₹ 5, ₹ 3.50, 150 gm is 125 gm ( )

29. If A:B = B:C = C:D = 5:6 then A:B:C:D = 125:150:180:216 ( )

30. If the first and third numbers of four positive number is continued proportion be 3 and 12 respectively then fourth number is 36 ( )

V. (a) Match the following

	Group A	Group B	
31.	If 15% of x = 20% of y then x:y	2.1	( )
32.	If 7:x = 17.5:22.5 then x =	4:9	( )
33.	If o.4:1.4 = 1.4: x then x = (–)	6:4:3	( )
34.	The compounded ratio of 2:3, 6:11 and 11:2	4:3	( )
35.	If 2A = 3B = 4C then A:B:C: __	9	( )

(b) Match the following

	Group A	Group B	
36.	The third proportional to (x^2-y^2) and (x-y) __	3:4:5	( )
37.	A fraction which bears the some ratio to $\frac{1}{27}$ that $\frac{3}{11}$ does $\frac{5}{9}$ is	$\frac{x-y}{x+y}$	( )
38.	If $\frac{a}{5}=\frac{b}{4}=\frac{c}{7}$ then $\frac{a+b+c}{c}$ = __	5:1	( )
39.	If $\frac{1}{3}$ A = $\frac{1}{4}$ B = $\frac{1}{5}$ C then A:B:C is	$\frac{1}{55}$	( )
40.	The ratio of third proportional to 12 and 30 and the Mean proportional of 9 and 25 is _	2	( )

Answers

II. Multiple choice questions

1. c 2. B 3. A 4. D 5. A 6. C 7. B 8. A 9. A 10. C

III. Fill in the blanks

11. ₹ 1320 12. 6:4:3 13. 1:1,50,000 14. 4 15. 7:9 16. 5:1

17. 21:40 18. 3:4:10 19. 15,24 20. ±9

IV. State whether the following statement are true (or) false

21. T 22. F 23. F 24. F 25. T 26. F

27. F 28. F 29. T 30. F

V. Match the following

(a)	31. (34)	32. (35)	33. (32)	34. (31)	35. (33)
(b)	36. (39)	37. (36)	38. (40)	39. (37)	40. (38)

VARIATIONS:

Direct variations:

A varies directly as B. that can be written as $A \alpha B$.

$A = KB \quad K = \frac{A}{B}$

Example: The area of a circle is directly proportional to radius of circle (i.e. if radius increases the area of a circle increases in the same ratio as radius) i.e.

$A \alpha B$(radius)

$A = Kr \quad K = \frac{A}{r}$

K is nothing but constant of proportional

Inverse variation:

A is said to be varies inversely as B. if A varies directly as the reciprocal of B.

i.e. $A \propto \frac{1}{B}$

$\therefore A = K\left(\frac{1}{B}\right)$

K = AB

Example: Speed is inversely proportional to time (t) i.e.

$S \propto \frac{1}{t}$

$\therefore \ S = K\left(\frac{1}{t}\right)$

K = st

Joint variation:

A is said to varies jointly as B, C, D…. if A varies directly as the product of B, C, D …. i.e

$A \propto (B, C, D \ldots \ldots \ldots)$

$\therefore \ A = K(B, C, D \ldots \ldots \ldots)$

Example: The volume of cuboid varies directly as the product of length (l), breadth (b), height(h)

i.e. $\therefore \ A \propto lbh$

V = K(lbh)

$K = \frac{v}{lbh}$

Direct Variation:

If two variable quantities A and B be so related that as A changes B also changes in the same ratio, then A is said to vary directly as (or simply vary) as B. This is symbolically denoted as $A \propto B$ (read as A varies as B)

Example:

The circumference of a circle = $2\pi r$, so circumference of a circle varies directly as the radius, for if the radius increase (or decrease), circumference also increases or decreases.

From the above definition, it follows that:

If A varies as B, then A = KB, where K is constant ($\neq 0$)

cor: $A \propto B$ then $B \propto A$. if $A \propto B$, then A = kB. or $B = \frac{A}{k}$ i.e., $B \propto A$.

Some Elementary Results:

(i) If $A \propto B$, then $B \propto A$

(ii) If $A \propto B$ and $B \propto C$, then $A \propto C$

(iii) If $A \propto B$ and $B \propto C$, then $A\text{-}B \propto C$

(iv) If $A \propto C$ and $B \propto C$, then $A\text{-}B \propto C$

(v) If $A \propto C$ and $B \propto C$, then $\sqrt{AB} \propto C$

(vi) If $A \propto B$, then $A^n \propto B^n$.

(vii) If $A \propto B$ and $C \propto D$, then $AC \propto BD$ and $\frac{A}{C} = \frac{B}{D}$

(viii) If $A \propto BC$, then $B \propto \frac{A}{C}$ and $C \propto \frac{A}{B}$

Solved examples:

1. If $a + b \propto a - b$, prove that

   (i) $a \propto b$ (ii) $a^2 + b^2 \propto a^2 - b^2$ (iii) $a^2 + b^2 \propto ab$

   Sol: (i) Since $a + b \propto a\text{-}b$, then a + b = k (a-b), k is constant of variation. Or, (k-1) a = (k+1) b

   or, $\frac{a}{b} = \frac{k+1}{k-1}$ or, a = $\frac{k+1}{k-1}$ b = m b. where m = $\frac{k+1}{k-1}$, a constant. Hence $a \propto b$.

   (ii) $\frac{a^2+b^2}{a^2-b^2} = \frac{m^2b^2+b^2}{m^2b^2-b^2} = \frac{m^2+1}{m^2-1}$ = a constant, $\therefore$ $a^2+b^2 \propto a^2-b^2$

   (iii) $\frac{a^2+b^2}{ab} = \frac{m^2b^2+b^2}{mb^2} = \frac{(m^2+1)\,b^2}{mb^2} = \frac{m^2+1}{m}$ a constant, $\therefore$ $a^2+b^2 \propto ab$

2. If $x+y \propto x\text{-}y$, prove that

   $Ax + by \propto px + qy$ where a, b, p, q are constants

   Sol: As $x+y \propto x\text{-}y$, so $\frac{x+y}{x-y}$ = k (a constant)

   Or, x+y = k(x-y) or, (k-1) x = (k+1)y or, $\frac{x}{y} = \frac{k+1}{k-1}$ y or, x = $\frac{k+1}{k-1}$y or x = my

   Now $\frac{ax+by}{px+py} = \frac{amy+by}{pmy+qy} = \frac{am+b}{pm+q}$ = a constant $\therefore$ $ax+by \propto px+qy$

3. If $x \propto y$, prove that $px + qy \propto ax + by$ where p, q, a, b are constants.

   As $x \propto y$, so x = ky

   Now $\frac{px+qy}{ax+by} = \frac{pky+qy}{aky+by} = \frac{pk+q}{ak+b}$ = k constant. Hence $px + qy \propto ax + by$

4. If $x^2 + y^2$ varies as $x^2 - y^2$, then prove that x varies as y.

   As $x^2 + y^2$ $x^2 - y^2$, so $x^2 + y^2 = k(x^2 - y^2)$, k is constant

   or, $(1-k)\, x^2 = (-k^2-1)y^2$

   or, $\frac{x^2}{y^2} = \frac{-k^2-1}{1-k}$ = a constant = m^2 (say)

   or, $x^2 = m^2y^2$ or, $x = my$ or, $x \propto y$.

5. If $(a+b) \propto (a-b)$ show that $(a^2+b^2) \propto ab$

   As $(a+b) \propto (a-b)$, so $a+b = k(a-b)$

   $(1-k)\, a = (-k-1)b$ or, $\frac{a}{b} = \frac{-k-1}{1-k}$ = m say or, a = mb

   Now $\frac{a^2+b^2}{ab} = \frac{m^2b^2+b^2}{mb^2} = \frac{(m^2+1)\,b^2}{mb^2} = \frac{m^2+1}{m}$ a constant, $\therefore\; a^2+b^2 \propto ab$

6. If the cost price of 12kg. of rice is ₹ 10, what will be the cost of 15 kg. of rice?

   Let A (= cost) = ₹ 10, B (= quantity) = 12 kg. Now $A \alpha B$ i.e., A = KB or, 10 = K. 12 or, K = $\frac{10}{12}$

   Now, we are to find A, when B = 15 kg.

   Again from A = KB, we have A = $\frac{10}{12}$. 15 = ₹12.50

7. A man can finish a piece of work, working 8 hours a day in 5 days. If he works now 10 hours daily, in how many days can he finish the same work?

   Let A (= days) = 5, B (= hours) = 8, it is clear that $A \propto \frac{1}{B}$

   i.e. A = K. $\frac{1}{B}$ or 5 = k. $\frac{1}{8}$ or k = 40

   to find A when B = 10, we have A= $40\frac{1}{10}$ = 4 days

One or two steps questions

1. If A varies inversely with B and if B = 3 then A = 8, then find B if A = 2
2. A is proportional to the square of B. If A = 3 then B = 16; find B if A = 5
3. A varies inversely with B and if B = 3 then A = 7. Find A if B = $2\frac{1}{3}$
4. If $x \propto y$ and x = 3, when y = 24, then find the value of y when x = 8.
5. A varies inversely with B and B = 10 when A = 2, find A when B = 4
6. If $y \propto \frac{1}{x^2}$ and x = 2 when y = 9, find y when x = 3
7. If A $A \propto B$, A=7 when B = 21. Find the relative equation between A and B.
8. If x varies inversely with y, s = 8 when y = 3, find y when x = 2
9. If p $A \propto q^2$ and the value of p is 4 when q=2, then find the value of q+1 when the value of p is 9.
10. If a + b $A \propto$ a-b, prove that a $a \propto b$
11. If x varies as y then show that x^2+y^2 varies x^2-y^2
12. If (a+b) varies as (a-b), prove that a^2+b^2 varies as b
13. If a+2b varies as a-2b, prove that a varies as b
14. X and y are two variable such that $x \propto y$. Obtain a relation between x and y if x = 20. Y = 4.

1.2 SIMPLE AND COMPOUND INTEREST

Simple Interest

Interest:

Interest is the additional money which is paid by the borrower to the lender on the principal borrowed. The additional money (or) interest is paid for the use of money by the borrower. Interest is usually denoted by I.

For example:

Y borrowed ₹ 500 from Z for a year and returned ₹ 550. Here ₹ 50 is paid additionally. This ₹ 50 is the interest.

Rate of interest per annum:

Rate of interest per annum is the interest paid yearly for every ₹ 100 It is denoted by $\frac{R}{100}$ (or) $\frac{r}{100}$.

Amount:

The sum of principal and interest paid is called on amount. It is denoted by A.

Simple Interest:

If the interest is calculated uniformly on the original principal through out the loan period it is called as simple interest. It is denoted by Simple Interest.

Formula:

Simple interest n the principal 'P' borrowed at the rate of r% p.a for a period of "t" years in usually given by S.I = $\frac{Prt}{100}$

For example:

Gopi borrowed ₹ 1200 from siva reddy at 9% interest p.a for 3 years.

Sol: P = 1200, r = 9%, t = 3

S.I = $\frac{Prt}{100} = 1200\times\frac{9}{100}\times 3 =$ ₹ 324

Important Relations to remembers:

1. S.I = $\frac{Prt}{100}$
2. A = P+S.I
3. r = $\frac{S.I\times 100}{Pt}$
4. t = $\frac{S.I\times 100}{Pr}$
5. P = $\frac{S.I\times 100}{rt}$
6. P = A-S.I
7. S.I = A-P
8. A = P $\left(1+\frac{rt}{100}\right)$

Illustration:

1. Amit deposited ₹ 1200 to a bank at 9% interest p.a. find the total interest that he will get at the end of 3 years.

   Here p = 1200, I = $\frac{9}{100}$ = 0.09, n = 3, I = ?

   L = p. i.n = 1200 x 0.09 x 3 = 324.

   Amit will get ₹ 324 as interest.

2. Sumit borrowed ₹ 7500 at 14.5% p.a for $2\frac{1}{2}$ year. Find the amount he had to pay after that period

   P = 7500, i = $\frac{14.5}{100}$ = 0.145, n = $2\frac{1}{2}$ = 2.5A = ?

   A = P(1+in) = 7500 (1+ 0.145 x 2.5) = 7500 (1+ 0.3625)

   = 7500 x 1.3625 = 10218.75

   Reqd. amount = ₹ 10218.75

3. What sum of money will yield ₹ 1407 as interest in $1\frac{1}{2}$ year at 14% p.a simple interest.

   Here S.I = 1407, n = 1.5, I = 0.14, P = ?

   S.I = P. i.n or,1407 = p x 0.14 x1.5

   Or, p = $\frac{1407}{1.14 \times 1.5} = \frac{1407}{0.21}$ = 6700

   Required amount = ₹ 6700

4. A sum of ₹ 1200 was lent out for 2 years at S.I. The lender got ₹ 1536 in all. Find the rate of interest p.a.

   P = 1200, a = 1536, n = 2, i = ?

   A = P(1+ni) or, 1536=1200 (1+2I) = 1200+ 2400 1

   Or, 2400i = 1536 -1200 = 336 or, i = $\frac{336}{2400}$ = 0.14

   Required rate = 0.14x100 =14%

5. At what rate percent will a sum, become double of itself in $5\frac{1}{2}$ years at simple interest?

   A = 2P, P = principal, n = $5\frac{1}{2}$, i = ?

   A = P(1+ni) or, 2P = P$\left(1+\frac{11}{2}i\right)$

   or, 2 = 1+$\frac{11}{2}$ i or, i = $\frac{2}{11}$

   or, r = $\frac{2}{11}$ × 100 = 18.18 (approx); required. rate = 18.18%

6. In a certain time ₹1200 becomes ₹1560 at 10% p.a simple interest. Find the principal that will become ₹ 2232 at 8% p.a in the same time.

   In 1st case: P = 1200, A = 1560, i = 0.10, n = ?

   1560 = 1200 (1+n(.10)) = 1200+120n

   Or, 120n = 360 or, n = 3

   In 2nd case: A = 2232, n = 3, i = 0.08, p = ?

$2232 = P\ (1+3x0.08) = P(1+0.24) = 1.24\ P$

Or, $P = \frac{2232}{1.24} = 1800$

7. A Sum of money amount to ₹ 2600 in 3 years and to ₹ 2900 in $4\frac{1}{2}$ years at simple interest. Find the sum and rate of interest.

Amount in $4\frac{1}{2}$ years = 2900

Amount in 3 years= 2600

S.I for $1\frac{1}{2}$ yrs. = 300

S.I for 1 year $= \frac{300}{1\frac{1}{2}} = 300 \times \frac{2}{3} = 200$

and S.I for 3 years = 3 x 200= 600

principal = 2600 - 600 = ₹ 2000

P = 2000, A = 2600, n = 3, i = ?

2600 = 200 (1+3 i) = 2000 + 6000 i

Or, 6000 i = 600 or $i = \frac{600}{6000} = \frac{1}{10}$ or, $r = \frac{100}{10} = 10$

required rate = 10%

Alternatively. 2600 = P(1+3 i)....... (i), 2900 = p(1+4.5 i) .... (ii)

Dividing (ii) by (1), $\frac{2900}{2600} = \frac{p(1+4.5\,i)}{p(i+3\,i)} = \frac{1+4.5\,i}{1+3\,i}$

Or, $\frac{29}{26} = \frac{1+4.5\,i}{1+3\,i}$ or, i = 0.10 (no reduction)

Or, r = 0.10 x 100 = 10%

Form (i) 2600 = P (1+3 x 0.10) = P(1+0.30) = p(1.30)

$\therefore P = \frac{2600}{1.30} = 2000$

8. Divide ₹ 2760 in two parts such that simple interest on one part at 12.5% p.a for 2 years is equal to the simple interest on the other part at 12.5% p.a for 3 years.

Investment in 1st case = ₹ X (say)

Investment in 2nd case = ₹ (2760-x)

Interest in 1st case $= x \times \frac{10}{100} \times 2 = \frac{x}{5}$

Interest 2nd case $= (2760 - x) \times \frac{12.5}{100} \times 3 = 1035 - \frac{3x}{8}$

By question, $\frac{x}{5} = 1035 - \frac{3x}{8}$ or $\frac{x}{5} + \frac{3x}{8} = 1035$

Or, $\frac{8x + 15x}{40} = 1035$ or, $\frac{23x}{40} = 1035$ or, x = 1800

$\therefore$ investment in 2nd case = ₹ (2760 – 1800) = ₹ 960

9. A person borrowed ₹ 8,000 at a certain rate of interest for 2 years and then ₹ 10,000 at 1% lower than the first. In all he paid ₹ 2500 as interest in 3 years. Find the two rates at which he borrowed the amount.

   Let the rate of interest = r, so that in the 2nd case, rate of interest will be (r-1). Now $800 \times \frac{r}{100} \times 2 + 10{,}000 \times \frac{(r-1)}{100} \times 1 = 2500$

   Or 160r+ 100 r-100 = 2500 or, r=10

   In 1st case rate of interest = 10% and in 2nd case rate of interest = (10-1)= 9%

   Calculate of interest on deposits in a bank: Bank allow interest at a fixed rate on deposits from a fixed day of each month up to last day of the month. Again interest may also be calculated by days.

Practice Question

1. How much interest will be earned on ₹ 2,000 at 6 % simple interest for 2 years?
2. X deposited ₹ 50,000 in a bank for two years with the interest rate of 5.5% p.a. How much interest would she earns/ What would be the final value of the deposit?
3. Rahul deposited ₹ 1,00,000 in his bank for 2 years at simple interest rate of 6%. How much interest would he earns? How much would be the final value of deposit?
4. Find the rate of interest if the amount owed after 6 months is ₹1,050, borrowed amount being ₹1,000.
5. Rahul invested ₹ 70,000 in a bank at the rate of 6.5% p.a S.I. He received ₹ 85,925 after the end of term. Find out the period for which sum was invested by rahul.
6. Kapil deposited some amount in a bank for 71/2 years at the rate of 6% p.a. simple interest. Kapil received ₹ 1,01,500 at the end of term. Compute initial deposit of Kapil.
7. A sum of ₹ 46,875 was lent out at simple interest and at the end of 1 year 8 months the total amount was ₹ 50,000. Find the rate of interest percent per annum.
8. What sum of money will produce ₹ 28,600 as an interest in 3 years and 3 months at 2.5% p.a simple interest?
9. In what time will ₹ 85,000 amount to ₹ 1,57,675 at 4.5% p.a?

I. Choose the correct Answer

1. At what rate of Simple Interest will ₹ 1000 amount to ₹ 1200 in 2 years?

   (a) 8% (b) 10% (c) 9% (d) 7½%

2. In what time will ₹ 2000 amount to ₹ 2600 at 5% S.I ?

   (a) 10 (b) 9 (c) 7 (d) 6 yrs

3. At what rate p.a S.I, will a sum of money double itself in 25 yrs?

   (a) 4% (b) 3% (c) 5% (d) 6%

4. What sum of money will amount to ₹ 5200 in 6 yrs at the same rate of interest at which ₹ 1706 amount to ₹ 3412 in 20 yrs.

   (a) 3900 (b) 4000 (c) 4400 (d) 3800

5. A sum of money becomes double in 20 yrs at S.I. in how many years will it be triple ___

   (a) 40 (b) 35 (c) 38 (d) 42

II. Fill in the blanks

6. A certain sum of money at S.I amounts to ₹ 500 in 3 years and to ₹ 600 in 5 yrs then the principal is____
7. In ___ time will be the Simple interest on ₹ 900 at 6% be equal to S.I on ₹ 540 for 8 yrs at 5%
8. Due to fall in the rate of interest from 12% to 10½% p.a money lender's yearly income diminished by ₹ 90 the capital _____
9. A sum was put at S.I at a certain rate for 2 years. Had it been put at 2% higher rate, it would have fetched ₹ 100 more the sum is _____
10. At ____ rate percent will be Simple interest be equal to principal amount in 10 yrs.

III. State whether the following statement are true (or) false

11. A sum of money amounts to ₹ 720 in 2 years and ₹ 783 in 3 years the rate of interest is 12% ( )
12. The SA.I at x% for x years will be ₹ X on a sum of x ( )
13. The S.I on a sum of money at 8% p.a for 6years is half the sum. The sum is ₹ 640 ( )
14. If the interest on ₹ 1200 more than the interest on ₹ 1000 by ₹ 50 in 3years the rate of interest is $8^1/_3$% ( )
15. The rate of S.I p.a a sum of money grows to one and a half times itself in 8 yrs is 6½% ( )

IV. Answer the following in one (or) two steps

16. If the amount of ₹ 500 in three years is ₹ 575, what will be the amount of ₹ 700 in 5 years?
17. A certain sum of money becomes there times in 12 yrs by S.I. The rate of interest?

KEY

I. 1. b 2. D 3. A 4. B 5.a

II. Fill in the blanks

6. ₹ 500 7. 4 8. ₹ 6000 9. ₹ 11000 10. 10%

III. True or false

11. T 12. F 13. T 14. T 15. T

IV. 16. 875 17. $16^2/_3$%

Compound Interest:

If the interest is added after certain period of time (say a yearly, half yearly, quarterly and monthly etc) to the principal so that the amount at the end of the period becomes the principal for the next time period then the total interest paid over all the time period is called the compound interest. It is usually denoted by C.I.

Formula:

Compound Interest = Amount – principal

When amount is calculated by the following formula $A = P\left(1+\frac{r}{100}\right)^n$

Where P = principal

r = rate of compound interest

n = no. of payment period $\quad\quad$ $i = \frac{r}{100}$ = rate of interest of ₹ 1 per year

Compound can also be calculated directly

Compound Interest = $P\left[\left(1+\frac{r}{100}\right)^n - 1\right]$

Short cut methods to solve special type of problems:

Type – I:- When the interest is compounded half-yearly then amount is given by $A = P\left(1+\frac{r}{200}\right)^{2n}$

Type – II :- When the interest is compounded quarterly then amount is given by $A = P\left(1+\frac{r}{400}\right)^{4n}$

Type – III:- When the rate of interest for 1st, 2nd and 3rd years are $r_1\%$, $r_2\%$ and $r_3\%$ respectively and vice versa then amount is given by

$$A = P\left(1+\frac{r_1}{100}\right)\left(1+\frac{r_2}{100}\right)\left(1+\frac{r_3}{100}\right)\ldots\ldots\ldots\ldots\ldots$$

Type – IV:-If a principal becomes P times in 'x' years the rate of compound interest is given by

$r = 100\,(P^{1/x} - 1)$.

Type – V:- if the time is given in mixed fraction say $x\frac{k}{x}$ years and interest is compounded annually then amount is given by

$$A = P\left(1+\frac{i}{100}\right)^n \left(1+\frac{\frac{k}{m}R}{100}\right)$$

Effective rate of interest :

When an amount is compounded once yearly at a certain rate of interest the interest so calculated is termed as nominal interest. However when the same amount is compounded more than once a year (say monthly, quarterly, half yearly) then the rate of interest actually exceeds the nominal interest rate p.a. This exceeded rate of interest is called as effective rate of interest and is denoted by r_e and is defined as $r_e = 100\left(1+\frac{i}{m}\right)^m - 1$

Where i = rate of compounded interest

m = no. of terms the interest is compounded in a year

Illustrations

1. Find the compound interest on ₹ 1,000 for 4 years at 5% p.a

   Here p = ₹ 1000, n = 4, i = 0.05, A = ?

   We have $A = P(1+n)^n$

   $A = 1000\ (1+ 0.05)^4$

   Or log A = log 1000 + 4 log (1+ 0.05) = 3+ 4 log (1.05) = 3+ 4 (0.0212) = 3+ 0.0848 = 3.0848

   A = antilog 3.0848 = 1215

   C.I = ₹ 1215 – ₹ 1000 = ₹ 215

2. In what time will a sum of money double itself at 5% p.a C.I.

   Here, P = P, A = 2p, i = 0.05, n = ?

   $A = P(i+i)^n$ or, $2p = p(1+0.05)^n = P(1.05)^n$

   Or, $2 = (1.05)^n$ or log 2 = nlog 1.05

   $\therefore n = \frac{\log\ 2}{\log\ 1.05} = \frac{0.3010}{0.0212} = 14.2$ years (Approx)

   ∴ (anti-logarithm table is not required for finding time).

3. The difference between simple and compound interest on a sum put out for 5 years at 3% was ₹ 46.80. Find the sum.

   Let P = 100, i = .03, n = 5, from $A = P(1+n)^n$

   $A = 100\ (1+.03)^5 = 100\ (1.03)^5$

   Log A = log 100 + 5 log (1.03) = 2 + .0640 = 2.0640

   ∴ A = anti log 2.0640 =115.9

   C. I = 115.9 – 100 = 15.9

   Again S. I = 3×5 =15

   ∴ difference 15.9 – 15 = 0.9

Diff.	capital	$x = 100 \times \frac{46.80}{0.9} = 5,200$
0.9	100	
46.80	x	

   ∴ original sum = ₹ 5,200

4. What is the present value of ₹1,000 due in 2 years at 5% compound interest, according as the interest it paid (a) yearly (b) half-yearly?

   (a) Here A = ₹ 1,000, $i = \frac{5}{100} = 0.05$, n = 2, p = ?

   $A = P(1+n)^n$ or $1000 = P(1+0.5)^n = P(1.05)^2$

   $\therefore P = \frac{1000}{(1.05)^2} = \frac{1000}{1.1025} = 907.03$

   ∴ present value = ₹ 907.03

(b) Interest per unit per half-yearly $\frac{1}{2} \times 0.05 = 0.025$

From $A = p\left(1+\frac{i}{2}\right)^{2n}$ we find

$1{,}000 = p\left(1+\frac{0.05}{2}\right)^{2\times2} = p\ (1+.025)^4 = P\ (1.025)^4$

Or, $P = \frac{1000}{(1.025)^4}$

$\therefore$ log p = log 1000 – 4 log (1.025) = 3 – 4 (0.0107) = 3 – 0.0428 = 2.9572

$\therefore$ P = antilog 2.9572 = 906.1

Hence the present amount = ₹ 906.10

5. A sum of money invested at C.I. payable yearly amounts to ₹ 10,816 at the end of the second year and to ₹ 11,248.64 at the end of the third year. Find the rate of interest and the sum,

Here A_1 = 10,816, n = 2, and A_2 = 11,248.64, n = 3

From $A = P(1+i)^n$ we get,

$10{,}816 = P(1+1)^2$ ….. (i) and $11{,}248.64 = P(1+1)^3$……… (ii)

Dividing (ii) by (i) $\frac{11{,}248.64}{10{,}816} = \frac{P(1+i)^3}{P(1+i)^3}$ or, $(1+1) = \frac{11{,}248.64}{10{,}816}$

or $i = \frac{11{,}248.64}{10{,}816} - 1 = \frac{432.64}{10{,}816}$ = .04, r = i × 100 = .04 × 100 = 4 $\therefore$ required rate = 4%

Now from (i) $p = \frac{10{,}816}{(1+.04)^2} = \frac{10{,}816}{(1.04)^2}$

Log P = log 10,816 – 2log (1.04) = 4.034 – 2(0.170) = 4.034 – .040 = 4.00

$\therefore$ P = antilog 4.000 = 10,000 $\therefore$ required sum = ₹ 10,000

Practices Question of Compound Interest

1. On what sum will the compound interest at 5% per annum for two years compounded annually be ₹ 1,640?
2. What annual rate of interest compounded annually doubles an investment in 7 years Given that 2 1/7 = 1.104090.
3. When will ₹ 8,000 amount to ₹ 8,820 at 10% per annum interest compounded half –yearly?
4. Find the rate percent per annum if ₹ 2,000 amount t0 ₹ 2,31,525 in 1 ½ year interest being compounded half – yearly.
5. A certain sum invested at 4% per annum compounded semi- annually amounts to ₹ 7830 at the end of one year. Find the sum.
6. 1600 invested at 105 p.a compounded semi- annually amounts to ₹ 18,522. Find the time period of investment.
7. A person opened an account on Apirl, 2011 with a deposit of ₹ 800. The account paid 6% interest compounded quarterly. On October 1, 2011 he closed the account and added enough additional money to invest in a 6 months time – deposit for ₹ 1000. Earning 6% p.a compounded monthly.

   (a) How much additional amount did the person invest on October, 1?

(b) What will be the maturity value of his deposit on April 1 2012?

(c) How much total interest is earned?

Given that $(1+i)^n$ is 1.03022500 for i=1 ½%, n= 2 and $(1+i)^n$ is 1.03037751 for i= ½ % and n=6.

8. Find the amount of compound interest and effective rate of interest if an amount of ₹ 20,000 is deposited in a bank for one year at the rate of 8% per annum compounded semi annually.

9. Which is a better investment 3% per year compounded monthly or 3.2% per year simple interest? Given that $(1+ 0.0025)^{12}= 1.0304$.

I Choose the correct Answer

1. The compound interest on ₹ 500 for 2 years at 10% p.a

   (a) 100 (b) 105 (c) 110 (d) 120

2. Compound interest on ₹ 1000 at 8% p.a compound interest half yearly for 2 years.

   (a) 169.90 (b) 196.60 (c) 175.10 (d) 199.40

3. An amount of money lent at C.I amounts to ₹ 1323 is 2 years at 5% C.I. The sum is ____

   (a) 1100 (b) 1300 (c) 1000 (d) 1200

4. What will ₹ 6250 amount to at C.I for 3years at 4% when C.I is reckoned yearly?

   (a) 7020.20 (b) 7030.40 (c) 7100 (d) 6990.90

5. Compute C.I on ₹ 2500 for 1 year at 12% compounded six monthly

   (a) 309 (b) 390 (c) 300 (d) 290

II. Fill in the blank

6. I lent ₹ 4000 for 9 months at 12% p.a If C.I is reckoned quarterly what will I get after 9 months _____

7. The C.I on ₹ 5000 for 3 years if the rate of interest is 5%, 6% and 7% for the 1st, 2nd and 3rd years respectively is ____

8. The difference between S.I and C.I o ₹ 4000 for 2 ½ years @ 10% p.a is _____

9. The sum of money will amount to ₹ 6050 in 2 years at 10% p.a C.I is _____

10. The C.I on a certain sum of money for 2 years at 8% p.a compounded annually is ₹ 1040. The sum is ____

III. State whether the following statements are true or false

11. The C.I on a certain sum of money for 1year at 8% p.a compounded quarterly is ₹ 824 then the sum is ₹ 10,000 ( )

12. In 3 years ₹ 1600 amount to ₹ 1936 at 10% p.a C.I ( )

13. The C.I on ₹ 5000 for 2 years at a certain rate of interest p.a amounts to ₹ 1050 then the rate of interest is 10% ( )

14. In 3 years the population of a village change form 15625 to 17576 if the rate of interest is 4% p.a ( )

15. The difference between S.I and C.I on ₹ 1000 for 1 years at 4% payable quarterly is 0.40 ( )

IV. Answer the following problems is one (or) two steps

16. In what time will ₹ 2000 lent at 5% p.a fetch ₹ 2050 at C.I?
17. What sum at C.I will amount to ₹ 650 at the end of the year and to ₹ 676 at the end of the second year?
18. The interest on a sum for the first and second years is ₹ 225 and ₹ 238.50 respectively. Find the rate of interest and the sum?

Compound Interest Key

I.	1. b	2. A	3. D	4. B	5. A
II.	6. ₹ 4370.91	7. ₹ 954.55	8. ₹ 82	9.₹ 5000	10. ₹ 6250
III.	11. T	12. F	13. T	14. T	15. F
IV.	16. 2 years	17.₹ 625	18.₹ 3750		

1.3 ARITHMETIC PROGRESSION AND GEOMETRIC PROGRESSION

PROGRESSIONS

Sequence:

Sequence is an arrangement of numbers in a definite order/law.

Ex. 1:

6,11,16…….. is a sequence such that every term is 6^5 more than the preceding term.

Ex. 2:

2,4,6…… is a sequence such that every term is 2 more than the preceding term.

Note:

Generally the terms of a sequence is denoted by t_1, t_2,t_3…………

Series:

The sum of terms of a sequence is called series.

Note:

A sequence has finite numbers is called a finite sequence and it has infinite number of elements is called a infinite sequence.

Arithmetic progression:

If certain quantities are increased or decreased by same constant then the quantities form a series then the series is called arithmetic progression. Simply it can be written as A.P. Here the same constant is called common difference of an arithmetic progression.

Ex. 1: 2+4+6…… is an A.P

Ex. 2: 9+6+3+0-3…… is an A.P

Note:

The general form of an A.P is a+(a+d)+(a+2d)+……….

Important Notations:

The first term of an A.P is denoted by 'a' and common difference is denoted by 'd' and last term is denoted by 'l' and 4th term of an A.P is denoted by 4th and finally the sum of n-terms of an A.P is denoted by S_n.

Find the nth term of an A.P:

The general form of an A.P is (a) + (a+d)+(a+2d) + ………..

Now

T_1 = a = a+(1-1) d

T_2 = a = d = a+(2-1)d.

T_3 = a+2d = a+(3-1)d

…………………………

……………………………

T_n = a+(n-1)d

t_n of an A.P = a+(n-1)d

Find the sum of n terms of an A.P:

We know that, the general form of an A.P is a+(a+d)+(a+2d)+ …… +a+(n-1)d.

Let

S_n = a+(a+d) + (a+2d) + ………..+ a+(n-1)d. .... (1)

S_n = {a+(n-1)d} + {a+(n-2)d} + ………..+ a. .... (2)

On adding the both equations,

$2S_n$ = n[2a+(n-1)d]

S_n= n/2[2a+(n-1)d]

Note:

l = t_n = a+(n-1)d

$S_n = \frac{n}{2}$ [2a+(n-1)d]

$= \frac{n}{2}$ [a+a+(n-1)d]

$= \frac{n}{2}$ [a+l]

A.M of two quantities a and b:

Let

'x' be the A.M of two quantites a and b. So that a, x, b are in A.P

$t_2 - t_1 = t_3 - t_2$

x-a = b-x

2x = a+b

$X = \frac{a+b}{2}$

n arithmetic means between two quantities a and b:

Let

$X_1; x_2; x_3 \ldots\ldots x_n$ be the n arithmetic means between two quantities a and b.

So that a, $x_1, x_2, x_3 \ldots\ldots\ldots x_n$ b are in A.P having (n+2) terms.

t_n of an A.P = a+(n-1)d

$\therefore\ b = t_{n+2}$

b = a+(n+2-1)d

b-a = (n+1)d

$\therefore\ d = \frac{b-a}{n+1}$

$\therefore\ x_1 = a+d = a+\frac{b-a}{n+1} = \frac{an+a+b-a}{n+1} = \frac{an+b}{n+1}$

Similarly

We can find $x_2, x_3, \ldots\ldots\ldots$

Formulae:

1. Sum of first n - natural numbers

   i.e. $1+2+3+\ldots\ldots+n = \frac{n(n+1)}{2}$

2. Sum of first n-odd natural numbers

   i.e. $1+3+5+7+\ldots\ldots + (2n-1)$

   $S_n = \frac{n}{2}\{2\times1+(n-1)2\}$

   $= \frac{n}{2}\{\cancel{2}+2n-\cancel{2}n\}$

   $= n^2$

3. Sum of first n-even natural numbers i.e

   $S_n = 2+4+6+\ldots..+2n$

   $= \sum 2n$

   $= 2.\sum n$

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

   $= n^2+n$

4. Sum of squares of first n-natural numbers

   i.e $S_n = 1^2+2^2+3^3+\ldots\ldots\ldots+n^2$

   $S_n = \frac{n(n+1)(2n+1)}{6}$

5. Sum of cubes of first n-natural numbers

   i.e. $S_n = 1^3+2^3+3^3+\ldots\ldots\ldots+n^3$

$S_n = \frac{n^2 (n+1)^2}{4}$

7. T_n of an A.P when S_n is given

   $t_n = s_n - s_{n-1}$

Illustrations:

1. If the 7th and 11th terms of a A.P are (-39) and 5 respectively then find the 20th term and the sum of first 20 term

Solution:

Let the 1st term and the common difference of the A.P be a and b respectively. Then

t_7 = 7th term = a+(7-1)d = a+6d

t_{11} = 11th term = a+(11-1)d = a+10d

By the given conditions, we get,

a+6d = – 39…………………………(1)

a+10d = 5 ……………….............(2)

Solving (1) and (2), we get, a = –105, d = 11

$\therefore$ t_{20} = 20th term = a+(20-1) d = a+19d = –105+19x11

= –105+209 = 104.

S_{20} = sum of first 20 terms = $\frac{n}{2}\{2a+(n-1)d\}$ = 10 {2x(-105)+19x11}

= 10(-210+209) = 10 ×(-1) = –10.

2. If the first term of an A.P is -5 and the last term is 25; the total number of terms is 10, then find the sum of all terms.

Solution:

Here a = –5, n = 10, t_{10} = 25. Let the common difference be d.

$\therefore$ t_{10} = a+(10-1)d = –5+9d. So -5+9d = 25 or, d = $\frac{10}{3}$.

Sum of all terms = $S_{10} = \frac{10}{2}$ {2a+(10-1)d} = $5\left[2(-5)+9\times\frac{10}{3}\right]$ = 100.

3. If the sum of first n terms of an A.P is n^2, then find its common difference.

Solution:

Let the sum of first n terms b S_n. $\therefore$ $S_n = n^2$. Putting n = 1, 2, 3, 4

We get, $S_1 = 1^2 = 1$, $S_2 = 2^2 = 4$, $S_3 = 3^2 = 9$, $S_4 = 4^2 = 16$

$\therefore$ t_1 = first term = S_1=1

t_2 = second term = S_2-S_1 = 4-1 = 3

t_3 = third term = S_3-S_2=9-4=5

t_4 = fourth term = S_4-S_3 = 16-9 = 7.

$\therefore$ The sequence is 1, 3, 5, 7 ……… So the common difference d = 3-1 = 2.

4. IF K^2-2k+5,3k^2+K+4, 4K^2+2k+11, are in A.P., then find the value of K.

Solution:

$\therefore$ K^2-2k+5,3k^2+K+4, 4K^2+2k+11, are in A.p., then

$(3k^2+k+4)-(k^2-2k+5)=(4k^2+2k+11)-(3k^2+k+4)$

Or, $2k^2+3k-1=k^2+k+7$

Or, $K^2+2k-8=0$ or $k^2+4k-2k-8=0$

Or, $k(k+4)-2(k+4)=0$ or $(k+4)(k-2)=0$

So either k-2=0

Or, K+4=0

$\therefore$ k = 2 or -4

5. If $\frac{1}{q+r}, \frac{1}{r+p}, \frac{1}{p+q}$, are in A.P., then show that p^2, q^2, r^2 are in A.p

Solution:

$\therefore \frac{1}{q+r}, \frac{1}{r+p}, \frac{1}{p+q}, \quad \therefore \frac{1}{r+p} - \frac{1}{q+r} = \frac{1}{p+q} - \frac{1}{r+p}$

Or $\frac{q+r-r-p}{(r+p)(q+r)} = \frac{r+p-q-p}{(p+q)(r+p)}$ or $\frac{q-p}{q+r} = \frac{r-q}{p+q}$ or, $(q-p)(p+q) = (r-q)(q+r)$

Or $q^2-p^2 = r^2-q^2$. This implies that p^2, q^2, r^2 are in A.P

6. If a,b,c are the p-th, q-th and r-th terms of an a.P., then show that $a(q-r)+b(r-p)+c(p-q)=0$.

Solution:

Let the first term and commond difference of the A.P be x and y respectively. Then

a = pth term = x+(p-1)y; b = qth term = x+(q-1)y; c = r-th term = x+(r-1)by

$\therefore$ L.H.S = a(q-r)+b(r-p)+c(p-q)

= {x+(p-1)y} (q-r)+{x+(q-1)y} (r-p)+{x+(r-1)y}(p-q)

= x(q-r+r-p+p-q)+y{(p-1)(q-r)+(q-1)(r-p)+(r-1)(p-q)}

= x.o+y(pq-pr-q+r+qr-pq-r+pr-qr-p+q)

= o+y.o = 0 R.H.S

7. If the sum of four integers in A.P is 48 and their product is 15120, then find the numbers.

Solution:

Let the four integers in A.P be a-3d,a-d,a+d,a+3d.

By the given condition, a-3d+a-d+a+d+a+3d = 48

Or 4a = 48

Or, a = 12.

Again, (a-3d)(a-d)(a+d)(a+3d) = 15120.

$\therefore$ $(a^2-d^2)(a^2-9d^2)$ = 15, 120

Or, $(12^2-d^2)(12^2-9d^2)=15120$

Or, $(144-d^2)\ (144 - 9d^2)=15120$

Or, $(144 - d^2).9(16-d^2) = 1680$

Or, $(144 - d^2)(16-d^2) = 1680$

Or, $2304 -160d^2+624 = 0$

Or, $(d^2 - 4)(d^2 -156) = 0$

Either $d^2 - 4 = 0$

Or, $d^2 -156 = 0$

$\therefore\ d^2 = 4$

Or, $d \pm 2$

If $d^2 =156$ or $d = \pm\sqrt{156}$

Which is not possible because the numbers are not integers.

$\therefore\ d \pm 2$.

So when a = 12, d = 2, then the numbers are

$a - 3d = 12-3.2 = 6$

$a + d = 12-2 = 10$

$a + d = 12+2 = 14$

$a + 3d = 12+3.2 = 18$

When a = 12, and d = -2, then the numbers are

$a - 3d = 12-3-2 =12+6 = 18$

$a - d = 12-(-2) =12+2 = 14$

$a + d = 12+(-2) =10$

$a + 3d = 12+3(-2) =12-6 = 6$

$\therefore$ The numbers are 6,10,14,18 or 18,14,10,6

8. Find the sum to n terms of the series: 1 x 3+3x5+5x7+7x9+………………….

Solution:

Let the required sum be S and the nth term of the series be t_n.

$\therefore\ t_n$ = (n^{th} term of the series 1+3+5+7+……….) × (nth term of the series 3+5+7+9+……..)

$= \{1+(n-1)x2\}\{3+(n-1)x2\} = (1+2n-2)(3+2n-2)$

$= (2n-1)(2n+1) = 4n^2 -1.$

So, $t_n = 4n^2 -1$

Now putting n = 1, 2, 3……n, we get,

$t_1 = 4.1^2 -1$

$t_2 = 4.2^2 -1$

$t_3 = 4.3^2 -1$

$t_n = 4.n^2 - 1$

$S = 4(1^2+2^2+3^2+4^2+n^2)-(1+1+1 \ldots\ldots\ldots\ldots \text{ to n times})$

$$= 4\frac{n(n+1)(2n+1)}{6} - 1 \times n = \frac{2}{3}n(n+1)(2n+1) - n = \frac{2n(n+1)(2n+1) - 3n}{3}$$

$$= \frac{n}{3}\{2(n+1)(2n+1) - 3\} = \frac{n}{3}(4n^2 + 6n + 2 - 3) = \frac{n}{3}(4n^2 + 6n - 1).$$

$\therefore$ The required sum is $n/3\ (4n^2+6n-1)$.

Practice Questions:

1. Find the 7^{th} term of the A.P 8, 5, 2, -1, -4, ...
2. If 5^{th} and 12^{th} terms of an A.P are 14 and 35 respectively find the A.P
3. Which term of the AP $\frac{3}{\sqrt{7}}, \frac{4}{\sqrt{7}}, \frac{5}{\sqrt{7}} \ldots\ldots\ldots\ldots$ is $\frac{17}{\sqrt{7}}$?
4. Divide 69 into three parts which are in A.P and are such that the product of the 1^{st} two parts is 483.
5. Find the arithmetic mean between 4 and 10.
6. Insert 4 arithmetic means between 4 and 324

Geometric Progression:

If certain quantities are multiplied or divided by the same constants then the series is called geometric progression simply it can be written as G.P.

Here

The same constant is called common ratio of G.P and is denoted by ‘r’

Ex. 1: 4, 16, 64……….. are in G.P

With common ratio r = 4.

Ex. 2: 9, 6, 4………. Are in G.P

With common ratio r = 2/3

Note:

The general form of G.P is a, ar, ar^2 ………..

Important notations:

The first term of G.P is denoted by ‘a’ and common ratio is denoted by ‘r’ and n^{th} term is denoted by t_n. and sum of n terms is denoted by ‘s’ and the sum of infinite no. of terms is denoted by $S\ \alpha$

Find the n^{th} term of a G.P :

General form of G.P is a,ar,ar^2……

$t_1 = a = ar^{1-1}$

$t_2 = ar = ar^{2-1}$

$t_3 = ar^2 = ar^{3-1}$

………………………..

…………………………..

$t_n = ar^{n-1}$

$t_n = ar^{n-1}$

Find the sum of n term of a G.P:

Let $S = a + ar + ar^2 + \ldots\ldots\ldots + ar^{n-1}$

(-) $\underline{Sr = ar + ar^2 + \ldots\ldots\ldots\ldots\ldots + ar^{n-1} + ar^n}$

$S(1-r) = a - ar^n$

$= a(1-r^n)$

$$S = \frac{a(1-r^n)}{(1-r)} \text{ if } (r<1)$$

$$S = \frac{a(r^n-1)}{r-1} (r>1)$$

Sum of infinite no. of terms in a G.P = $S_\infty = \frac{a}{1-r}$ if $|r| < 1$ or $-1 < r < 1$

Geometric mean:

If 'g' is the G.M between two quantities a and b then a, g, b are in G.P

$$\Rightarrow \frac{g}{b} = \frac{b}{g}$$

$g^2 = ab$

$g = \sqrt{ab}$

Let n – geometric means between two quantities a and b

$g_1, g_2, g_3, \ldots\ldots\ldots\ldots\ldots g_n$ be n geometric means between two quantities a and b then a, g_1, g_2, g_3, ………….. g_n, b are in G.P having (n+2) terms

b = last term

= $(n+2)^{th}$ term

$= ar^{n+2-1}$

$= ar^{n+1}$

$\frac{b}{a} = r^{n+1}$

$$r = \left(\frac{b}{a}\right)^{1/n+1}$$

Properties of G.P:

1. If all terms of G.P be multiplied or divided by same non-zero number, then it remains G.P with Common ratio.
2. The reciprocals of the term of a sign G.P form a G.P
3. If each term of a G.P be raised to the same power the result series also from a G.P relation between AM $\in$ G.M

   $= A = \frac{a+b}{2}$, $G\sqrt{ab}$ when a&b are any two

= A > G

= The quadratic equation having a, b as it, roots is $x^2 - 2ax + g^2 = 0$

Illustrations:

1. If x+9, x-6, 4 are in G.P., then find x

Solution:

X+9, x-6, 4 are in G.P., $\frac{x-6}{x+9} = \frac{4}{x-6}$ or $(x - 6)^2 = 4\ (x+9)$

Or, x^2-12x+36 = 4x+36

Or x^2-16x = 0

$\therefore$ x = 0, 16.

2. If a,4 b are in A.p and a, 2, b are in G.P., then rove that $\frac{1}{a}+\frac{1}{b} = 2$

$\therefore$ a, 4, b are in A.P., then 4-a = b-4

Or, a+b = 8 .....(1)

Again a, 2, b are in G.P., then $\frac{2}{a} = \frac{b}{2}$

Or, ab = 4 ......(2)

Dividing (1) by (2), we get,

$\frac{a}{ab} + \frac{b}{ab} = \frac{8}{4}$

Or, $\frac{1}{b} + \frac{1}{a} = 2$

Or, $\frac{1}{a} + \frac{1}{b} = 2$

3. The product of 3 consecutive terms in G.P is $\frac{27}{8}$. Find the middle term.

Solution:

Let the 3 consecutive terms in G.P be $\frac{a}{r}$, a, ar

$\therefore \frac{a}{r}$.a. ar = $\frac{27}{8}$

Or, $a^3 = \frac{27}{8}$

Or, $a = \frac{3}{2}$

The middle term = $a = \frac{3}{2}$

4. In a G.P., the sum of first three terms is to the sum of first six terms is equal to 125: 152. Find the common ratio.

Solution:

Let the first term of the G.P be a and the common ratio be r.

$\therefore$ The sum of first 3 terms = $\frac{a\,(r^3-1)}{r-1}$ and the sum of first six terms = $\frac{a\,(r^6-1)}{r-1}$

By the given condition, we get,

$$\frac{a(r^3-1)}{r-1}:\frac{a(r^6-1)}{r-1}=125:152 \; (a \neq 0, r \neq 1)$$

Or, $\frac{r^3-1}{r^6-1}=\frac{125}{152}$

Or, $\frac{r^6-1}{r^3-1}=\frac{152}{125}$

Or, $\frac{(r^3-1)(r^3+1)}{r^3-1}=\frac{152}{125}$

Or, $r^3+1 = \frac{152}{125}$

Or, $r^3 = \frac{27}{125}$

Or, $r = \frac{3}{5}$

5. If 5^{th} and 2^{nd} terms of a G.P are 81 and 24 respectively, then find the series and the sum of first eight terms.

Solution:

Let the first term and the common ratio of the G.P be a and r respectively.

$\therefore$ $t_5 = 5^{th}$ term $= ar^4$ and $t_2 = 2^{nd}$ term = ar

So $ar^4 = 81$ ………..(1) and $ar = 24$ .... (2)

Dividing (2) from (1)

$\frac{ar^4}{ar}=\frac{81}{24}$

Or, $r^3 = \frac{27}{8}$

Or, $r = \frac{3}{2}$

From (2), we get, $a \times \frac{3}{2} = 24$

Or, a = 16.

$\therefore$ The G.P is : 16, $16\times\frac{3}{2}$ = 24, $24\times\frac{3}{2}$ = 36, $36\times\frac{3}{2}$ = 54, ……………..

= 16,24,36,54……

Sum of first 8 terms $=\frac{a(r^8-1)}{r-1}=\frac{16\left\{\left(\left(\frac{3}{2}\right)^8-1\right)\right.}{\frac{3}{2}-1}=\frac{16\left(\frac{6561}{256}-1\right)}{\frac{1}{2}}=\frac{6305}{8}=788\frac{1}{8}$

6. Which term of the sequence 1,3,9,27,………. Is 6561?

Solution:

Let nth term of the given sequence be 6561.

$\therefore$ $t_n = ar^{n-1} = 1.3^{n-1}$ (Here a = 1^{st} term = 1 and the common ratio = r = 3).

$\therefore$ $3^{n-1} = 6561 = 3^8$

Or, n-1= 8

Or, n = 9.

$\therefore$ 9^{th} term of the sequence is 6561.

7. If a,b,p be the first term, nth term and the product of first n terms of a G.P., then prove that $P^2 = (ab)^n$

Solution:

Let the common ratio be r.

$\therefore$ $t_n = n^{th}$ term $= ar^{n-1}$

So $b = ar^{n-1}$................ (1)

The sequence is a, ar, ar^2, ar^3,............

$\therefore$ P = The product of first n terms = $a \times ar \times ar^2 \times ....\times ar^{n-1}$

$= an \times r^{1+2+ .... + n-1}$

$= a^n\, r^{n(n-1)/2}$

$p^2 = \left[a^n r^{\frac{n(n-1)}{2}}\right]^2$

$= a^{2n}\, r^{n(n-1)}$

$= \left(a^2 r^{n-1}\right)^n$

$= \left(a \cdot ar^{n-1}\right)^n$

$= ab^n$

8. The sum of 3 numbers in a G.P is 35 and their product is 1000. Find the numbers.

Solution:

Let the 3 numbers in G.P be $\frac{a}{r}$, a,ar

$\therefore \frac{a}{r} + a + ar = 35$ or $a\left(\frac{1}{r} + 1 + r\right) = 35$ ..... (1)

$\frac{a}{r}$.a. ar = 1000 or, a^3 = 1000 or, $a^3 = 10^3$ or a = 10

From (1), we get, $10\left(\frac{1}{r} + 1 + r\right) = 35$ or, $2\left(\frac{1 + r + r^2}{r}\right) = 7$ or $2 + 2r + 2r^2 = 7r$

Or, $2r^2 - 5r + 2 = 0$

Or, $2r^2 - 4r - r + 2 = 0$

Or, 2r(r-2) -1(r-2) = 0

Or, (r-2) (2r-1) = 0

$\therefore$ $r = \frac{1}{2}$ or, 2

When a = 10, $r = \frac{1}{2}$, the numbers are $\frac{a}{r} = \frac{10}{2} = 5$, a = 10, ar = 10.2 = 20.

$\therefore$ The numbers are either 20,10,5 or 5,10,20.

9. Find the sum to n terms of the series: 0.6+0.66+0.666+0.6666 + ……….

Solution:

Let the required sum be S.

$\therefore$ S = 0.6 + 0.66 + 0.666 + 0.6666 +……. To n terms

= 6 (.1+.11+.111+.1111+……..ton terms)

= $\frac{6}{9}$ (0.9 + 0.99 + 0.999 + 0.9999 + ….. ton terms)

= $\frac{6}{9}$ (1-0.1) + (1-0.1) + (1-0.0001) +……. Ton terms)

= $\frac{6}{9}$ {(1+1+1+ ... .. ton n times) – (0.1 + 0.01 + 0.001 + 0.0001 + ... to n terms)}

$$= \frac{6}{9}\left[1\times n - \times 1\left(\frac{1-(x-1)^n}{1-x1}\right)\right] = \frac{6}{9}\left(n - \frac{\frac{1}{10}\left(1-\left(\frac{1}{10}\right)^n\right)}{\frac{9}{10}}\right)$$

$$= \frac{6}{9}\left(n - \frac{\cdot 1\{1-(0\cdot 1)^n\}}{1-0\cdot 1}\right)$$

Practice Questions:

1. If a,ar,ar^2, ar^3,……. Be in G.P. Find the common ratio.

2. Which term of the progression 1,2,4,8,……. Is 256?

3. Insert 3 geometric m4eans between 1/9 and 9

4. Find the G.P where 4th term is 8 and 8th term is 128/625

5. Find the sum of 1=2+4+8+…… to 8 terms.,

6. Find three numbers in GP. Whose sum is 19& Product is 216?

7. Find the sum of the series $-\log_{n\to\infty}\frac{1}{6}+\frac{1}{6^2}+\frac{1}{6^3}+\cdots\ldots\ldots\frac{1}{6^n}$

8. 14.If x = $1+a+a^2+a^3+$……… to ∞; y = $1+b+b^2+b^3+$………. to∞; Find the sum of the series $1+ab+(ab)^2+(ab)^3$; where, $|a|<1$ and $|b| <1$.

9. Three numbers are in A.P and their sum is 15. If 1,3,9 are added to them respectively, they form a G.P. Find the numbers.

10. Find the sum of the series 6,27,128,629

11. Find the sum to be terms of the series 3+33+333+…….

12. Find the sum of n terms of the series 0.7+0.77+0.777 + ….. to n terms

Objective type questions (A.P and G.P)

1. In an A.P., first term = a, difference is d, last term is l, find the term before last

   (a) l-a (ii) l-d (iii) l+(n-1)d (iv) a+ (n-1)d

2. Fill up the gaps:- -14, -10,-6,-,-,6

   (i) 2,-2 (ii) 4,-2 (iii) 2,-4 (iv) -2,2

3. Fill up the gaps: 2,-,-,-,-,17(A.P)

   (i) 5,8,11,14 (ii) 4,7,10,13 (iii) -5,-8,-11,-14 (iv) none of these

4. Evaluate: 1+2+3+…….. + (n-1)

   (i) $\frac{n}{2}$ (ii) n(n-1) (iii) $\frac{n(n-1)}{2}$ (iv) $\frac{(n-1)}{2}$

5. If the terms of an A.P series be added or subtracted by a constant number, then the resultant series will be in what series

   (a) A.D (ii) G.P (iii) none

6. A boy saves 1P to-day, 2P tomorrow, 3P day after tomorrow. How much he can save in 12 days?

   (i) 68 (ii) 70 (iii)78 (iv) 87 (paise)

7. The first term of an A.P is 30 and 3rd term is 26, find the 8th term of that series

   (a) 12 (ii) 16 (iii) 20 (iv)24

8. In an A.P a = 6, d = 2, find t_7

   (i) 12 (ii) 3 (iii) 20 (iv) 18

9. Of the series 2,5,8,………..; find the 10th term.

   (i) 27 (ii) 29 (iii) 31 (iv) 34

10. If the terms -1+2x, 5,5+x from an A.P., then find x

    (i) 2 (ii) 3 (iii) 4 (iv) 5

11. Whatever be the value of x, the terms x,x+3 and x+6 will be in what series?

    (i) A.P (ii) G.P (iii) none

12. Find the 8th term of the series 4,-8,16,-32……

    (i) -512 (ii) 512 (iii)-521 (iv) 521

13. A.M and G.M of two positive integers a and be (a<b) are respectively 5 and 4; find a and b.

    (i) 2,8 (ii) 6,4 (iii) 8,2 (iv) 4,6

14. Find G.M of 2 and 6

(i) ± 4 (ii) $\pm\sqrt{3}$ (iii) ± 2 (iv) $\pm 2\sqrt{3}$

15. If 2,x,50 are in G.P find x

(i) ± 10 (ii) ± 8 (iii) ± 9 (iv) none

16. Find G.M of the numbers 3,-,-,24

(i) 4,12 (ii) 6,12 (iiii) 6,8 (iv) none

17. Find the sum of the series 2+1+

(i) $2\frac{1}{8}$ (ii) 4½ (iii) 3½ (iv) 4

18. Find the sum of : 0.9+0.81+0.729+ …………

(i) 8 (ii) 9 (iiii) 10 (iv) 11

19. Which term of the series 2,4,8,16,………… is 2048?

(i)10 (ii) 11 (iii) 12 (iv) 14

20. Of a G.P series the first two terms are 3 and 1, find 7th term

(i) $\frac{1}{3^5}$ (ii) 3^4 (iii) $\frac{5}{2}$ (iv) $\frac{1}{3^2}$

21. The product of three terms in G.P is 125/8; find G.M

(i) $\frac{2}{5}$ (ii) $\frac{3}{5}$ (iii) $\frac{5}{2}$ (iv) $\frac{5}{3}$

22. Of a series nth term is 2n+5. Find the nature of the series

(i) G.P (ii) A.P (iii) None of these

23. If the nth term of a series be 2n+5, find whether the series is in A.P or G.P

(i) A.P (ii) G.P

24. The product of three terms in G.P is 1000, what is its middle term?

(i)12 (ii) 14 (iii) 16

25. If the sum of three numbers in A.P is 18, then what is the middle term?

(i) 6 (ii)4 (iii) 8 (iv) 10

26. The sum of three numbers in A.P is 15 and their product is 80. Find the numbrs (ascending order)

(i) 2,5,8 (ii) 8,5,2 (iii) 1,4,7 (iv) none of these

27. If -1+2x, 5,5+x are in A.P. Find x

    (i) 4 (ii) 3 (iii) 2 (iv) 5

28. Find the sum of $\frac{1}{2}+\frac{1}{3}+\frac{1}{6}+\cdots\ldots$ to 10^{th} term

    (i) $\frac{5}{2}$ (ii) $-\frac{5}{2}$ (iii) $\frac{2}{5}$ (iv) $-\frac{2}{5}$

29. In a G.P., t_3=3, t_6=81, find first term

    (i) 3 (ii) $-\frac{1}{3}$ (iii) -3 (iv) $\frac{1}{3}$

30. The 2^{nd} term of a G.P is b and common ratio is r. If the product of the first three terms is 64, find the value of b.

    (a) 4 (ii) 3 (iii) 5 (iv) 6

31. The first and last terms of an A.P are -4 and 146 and the sum is 7171. Find the number of terms

    (a) 98 (ii) 100 (iii) 101 (iv) 110

32. You save 1P to-day, 2p to-morrow and 3p day after tomorrow. How much you can save in 1 year (1 year = 365 days)?

    (i) 696.70 (ii) 667.95 (iii) 766.94 (iv) 676.75

33. A man puts by ₹ 5 in the first month and ₹ 2 more in every succeeding month. How much should he save at the end of 5 years?

    (i) 3840 (ii) 3480 (iii) 3540 (iv) 3804

34. The 3^{rd} and 6^{th} terms of a G.P are 3 and 81 respectively, find the common ratio

    (a) 1 (ii) 2 (iii) 3 (iv) 4

35. A.M of two integral numbers exceeds their G.M by 2 and the ratio of the numbers is 1:4. Find the numbers.

    (i) 5,20 (ii) 1,4 (iii) 2,8 (iv) 4,16

36. The sum of the first 5 and first 10 terms of a G.P. are respectively 16 and 3904. Find the common ratio.

    (i) 2 (ii) 3 (iii) 4 (iv) 5

Answers:

1. (i)	2. (iv)	3. (i)	4.(iii)	5.(i)	6.(iii)	7. (ii)	8.(iv)
9. (ii)	10.(i)	11.(iii)	12.(i)	13.(iii)	14.(iv)	15.(i)	16.(ii)
17.(iv)	18.(ii)	19.(ii)	20.(i)	21.(iii)	22.(ii)	23.(i)	24.(iv)
25.(i)	26.(i)	27.(iii)	28.(ii)	29.(iv)	30.(i)	31.(iii)	32.(ii)
33.(i)	34.(iii)	35.(iv)	36.(ii)				

Study Note - 2

ALGEBRA

This Study Note includes

2.1 SET THEORY

SET: - Set is a well defined collection of distinct objects. The objects of the set are called its elements (or) members. Sets are generally denoted by capital letters A, B, C.... and the elements of the sets are denoted by letters in lower case. i.e., a, b, c.... If an element x belongs to the Set A it is denoted by $x \in A$. If x is not an element of A it is denoted by $x \notin A$.

Sets are generally represented by two methods:

(i) **Roster method (or) Tabular method**: In this method, a set is described by actually listing out the elements.

Example: -The set of all even natural numbers less than 20 is represented by {2,4,6,8,10,12,14,16,18}

(ii) **Set Select or builder method**: In this method, a set is described by characterising property. For example the set of all even natural numbers less than 20 is represented by $\{x/x \in N\ x < 20\ d\ x \text{ is an even number}\}$

Types of Sets:

1. **Finite set:** A Set 'A' is said to be finite if it consisting of finite number of elements. If a set consists m elements then no. of elements in A is denoted by n(A) i.e. n(A) = m. For example: - Set of voters in Vijayawada city is a finite set.
2. **Infinite Set:** A set having an infinite no. of elements is called an infinite set. For example: the set of stars in the sky is an infinite set.
3. Null set (or) empty set (or) avoid set: -A set which contains no elements is called null set. It is denoted by ϕ (or) { } Examples:
    1. $\{x/x \text{ is a perfect square and } 5 < x < 9\}$ is an empty set
    2. $\{x/x \text{ is a real number and } x^2 < 0\}$ is a Null set

Note:

(i) $\phi \neq \{\phi\}$ ($\because$ $\{\phi\}$ is a set whose element is ϕ)

(ii) $\phi \neq \{0\}$ ($\because$ $\{0\}$ is a set whose element is 0)

4. **Equal sets**: - Two sets A and B are said to be equal sets if they have the same elements (i.e. every element of set A is an element of B and also every element of B is an element of A then A = B)

    For example: (i) A = {a, b, c}, b = {b, c, a} are equal sets (ii) A = {1, 2, 3, 4} B = {1, 1, 2, 3, 4, 4, 3} are also equal Note: - The order of writing elements (or) repetition of elements does't change the nature of set

5. **Equivalent sets**: Two finite sets A and B are said to be equivalent if 0(A) = 0(B) (or) n(A) = n(B).

   The symbol '0' is used equivalent.

   Ex: - Lt A = {a, e, i, o , u} then n (A) = 5

   And B = {1, 2, 3, 4, 5} then n (B) = 5

   $\therefore$ A 0 B ($\because$ n (A) = n (B))

6. **Sub set**: - If every element of set A is also an element of set B. Then A is called a subset of B. Also B is said to be a super set of A. Thus we write A is contained in B as A ≤ B and B contains A as B ≥ A. If A is not a subset of B then it is denoted by $A \not\leq B$.

   Ex: - at A = {1, 2, 3}, B = {1, 2, 3, 4, 5} then A ≤ B and B ≥ A.

   Note 1: Every set is a subset of itself

   Note 2: Empty set is a subset of every set i.e. ∅ ≤ A

   Note 3: Every set has atleast two subsets. i.e., the empty set and the set itself

   Note 4: A ≤ B, B ≤ C ⇒ A ≤ C

   Note 5: A ≤ B and B ≤ A ⇒ (A = B)

   Note 6: If A ≤ ∅ then A = ∅

   Formula: Number of subsets of a set A containing n elements is 2^n

7. **Proper Subset**: - If each and every element of a set A are the elements of B and = 7 at least one element of B that does't belongs to A then the set A is said to be a proper subset of B (or) B is called super set of A. Symbolically. We may write A < B.

   Ex: - If B = {a, b, c} then the proper subsets are ∅, {a}, {b}, {a, b} {b, c}, {c} {a, c}

   Note: (1) A set is not a proper subset of itself

   (2) ∅ is not a proper subsets of set A containing n elements is 2^n-1

8. **Power Set**: The set of all the subsets of a given set A is called the power set of A and is denoted by P(A)

   Ex: If A = {1, 2, 3} then

   P(A) = {∅, {1}, {2}, {3}, {1, 2}, {2, 3}, {1, 2}, {2, 3}, {1, 3}, {1, 2, 3}}

   Note1: If no. of elements in Set A is 3 then no. of elements in P2^3 = 28

   Formula: - If a set 'A' has n elements the p (A) will have 2^n elements.

   Universal Set: - If all the sets under particular consideration are the subsets of a fixed set 'S' then the set 'S' is called universal set. Universal sets are denoted by U (or) μ.

   Ex 1: The set of all the stars in the sky is an universal set

   Ex 2: A set of integers is an universal set for the set of even (or) odd integers.

Cardinality (or) order of a set: -The no. of distinct elements in a set is called the cardinality (or) order of the set. If a finite set A has n distinct elements, then the cardinality of the set A is n and is denoted by 0(A) (or) n(A)

Ex: - If A = {a, b, c} then n(A) = 3

B = {1, 2, 3} then 0 (B) = 3

$\therefore$ n (A) = 0 (B)

Note: – The order of empty set is zero.

Disjoint sets: Two sets A and B are said to be disjoint sets if they have no element in common i.e., $A \cap B = \varnothing$.

Ex: at A = {1, 3, 5}, B = {2, 4 }

$\therefore$ $A \cap B = \varnothing$. Hence A & B are disjoint sets

Basic Operation on sets: - (By using Venn diagram)

(i) **Union of sets:** - The union of sets A and B is the set of all the elements which belong to A (or) B (or) both and is written as AUB. In set builder from the union of sets can be represented as AUB = {x/x $\in$ A (or) x $\in$ B}

Ex: - at A = {1, 2, 3, 4} and B = {2, 4, 6, 8}

Union of sets may be illustrated more clearly by using Venn-diagram.

Properties: (1) Union of sets satisfies commutative property i.e., AUB = BUA

(2) Union of sets satisfies Associative property i.e., (AUB)UC = AU(BUC)

(ii) **Intersection of sets**: - The intersection of two sets A and B is the set consisting of all the common elements of A & B and is written as $A \cap B$.

In set builder form intersection of two sets can be represented as

$A \cap B = \{x / x \in A \text{ and } x \in B\}$

Ex: - at A = {1, 2, 3, 4} and B = {2, 4, 6, 8}

$A \cap B = \{2, 4\}$

Intersection of sets may be illustrated more clearly by using Venn diagram.

Difference of Sets: - The difference of two sets A and B is the set of elements which belongs to a but which does't belongs to B. It is written as A ~ B (or) A- B.

In set builder form difference of sets can be represented as $A \sim B = \{x / x \in A \text{ and } x \notin B\}$

$B \sim A = \{x / x \in B \text{ and } x \notin A\}$

Ex: Let A = {1, 2, 3, 4, 5} and B = {5, 7, 9}

Difference of sets can be illustrated more clearly by using Venn diagram.

Compliment of a set: IF U is the universal set and A is a subset of U. Then compliment of A is the set which contains those element of U. Which are not contained in A and denoted by A^1 (or) A^c (or) If A is a set then U-A is called compliment of A. In set builder form compliment of a set can be represented as $A^c = \{x / x \in U \text{ and } x \notin A\}$

Eg: - If U = {1, 2, 3, ...} and A = {2, 4, 6, 8....} then A^c = U –A = {1, 3, 5......}

Compliment of a set can be illustrated more clearly by using Venn diagram.

Note: 1) $(A^c)^c = A$

2) $AUA^1 = U$

3) $A \cap A^1 = Q$

Symmetric difference of two sets: The symmetric difference of two sets A and B denoted by $A \Delta B$ and is defined as $A \Delta B$ = (A-B) U (B-A)

Ex: - If A = {1, 2, 3, 4, 5} & B = {1, 3, 5, 7, 9}

A Δ B = (A – B) U (B – A)

= {2, 4} U {7. 9}

= {2, 4, 7, 9}

Note 1: A Δ B can also be written as

A Δ B = (AUB) - A ∩ B

Illustrations:

1. Rewrite the following examples using set notation
   (i) First ten even natural numbers.
   (ii) Set of days of a week.
   (iii) Set of months in a year which have 30 days.
   (iv) The numbers 3, 6, 9, 12, 15
   (v) The letters m, a, t, h, e, m, a, t, I, c, s

Solution:

(i) A = {2, 4, 6, 8, 10, 12, 14, 16, 18, 20}
= {x : x is an even integer and $2 \leq x \leq 20$}

(ii) A = {Sunday, Monday ……. Saturday}
= {x : x is a day in a week}

(iii) A = {April, June, September, November}
= {x : x is a month of 30 days} (Selector)

(iv) A = {x : x is a positive number multiple of 3 and $3 \leq x \leq 15$}

(v) A = {x : x is a letter in the word mathematics}

2. Write the following set in roster form.
   (i) A = {x : x is an integer, $-3 \leq x \leq 7$}
   (ii) B = {x : x is an integer, $4 < x \leq 12$}

Solution:

(i) A = {-3, -2, -1, 0, 1, 2, 3, 4, 5, 6}
(ii) B = {6, 8, 10, 12}

3. Represent the following sets in a selector method:
   (i) all numbers less than 15
   (ii) all even numbers

Solution:

Taking R to be the set of all real numbers in every case:

(i) {x : $x \in R$ and $x < 15$}

(ii) {x : $x \in R$ and x is a multiple of 2}

4. A = {1, 2, 3, 4, 6, 7, 12, 17, 21, 35, 52, 56}, B and C are subsets of A such that B = {odd numbers}, C = {prime numbers}. List the elements of the set $\{x : x \in B \cap C\}$

Solution:

$B \cap C = \{1, 3, 7, 17, 21, 35\} \cap \{2, 3, 7, 17\} = \{3, 7, 17\}$

Therefore, reqd. list = {3, 7, 17}

5. If A = {1, 2, 3, 4, 5}, B = {2, 4, 5, 8}, C = {3, 4, 5, 6, 7}, find $A \cup (B \cup C)$.

Solution:

$B \cup C = \{2, 3, 4, 5, 6, 7, 8\}$, $A \cup (B \cup C) = \{1, 2, \ldots\ldots.. 7, 8\}$

6. If A = {1, 2, 3} and B = {2, 3, 4}; find $(A – B) \cup (B – A)$

Solution:

$A – B = \{1\}$, $B – A = \{4\}$, $(A- B) \cup (B – A) = \{1, 4\}$

7. If S is the set of all prime numbers, $M = \{x : 0 \leq x \leq 9\}$ exhibit

(i) $M – (S \cap M)$ (ii) $M \cup N$, $N = \{0, 1, 2 \ldots\ldots.. 20\}$

Solution:

$S = \{2, 3, 5, 7, 11, 13, \ldots\ldots..\}$, $M = \{0, 1, 2, \ldots.. 8, 9\}$

(i) $S \cap M = \{2, 3, 5, 7\}$

(ii) M U N = {0, 1, ….. 20}

8. In a class of 100 students, 45 students read Physics, 52 students read Chemistry and 15 students read both the subjects. Find the number of students who study neither Physics nor Chemistry.

Solution:

We know $n (A \cup B) = n (A) + n (B) - n (A \cap B)$. Let A indicates Physics, B for Chemistry. Now

$n (A) = 45$, $n (B) = 52$, $n (A \cap B) = 15$

So, $n (A \cup B) = 45 + 52 – 15 = 82$

We are to find $n (A' \cap B') = 100 – n (A \cup B) = 100 – 82 = 18$

9. In a survey of 1000 families it is found that 454 use electricity, 502 use gas, 448 use kerosene, 158 use gas and electricity, 160 use gas and kerosene and 134 use electricity and kerosene for cooking. If all of them use at least one of the three, find how many use all the three fuels.

Solution:

Let us take E for electricity, G for gas, K for kerosene.

Now $n (E) = 454$, $n (G) = 502$, $n (K) = 448$

$n (G \cap E) = 158$ $\quad$ $n (G \cap K) = 160$ $\quad$ $n (E \cap K) = 134$, $\quad$ $n (E \cap G \cap K) = ?$

$n (E \cup G \cup K) = 1000$

Again $n (E \cup G \cup K) = n (E) + n (G) + n (K) – n (E \cap G)$

$\qquad - n (G \cap K) – n (K \cap E) + n (E \cap G \cap K)$

or, $1000 = 454 + 502 + 448 – 158 – 160 – 134 + n (E \cap G \cap K)$

$\qquad = 952 + n (E \cap G \cap K)$

Or, $n (E \cap G \cap K) = 1000 – 952 = 48$.

Practice Questions

1. (i) If $n(A) = 20$, $n(B) = 12$, $n(A \cap B) = 4$, find $n(A \cup B)$ [Ans. 28]

   (ii) H $n(A) = 41$, $n(B) = 19$, $n(A \cap B) = 10$, find $n(A \cup B)$ [Ans. 50]

   (iii) If $n(A) = 12$, $n(B) = 20$, and A C B, find $n(A \cup B)$ [Ans. 20]

   (iv) If $n(A) = 24$, $n(B) =18$ and B c A, find $n(A \cup B)$ [Ans. 24]

   [Hints. $n(A \cap B) = n(B)$ as B C A

   $n(A \cup B) = n(A) + n(B) - n(A \cap B) = n(A) + n(B) - n(B) = n(A) =$ etc.]

2. In a class 60 students took mathematics and 30 took physics. If 17 students were enrolled in both the subjects, how many students all together wee in the class, who took mathematics or physics or both. [Ans.73]

3. In a class 52 students, 20 students play football and 16 students play hockey. It is found that 10 students play both the game. Use algebra of sets to find out the number of students who play neither. [Ans.24]

4. In a class test of 45 students, 23 students passed in paper first, 15 passed in paper first but not passed in paper second. Using set theory results, find the number of students who passed in both the papers and who passed in paper second but did not pass in paper first. [8;22]

5. In a class of 30 students, 15 students have taken English, 10 students have taken English but not French. Find the numbers of students have taken: (i) French, and (ii) French but not English. [Ans.20,15]

6. In a class test of 70 students, 23 and 30 students passed in mathematics and in statistics respectively and 15 passed in mathematics but not passed in statistics. Using set theory result, find the number of student who passed in both the subjects and who did not pass in both the subjects. [Ans.8;25]

7. In a survey of 100 students it was found that 60 read Economics, 70 read mathematics, 50 read statistics, 27 read mathematics and statistics, 25 read statistics and Economics and 35 read mathematics and Economics and 4 read none. How many students read all there subjects?

   [hints: refer solved problem no.3] [Ans. 3]

OBJECTIVE TYPE QUESTIONS:

I. Choose the correct answer:

1. Set of even positive integers less than equal to 6 by selector method:

   (a) $\{x/x<6\}$ (b) $\{x/x = 6\}$ (c) $\{x/x \le 6\}$ (d) None

2. By Roster method, to express integers greater than 5 and less than (or) equal to 8

   (a) {5, 6, 7} (b) {5, 6, 7, 8} (c) {8} (d) {6, 7, 8}

3. Write the set containing all days of the week beginning with 5.

   (a) {Sunday, Monday} (b) {Saturday, Sunday} (c) (Friday, Saturday} (d) None

4. State Whether the following statements are correct:

   (i) $\{1, 2, 3\} = \{2, 3, 4\}$ (ii) $\{1, 2, 3\} = \{1, 1, 2, 2, 3, 3\}$ (iii) $\{1, 2, 3\} \le \{3, 2, 1\}$

   (iv) $\varnothing \in \{1, 2, 3\}$ (v) $4 \notin = \{1, 2, 3\}$

   (a) (i) (ii) (b) (ii) (iii) (c) (iii) (iv) (d) (v) (iii)

5. From the sets given below, pair the equal sets

(i) A= {1, 2, 34}   (ii) B = {p, q, r, s}   (iii) C = {1, 4, 9, 16}

(iv) D = {x, y, z, w}   (v) E = {16, 1, 4, 9}   (vi) F = {4, 2, 3, 1}   (vii) G= {r, p, q, s}

(a) (i) (ii)   (b) (i) (v)   (c) (i) (vi)   (d) (ii) (vii)

6. From the given sets pair the equivalent sets:

(i) A= {4, 5, 6, 7}   (ii) B = {0, Δ, $\square$}   (iii) C = {a, b}

(iv) D = {5}   (v) E = {4, ϑ}   (vi) {1, 2, 3}

(a) (ii) (i)   (b) (iii) (v)   (c) (vi) & (ii)   (d) (iv) (v)

7. Find which one of the following is a Null set

(i) A = {x/x < x}   (ii) B = {x/x + 2 = 2}   (iii) {x/x is a positive number less than 0}

(a) (i)   (b) (ii)   (c) (iii)   (d) None

8. Which one of the following is a singleton set?

(i) $\{x/x^2 = x, x \in R\}$ (ii) $\{x/x^2 = -1, x \in R\}$ (iii) {x/2x = 0}   (iv) $\{x/3x + 2 = 0\ x \in N\}$

(a) (i)   (b) (ii)   (c) (iii)   (d) (iv)

II. **Fill in the blanks**:

9. If A = {1, 2, 3, 4}, B = {2, 4, 6}, then A Δ B is ____

10. If A and B are two sets then A $\cap$ (B-A) is ________

11. If A = {1, 2, 3}, B = {2, 3, 4}, C = {1, 2, 5, 6) then AU(B $\cap$ C) is ______

12. IF A and B are two sets then A $\cap$ B = A $\cup$ B if and only if ______

13. If A and B are two disjoint sets then n(A $\cup$ B) is equal to ______

14. A has 2 elements, B has 4 elements and A<B than A $\cap$ B has ______ elements

15. If A and B are the two sets of positive and negative integers respectively then AUB is

III. **State whether the following statements are True (or) False**:

16. The statement $(A \cap B)^1 = A^1UB^1$ is true (or) False ( )

17. If the set A has 4 elements, B has 3 elements then the number of elements in A x B is 12 elements ( )

18. The Statement {2} $\in$ {2, 3, 5} is true (or) False ( )

19. The statement {1} C {1, 2, 3} is true (or) False ( )

20. The statements "Equivalent sets are always equal is True (or) False ( )

IV. **Match the following**:

Group A		Group B
21. $\{x/x \in N, 2x = 5\}$		(A) 6
22. $(A^c)^c$ =__________		(B) A^1UB^1
23. A has 4 elements and B has 6 elements such that A<B then no. of elements in AUB is ______		(C) A
24. $(AB)^1$ = ______		(D) C
25. If AUB = AUC then B = ____		(E) Null

V. Answer the following in one (or) two sentences:

26. If A = {1, 2, 3} and B = {2, 4} then find A x B
27. In a group of 63 persons, 24 persons take wheat but not rice, 37 persons take wheat then find the number of persons taking rice but not wheat?
28. In a class each student plays either cricket (or) football. If 50 students play football, 30 students play cricket while 15 students play both then find the no. of students in a class.
29. If A = {-1, 1} B = {0, 2} C = {0, 1} then find (A x B) $\cap$ (B x C)
30. If A = {4, 3, 6, 5}; B = {7, 5, 8} C = {5, 1 , 6, 2} then find (A $\cap$ C)

Set Theory Key:

I. 1)

II. 9) {1, 3, 6} 10) φ 11) {1, 2, 3} 12) A = B 13) n (A)+ n (B) 14) 2

15) Set of all integers except zero

III. 16) T 17) T 18) F 19)T 20) F

IV. 21) Null 22) A 23)6 24)A^1UB^1 25) C

2.2 INDICES AND LOGARITHMS

Indices

Index:

If a number 'a' is taken and added three times then we say that a+a+a = 3a. Instead of addition the same number is multiplied three times then we say that.

$a \times a \times a = a^3$

here

3 is called index

A is called base

Root of a number:

If $x^n = a$ then,

X is called nth root of a

i.e.

$x = \sqrt[n]{a}$

Ex: $2^4 = 16$, $2 = \sqrt[4]{16}$.

Basic rule on Indices:

1. $a^m \times a^n = a^{m+n}$

   Ex: $5^3 \times 5^2 = 5^{3+2} = 5^5$

2. $\dfrac{a^m}{a^n} = a^{m-n}$

   Ex: $\dfrac{5^3}{5^2} = 5^{3-2} = 5$

3. $\left(\dfrac{a}{b}\right)^m = \dfrac{a^m}{b^m}$

   Ex: $\left(\dfrac{6}{5}\right)^2 = \dfrac{6^2}{5^2}$

4. $(ab)^m = a^m b^m$

   Ex: $(5 \times 4)^2 = 5^2 \times 4^2 = 25 \times 16$.

5. $\dfrac{1}{a^m} = a^{-m}$

   Ex: $\dfrac{1}{5^2} = 5^{-2}$

6. $\sqrt[n]{a} = a^{1/n}$

   Ex: $\sqrt[5]{6} = 6^{1/5}$

7. $a^0 = 1$

   Ex: $6^0 = 1$

8. The basis of two equal numbers are equal then powers also to be equal i.e.

$2^2 = 2^x$

$X = 2$

9. The powers of two equal numbers are equal then basis are also equal when they are of same sign.

Ex: $x^2 = 2^2$

$X = 2$ (when $x > 0$).

Illustration:

1. Express as positive indices –

(i) $x^{-1/3}$ (ii) $2x^{-2/5}.\,a^{-1/3}$ (iii) $x^{2/3} \div a^{-1/5}$

Solution:

(i) $x^{-1/3} = \dfrac{1}{x^{-1/3}}$ (ii) Expr. $= \dfrac{2}{x^{2/5}}\cdot\dfrac{1}{x^{1/3}}$ (iii) Expr. $= \dfrac{2/3}{1/5} = x^{2/3}\cdot a^{1/5}$

2. Find the value of (i) $\left(\dfrac{1}{81}\right)^{-\frac{3}{4}}$ (ii) $243^{-\frac{1}{5}}$

(i) $\left(\dfrac{1}{81}\right)^{-\frac{3}{4}} = (81^{-1})^{-3/4} = \{(3^4)^{-1}\}^{-3/4} = 3^3 = 27$

(ii) $243^{-1/5} = (3^5)^{-1/5} = 3^{-1} = \dfrac{1}{3}$

3. Show that $\dfrac{2\left(3^{n+1}\right)+7\left(3^{n-1}\right)}{3^{n+2}-2\left(\dfrac{1}{3}\right)^{1-n}} = 1$ (ICWA (F) June 98)

Solution:

$2(3^{n+1}) + 7(3^{n-1}) = 2.3^n.3 + 7.3^n.3^{-1} = 3^n\left(2.3+\dfrac{7}{3}\right) = 3^n\left(6+\dfrac{7}{3}\right) = 3^n.\dfrac{25}{3}$

$3^{n+2} - 2\left(\dfrac{1}{3}\right)^{1-n} = 3^n.3^2 - 2\left(3^{-1}\right)^{1-n} = 3^n.3^2 - 2.3^{n-1} = 3^n.3^2 - 2.3^n.3^{-1}$

$3^n\left(3^2 - 2.\dfrac{1}{3}\right) = 3^n\left(9-\dfrac{2}{3}\right) = 3^n.\dfrac{25}{3}$

$\therefore$ Expr. $= \dfrac{3^n.25/3}{3^n.25/3} = 1$

4. Simplify $\dfrac{\left(P+\dfrac{1}{q}\right)^p\left(P-\dfrac{1}{q}\right)^q}{\left(q+\dfrac{1}{p}\right)^p\left(q-\dfrac{1}{p}\right)^q}$

Solution: $\dfrac{\left(\dfrac{pq+1}{q}\right)^p\left(\dfrac{pq-q}{q}\right)^q}{\left(\dfrac{pq+1}{p}\right)^p\left(\dfrac{pq-1}{p}\right)^q} = \dfrac{(pq+1)^p}{q^p}\times\dfrac{(pq-1)^q}{q^q}\times\dfrac{p^p}{(pq+q)^p}\times\dfrac{p^q}{(pq-1)^q} = \left(\dfrac{p}{q}\right)^{p+q}$

5. Show that $\left(\frac{x^a}{x^b}\right)^{a^2+ab+b^2} \times \left(\frac{x^b}{x^c}\right)^{b^2+bc+c^2} \times \left(\frac{x^c}{x^a}\right)^{c^2+ca+a^2} = 1$

Solution:

$$\text{L.H.S} = \left(x^{a-b}\right)^{(a^2+ab+b^2)} \times \left(x^{b-c}\right)^{(b^2+bc+c^2)} \times \left(x^{c-a}\right)^{(c^2+ca+a^2)}$$

$$x^{(a-b)\left(a^2+ab+b^2\right)} \times x^{(b-c)\left(a^2+bc+c^2\right)} \times x^{(c-a)\left(c^2+ca+a^2\right)}$$

$$= x^{a^3-b^3} \times x^{b^3-c^3} \times x^{c^3-a^3} = x^{a^3-b^3+b^3-c^3+c^3-a^3} = x^0 = 1$$

6. Find the value of (a+ b+c) if $x^{1/a} = y^{1/b} = z^{1/c}$ satisfying xyz = 1

Solution:

Let $x^{1/a} = y^{1/b} = z^{1/c} = k$; $x = k^a$, $y = k^b$, $z = k^c$

Now xyz =1 or , $k^a.\ k^b.k^c = 1$ or $k^{(a+b+c)} = 1 = k^0$ or a+b+c = 0

7. Solve: $2^{x+3}+2^{x+1} = 320$ or , $2^{x+1}(2^2+1) = 320$ or $2^{x+1} = \frac{320}{5} = 64$

Or $2^{x+1} = 2^6$ or, x = 5

8. If $2^{x+4} - 2^{x+2} = 3$, find x^7

Solution: $2^{x+4} - 2^{x+2} = 3$ or, $2^x(2^4 - 2^2) = 3$ or, $2^x.12 = 3$

Or, $2^x = \frac{1}{4} = 4^{-1} = 2^{-2}$ or, x = –2

Now, $x^7 = (-2)^7 = -128$

Practice problems:

1. Express the following in single positive index

(i) $(x^{-3/4})^{5/3}$

2. $\frac{(81)^n.3^5 - 3^{4n-1}.243}{9^{2n}.3^3} - \frac{4.3^n}{3^{n+1}-3^n}$

3. Show that $\left(\frac{x^b}{x^c}\right)^a \times \left(\frac{x^c}{x^a}\right)^b \times \left(\frac{x^a}{x^b}\right)^c = 1$

4. Show that $\left(\frac{x^m}{x^n}\right)^{m+n-l} \times \left(\frac{x^n}{x^l}\right)^{n+l-m} \times \left(\frac{x^l}{x^m}\right)^{l+m-n} = 1$

5. Show that $\left(\frac{x^{a^2+b^2}}{x^{-ab}}\right)^{a-b} \times \left(\frac{x^{b^2+c^2}}{x^{-bc}}\right)^{b-c} \times \left(\frac{x^{c^2+a^2}}{x^{-ca}}\right)^{c-a} = 1$

6. Show that $\frac{1}{1+x^{b-a}+x^{c-a}} + \frac{1}{1+x^{a-b}+x^{c-b}} + \frac{1}{1+x^{a-c}+x^{b-c}} = 1$

7. If $2^x = 3^y = 6z$, show that $z = \frac{xy}{x+y}$

8. If $a^x = b^y = c^c$ and $b^2 = ac$, prove that $\frac{1}{x}+\frac{1}{z}=\frac{2}{y}$

9. If $x^a = y^b = (xy)^c$ show that ab = c(a+b)

10. Show that $x^3 - 6x - 6 = 0$ if $x = \sqrt[3]{2} + \sqrt[3]{4}$

11. Solve $\left(\sqrt{5}\right)^{4x-4} - 5^{2x-3} = 20$

I. Choose the correct Answers

1. If $a^x = b^y = c^z$ and $b^2 = ac$ then $xy+yz =$ _______

   (a) xz (b) – xz (c) 2xz (d) none of the these

2. If $\dfrac{\left(P+\frac{1}{q}\right)^p \left(P-\frac{1}{q}\right)^q}{\left(q+\frac{1}{p}\right)^p \left(q-\frac{1}{p}\right)^q} = \left(\dfrac{p}{q}\right)^x$ then the value of x is _____

   (a) p-q (b) p+q (c) q-p (d) none of these

3. The digit in the unit place of $(2\times 4^x)^2 + 1$ (where x is a positive integer

   (a) 1 (b) 5 (c) 3 (d) none of these

4. If $\dfrac{\left(2^{x+1}\right)^y . \left(2^{2x}\right)^{2x}}{\left(2^{y+1}\right)^{x.} . 2^{2y}}$

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

5. If $3^x = 5^y = (225)^z$ then z = _____

   (a) $\dfrac{xy}{x+y}$ (b) $2\dfrac{xy}{x+y}$ (c) 2(x+y) (d) none of these

II. Fill in the blanks

6. If $a^{1/3}+b^{1/3}+c^{1/3} = 0$ then $(a+b+c)^3 =$ ____

7. If $y = x^{1/3}-x^{-1/3}$ theny^{3+3y} is _______

8. If $a = 2+\sqrt[3]{2}+\sqrt[3]{4}$ then a^3-6a^2+6a is _____

9. If $64^x = 2\sqrt{2}$ then x =______

10. If x = 8, y = 27 then the value of $(x^{4/3}+y^{2/3})^{1/2}$ is ____

III. State whether the following statement are true or false

11. If $9\times 81^x = \dfrac{1}{27^x - 3}$ then the value of x is ( )

12. If x = $x = \sqrt[3]{\sqrt{2}+1} - \sqrt[3]{\sqrt{2}-1}$ then the value of x^3+3x is 2 ( )

13. If $x = 5+ 2\sqrt{6}$ and xy = 1 then $\dfrac{1}{x^2}+\dfrac{1}{y^2}$ is 89 ( )

14. If $2^{x+2y} = 2^{2x-y} = \sqrt{8}$ then $x = \dfrac{9}{10}$ and $y = \dfrac{10}{3}$

15. If $x = 2 + \sqrt{5}$ then $x^3 + 3x^2 - 29x$ is 7 ( )

IV. Match the following

Group A	Group *B*	
16. If $a=\frac{1}{2+\sqrt{3}}$ and $b=\frac{1}{2-\sqrt{3}}$ then the value of $2a^2+3ab-2b^2$ is	3	( )
17. If $x=7+4\sqrt{3}$ then $\sqrt{x}+\frac{1}{\sqrt{x}}=-$	5	( )
18. The value of $\sqrt{6+\sqrt{6+\sqrt{6+.........}}}$ to ∞	0	( )
19. If $\frac{\left(x-\sqrt{24}\right)\left(\sqrt{75}+\sqrt{50}\right)}{\sqrt{75}-\sqrt{50}}$ = 1 then the value of x is	3-16$\sqrt{3}$	( )
20. If $x^a = y^b = z^c$ and xyz =1 then $\frac{1}{a}+\frac{1}{b}+\frac{1}{c}=$ __	4	( )

V. Answer the question in on (or) two steps

21. Simplify $\frac{a^{a+b}\cdot x^{a-b}\cdot x^{m-2a}}{x^{m-a}}$ Ans: x^a
22. Show : $2^{x=3} + 2^{x+1} = 320$ Ans: 5
23. If $\left(\sqrt{10}\right)^{3a}$ = 16 then the value of 10^{-3a} Ans: 1/256
24. Find the simple value of 4 x $8^{-2/3}$ Ans: 1
25. If $3^x = 2^{-x}$ then the value of x Ans: 0

Indices Key:

II. Choose the correct answer

1. C 2. B 3. B 4. D 5. D

III. Fill in the blanks

6. 27abc 7. $x-\frac{1}{x}$ 8. 2 9. ¼

IV. State whether the following statements are true (or) false

11. T 12. T 13. F 14. F 15. T

V. Match the following

16. (18)
17. (19)
18. (20)
19. (16)
20. (17)

Logarithm:

Let N, a (a ≠ 1) be two positive real numbers and some real number x such that $a^x = N$ then $x = \log_a N$ and is read as logarithm of N to the base a.

$$a^x = N \Leftrightarrow x = \log_a N$$

Note:

Logarithms are defined for only positive real numbers

Basic rules of logarithms or basic formulae:-

1. The logarithm of a number which is not equal to 1 is its self as a base is unity.

   Ex: $a = a \Leftrightarrow 1 = \log_a^a$

   Ex: $5^1 = 5 \Leftrightarrow \log_5 5 = 1$

2. The logarithm of one to any base is zero (a ≠ 1)

   Ex: $a^0 = 1 \Rightarrow 0 = \log_a^1$ Ex: $10^0 = 1 \Rightarrow 0 = \log_{10}^1$

3. The logarithm of same number for different bases are different i.e.,

   Ex: $64 = 4^3$ Ex: $64 = 8^2$

   $\log_4^{64} = 3$ $\log_8^{64} = 2$

4. Prove that $\log_a^m + \log_a^n = \log_a mn$

   Let $\log_a^m = p \Rightarrow m = a^p$

   $\log_a^n = q \Rightarrow n = a^q$

   $\log_a^{mn} = r \Rightarrow mn = a^r$

   $mn = a^r$

   $\Rightarrow a^{p+q} = a^r$

   $\Rightarrow p + q = r$

   $\Rightarrow \log_a^m + \log_a^n = \log_a^{mn}$

   e.g: – $\log_{10}^5 + \log_{10}^{10} = \log_{10}^{50}$

5. Prove that $\log_a^m = p \Rightarrow m = a^p$

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

   Let $\log_a^n = q \Rightarrow n = a^q$

   at $\log_a\left(\frac{m}{n}\right) = r \Rightarrow \frac{m}{n} = a^r$

   $\Rightarrow \frac{a^p}{a^q} = a^r$

   $\Rightarrow a^{p-q} = a^r$

$\Rightarrow p - q = r$

$\log_a^m - \log_a^n = \log^{\frac{m}{n}}$

e.g: – $\log_{10}^{50} - \log_{10}^{10} = \log_{10}^{\left(\frac{50}{10}\right)} = \log_{10}^{5}$

6. Prove that $\log_x a^m = m\ \log_x^a (a > 0)$

Proof :- Let $\log_x^a = p$

$a = x^p$

$a^m = (x^p)^m$

$a^m = x^{pm}$

$p_m = \log_x^{a^m}$

$\Rightarrow m \log_x^a = \log_x^{a^m}$

$\Rightarrow \log_x^{a^m} = m\ \log_x^a$

Ex: $\log_2^{2^4}$

$4\ \log_2^2 = 4$

7. Prove that $\log_a^y . \log_x^a = \log_x^y$

Proof:- Let $\log_a^y = p \Rightarrow y = a^p$

$\log_x^a = q \Rightarrow a = x^q$

$\log_x^y = r \Rightarrow y = x^r$

$y = a^p = (x^q)^p$

$y = x^{pq}$

$pq = \log_x^y$

$\Rightarrow \log_a^y \cdot \log_x^a = \log_x^y$

This rule is used to change the base of logarithm

Note:-

$\Rightarrow \log_b^a = \dfrac{1}{\log_a^b}$

$\Rightarrow a^{\log_a^N} = N$

$\Rightarrow \log_b^a = \dfrac{\log_{10}^a}{\log_{10}^b}$

$\Rightarrow \log_{b^n}^a = \dfrac{1}{n}\log_b^a$

$\Rightarrow \log_{b^n}^{a^m} = \dfrac{m}{n}\log_b^a$

$\Rightarrow x^{\log_a^y} = y^{\log_a^x}$

Illustration:

1. Find the logarithm of 2025 to the base $3\sqrt{5}$

Solution:

Let x be the required number; then $\left(3\sqrt{5}\right)^x = 2025 = 3^4.5^2 \Rightarrow \left(3\sqrt{5}\right)^4$

$\therefore x = 4$ $\therefore$ 4 is the required number

2. The logarithm of a number of the base $\sqrt{2}$ is k. What is its logarithm to the base $2\sqrt{2}$?

Solution:

Let $\left(\sqrt{2}\right)^k$ = N.

Since = $2\sqrt{2} = 2.2^{1/2} = 2^{3/2}$

So $\sqrt{2} = (2^{\frac{3}{2}})^{\frac{1}{3}} = \left(2\sqrt{2}\right)^{\frac{1}{3}}$

$\left(2\sqrt{2}\right)^{k/3}$ = N.

$\therefore$ the required number is $\frac{k}{3}$

3. Find the value of $\log_2(\log_2\{\log_3(\log_3^{27})^3\})\}]$.

Solution:

Given expression

$= \log_2[\log_2\{\log_3(\log_3 3^9)\}] = \log_2[\log_2\{\log_3(9\log_3 3)\}]$

$= \log_2[\log_2\{\log_3 9\}]$ (as $\log_3 3=1$)

$= \log_2\{\log_2\{\log_3 3^2\}] = \log_2 \log_2^2 \log_3^3 = \log_2[\log_2 2] = \log_2 1 = 0$

4. If $\log_2 x + \log_4 x + \log_{16} x = \frac{21}{4}$ find x

Solution: $\log_2^{16} \times \log_{16}^x + \log_4^{16} \times \log_{16}^x + \log_{16}^x = \frac{21}{4}$

Or $\log_{16}^x (4+2+1) = \frac{21}{4}$

Or, $\log_{16}^x = \frac{3}{4}$

Or, $x = (16)^{\frac{3}{4}} = 8$

5. If P = $\log_{10} 20$ and q = $\log_{10} 25$, find x and such that 2 $\log_{10}$ (x+1) = 2p-q (ICWA(F) Dec 2003)

Solution:

$2p-q = 2\log_{10} 20 - \log_{10} 25 = \log_{10}(20)^2 - \log_{10} 25$

$\text{Log}_{10} 400 - \log_{10} 25 = \log_{10}\frac{400}{25} = \log_{10} 16$

Now, $2\log_{10}(X+1) = \log_{10} 16$ or, $\log_{10}(X+1)^2 = \log_{10} 16$ or, $(x+1)^2 = 16$

$= (\pm 4)^2$ or $x+1 = \pm 4$

$\therefore x = 3, -5$

6. If $x = \log_{2a} a$, $y = \log_{3a} 2a$, $z = \log_{4a} 3a$, show that : $xyz+1 = 2yz$

Solution:

L.H.S = $\log_{2a} a . \log_{3a} 2a . \log_{4a} 3a + 1$

$= (\log_{10} a \times \log_{2a} 10).(\log_{10} 2a \times \log_{3a} 10).(\log_{10} 3a \times \log_{4a} 10)+1$

$$= \frac{\log_{10} a}{\log_{10} 2a} \times \frac{\log_{10} 2a}{\log_{10} 3a} \times \frac{\log_{10} 3a}{\log_{10} 4a} + 1$$

$$= \frac{\log_{10} a}{\log_{10} 4a} + 1 = \log_{4a} a + \log_{4a} 4a = \log_{4a}(a.4a) = \log_{4a} 4a^2$$

R.H.S.= $2\log_{3a} 2a . \log_{4a} 2a = \log_{4a}(2a)^2 = \log_{4a} . 4a^2$

Hence the result $\log_{\frac{3a}{4a}}$

7. Show that $\log_3 \sqrt{3\sqrt{3\sqrt{3........\infty}}} = 1$.

Solution:

Let, $x = \sqrt{3\sqrt{3\sqrt{3......}}}$ or $x^2 = \sqrt{3\sqrt{3\sqrt{3........\infty}}}$ (squaring both sides)

Or $x^2 = 3x$ or $x^2 - 3x = 0$ or $x(x-3) = 0$ or $x-3 = 0$ (as $x \neq 0$).

$\therefore x = 3$. $\therefore$ given expression $= \log_3 3 = 1$

Practice problems

1. If $\frac{1}{2}\log_3 M + 3\log_a N = 1$, express M in terms of N. Ans: $m = 9^{-6}_{N}$
2. If $a^2 + b^2 = 7ab$, show that:
   (i) $2\log(a-b) = \log 5 + \log a + \log b$
   (ii) $2\log(a+b) = \log 9 + \log a + \log b$.
3. if $a^2 + b^2 = 23ab$, show that $\log \frac{1}{5}(a+b) = \frac{1}{2}\{\log a + \log b\}$
4. If $a = b^2 = c^3 = d^4$, prove that $\log_a(abcd) = 1 + \frac{1}{2} + \frac{1}{3} + \frac{1}{4}$
5. $(1+\log_n^m)(\log_{mn}^x) = \log_n^x$ (ICWA(F) june 2000)
6. Show that :- (i) $16\log\frac{16}{15} + 12\log\frac{25}{24} + 7\log\frac{81}{80} = \log^5$
7. Prove that $x^{\log y - \log z} \times y^{\log z - \log x} \times z^{\log x - \log y} = 1$.
8. Prove that (i) $\dfrac{\log\sqrt{27} + \log 8 + \log\sqrt{1000}}{\log 120} = \dfrac{3}{4}$ (ICWA(F) june 2000)
9. If $\dfrac{\log x}{y-z} = \dfrac{\log y}{z-x} = \dfrac{\log z}{x-y}$, show that $xyz = 1$
10. If $\log_a bc = x$, $\log_b ca = y$, $\log_c ab = z$. prove that $\dfrac{1}{x+1} + \dfrac{1}{y+1} + \dfrac{1}{z+1} = 1$
11. If $\dfrac{\log x}{l+m-2n} = \dfrac{\log y}{m+n-2l} = \dfrac{\log z}{n+l-2m}$, show that $xyz = 1$
12. If $p = a^x$, $q = a^y$ and $a^4 = (p^{4y}.q^{4x})^x$ prove that $xyz = \dfrac{1}{2}$
13. Prove that $\dfrac{\log 3\sqrt{3} + \log\sqrt{8} - \log\sqrt{125}}{\log 6 - \log 5} = \dfrac{3}{2}$

I. Choose the correct Answer

1. $\left[\log\left(\frac{a^2}{bc}\right)+\log\left(\frac{b^2}{ac}\right)+\log\left(\frac{c^2}{ab}\right)\right]$ is equal to:

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

2. $\left(\log_{b^a}\times\log_{c^b}\times\log_{a^c}\right)$ equal to :

(a) 0 (b) 1 (c) abc (d) a+b+c

3. $\left[\frac{1}{\log_{xy}(xyz)}+\frac{1}{\log_{yz}(xyz)}+\frac{1}{\log_{zx}(xyz)}\right]$ is equal to :

(a)1 (b) 2 (c) 3 (d) 4

4. $\left[\frac{1}{(\log_a bc)+1}+\frac{1}{(\log_b ca)+1}+\frac{1}{(\log_c ab)+1}\right]$ is equal to:

(a) 1 (b) 2 (c) 3/2

5. If $\log_2(\log_3(\log_2 x)=1$, then x is equal to :

(a) 512 (b) 128 (c) 12 (d) 0

6. $(\log_5 3)\text{x}(\log_3 625)$ equal:

(a) 1 (b) 2 (c) 3 (d) 4

7. $(\log_5 5)\,.\,(\log_4 9)\,(\log_3 2)$ is equal to:

(a) 2 (b) 1 (c) 5 (d) 3/2

8. If $\log_{10} 2 = 0.3010$ and $\log_{10} 3 = 0.4771$, then the value of $\log_{10} 1.5$ is

(a) 0.7161 (b) 0.1761 (c) 0.7116 (d) 0.7611

9. If $(\log_{10} 2) = 0.3010$, then $\log_2 10$ is:

(a) 0.3322 (b) 3.2320 (c) 3.3222 (d) 5

10. The value of $\left[\frac{1}{(\log_3 60)}+\frac{1}{(\log_4 60)}+\frac{1}{(\log_5 60)}\right]$ is:

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

II. Fill in the blanks

1. If $\log_x 0.1 = -\frac{1}{3}$ then the value of x is :______

2. If $\log_{32} x = 0.8$ then x is equal to :__________

3. If $\log_4 x + \log 2\, x = 6$, then x is equal to : ____________

4. If $\log_8 x + \log_8 \frac{1}{6} = \frac{1}{3}$ then the value of x is :____________

5. If log 2 = 0.30103, then the number of digits in 4^{50} is: ________

6. If log 2 = 0.30103, then the number of digits in 5^{20} is: ______

7. The value of $\log_{(-1/3)} 81$ is equal to: _______

8. The value of $\log_{2\sqrt{3}}(1728)$ is equal to: _____

9. The value of $\log_2(\log_5 625)$ is: ______

10. The value of $\left(\frac{1}{3}\log_{10}125 - 2\log_{10}4 + \log_{10}32\right)$ is : ______

IV. Match the following

(a) Group A	(b) Group B	
1. The of $\log_{343}$	-4	( )
2. The value of $\log_5\left(\frac{1}{125}\right)$	10	( )
3. The value of $\log_{\sqrt{2}}^{32}$	$\frac{1}{3}$	( )
4. The value of $\log_{0.1}(0.0001)$	-3	( )
5. The value of $\log_{10}(0.001)$	-3/2	( )

(a) Group A	(b) Group B	
6. If $\log_3^x = -2$ then x =	625	( )
7. If $\log_8^x = -2/3$ then x =	$\frac{1}{9}$	( )
8. The value of $\log_x\left(\frac{1}{25}\right) = -\frac{1}{2}$ then x	$\sqrt{2}$	( )
9. If $\log_{10000}^x = -\frac{1}{4}$ then x = -	1/4	( )
10. If $\log_4^x = \frac{1}{4}$ then x = -	$\frac{1}{10}$	( )

III. State whether the following statements are true (or) false

1. Log3 + log 5 is log 15 ( )
2. The value of $\log_2 \log_2 \log_3^{81}$ is 2 ( )
3. The logarithm of 324 to base $\frac{1}{3\sqrt{2}}$ is -4 ( )
4. The logarithms with base 10 are called Natural logarithm ( )
5. The logarithms with base e are called comm. Logarithm ( )
6. The logarithm of one to any base is zero ( )
7. The logarithm of a number which is not equal to one is itself as base is zero ( )
8. The integral part of the value of logarithm of a number is called characteristic ( )
9. The decimal part of the value of logarithm of a number is called mantissa ( )
10. The logarithm of a some number for different bases are different ( )

Key:

I. Choose the correct answer

1. a	2. B	3. B	4. A	5. A
6. D	7. B	8. b	9. C	10. B

II. Fill in the blanks

1. 1000	2. 16	3. 16	4. 12	5. 31
6. 14	7. -4	8. 6	9. 2	10. 1

III. State whether the following statements are true (or false)

1. T	2. F	3. T	4. F	5. F
6. T	7. F	8. T	9. T	10.T

IV. Match the following:

(1) 1/3	(2) -3	(3) 10	(4) -3/2	(5) -3
(6) 1/9	(7) ¼	(8) 625	(9) 1/10	(10) $\sqrt{2}$

2.3 PERMUTATIONS AND COMBINATIONS

PERMUTATIONS

Fundamental principle of counting:

If there are m way of doing a thing and for each of m ways there are associated n ways of doing the second thing then total no. of ways of doing both the things will be mn ways.

Example:

Suppose 6 subjects are to thought in 4 periods for the first period we can put any one of the 6 subjects i.e. there are 6 ways of filling first period.

For the second period they are left with remaining 5 subjects and hence there are 5 ways of filing second period.

$\therefore$ No. of ways in which the first two periods can be filled is $6 \times 5 = 30$ ways

Factorial:-

The product of first n natural numbers is nothing but n factorial and is denoted by n! or $\underline{|n}$.

i.e.

$n! = 1\times2\times\ldots\ldots\ldots\times n$

(or)

$= n\times(n-1)\times(n-2)\times\ldots\ldots\times2\times1.$

Note:

Factorials are defined only for positive integers.

Properties:-

1. $0! = 1$
2. $1! = 1$
3. $n! = n(n-1)(n-2)\ldots\ldots\ldots.3.2.1.$
4. $2n! = 2n(2n-1).(2n-2)\ldots\ldots\ldots3.2.1.$
5. $3n! = 3n(3n-1)\ (3n-2)\ \ldots\ldots\ldots3.2.1.$

Ex: $5! = 5\times4\times3\times2\times1 = 120$

$6! = 6\times5! = 120\times6 = 720$

$7! = 7\times6! = 720\times7 = 5040.$

Permutation:

The different arrangement that can be made out of given no. of things by taking some or all at a time is called permutation and is denoted by ^np_r and is defined as

$$^np_r = \frac{n!}{(n-r)!}$$

Here

n stands given total No. of things

P stands for permutation.

r stands for no. of things taken at a time.

Points to be remembered:

1. The no. of permutations of n different things taken all at a time is ${}^{n}p_{n} = n!$
2. The no. of permutations of n different things taken r at a time when particular things is always included in each arrangement is $r.{}^{n-1}p_{r-1}$.
3. The no. of permutations of n different things taken r at a time when particular things q always included in each arrangement is

   ${}^{r}p_{q}.{}^{n-q}p_{r-q}$
4. The no. of permutations of n different things taken r at a time when particular things q never comes in each arrangement is.

   ${}^{n-q}p_{r}$
5. If there are n things of which p are one kind and q are second kind and r third kind and remaining all the objects are different then total no. of arrangements is $\frac{n!}{p!q!r!}$
6. The no. of permutations of (n) different things taken r at a time when each may be repeated any no.of times in each arrangement is n^{r}

Circular permutation:

When the objects are arranged in a closed curve i.e. a circle then the permutation is called circular permutation.

- The no. of circular permutation of n different things taken all at a time is (n-1)!.
- The no. of circular permutations of n thing when clock wise and Anti-clock wise arrangements are not different is $\frac{(n-1)!}{2}$

Illustration

1. Find the values of (i) ${}^{7}P_{5}$ (ii) ${}^{7}P_{1}$ (iii) ${}^{7}P_{0}$ (iv) ${}^{7}P_{7}$

   (i) ${}^{7}P_{5} = \frac{7!}{(7-5)!} = \frac{7!}{2!} = \frac{7.6.5.4.3.2.1}{2.1} = 4.6.5.4.3 = 2520$

   (ii) ${}^{7}P_{1} = \frac{7!}{(7-1)!} = \frac{7!}{6!} = 7 \times \frac{6!}{6!} = 7$

   (iii) ${}^{7}P_{0} = \frac{7!}{(7-0)!} = \frac{7!}{7!} = 1$

   (iv) ${}^{7}P_{7} = \frac{7!}{(7-7)!} = \frac{7!}{0!} = \frac{7!}{1!} = 7.6.5.4.3.2.1. = 5040$

2. If ${}^{n}P_{2} = 110$, n.

   ${}^{n}P_{2} = 110$ or $\frac{n!}{(n-2)!} = 110$ or, $\frac{n(n-1)(n-2)!}{(n-2)!} = 110$

   or, $n(n-1) = 110 = 11 \times 10 = 11 \times (11-1)$ $\therefore$ n = 11

3. Solve for n given ${}^{n}p_4 = 30 \times {}^{n}p_2$

Solution: ${}^{n}p_4 = 30 \times {}^{n}p_2$ or, $\frac{n!}{(n-4)!} = 30 \times \frac{n!}{(n-2)!}$

Or, $\frac{n(n-1)(n-2)(n-3)(n-4)!}{(n-4)!} = 30 \times \frac{n(n-1)(n-2)!}{(n-2)!}$

Or, n(n-1)(n-2)(n-3) = 30×n(n-1) or, (n-2)(n-3) = 30

Or $n^2 - 5n - 24 = 0$ or, (n–8) (n+3) = 0

Or, n = 8, -3 (inadmissible)

4. Solve for n given $\frac{n_{P5}}{n_{P3}} = \frac{2}{1}$

Solution:

$\frac{n_{P5}}{n_{P3}} = \frac{2}{1}$ or, $n_{P5} = 2 \times n_{P3}$

or, $\frac{n!}{(n-5)!} = 2 \times \frac{n!}{(n-3)!}$ or, $1 = 2 \times \frac{1}{(n-3)(n-4)}$

$n^2 - 7n + 12 = 2$

$n^2 - 7n + 10 = 0 \quad \therefore n = 5, 2$

5. In how many ways 6 books out of 10 different books can be arranged in a book-self so that 3 particular books are always together?

A first 3 particular books are kept outside. Now remaining 3 books out of remaining 7 books can be arranged in ${}^{7}p_3$ ways. In between these three books there are 2 places and at the two ends there are 2 places i.e. total 4 places where 3 particular books can be placed in ${}^{4}P_1$ ways. Again 3 particular books can also be arranged among themselves in 3! Ways.

Hence, required no. of ways ${}^{7}P_3 \times {}^{4}P_1 \times 3! = \frac{7!}{4!} \times \frac{4!}{3!} \times 3! = 7.6.5.4 = 3.2.1 = 5040$

6. In how many ways can be letter of the word table be arranged so that the vowels are always (i) together (ii) separated?

(i) In the word there are 2 vowels, 3 consonants all different. Taking the 2 vowels (A, E) as one letter we are to arrange 4 letters (i.e 3 consonants + 1) which can be done in 4! Ways. Again 2 vowels can be arranged among themselves in 2! Ways. So no. of Ways = 41×21 = 48 Ways

(ii) Hence, required numbers of ways = 5! – 48 = 120 – 48 = 72.

7. Find how many ways can be letters of the PURPOSE be rearranged –

(i) keeping the positions of the vowels fixed;

(ii) without changing the relative order to the vowels and consonants.

(i) In the word, there are 3 vowels and 4 consonants. Since the positions of all vowels fixed are to rearrange only 4 consonants, in which there are 2P, so the arrangement is

$\frac{4!}{2!} = \frac{4 \times 3 \times 2!}{2!} = 4 \times 3 = 12$

(ii) The relative order of vowels and consonants unaltered means that vowel will take place of vowel and consonant will take place of consonant. Now the 3 vowels can be arranged among themselves in 3! Ways, while 4 consonants with 2P can be arranged in

$$\frac{4!}{2!} = \frac{4 \times 3 \times 2!}{2!} = 4 \times 3 = 12 \text{ ways}$$

So total number of ways of rearrangement in which the given arrangement is included = 3! ×12 = 6×12 = 72

Hence, required number of arrangement = 72-1=71

8. How many numbers between 5000 and 6000 can be formed with the digits 3,4,5,6,7,8? The number to be formed will be of 4 figures, further digit 5 is to be placed in 1st place (from left). Now the remaining 3 places can be filled up by the remaining 5 digits in ${}^5P_3$ ways.

   Hence, required no. $5P_3 \times \frac{5!}{2!}$ = 60.

9. In how many ways can be letters of the word SUNDAY be arranged? How many of them do not begin with S? How many of them do not begin with S, but end with Y?

   There are 6 letters in the word SUNDAY, which can be arranged in 6! = 720 ways.

   Now placing S in first position fixed, the other 5 letters can be arrange in (5)! =120 ways.

   The arrangements of letters that do not begin with S = (6)! – (5)! = 720 –120 = 600 ways

   Lastly. Placing Y in the last position, we can arrange in (5)! = 120 ways and keeping Y in the last position and S in the first position, we can arrange in (4)! = 24 ways.

   Hence, the required no. of arrangements = (5)! -4! = 120 – 24 = 96 ways.

   (problems regarding ring or circle)

10. In how many ways 8 boys can form a ring?

Keeping one boy fixed in any position, remaining 7 boys can be arranged in 7! Ways.

Hence, the required on. of ways = 7! = 7.6.5.4.3.2.1 = 5040.

11. In how many ways 8 different beads can be placed in necklace?

    8 beads can be arranged in 7! Ways. In this 7! Ways, arrangements counting from clockwise and anticlockwise are taken different, But necklace obtained by clockwise permutation will be same as that obtained from anticlockwise. So total arrangement will be half of 7!.

    Hence, required no. of ways = ½ ×7! = ½ ×5040 = 2520

12. In how many ways 5 boys and 5 girls can take their seats in a round table, so that no two girls will sit side by side.

    If one boy takes his seat anywhere in a round table, then remaining 4 boys can take seats in 4! = 24 ways. In each of these 24 ways, between 5 boys, if 5 girls take their seat then no two girls will be side by side. So in this way 5 girls may be placed in 5 places in 5! = 120 ways.

    Again the first boy taking seat, may take any one of the 10 seats. i.e. he may take his seat in 10 ways.

    Hence, required number ways = 24×120×10 = 28800

Practice question:

1. Find the value of (i) $^{10}P_2$ (ii) $^{10}P_0$ (iii) $^{10}P_{10}$
2. Find the value of n: $^{n}P_4 = 10\times {}^{n-1}P_3$
3. Find the value of r: (i) $^{11}P_{3r} = 110$ (ii) $^{7}P_r = 2520$
4. Find if $^{n}P_3 : {}^{n+2}P_3 = 5:12$ (ICWA (F)- june – 2005)
5. Prove the "CALCUTTA" is twice of "AMERICA" in respect of number of arrangements of letters.

Objective one (or) two steps questions

1. There are 20 stations on a railway line. How many different kinds of single first-class tickets must be printed so as to enable a passenger to go from one station to another? (Ans: 380)
2. Four travelers arrive in a town where there are six hotels. In how any ways can they take their quarters each at a different hotel? (Ans: 360)
3. In how many ways can 8 mangoes of different sizes be distributed amongst 8 boys of different ages so that the largest one is always given to the youngest boy? (Ans: 5040)
4. Find the number of different number of 4 digits that can be formed with the digits 1,2,3,4,5,6,7; the digits in any number being all different and the digit in the unit place being always 7. (Ans: 120)
5. How many different odd numbers of 4 digits can be formed with the digits 1,2,3,4,5,6,7; the digits in any number being all different? (Ans: 480)
6. Find the number of arrangements that can be made out of the letters of the following words:

   (a) COLLEGE

   (b) MATHEMATICS (Ans: (a) 1260; (b) 49, 89, 600)
7. In how many ways can the colours of rainbow be arranged, so that the red and the blue colours are always together?
8. In how many ways 3 boys and 5 girls be arranged in a row so that all the 3 boys are together? (Ans: 4320)
9. Find how many words can be formed of the letters in the word FAILURE so that the four vowels come together. (Ans;576)
10. In how many ways can be colours of the rainbow be arranged so that red and blue colours are always separated? (Ans: 3600)
11. In how many ways can be 5 boys from a ring? (Ans: 24)
12. In how many ways 5 different beads be strung on a necklace? (Ans: 12)

Objective Type Questions

I. Choose the correct Answer

1. If $^{11}P_r = 110$ then the value of r is:

   (a) 2 (b) 10 (c) 4 (d) none of these

2. If $^{n-1}P_3 : {}^{n+1}P_3 = 28:55$ then n :-

   (a) 6 (b) 8 (c) 10 (d) 12

3. If ${}^{m+n}P_2 = 42$, and ${}^{m-n}P_2 = 6$ then the values of m and n are :____
   (a) m = 6, n = 2 (b) m = 5 n = 2 (c) m = 6, n = 1 (d) none of these
4. If ${}^{9}P_5 + 5.{}^{9}P_4 = {}^{10}P_r$ then the value of r is
   (a) 3 (b) 4 (c) 5 (d) none of these
5. If ${}^{n}P_3 = 120$ then n:
   (a) 8 (b) 4 (c) 6 (d) none of these

II. Fill in the blanks:

6. The number of ways in which the letters of the word "VOWEL" can be arranged so that the letters O, E occupy even places is _____
7. 5 letters can be posted in 4 letter boxes in _______.
8. 3 distinct prizes can be distributed among 10 boys (any boy can get more than once) in ___
9. Total number of ways in which the letters of word strange can be arranged so that the vowels may appear in the odd places is _____.
10. The number of six letter word that can be formed using the letter of the word "assist" in which 5's alternate with other letters is ______.

III. State the following statements true (or) false.

11. The number of permutations if the letter in the word "BANANA" is which two letters N do not come together is 60. ( )
12. There 11 distinct books, among them 6 books can be arranged in a shelf. The number of arrangements so that 3 particular books will be always side by side is 8064. ( )
13. The number of different number of 6 digits (without repetition) can be formed form the digits 3,1,7,0,9,5 is 120 ( )
14. The total number of arrangements of the letters in the expression $x^3y^2z^4$ when written in full length is 1260. ( )
15. The number of different words that can be formed form the letter of the work "TRIANGLE" so that no vowels one together is 36000. ( )

IV. Matching the following

	Group A	Group B	
16.	${}^{n}c_r + {}^{n}c_{r-1}$ is equal to	${}^{n}c_r$	( )
17.	0 =	5	( )
18.	${}^{n}P_r$/ r!	4	( )
19.	If ${}^{n}P_3 = 60$ then n =	${}^{n+1}c_r$	( )
20.	If (n+1)! = 20(n-1)! Then n =	1	( )

Permutations Key

I. Choose the correct Answer

1. a	2. C	3.b	4.c	5.c

II. Fill in the blanks

6. 12	7. 1024 way	8. 1000 ways	9. 1440	10. 12

III. State whether the following statement are true (or) false

11. F	12. T	13. F	14. T	15. F

IV. Match the following

16. ${}^{n+1}c_r$	17. 1	18. ${}^{n}c_r$	19. N = 5	20. 4

Combinations:

Each of different groups of selection that can be made by taking some or all of no. of things at a time is called combination it is denoted by ${}^nc_r$ and is defined as ${}^nc_r = \frac{n!}{(n-r)!r!}$

Here n = given total no. of things

C = combination or selection

R = no. of things taken at a time

Example: In how many ways to form a committee of 2 boys from 4 boys

Solution: ${}^4c_2 = \frac{4!}{(4-2)!2!} = \frac{4\times3\times2}{4} = 6$

Relation between ${}^np_r$ and ${}^nc_r$:

$\therefore {}^np_r = \frac{n!}{(n-r)!}$

$\frac{n_{Pr}}{r!} = \frac{n!}{(n-r)!r!}$

$\frac{n_{Pr}}{r!} = {}^nc_r$, ${}^nP_r = r!\,{}^nc_r$

Properties:

1. ${}^nc_n = 1$
2. ${}^nc_0 = 1$
3. ${}^nc_r = {}^nc_{n-r}$

Proof:-

R.H.S

$$^{n}C_{n-r} = \frac{n!}{(n-n+r)!(n-r)!} = \frac{n!}{(n-r)!r!} = ^{n}C_{r} = R.H.S$$

4. $^{n}C_{r} = ^{n}C_{s}$

   Or, n = n + s = n = r+s (or) r = s

5. **Pascal's law:-**

   If n and r be two non-negative integers such that $1 \le r \le n$. then show that

   $^{n}C_{r-1} + ^{n}C_{r} = ^{n+1}C_{r}$

L.H.S :

$^{n}C_{r-1} + ^{n}C_{r}$

$$= \frac{n!}{(n-r+1)(n-r)!(r-1)!} + \frac{n!}{(n-r)!\,r(r-1)!}$$

$$= \frac{n!}{(n-r)!\,(r-1)!}\left[\frac{1}{n-r+1} + \frac{1}{r}\right]$$

$$= \frac{n!}{(n-r)!\,(r-1)!}\left(\frac{\cancel{r}+n-\cancel{r}+1}{(n-r+1)\,r}\right) = \frac{n!\,(n+1)}{(n-r+1)!\,r!} = ^{n+1}C_{r}$$

Formulae:

- No.of combinations of n dissimilar things taken r at a time when 'p' particular things always occurs is $^{n-p}C_{r-p}$
- No. of combinations of n dissimilar things taken 'r' at a time when P particular things never comes is $^{n-p}C_{r}$
- Total No. of combinations of n different things taken 1,2,3……n things at a time is $2^{n}-1$.

  i.e. $^{n}C_{1} + ^{n}C_{2} + ^{n}C_{3} + ^{n}C_{4} + \ldots\ldots + ^{n}C_{n} = 2^{n}-1$.

  Total no. of combinations of n different things taken one or more things at a time is $2^{n}-1$.

Grouping:

(A) If, it is required to form two groups out of (m + n) things, ($m \neq n$) so that one group consist of m things. Now formation of one group represents the formation of the other group automatically. Hence the number of ways m thing can be selected from (m + n) things.

$$= ^{m+n}C_{m} = \frac{(m+n)!}{m!\,(m+n-m)!} = \frac{(m+n)!}{m!n!}$$

Note:

1. If m=n, the groups are equal and in this case the number of different ways of subdivision $= \frac{2m!}{(m!)^{2}}$ since two groups can be interchanged without getting a new subdivision.
2. If 2m things be divided equally amongst 2 person, then the number of ways $\frac{2m!}{m!m!}$

(A) Now (m+n+p) things (m≠n≠p), to be divided in to three groups containing m,n,p things respectively.

M things can be selected out of (m+n = p) things in ${}^{m+n+p}C_m$ ways, then n things out of remaining

(n+p) things in ${}^{n+p}C_m \times {}^{n+p}C_n = \frac{(m+n+p)}{m!(n+p)!}!\frac{(n+p)!}{n!p!} = \frac{(m+n+p)!}{m!n!p!}$

Note:

1. if now m = n = p, the groups are equal and in this case, the different ways of subdivision = $\frac{(3m)!}{m!m!m!} \times \frac{1}{3!}$ since the three groups of subdivision can be arranged in 3! Ways.

Note 2:

If 3m things are divided equally amongst three persons, the number of ways = $\frac{(3m)!}{m!m!m!}$

Illustrations:

1. In how many ways can be college football team of 11 players be selected from 16 players?

   The required number = ${}^{16}C_{11} = \frac{16!}{11!(16-11)!} = \frac{16!}{11!5!} = 4,368$

2. From a company of 15 men, how many selections of 9 men can be made so as to exclude 3 particular men?

   Excluding 3 particular men in each case, we are to select 9 men out of (15-3) men. Hence the number of selection is equal to the number of combination of 12 men taken 9 at a time which is equal to

   ${}^{12}C_9 = \frac{12!}{9!3!} = 220$

3. There are seven candidates for a post. In how many ways can a selection of four be made amongst them, so that:

   1. 2 persons whose qualifications are below par are excluded?
   2. 2 persons with good qualifications are included?
   3. Excluding 2 persons, we are to select 4 out of 5(=7-2) candidates. Number of possible selections = ${}^5C_4 = 5$.
   4. In this case, 2 persons are fixed, and we are to select only 2 persons out of (7-2), i.e. 5 candidates. Hence the required number of selection $= {}^5C_4 = 10$.

Committee from more than one group:

4. In how many ways can a committee of 3 ladies and 4 gentlemen be appointed from a meeting consisting of 8 ladies and 7 gentlemen? What will be the number of ways if Mrs. X refuses to serve in a committee having Mr. Y as a member?

   1st part. 3 ladies can be selected from 8 ladies in ${}^8C_3 = \frac{8!}{3!\,5!} = 56$ ways and

   4 gentlemen can be selected from 7 gentlemen in ${}^7C_4 = \frac{7!}{4!\,3!} = 35$ ways

   Now, each way of selecting ladies can be associated with each way of selecting gentlemen. Hence, the required no. of ways = 56 x 35 = 1960.

2nd part: if both Mrs. X and Mrs. Y are members of the committee then we are to select 2 ladies and 3 gentlemen from 7 ladies and 6 gentlemen respectively. Now 2 ladies can be selected out of 7 ladies in ${}^7C_2$ ways, and 3 gentlemen can be selected out of 6 gentlemen in ${}^6C_3$ ways.

Since each way of selecting gentlemen can be associated with each way of selecting ladies.

Hence, No. of ways = ${}^7C_2 \times {}^6C_3 = \frac{7!}{2!\ 5!} \times \frac{6!}{3!\ 3!} = 420$

Hence, the required no. of different committees, not including Mrs. X and Mr. Y = 1960 – 420 =1540.

5. From 7 gentlemen and 4 ladies a committee of 5 is to be formed. In how many ways can this be done to include at least one lady?

(i) 1 lady and 4 gentlemen

(ii) 2 ladies and 3 gentlemen

(iii) 3 ladies and 2 gentlemen

(iv) 4 ladies and 1 gentleman

For (i), 1 lady can be selected out of 4 ladies in ${}^4C_1$ ways and 4 gentlemen can be selected from 7 gentlemen in ${}^7C_4$ ways. Now each way of selecting lady can be associated with each way of selecting gentlemen. So 1 lady and 4 gentlemen can be selected in ${}^4C_1 \times {}^7C_4$ ways.

Similarly,

Case (ii) can be selected in ${}^4c_2 \times {}^7c_3$ ways.

Case (iii) can be selected in ${}^4c_3 \times {}^7c_2$ ways.

Case (iv) can be selected in ${}^4c_4 \times {}^7c_1$ ways.

Hence the total number of selections, in each case of which at least one lady is included

$= {}^4C_1 \times {}^7C_4 + {}^4C_2 \times {}^7C_3 + {}^4C_3 \times {}^7C_2 + {}^4C_4 \times {}^7C_1$

$= 4 \times 35 + 6 \times 35 + 4 \times 21 + 1 \times 7$

$= 140 + 210 + 84 + 7 = 441.$

6. In how many ways can a boy invite one or more of 5 friends?

The number of ways = ${}^5C_1 + {}^5C_2 + {}^5C_3 + {}^5C_4 + {}^5C_5 = 2^5 - 1 = 32 - 1 = 31.$

7. In a group of 13 workers contains 5 women, in how many ways can a subgroup of 10 workers be selected so as to include at least 6 men?

In the given group there are 8(=13-5) men and 5 women in all. Possible cases of forming the subgroup of 10 workers.

(i)	6	4	${}^8C_6 \times {}^5c_4$	= 28×5 =140
(ii)	7	3	${}^8C_7 \times {}^5C_3$	= 8×10 = 80
(iii)	8	2	${}^8C_8 \times {}^5C_2$	= 1×10 = 80

$\therefore$ required no of ways = 230

8. In how many ways 15 things be divided into three groups of 4,5,6 things respectively.

The first group can be selected in ${}^{15}C_4$ ways:

The second group can be selected in ${}^{(15-4)}C_5 = {}^{11}C_5$ ways;

And lastly the third group in ${}^6C_6$ =1 way.

Hence the total number of ways $= {}^{15}C_4 \times {}^{11}C_5 = \frac{15!}{4!\ 11!} \times \frac{11!}{5!\ 6!} = \frac{15!}{4!\ 5!\ 6!}$

9. A student is to answer 8 out of 10 questions on an examination:

(i) How many choice has he?

(ii) How many if he must answer the first three questions?

(iii) How many if he must answer at least four of the first give questions?

(i) the 8 questions out of 10 questions may be answered in ${}^{10}C_8$

Now ${}^{110}C_8 = \frac{10!}{8!2!} = \frac{10 \times 9 \times (8)!}{8!2!} = 5 \times 9 = 45$ ways

(ii) The first 3 questions are to be answered. So there are remaining 5(= 8-3) questions to be answered out of remaining 7(=10-3) questions which may be selected in ${}^{7}C_5$ ways.

Now, ${}^{7}C_5 = 21$ ways

(ii) Here we have the following possible cases:

(a) 4 questions from first 5 question (say, group A), then remaining 4 questions from the balance of 5 questions (say, group B).

(b) Again 5 questions from group A, and 3 questions from group B.

For (a), number of choice is ${}^{5}C_4 \times {}^{5}C_4 = 5 \times 5 = 25$

For (b), number of choice is ${}^{5}C_5 \times {}^{5}C_3 = 1 \times 10 = 10$

Hence, required no. of ways = 25+10=35.

10. Given n points in space, no three of which are collinear and no four coplanar, for what value of no will the number f straight lines be equal to the number of planes obtained by connecting these points?

Since no three points, are collinear, the number of lines = number of ways in which 2 points can be selected out of n points

$= {}^{n}C_2 = \frac{n(n-1)}{2}$ lines

Again since three non-collinear points define a space and no four of the points are coplanar; the number of planes = number of ways in which 3 points can be selected out of n points

$= {}^{n}C_3 = \frac{n(n-1)(n-2)}{6}$

Now, we have $= \frac{n(n-1)}{2} = \frac{1}{6}n(n-1(n-2))$ or, 6 = 2(n – 2) Hence, n = 5

Practice questions:

1. Out of 16 men, in how many ways a group of 7 men may be selected so that:
   (i) particular 4 men will not come,
   (ii) particular 4 men will always come? (Ans: 729, 220)

2. A person has got 15 acquaintances of whom 10 are relatives. In how many ways may be invite 9 guests so that 7 of them would be relatives?

3. A question paper is divided in three groups A, B and C each of which contains 3 questions, each of 25 marks. One examinee is required to answer 4 questions taking at least one from each group. In how many ways he can choose the questions to answer 100 marks. (ICWA(F) Dec. 2004
(Ans: 8)

Hints: $({}^{3}c_1 \times {}^{3}c_1 \times {}^{3}c_2) + ({}^{3}c_1 \times {}^{3}c_2 \times {}^{3}c_1) + ({}^{3}c_2 \times {}^{3}c_1 \times {}^{3}c_1)$ etc.

4. Out of 5 ladies and 3 gentlemen, a committee of 6 is to be selected. In how many ways can this be done : (i) when there are 4 ladies, (ii) when there is a majority of ladies ? (I C W A (F) Dec. 2006) (Ans: 15,18)
5. A cricket team of 11 players is to be selected from two groups consisting of 6 and 8 players respectively. In how many ways can the selection be made on the supposition that the group of six shall contribute no fewer than 4 players? (ANS: 344)
6. There are 5 questions in group A, 5 in group B and 3 in C. In how many ways can you select 6 questions taking 3 from group A, 2 from B, and 1 from group C. (Ans: 180)
7. A question paper is divided into three groups A, B C which contain 4,5 and 3 questions respectively. An examinee is required to answer 6 questions taking at least 2 from A, 2 from B, 1 from group C. In how many ways he can answer. (ICWA(F) June 2007) (Ans: 480)
8. In how many ways can a person choose one or more of the four electrical appliances; T.V., Refrigerator, Washing machine, radiogram? (ICWA (F) June 1979) (Ans: 15)
9. In how many way can 15 things be divided into three groups of 4,5,6 things respectively?
10. Out of 10 consonants and 5 vowels, how many different words can be formed each consisting 3 consonants and 2 vowels.

Objective questions:

I. One or two steps problems:

1. If ${}^{18}c_r = {}^{18}c_{r+2}$ find the value of ${}^{r}c_5$ (Ans: 56)
2. If ${}^{n}p_r = 336$, ${}^{n}c_r = 56$, find n and r (Ans: 8, 3)
3. ${}^{2n}c_3 : {}^{n}c_2 = 44:3$, find n (Ans: 6)
4. Pove that ${}^{10}P_{10} \times {}^{22}P_{12} = {}^{22}P_{10}$ (ICWA (F) Dec 2006)
5. If $x \neq y$ and ${}^{11}c_x = {}^{11}P_y$, find the value of (x+y) (Ans: 11)
6. If ${}^{r}c_{12} = {}^{r}c_8$ find ${}^{22}C_r$ (Ans: 231)

(ICWA(F) june 2005

II. Choose the correct Answer:

1. If ${}^{n}C_{12} = {}^{n}C_8$ then n:___

   (a) 20 (b) 12 (c) 6 (d) none of these
2. If ${}^{8}C_r - {}^{7}C_3 = {}^{7}c_2$ then r= ____

   (a) 3 (b) 4 (c) 2 (d) 6
3. If ${}^{15}C_r : {}^{15}C_{r-1} = 11:5$ then r= ______

   (a) 4 (b) 5 (c) 6 (d) 7
4. A man has 6 friends. The total number of ways so that he can invite one (or) more of his friends is equal to

   (a) 64 (b) 60 (c) 720 (d) 63
5. Every body in a room shakes hands with every body else. The total number of hands shakes is 66. The total number of person in the room is ___

   (a) 11 (b) 12 (c) 10 (d) 14

III. Fill in the blanks

6. These are 10 lamps in a room. Each one of them can be switched on independently. The number of ways in which the hall can be illuminated is ______

7. These are 11 questions in a examination paper of mathematics. A candidate has to answer 6 questions of which the question under 1 is compulsory. The total number of selections of his answering in questions is __

8. There are two groups in a question paper, each group contains 7 questions. A candidate has to answer questions but taking not more than 5 from any group. The total number of selections of 9 questions is ___

9. There are 10 points in plane and among them 4 are collinear. The total number of triangles formed by joining them is ______

10. Out of 18 points in plane, no three are in the same straight line except 5 points which are collinear. Then the number of straight lines obtained by joining them is _____

IV. Statement whether the following statements are true (or) false

11. A polygon has 44 diagonals then the number of its sides are 8 ( )

12. The total number of 9 digits numbers which have all different digits id 9 x 9 ( )

13. There are 8 questions in an examination paper and each question has an alternative. The number of ways in which a student can give his answer is 6561 ( )

14. In a foot ball competition there were 153 matches. A match occurs between two terms. The toal numbers of teams took part in the competition is 18 ( )

15. If ${}^{n}C_{n}$ =1 then 0! =1 ( )

V. Match the following

	Group A	Group B
16.	If ${}^{n}p_{3}= 2\ {}^{n-1}p_{3}$ then n =	8 ( )
17.	If ${}^{n}c_{n-2}= 21$ then n =	2 ( )
18.	If ${}^{n}p_{2}= 56$ then n =	6 ( )
19.	$4p_{2} \div 4c_{2}$	3 ( )
20.	If ${}^{n}p_{r} = 210\ {}^{n}c_{r} = 35$ then r =	7 ( )

Combinations KEY

II. Choose the correct Answer

1. A	2. A	3. B	4. D	5. B

III. Fill in the blanks

6. 1023	7.252	8. 1470	9. 116	10. 144

III. State whether the following statements are true or false

11.F	12.T	13.F	14.T	15.T

IV. Match the following

16. 6	17.7	18. 8	19. 2	20. 3

2.4 QUADRATIC EQUATIONS

QUADRATIC EQUATIONS

Quadratic Expression:

An expression which is of the form $ax^2 + bx + c$ is called quadratic expression here a,b,c, $\in R$

Ex: $2x^2 + 4x + 7$ is a quadratic expression.

Ex: $(2+i)x^2 - 3ix + 6$ is also quadratic expression.

Quadratic equation:

An equation which is of the form $ax^2 + bx + c = 0$ is called quadratic equation.

Ex: $2x^2 + 6x = 5x + 4$

$2x^2 - x - 4 = 0$

Roots of a quadratic equation:

If α is a complex number is called solution or root of the quadratic equation.

$ax^2+bx+c = 0$

$\rightarrow a\alpha^2 + b\alpha + c = 0.$

Solution methods of quadratic equation:

1. **Factorization method**
2. **Completing square method:**

Find the roots of the quadratic equation $ax^2+bx+c = 0$

Given

Quadratic equation $ax^2+bx+c = 0$ through multiplied by 4a.

$4a^2x^2+4abx+4ac = 0$

$4a^2x^2+4abx = -4ac.$

Adding b^2on b,s

$4a^2x^2+4abx+b^2 = b^2-4ac.$

$(2ax)^2+2(2abx)+b^2 = b^2-4ac.$

$(2ax+b)^2 = b^2-4ac.$

$2ax+b = \pm\sqrt{b^2 - 4ac}$

$2ax = -b \pm \sqrt{b^2 - 4ac}$

$= \dfrac{-b \mp \sqrt{b^2 - 4ac}}{2a}$

$\therefore \dfrac{-b + \sqrt{b^2 - 4ac}}{2a}, \dfrac{-b - \sqrt{b^2 - 4ac}}{2a}$ are the roots of given quadratic equation.

Note: here b^2-4ac is called the discriminate and it describes the nature of the roots of quadratic equations and is denoted by Δ.

i.e $\Delta = b^2 - 4ac$.

Nature of roots:

1. If $b^2 - 4ac > 0$ then the roots are real and unequal.
2. If $b^2 - 4ac = 0$ then the roots are real and equal.
3. If $b^2 - 4ac < 0$ then the roots are imaginary or conjugate complex numbers.
4. If $b^2 - 4ac > 0$ and perfect square then the roots are rational and unequal.
5. If $b^2 - 4ac > 0$ but not a perfect square the roots are irrational numbers.
6. If one root of a quadratic equation $ax^2+bx+c = 0$ is a+Ib then the other root will be a-Ib.
7. If one root of quadratic equation $ax^2+bx+c = 0$ is $a+i\sqrt{b}$ then the other root is $a-i\sqrt{b}$.

To find the quadratic equation having roots $\propto_1$ β:

Required Quadratic equation is $(x-\propto)(x-\beta)=0$

$$= x^2 - x(\propto + \beta) + \propto\beta = 0$$

$$= x^2 - sx + p = 0.$$

Here

S- sum of roots

P- product of roots.

Relation between roots and co-efficient of quadratic equation $ax^2 + bx + c = 0$

Let $\propto, \beta$ be the roots of a quadratic equation $ax^2 + bx + c = 0$. Now

$$\propto = \frac{-b+\sqrt{b^2-4ac}}{2a} \quad \beta = \frac{-b-\sqrt{b^2-4ac}}{2a}$$

Case (i): Sum of the roots

$$\propto + \beta = \frac{-2b}{2a} = \frac{-b}{a} = \frac{-\text{ coefficient of x}}{\text{constant term}}$$

Case (ii): Product of the roots

$$\propto\beta = \left(\frac{-b+\sqrt{b^2-4ac}}{2a}\right)\left(\frac{-b-\sqrt{b^2-4ac}}{2a}\right)$$

$$= \frac{b^2-b^2+4ac}{4a^2}$$

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

Special cases:

If the constant term is vanishes then one of the root is zero.

Explanation:-

$\therefore$ c = 0

and $\propto\beta = \frac{c}{a} = \frac{0}{a} = 0$

$\propto\beta = 0$

$\propto = 0$ or $\beta = 0$

If the coefficient of x is vanishes then the roots are equal in magnitude and opposite in sign.

Explanation:-

$\therefore$ b = 0

and $\propto + \beta = \frac{-b}{a} = \frac{0}{a} = 0^3$

$\propto = -\beta$ or $\beta = -\propto$

Illustrations:

1. Solve x^2-7x+12=0

Solution:

This can be expressed as x^2-3x-4x+12 = 0 or, x(x-3)- 4(x-3) = 0

Or, (x-3) (x-4) = 0 Hence, x = 3, 4

Alternatively , x = $\frac{-(-7) \pm \sqrt{49-48}}{2}$. Hence , x = 4, 3. (here a = 1, b = −7, c = 12)

2. Solve $x^4 - 10x^2 + 9 = 0$

Solution:

Taking x^2 = u, we get u^2- 10u+9 = 0

Or, (u-9)(u-1) = 0; either (u-9) = 0 or, (u-1) = 0 Hence, u = 9,1.

When U = 9, x^2 = 9 or x = ±3

Again, U = 1, x^2 = 1 or x = ±1

Here the power of x is 4, so get four values of x.

3. Solve $(1+x)^{1/3} + (1-x)^{1/3} = 2^{1/3}$

Soution:

We get $(1+x) + (1-x) + 3(1+x)^{1/3}(1-x)^{1/3}.\{(1+x)^{1/3} + (1-x)^{1/3}\} = 2$ (cubing sides)

Or, $2+3(1-x^2)^{1/3}.\ 2^{1/3} = 2$ or, $3(1-x)^{1/3}.\ 2^{1/3} = 0$

Or, $(1-x^2)^{1/3} = 0$, as $3.2^{1/3} = 0$

Or, $1-x^2 = 0$ (cubing again)

Or $x^2 = 1$ $\therefore$ x = ±1

4. Solve $\frac{6-x}{x^2-4}=\frac{x}{x+2}+2$

Solution:

Multiplying by the L.C.M or the denominators, we find:

$6-x = x(x-2) + 2(x^2-4)$ or, $3x^2-x-14 = 0$

Or, $(3x-7)(x+2)=0$

$\therefore$ either $3x-7=0$ or $x+2=0$

$\therefore\ x=\frac{7}{3}$ or, -2

Now $x = -2$ does not satisfy the equation $x=\frac{7}{3}$ is the root of the equation

5. Solve $4^x-3.2^{x+2}+2^5 = 0$.

Solution:

Here $2^{2x}-3.2^x.2^2+32 = 0$

Or, $(2^x)^2-12.2^x+32 = 0$ or, $U^2-12u+32 = 0$ (taking $u = 2^x$)

Or, $(u-4)\ (u-8) = 0 \quad \therefore\ u = 4$ or $x^2 = 4 = 2^2$ or $x = 2$

Again $u = 8$, $2^x = 8 = 2^3$ or $x = 3$

6. Solve $x^2+7x+\sqrt{x^2+7x+9} = 3$

Solution:

Adding 9 to both sides, we have $x^2+7x+9+\sqrt{x^2+7x+9} = 12$.

Now putting $u = \sqrt{x^2+7x+9}$, the equation reduces to

$U^2 = u + 12$ or $U^2 + U -12 = 0$

Or, $U^2 + 4U - 3U - 12 = 0$ or, $u(u+4) -3(u+4) = 0$

Or, $(u-3)\ (u+4) = 0\ u = 3, -4$

Since is not negative, we reject the value – 4 for u

$u = 3$

$\sqrt{x^2+7x+9} = 3$

Or, $x^2 + 7x + 9 = 9$

Or, $x^2 + 7x = 0$

Or, $x^2 = -7x$

Or, $x = -7$

$u = -4$

$\sqrt{x^2+7x+9} = -4$

Or, $x^2 + 7x + 9 = 16$

Or, $x^2 + 7x - 7 = 0$

$x = \frac{-7\pm\sqrt{49+28}}{2}$

7. Solve $(x^2+3x)^2 + 2(x^2 + 3x) = 24$

Solution:

Let $x^2 + 3x = u$, so that equation becomes $u^2+2u = 24$ or, $u^2+2u - 24 = 0$

Or, $u^2+6u-4u-24 = 0$ or $u(u+6)-4(u+6) = 0$

Or, $(u+6)(u-4) = 0$ or, $u = -6$, $u = 4$

For $u = -6$, $x^2+3x = -6$ or, $x^2+3x+6 = 0$

$x = \frac{-3 \pm \sqrt{9-24}}{2} = \frac{-3 \pm \sqrt{-15}}{2}$, rejected as values are not real

For $u = 4$, $x^2+3x = 4$ or $x^2+3x-4 = 0$

Or $(x+4)(x-1) = 0$ or, $x = -4, 1$.

8. Solve $2\left(x+\frac{1}{x}\right)^2 - 7\left(x+\frac{1}{x}\right) + 5 = 0$

Expression is $2u^2-7u+5 = 0$ or $2u^2-5u-2u+5 = 0$, $x + \frac{1}{x} = u$

Or, $u(2u-5)-1(2u-5) = 0$ or, $(2u-5)(u-1) = 0$

Either $2u-5 = 0$ or, $u-1 = 0$ i.e., $u = \frac{5}{2}$

$x + \frac{1}{x} = \frac{5}{2}$ or $\frac{x^2+x}{x} = \frac{5}{2}$

$2x^2 + 2x = 5x$

Or, $2x^2+2 = 5x$ or, $2x^2-5x+2 = 0$ or $2x^2-4x-x+2 = 0$

Or $2x(x-2) - 1(x-2) = 0$ or $(x-2)(2x-1) = 0$

Either $x-2 = 0$ or, $2x-1 = 0$ i.e., $x = 2, \frac{1}{2}$

Again for $u = 1$, we get $x + \frac{1}{x} = 1$ or, $\frac{x^2+1}{x} = 1$ or $x^2 + 1 = x$

Or, $x^2 - x + 1 = 0$

$\therefore \frac{1 \pm \sqrt{1-4}}{2} = \frac{1 \pm \sqrt{1-3}}{2}$ (the values are rejected as they are not real)

Practice questions

1. $6x^2-11x-10 = 0$	Ans: (5/2, -2/3)
2. $(2x-1)^{1/3} = (6x-5)^{1/3}$	Ans: 1
3. $\left(x+\frac{1}{x}\right) = \frac{10}{3}$	Ans: (3, 1/3)
4. $4x + \frac{4}{x} = 17$	Ans: (4, 1/4)
5. $\frac{1}{x+a+b} = \frac{1}{x} + \frac{1}{a} + \frac{1}{b}$	Ans: (-a, -b)
6. $\frac{x}{3} + \frac{3}{x} = 4\frac{1}{4}$	Ans: (12, 3/4)

7. $\frac{7x}{6}-\frac{30}{x}=\frac{x}{3}$ Ans: (±6)

8. $3^x+\frac{1}{3y}=\frac{10}{3}$ Ans: (±1)

9. $2x^{-1}+x^{-1/2} = 6$ Ans: (1/4, 4/9)

10. $2^{x-2}+2^{3-x} = 3$ Ans: (2, 3)

11. $\sqrt{\frac{x}{1-x}}+\sqrt{\frac{1-x}{x}}=\frac{13}{6}$ Ans: (9/13, 4/13)

I. Choose the correct answer

1. $3x^2+6x+3 = 0$ then the roots of the equations are –

   (a) (3,3) (b) (-1,-1) (c) (2,4) (d) (4,1)

2. If $16x^2-8x+1 = 0$ when $x=\frac{1}{4}y$ Find the value of y___

   (a) ¼ (b) 1 (c) 2 (d) -1/4

3. If the roots of the equations $\frac{3}{4}x^2+9x+c^3=0$ are equal then c is equal to ___

   (a) 5 (b) 3 (c) 8 (d) 5

4. If $(3-\sqrt{3})$ is one of the roots of an equations then the equation is ___

   (a) $x^2-2x-3 = 0$ (b) $x^2-3x-1 = 0$ (c) $x^2-4x+2 = 0$ (d) $x^2-6x+6 = 0$

5. If one root of the equation $x^2-3x+m = 0$ exceeds the other by 5 then the value of M is equal to __

   (a) -6 (b) -4 (c) 12 (d) 18

Fill in the blanks

6. If the equations $x^2+7x+12 = 0$ and $x^2+mx+5 = 0$ have common roots the value of m is equal to ___
7. The least positive value of m for which the equation $x^2+mx+4 = 0$ has real roots ____
8. The value of M if one root is 2 of $f(x) = 2x^2+4x-6 = 0$. _____
9. The value of M for which the difference between the roots of the equation $x^2+mx+8 = 0$ is 2 are ____
10. If p,q are the roots of the equation $f(x) = 6x^2+x-2$ the value of $\frac{p}{q}-\frac{p}{q}=$

State whether the following statements are true of false

11. If the roots of the equations $2x^2+8x+c = 0$ are equal then c = 8 ( )
12. The g.c.d of the equations $2x^2-x-1= 0$ and $4x^2+8x+3 = 0$ is 3x+1 ( )
13. $X^2-4x-1= 0$ is the quadratic equation whose roots are $2+\sqrt{5}$ and $2-\sqrt{5}$ ( )
14. The roots of the equation $(x-4)^2(x-2)(x+4)$ are 4,4,2,-2 ( )

15. The degree of the equation $3x^5+xyz^2+y^3$ in 3 ( )

IV. Match the following

Group A	Group B1	
16. If p,q are the roots of the equation $x^2+k+1 = 0$ then $\frac{1}{p}+\frac{1}{q}=$______	$4x/x^2-1$	( )
17. The degree of the equation $4x^2+xyz^2+xy^3+yz^5$ is	3/2	( )
18. If $4x^2 - 8x + 3 = 0$ then $x = \frac{1}{2}y$ find the value of y = _____	6	( )
19. For what value of K the equations $3x+2y = 6$, $(k+1)x+4y = (2k+2)$ have infinite solutions	-1	( )
20. If $A = \frac{x+1}{x-1}$ then $A = \frac{1}{A}$ = _______	5	( )

Quadratic Equations KEY

I. Choose the correct answer

1. b	2. A	3. B	4. D	5. b

II. Fill in the blanks

6. 21/4, 14/3	7. 4	8. -1	9. ± 6	10. -7/25

III. State whether the following statements are true of false

11. T	12. F	13. T	14. F	15. F

IV. Match the following

16. -1	17.6	18. 3/2	19. 5	20. $4x/x^2-1$

Section B
Fundamentals of Business Statistics
(Syllabus - 2016)

Study Note - 3

STATISTICAL REPRESENTATION OF DATA

This Study Note includes

3.1 INTRODUCTION TO STATISTICS

The word 'Statistics' has been derived from the Latin word 'Status' which means a political state. It has also its root either to the Italian word 'Statista' or the German word 'Statistik' each one of which means a political state. For several decades, the word 'statistics' was associated solely with the display of facts and figures pertaining to the economic, demographic and political situations prevailing in a country, usually, collected and brought out by the local governments.

Statistics is a tool in the hands of mankind to translate complex facts into simple and understandable statements of facts.

Meaning and definition of Statistics:

Meaning of statistics: The word Statistics is used in two different senses - Plural and singular. In its plural form, it refers to the numerical data collected in a systematic manner with some definite aim or object in view such as the number of persons suffering from malaria in different colonies of Delhi or number of unemployed girls in different states of India and so on. In Singular form, the word statistics means the science of statistics that deals with the principles, devices or statistical methods of collecting, analyzing and interpreting numerical data.

Thus, 'statistics' when used in singular refers to that branch of knowledge which implies Applied Mathematics.

The science of statistics is an old science and it has developed through ages. This science has been defined in different ways by different authors and even the same author has defined it in different ways on different occasions.

It is impossible to enumerate all the definitions given to statistics both as "Numerical Data i.e., Plural Form: and "Statistical Methods, i.e., Singular Form". However, we have give below some selected definitions of both the forms.

Definitions of "Statistics in Plural Form or Numerical Data": Different authors have given different definitions of statistics. Some of the definitions of statistics describing it quantitatively or in plural form are:

"Statistics are the classified facts representing the conditions of the people in a state especially those facts which can be stated in number or in a table of numbers or in any tabular or classified arrangement.

This definition is narrow as it is confined only to the collection of the people in a state. But the following definition given by Secrist is modern and convincing. It also brings out the major characteristics of statistical data.

"By Statistics we mean the aggregate of facts affected to a marked extent by multiplicity of causes, numerically expressed, enumerated or estimated according to reasonable standards of accuracy collected in a systematic manner for a pre-determined purpose and placed in relation to each other"

This definition makes it clear that statistics (in plural form or numerical data) should possess the following characteristics.

I. Statistics are aggregate of facts

II. Statistics are affected by a large number of causes

III. Statistics are always numerically expressed

IV. Statistics should be enumerated or estimated

V. Statistics should be collected in a systematic manner

VI. Statistics should be collected for a pre-determined purpose

VII. Statistics should be placed in relation to each other.

Statistics as Statistical methods or Statistics in Singular Sense:

We give below the definitions of statistics used in singular sense, i.e., statistics as statistical methods.

Statistical methods provide a set of tools which can be profitably used by different sciences in the manner they deam fit. The term statistics in this context has been defined differently by different authors. A few definitions are given below:

"Statistics may be called the science of counting"

This definition covers only one aspect, i.e., counting, but the other aspects such as classification, tabulation, etc., have been ignored. As such, the definition is inadequate and incomplete

"Statistics may be defined as the collection, presentation, analysis and interpretation of numerical data"

This definition given by Croxton and Cowden is simple, clear and concise.

According to this definition, there are four stages – collection of data, and presentation of data, analysis of data, and interpretation of data. However, one more stage may be added and that is the organization of data. Thus, there are four stages:

1. **Collection and Organization of data**: Three are various methods for collecting the data such as census, sampling, primary and secondary data etc.

2. **Presentation of data**: The mass data collected should be presented in a suitable, concise form as the mass data collected is difficult to understand and analyse

3. **Analysis of data**: The mass data collected should be presented in a suitable, concise form for further analysis. Analysis includes condensation, summarisation conclusion, etc., through the means of measures of central tendencies, dispersion, skewness, kurtosis, correlation, regression, etc.

4. **Interpretation of data**: The last step is the drawing conclusions from the data collected as the figures do not speak for themselves.

Having briefly discussed some of the definitions of the term statistics and having seen their drawbacks we are now in a position to give a simple and complete definition of the 'Statistics' in the following words:

Statistics (as used in the sense of data) are numerical statements of facts capable of analysis and interpretation and the science of statistics is a study of the principles and methods used in the collection, presentation, analysis and interpretation of numerical data in any sphere of enquiry.

Importance and Scope of Statistics:

I. Statistics and Economics: According to Prof. Alfred Marshall, "Statistics are the straws out of which I like every other economist, have to make bricks." The following are some of the fields of economics where statistics is extensively used.

(a) **Consumption**: Statistical data of consumption enable us to find out the ways in which people in different strata of society spend their incomes.

(b) **Production**: The statistics of production describe the total productivity in the country. This enables us to compare ourselves with other countries of the world.

(c) **Exchange**: In the field of exchange, an economist studies markets, laws of prices are determined by the forces of demand and supply, cost of production, monopoly, competition, banking etc. A systematic study of all these can be made only with the help of statistics

(d) **Econometrics**: With the help of econometrics, economics has become exact science. Econometrics is the combination of economics, mathematics and statistics.

(e) **Public Finance:** Public finance studies the revenue and expenditure activities of a country. Budget, (a statistical document), fiscal policy, deficit financing, etc., are the concepts of economics which are based on statistics.

(f) **Input-Output Analysis**: The input-output analysis is based on statistical data which explain the relationship between the input and the output. Sampling, Time series, Index numbers, Probability, Correlation and Regression are some other concepts which are used in economic analysis.

II. Statistics and Commerce:

Statistical methods are widely applied in the solution of most of the business and trade activities such as production, financial analysis, costing, manpower, planning, business, market research, distribution and forecasting etc. A shrewd businessman always makes a proper and scientific analysis of the past records in order to predict the future course of the business conditions. Index numbers help in predicting the future course of business and economic events. Statistics or statistical methods help the business establishments in analysing the business activities such as:

(a) **Organization of Business**: Businessman makes extensive use of statistical data to arrive at the conclusion which guides him in establishing a new firm or business house

(b) **Production**: The production department of an organisation prepares the forecast regarding the production of the commodity with the help of statistical tools.

(c) **Scientific Management and Business Forecasting**: Better and efficient control of a business can be achieved by scientific management with the help of statistical data. "The success of businessman lies on the accuracy of forecast made". The successful businessman is one who estimate most closely approaches accuracy," said Prof. Boddington.

(d) **Purchase:** The price statistics of different markets help the businessman in arriving at the correct decisions. Raw material is purchased from those markets only where the prices are low.

III. Statistics and 'Auditing and Accounting': Statistics is widely used in accounting and auditing.

IV. Statistics and Economic Planning: According to Prof. Dickinson, "Economic Planning is making of major decisions – what and how much is to be produced, and to whom it is to be allocated – by the conscious decisions of a determinate authority on the basis of a comprehensive survey of economy as a whole. "The various documents accompanying preceding and following each of the eight Five Year Plans of India are a standing testimony to the fact that statistics is an indispensable tool in economic planning.

V. Statistics and Astronomy: Statistics were first collected by astronomers for the study of the movement of stars and planets. As there are a few things which are common between physical sciences, and statistical methods, astronomers apply statistical methods to go deep in their study. Astronomers generally take a large number of measurements and in most cases there is some difference between these several observations. In order to have the best possible measurement they have to make use of the technique of the law of errors in the form of method of least squares.

VI. Statistics and Meteorology: Statistics is related to meteorology. To compare the present with the past or to forecast for the future either temperature or humidity of air or barometrical pressures etc., it becomes necessary to average these figures and thus to study their trends and fluctuations. All this cannot be done without the use of statistical methods. Thus, the science of statics helps meteorology in a large number of ways.

VII. Statistics and Biology: The development of biological theories has been found to be closely associated with statistical methods. Professor Karl Pearson in his Grammar of Sciences has written, "The whole doctrine of heredity rests on statistical basis".

VIII. Statistics and Mathematics: Mathematics and Statistics have been closely in touch with each other ever since the 17th Century when the theory of probability was found to have influence on various statistical methods. Bowley was right when he said, "Acknowledge of Statistics is like knowledge of foreign language or of algebra: it may prove of use at any time under any circumstances".

Thus we observe that**:**

"Science without statistics bear no fruit, statistics without sciences have no Root".

IX. Statistics and Research: Statistical techniques are indispensable in research work. Most of advancement in knowledge has taken place because of experiments conducted with the help of statistical methods.

X. Statistics and natural sciences: Statistics finds an extensive application in physical sciences, especially in engineering physics, chemistry, geology, mathematics, astronomy, medicine, botany, meteorology, zoology, etc.

XI. Statistics and Education: There is an extensive application of statistics in Education. Statistics is necessary for formulation of policies to start new courses, infrastructure required for new courses consideration of facilities available for new courses etc.

XII. Statistics and Business: Statistics is an indispensable tool in all aspects of business. When a man enters business he enters the profession of forecasting because success in business is always the result of precision in forecasting and failure in business is very often due to wrong expectations. Which arise in turn due to faulty reasoning and inaccurate analysis of various cause affecting a particular phenomenon Boddington observes, "The successful businessman is the one, whose estimate most closely approaches the accuracy".

LIMITATIONS OF STATISTICS:

Statistics and its techniques are widely used in every branch of knowledge. W.I. King rightly says: "Science of statistics is the most useful servant, but only of great value to those who understand its proper use". The scope of statistics is very wide and it has great utility; but these are restricted by its limitations. Following are the important limitations of statistics:

1. **Statistics does not deal with individual item**: King says, "Statistics from the very nature of the subject cannot and never will be able to take into account individual cases". Statistics proves inadequate, where one wants to study individual cases. Thus, it fails to reveal the true position.

2. **Statistics deals with quantitative data**: According to Prof. Horace Secrist, "Some phenomenon cannot be quantitatively measured; honesty, resourcefulness, integrity, goodwill, all important in industry as well as in life, are generally not susceptible to direct statistical measurement".
3. **Statistical laws are true only on averages**. According to W.I. King, "Statistics largely deals with averages and these may be made up of individual items radically different from each other". Statistics are the means and not a solution to a problem.
4. **Statistics does not reveal the entire story**: According to Marshall, "Statistics are the straws, out of which, I like every other economist to have to make bricks. "Croxton says: "It must not be assumed that statistical method is the only method or use in research; neither should this method be considered the best attack for every problem".
5. **Statistics is liable to be misused**: According to Bowley, "Statistics only furnishes a tool though imperfect, which is dangerous in the hands of those who do not know its use and deficiencies".
   W.I. King states, "Statistics are like clay of which you can make a God or Devil as you please". He remarks, "Science of Statistics is the useful servant, but only of great value to those who understand its proper use".
6. **Statical data should be uniform and homogeneous**

STATISTICAL TOOLS USED IN ECONOMIC ANALYSIS:

The following are some of the important statistical techniques which are applied in economic analysis:

(a) Collection of data
(b) Tabulation
(c) Measures of Central Tendency
(d) Measures of Dispersion
(e) Time Series
(f) Probability
(g) Index Numbers
(h) Sampling and its uses
(i) Business Forecasting
(j) Tests of Significance and analysis of variance
(k) Statistical Quality Control

Collection of Data:

Data the information collected through censuses and surveys or in a routine manner or other sources is called a raw data. The word data means information (its literary meaning is given as facts). The adjective raw attached to data indicates that the information thus collected and recorded cannot be put to any use immediately and directly. It has to be converted into more suitable form or processed before it begins to make sense to be utilized gainfully. A raw data is a statistical data in original form before any statistical techniques are used to redefine, process or summarize it.

There are two types of statistical data:

(i) Primary data
(ii) Secondary data

   1. Primary Data: It is the data collected by a particular person or organization for his own use from the primary source
   2. Secondary data: It is the data collected by some other person or organization for their own use but the investigator also gets it for his use.

In other words, the primary data are those data which are collected by you to meet your own specific purpose, whereas the secondary data are those data which are collected by somebody else.

A data can be primary for one person and secondary for the other.

Methods of Collecting Primary Data:

The primary data can be collected by the following methods:

1. **Direct personal observation**: In this method, the investigator collects the data personally and, therefore, it gives reliable and correct information.
2. **Indirect oral investigation**: In this method, a third person is contacted who is expected to know the necessary details about the persons for whom the enquiry is meant.
3. **Estimates from the local sources and correspondence**. Here the investigator appoints agents and correspondents to collect the data
4. **Data through questionnaires**. The data can be collected by preparing a questionnaire and getting it filed by the persons concerned.
5. **Investigations through enumerators**. This method I generally employed by the Government for population census, etc.

Methods of Collecting Secondary data:

The secondary data can be collected from the following sources:

1. Information collected through newspapers and periodicals.
2. Information obtained from the publications of trade associations.
3. Information obtained from the research papers published by University departments or research bureaus or UGC.
4. Information obtained from the official publications of the central, state, and the local governments dealing with crop statistics, industrial statistics, trade and transport statistics etc.
5. Information obtained from the official publications of the foreign governments for international organizations. Like World Bank, ILO, IMF, etc.

Classification of Data: The process of arranging things in groups or classes according to their common characteristics and affinities is called the classification of data.

"Classification is the process of arranging data into sequences and groups according to their common characteristics or separating them into different but related parts – Secrist.

Thus classification is the process of arranging the available data into various homogenous classes and sub-classes according to some common characteristics or attribute or objective of investigation.

Requisites of a Good Classification:

The main characteristics of a good classification are:

1. It should be exhaustive
2. It should be unambiguous
3. It should be mutually exclusive
4. It should be stable
5. It should be flexible
6. It should have suitability
7. It should be homogeneous
8. It should be a revealing classification

9. It should be reliable
10. It should be adequate.

Advantages of classification of data:

(i) It condenses the data and ignores unnecessary details

(ii) It facilitates comparison of data

(iii) It helps in studying the relationships between several characteristics

(iv) It facilitates further statistical treatments

Types of Classification of Data:

There are four types of classification of data:

(i) Quantitative Classification

(ii) Temporal Classification

(iii) Spatial Classification and

(iv) Qualitative Classification

(i) **Quantitative Classification**: When the basis of classification is according to differences in quantity, the classification is called quantitative

A quantitative classification refers to classification that is based on figures: In other words, it is a classification which is based on such characteristics which are capable of quantitative measurement such as height, weight, number of marks obtained by students of a class.

(ii) **Temporal Classification**: When the basis of classification is according to differences in time, the classification is called temporal or chronological classification

(iii) **Spatial or Geographical Classification**: When the basis of classification is according to geographical location or place, the classification is called spatial or geographical

(iv) **Qualitative Classification**: When the basis of classification is according to characteristics or attributes like social status etc. is called qualitative classification.

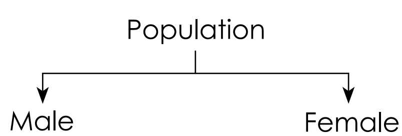
Classification according to attributes is a method in which the data are divided on the basis of qualities. (i.e., married or single; honest or dishonest; beautiful or ugly; on the basis of religion, viz., Hindu, Muslim, Sikh, Christian etc., known as attributes), which cannot be measured quantitatively.

Classification of this nature is of two types:

(i) Simple Classification or Two-Fold Classification

(ii) Manifold Classification

1. **Simple Classification or Two-fold Classification**: If the data are classified only into two categories according to the presence or absence of only one attribute, the classification is known as simple or two-fold classification or Dichotomous For example, the population of India may be divided into males and females; literate and illiterate etc.



Moreover, if the classification is done according to a single attribute it is also known as one way classification.

2. **Manifold Classification**: It is a classification where more than one attributes are involved.

MODE OF PRESENTATION OF DATA:

In this section we shall consider the following three modes of presentation of data

(a) Textual Presentation

(b) Tabular Presentation or Tabulation

(c) Diagrammatic presentation

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 are the examples of textual presentation.

Example1: In 1995, out of total of 2,000 students in a college, 1,400 were for graduation and the rest for post-graduation (P.G.) out of 1,400 Graduate students 100 were girls, however, in all there were 600 girls in the college. In 2000, number of graduate students increased to 1,700 out of which 250 were girls, but the number of P.G. students fall to 500 of which only 50 were boys. In 2005, out of 800 girls 650 were for graduation, where as the total number of graduates was 2,200. The number of boys and girls in P.G. classes was equal.

Merits and Demerits of Textual Presentation: The merit of this mode of presentation lies in its simplicity and even a layman can present and understand the data by this method. The observations with exact magnitude can be presented with the help of textual presentation. 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.

Tabular presentation or tabulation of data: Tabulation is a scientific process used in setting out the collected data in an understandable form

Tabulation may be defined as logical and systematic arrangement of statistical data in rows and columns. It is designed to simplify the presentation of data for the purposes of analysis and statistical inferences.

Secrist has defined tabulation in the following words:

"Tables are a means of recording in permanent form the analysis that is made through classification and by placing in juxtaposition things that are similar and should be compared".

The above definition clearly points out that tabulation is a process which gives classification of data in a systematic form and is meant for the purpose of making comparative studies.

Professor Bowley refers to tabulation as:

"The intermediate process between the accumulation of data in whatsoever form they are obtained, and the final reasoned account of the result shown by the statistics".

"Tabulation is the process of condensing classified data in the form of a table so that it may be more easily understood and so that any comparison involved may be more readily made".

Thus tabulation is one of the most important and ingenious devices of presenting the data in a condensed and readily comprehensible form. It attempts to furnish the maximum information in the minimum possible space, without sacrificing the quality and usefulness of the data.

Objectives of Tabulation:

The purpose of tabulation is to summarise lots of information in such a simple manner that it can be easily analysed and interpreted.

The main objectives of the Tabulation are:

1. To simplify the complex data.
2. To clarify the objective of investigation
3. Economise space.

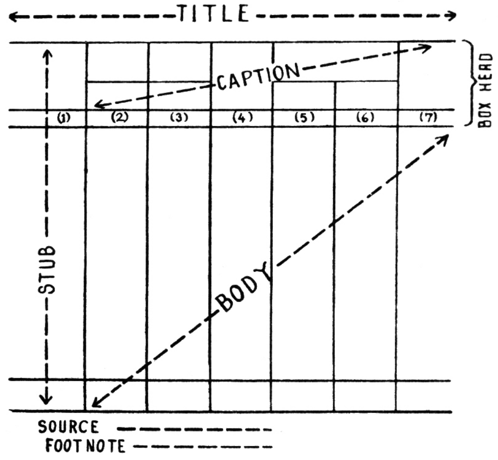
4. To facilitate comparison
5. To depict trend and pattern of data
6. To help reference for future studies.
7. To facilitate statistical analysis.
8. To detect errors and omissions in the data
9. To clarify the characteristics of data.

Essential Parts of a Statistical Table:

A good statistical table should invariably has the following parts:

1. **Table Number**: A table should be numbered for identification, especially, when there are a large number of tables in a study. The number maybe put at the centre, above the title or at the bottom of the table.
2. **Title of the table**: Every table should have a title. It should be clear, brief and self explanatory. The title should be set in bold type so as to give it prominence.
3. **Date**: The date of preparation of a table should always be written on the table. It enables to recollect the chronological order of the table prepared.
4. **Stubs or Row Designations**: Each row of the table must have a heading. The designations of the rows are called stubs or stub items. Stubs clarify the figures in the rows. As far as possible, the items should be considered so that they can be included in a single row.
5. **Captions or Column headings**: A table has many columns. Sub-headings of the columns are called captions or headings. They should be well-defined and brief.
6. **Body of the table**: It is the most vital part of the table. It contains the numerical information. It should be made as comprehensive as possible. The actual data should be arranged in such a manner that any figure may be readily located. Different categories of numerical variables should be set out in an ascending order, from left to right in rows and in the same fashion in the columns, from top downwards.
7. **Unit of Measurements**: The unit of measurements should always be stated along with the title, if this is uniform throughout. If different units have been adopted, then they should be stated along the stubs or captions.
8. **Source Notes**. A note at the bottom of the table should always be given to indicate the primary source as well as the secondary source from where the data has been taken, particularly, when there is more than one source.
9. **Foot Notes and References**: It is always placed at the bottom of the table. It is a statement which contains explanation of some specific items, which cannot be understood by the reader from the title, or captions and stubs.

Different Parts of Table



Difference between Textual and Tabular Presentation: The tabulation method is usually preferred to textual presentation as:

(i) It facilitates comparison between rows and columns

(ii) Complicated data can be represented using tabulation

(iii) Without tabulation, statistical analysis of data is not possible.

(iv) It is a must for diagrammatic representation.

3.2 DIAGRAMMATIC REPRESENTATION OF DATA

The representation of statistical data through charts, diagrams and picture is another attractive and alternative method. 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 presented 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.

In this chapter we shall consider the following three types of diagrams:

I. Line diagram chart;

II. Bar diagram;

III. Pie chart.

LINE CHART :

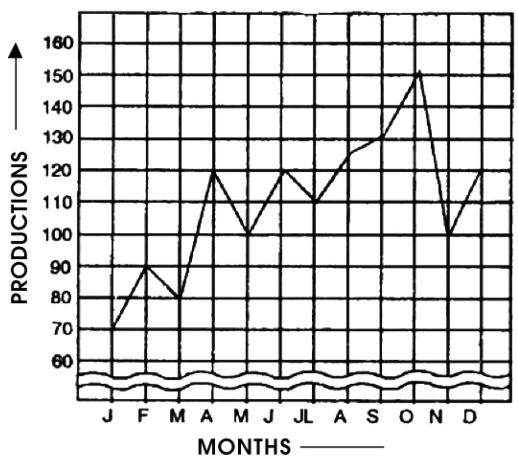
We take a rectangular axes. Along the abscissa, we take the independent variable (x or time) and along the ordinate the dependent variable (y or production related to time). After plotting the points, they are joined by a scale, which represents a line chart. The idea will be clear from the following example.

Example : Represent the following data by line chart.

The monthly production of motor cars in India during 2011-12

Jan	Feb	Mar	April	May	June	July	Aug	Sept	Oct	Nov	Dce
70	90	80	120	100	120	110	125	130	150	100	120

Graph showing production of motor cars.



BAR DIAGRAM:

The simplest type of graph is the bar diagram. It is especially useful in comparing qualitative data or quantitative data of discrete type. A bar diagram is a graph on which the data are represented in the form of bars. It consists of a number of bars or rectangles which are of uniform width with equal space between them on the x-axis. The length of the bar is proportional to the value it represents. It should be seen that the bars are neither too short nor too long. The scale should be clearly indicated and base line be clearly shown.

Bars may be drawn either horizontally or vertically. A good rule to use in determining the direction is that if the legend describing the bar can be written under the bars when drawn vertically, vertical bars should be used; when it cannot be, horizontal ones must be used. In this way, the legends can be read without turning the graph. The descriptive legend should not be written at the ends of the bars or within the bars, since such writing may distort the comparison. Usually the diagram will be more attractive if the bars are wider than the space between them.

The width of bars is not governed by any set rules. It is an arbitrary factor. Regarding the space between two bars, it is conventional to have a space about one half of the width of a bar.

The data capable of representation through bar diagrams, may be in the form of row scores, or total scores, or frequencies, or computed statistics and summarised figures like percentages and averages etc.

The bar diagram is generally used for comparison of quantitative data. It is also used in presenting data involving time factor. When two or more sets of data over a certain period of time are to be compared a group bar diagram is prepared by placing the related data side by side in the shape of bars. The bars may be vertical or horizontal in a bar diagram. If the bars are placed horizontally, it is called a Horizontal Bar Diagram. When the bars are placed vertically, it is called a Vertical Bar Diagram.

There are six types of Bar diagram:

(i) Simple Bar Diagram;

(ii) Multiple or Grouped Bar Diagram;

(iii) subdivided or Component Bar Diagram;

(iv) Percentage Subdivided Bar Diagram;

(v) Deviation or Bilateral Bar Diagram;

(vi) Broken Bars.

Simple Bar Diagram:

It is used to compare two or more items related to a variable. In this case, the data are presented with the help of bars. These bars are usually arranged according to relative magnitude of bars. The length of a bar is determined by the value or the amount of the variable. A limitation of Simple Bar Diagram is that only one variable can be represented on it.

Multiple or Grouped Bar Diagram:

A multiple or grouped bar diagram is used when a number of items are to be compared in respect of two, three or more values. In this case, the numerical values of major categories are arranged in ascending or descending order so that the categories can be readily distinguished. Different shades or colours are used for each category.

Sub-divided or Component Bar Diagram:

A component bar diagram is one which is formed by dividing a single bar into several component parts. A single bar represents the aggregate value whereas the component parts represent the component values of the aggregate value. It shows the relationship among the different parts and also between the different parts and the main bar.

Percentage Sub-divided Bar Diagram:

It consists of one or more than one bars where each bar totals 100%. Its construction is similar to the sub-divided bar diagram with the only difference that where as in the sub-divided bar diagram segments are used in absolute quantities, in the percentage bar diagram the quantities are transformed into percentages.

PIE DIAGRAM OR ANGULAR DIAGRAM:

A pie diagram is a circular graph which represents the total value with its components. The area of a circle represents the total value and the different sectors of the circle represent the different parts. The circle is divided into sectors by radii and the areas of the sectors are proportional to the angles at the centre. It is generally used for comparing the relation between various components of a value and between components and the total value. In pie diagram, the data are expressed as percentages. Each component is expressed as percentage of the total value. A pie diagram is also known as angular diagram.

The name pie diagram is given to a circle diagram because in determining the circumference of a circle we have to take into consideration a quantity known as 'pie' (written as π).

Method of Construction: The surface area of a circle is known to cover 2π radians or 360 degrees. The data to be represented through a circle diagram may therefore be presented through 360 degrees, parts or sections of a circle. The total frequencies or value is equated to 360° and then the angles corresponding to component parts are calculated (or the component parts are expressed as percentages of the total and then multiplied by 360/100 or 3.6). After determining these angles the required sectors in the circle are drawn. Different shades or colours of designs or different types of cross-hatchings are used to distinguish the various sectors of the circle.

Example: 120 students of a college were asked to opt for different work experiences. The details of these options are as under.

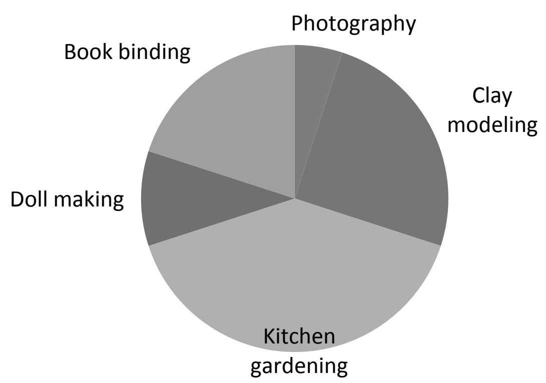
Areas of work experience	**No. of students**
Photography	6
Clay modeling	30
Kitchen gardening	48
Doll making	12
Book binding	24

Represent the above data through a pie diagram.

Solution:

The numerical data may be converted into the angle of the circles as given below:

Areas of work experience	**No. of students**	**Angle of the circle**
Photography	6	(6/120) x 360 = 18°
Clay modeling	30	(30/120) x 360 = 90°
Kitchen gardening	48	(48/120) x 360 = 144°
Doll making	12	(12/120) x 360 = 36°
Book binding	24	(24/120) x 360 = 72°
Total	120	360°



Areas of work experience opted by the students

3.3 FREEQUENCY DISTRIBUTION

TALLY BARS AND FREQUENCY:

In order to make the data easily understandable, we tabulate the data in the form of tables or charts. A table has three columns

(i) Variable

(ii) Tally marks

(iii) Frequency

(i) Variable: Any character which can vary from one individual to another is called a variable or a variate. For example, age, income, height, intelligence, colour etc. are variates. Some variates are measurable and others are not directly measurable. The examples of measurable variates are age, height, temperature, etc., where as colour and intelligence are the examples of those variates which cannot be measured numerically. Variables or observations with numbers as possible values are called quantitative variables, whereas those with names of places, quality, things etc., as possible values are called qualitative variables or attributes.

Variables are of two types i) Continuous; ii) Discontinuous or Discrete. Quantities which can take all numerical values within a certain interval are called continuous variables; But those variables which can take only a finite of values are called discrete variables; For example, number of students in a particular class, number of sections in a school etc.

(ii) Tally: It is a method of keeping count in blocks of five.

For example; | = 1; || = 2; ||| = 3; |||| = 4; ~~||||~~ = 5; ~~||||~~ | = 6 and so on.

Tally Bars: These are the straight bars used in the Tally.

Each item falling in the class interval, a stroke (vertical Bar) is marked against it. This stroke (Vertical Bar) is called the Tally Bar. Usually, after every four strokes (Tally Bar), in a class, the fifth item is marked by a horizontal or slanted line across the Tally Bars (Strokes). For example the frequency 5, 6, 7 is represented by ~~||||~~ , ~~||||~~ | , ~~||||~~ || respectively.

The above method of presentation of data is known as 'Frequency Distribution'. Marks are called variates. The number of students who have secured a particular number of marks is called Frequency of that variate.

In the first column of the table, we write all marks from lowest to highest. We now look at the first mark or value in the given raw data and put a bar (vertical line) in the second column opposite to it. We then, see the second mark or value in the given raw data and put a bar opposite to it in the second column. This process is repeated till all the observations in the given raw data are exhausted. The bars drawn in the second column are known as tally marks and to facilitate we record tally marks in bunches of five, the fifth tally marks is drawn diagonally across the first four.

For example, ~~||||~~ ||| = 8. We finally count the number of tally marks corresponding to each observation and write in the third column headed by frequency or number of students.

(iii) Frequency: The number of times an observation occurs in the given data is called the frequency of the observation.

Frequency Distribution: A frequency distribution is the arrangement of the given data in the form of a table showing frequency with which each variable occurs. In other words, Frequency distribution of a variable is the ordered set $\{x, f\}$, where f is the frequency. It shows all scores in a set of data together with the frequency of each score.

Types of frequency distributions:

Frequency distributions are of two types:

(i) Discrete Frequency Distribution

(ii) Grouped (or Continuous) Frequency Distribution

Discrete Frequency Distribution: The construction of discrete frequency distribution from the given raw data is done by the method of tally marks as explained earlier.

Construction of Discrete Frequency Distribution Table:

The frequency distribution table has three columns headed by

1. Variables (or classes)
2. Tally Mark or Bars
3. Frequency

The table is constructed by the following steps:

Step 1: Prepare three columns, viz., one for the variable (or classes), another for tally marks and the third for the frequency corresponding the variable (or class).

Step 2: Arrange the given data (or values) from the lowest to the highest in the first column under the heading variable (or classes)

Step 3: Take the first observation in the raw data and put a bar (or vertical line |) in the second column under Tally Marks opposite to it. Then take a second observation and put a tally marks opposite to it, continue this process till all the observations of the given raw data are exhausted. For the sake of convenience, record the tally marks in bunches of five, the fifth bar is placed diagonally crossing the other four (5 is represented by ~~||||~~) leave some space between each block of bars.

Step 4: Count the tally marks of column 2 and place this number opposite to the value of the variable in the third column headed by Frequency.

Step 5: Give a suitable title to the frequency distribution table so that it exactly conveys the information contained in the table.

SOME STATISTICAL TERMS:

Raw Data or Data: A raw data is a statistical data in original form before any statistical technique is applied to redefine process or summarize it.

Variable or Variate: Any character which can vary from one individual to another is called variable or variate. For example, age, income, height, intelligence, colour, etc., are variates. Some variates are measurable and others are not directly measurable. The examples of measurable variates are age, height, temperature, etc., whereas colour and intelligence are the examples of those variates which cannot be measured numerically. Variables or observations with numbers as possible values are called quantitative variables, whereas those with names of places, attributes, and things etc., as possible values are called qualitative variables.

Variables are of two types:

(i) Continuous;

(ii) Discontinuous or Discrete

Continuous Variable: A continuous variable is capable of assuming any value within a certain range or interval. The height, weight, age and temperature of any person can be expressed not only in integral part but also in fractions of any part. For example, the weight of a boy may 44.0 kg or 44.6 kg or 44.65 kg, similarly, his height may be 56 inches or 56.4 inches and age may be 10 years or 10.5 years. Thus, the height, weight, age or temperature etc. are continuous variables.

Discrete Variable: A discrete variable can assume only integral values and is capable of exact measurement. In other words, those variables which can take only a finite set of values are called discrete variables. For example: the number of students in a particular class, or the number of sections in a school, etc. are the examples of discrete variables. Discrete variables are also known discontinuous variables.

Continuous Series: When the continuous variables are arranged in the form of a series, it is called continuous series or exclusive series.

Discrete or Discontinuous Series: When the discrete variables are arranged in the form of a series, it is called a discrete or discontinuous series.

Array: An array is an arrangement of data in order of magnitude either in descending or ascending order.

Descending Order: When data is arranged from the highest value to the lowest value, the array so formed is in descending order.

Ascending order: When the data is arranged from the lowest value to the highest value, the array so formed is in ascending order.

Illustration: If the given data is 17, 7, 11, 5, 13, 9 then

Array in Ascending order: 5, 7, 9, 11, 13, 17.

Array in Descending order: 17,13,11,9,7, 5.

Range: It is the difference between the largest and the smallest number in the given data

The range of the data given in illustration is 17 – 5 = 12.

Class, Class-Interval and Class limits. If the observations of a series are divided into groups and the groups are bounded by limits, then each group is called a class. The end values of a class are called class limits. The smaller value of the two limits is called the lower limit and the higher value of the same is called the upper limit of the class. These two class limits are sometimes called the stated class limits.

Class Interval: The difference between the lower limit (L) and the upper limit (U) of the class is known as class interval (I).

Thus: I = U – L.

In other words, the range of a class is called its Class Interval.

Illustration: The given data is

Marks obtained	**Tally Marks**	**No. of Students**
1-10	𝍸 |	6
11-20	|||	3
21-30	𝍸 ||	7
31-40	||	2
Total		18

In the above data the classes are: 1-10, 11-20, 21-30, 31-40.

Class Interval: The range of the marks from 1 to 40 is grouped into four classes or groups viz: 1 – 10, 11, 20, 21-30, 31-40. Each group is known as class interval. The interval between one class and its adjacent class being 9. [as 10-1 = 9, 20 – 11 = 9, 30 – 21 = 9, etc.]

Class Limits: In the first class 1 – 10, its lower limit is 1 and upper limit is 10. Similarly, 31 is the lower limit and 40 is the upper limit of the class interval 31-40.

Actual Class Limit or Class Boundaries: In the illustration, there is a gap of 1 mark between the limits of any two adjacent classes. This gap may be filled up by extending the two limits of each class by half of the value of the gap. Thus

Lower class boundary = lower class limit – ½ of the gap

Upper class boundary = Upper class limit + ½ of the gap

The class boundary of the class 11 to 20 are

Lower class boundary = 11– ½ of 1 = 11 – 0.5 = 10.5

Upper class boundary = 20+ ½ of 1 = 20+ 0.5 = 20.5

In other words, the class boundaries are the limits up to which the two limits, (actual) of each class may be extended to fill up the gap that exists between the classes. The class boundaries of each class, so obtained are called the Actual class limits or True class limits.

True lower class limit = Lower class limit - ½ of the gap

True upper class limit = Upper class limit + ½ of the gap

Note: In the case of exclusive series True class limits are the same as class limits

Illustration:

Class Interval	Class Boundries
11-20	10.5-20.5
21-30	20.5-30.5
31-40	30.5-40.5

Class-mark or Mid-point or Mid-value: The central value of the class interval is called the mid-point or mid-value or class mark. It is the arithmetic mean of the lower class and upper class limit of the same class.

Mid-value of Class = $\dfrac{\text{Lower class limit+Upper class limit}}{2}$ (or)

Class mark = $\dfrac{\text{True Upper class limit} + \text{True Lower class limit}}{2}$

The class mark of the class 11-20 is $\dfrac{11+20}{2}$ = 15.5

Class Magnitude: It is the difference between the upper class boundary and the lower class boundary of the class. In the illustration the class magnitude of the class, 20.5 – 30.5 is (30.5 – 20.5) = 10.

Inclusive and Exclusive Series: In the above illustration, all the marks we considered were integers. Hence, it was possible for us to choose classes 11 to 20, 21 to 30 etc. there is a gap of 1 between the upper limit of a class and the lower limit of its next consecutive class, which has not created any difficulty. But there can be situations where the raw data is not in integers. For example, in the information regarding maximum temperature of the city or time required to solve a statistical problem is recorded in the data, it may contain fractions as well. In such cases, the consecutive classes have to be necessarily continuous. We have the following:

Inclusive Series: When the class-intervals are so fixed that the upper limit of the class is included in that class, it is known as inclusive method of classification, e.g., 0-5, 6-10, 11-15, 16-20.

In the inclusive series, the upper limit and lower limit are included in that class interval. For example, in illustration, the marks 11 and 20 are included in the class 11-20. It is a discontinuous series or inclusive series. In order to make it a continuous one, some adjustment with the class limits is necessary. The class limits are extended to class boundaries by the adjusting adjustment factor, which is equal to half of the difference between the upper limit of the one class and lower limit of the next class. The series so obtained is continuous and is known as exclusive series.

Exclusive or Continuous Series: In this series the upper limit of the class is the lower limit of the other class, the common point of the two classes is included in the higher class. For example, 10-15, 15-20, 20-25, …, represent a continuous series or the exclusive series. In this series, 15 is included in the class 15-20 and 20 is included in 20-30. Here the class intervals overlap and the upper limit of each class is treated as less than that limit and lower limit of each class actually represents exact value. Thus

When the class-intervals are so fixed that the upper limit of one class is the lower limit of the next class, it is known as Exclusive method of classification.

RELATIVE FREQUENCY AND PERCENTAGE FREQUENCY OF A CLASS INTERVAL:

Relative Frequency: Frequency of each class can also be expressed as a fraction of percentage terms. These are known as relative frequencies. In other words, a relative frequency is the class frequency expressed as a ratio of the total frequency, i.e.,

Relative frequency = $\dfrac{\text{Class frequency}}{\text{Total frequency}}$

Percentage Frequency: Percentage frequency of a class interval may be defined as the ratio of the class frequency to the total frequency expressed as a percentage.

$$\text{Percentage frequency} = \frac{\text{Class frequency}}{\text{Total frequency}} \times 100$$

3.4 GRAPHICAL REPRESENTATION OF FREQUENCY DISTRIBUTION

The graphs of frequency distribution are designed to present the characteristic features of a frequency data. They facilitate comparative study of two or more frequency distributions regarding their shape and pattern.

The most commonly used graphs are:

1. Histogram
2. Frequency Polygon
3. Frequency Curve
4. Cumulative Frequency Curve or Ogive.

HISTOGRAM (when C.I. are equal)

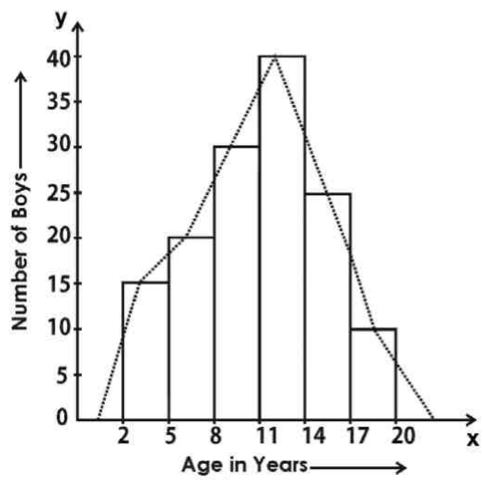
Let us consider a frequency distribution having a number of class intervals with their respective frequencies. The horizontal axis is marked to represent the C.I. and on these markings rectangles are drawn by taking the C.I. as breadth and corresponding frequencies as heights. Thus a series of rectangles are obtained whose total area represents the total of the class frequencies. The figure thus obtained is known as histogram.

It may be noted here that C.I. must be in continuous form. Even if this is not given, then the discrete C.I. must be transferred to class boundaries and hence to draw the histogram.

Example : Draw a histogram of the following frequency distribution showing the number of boys in the register of a school.

Age (in years)	No. of boys (in '000)
2–5	15
5–8	20
8–11	30
11–14	40
14–17	25
17–20	10

C.I. given are in class boundaries.



Histogram

Histogram (when C.I. are uneqal) : If the C.I. are unequal the frequencies must be adjusted before constructing the histogram. Adjustments are to be made in respect of lowest C.I. For instance if one C.I. is twice as wide as the lowest C.I., then we are to divided the height of the rectangle by two and if again it is three times more, then we are to divide the height of the rectangle by three and so on.

Aliter (with the help of frequency density) :

If the width of C.I. are euqal the heights of rectangles will be proportional to the corresponding class frequencies. But if the widths of C.I. are unequal (i.e. some are equal and others are unequal), then the heights of rectangles will be proportional to the corresponding frequency densities (and not with the class frequencies)

$$\text{Frequency density} = \frac{\text{Class frequency}}{\text{Width of C.I.}}$$

MULTIPLE CHOICE QUESTIONS

1. Statistics is applied in
   (a) Economics
   (b) Business management
   (c) Commerce and industry
   (d) All these
2. Statistics is concerned with
   (a) Qualitative information
   (b) Quantitative information
   (c) a or b
   (d) Both a and b
3. Which of the following statements is false?
   (a) Statistics is derived from Latin word 'status'
   (b) Statistics is derived from Italian word "Statista"
   (c) Statistics is derived from French word "Statistik"
   (d) None of these
4. The colour of a flower is an example of
   (a) An attribute
   (b) A variable
   (c) A discrete variable
   (d) A continuous variable
5. A Quantitative characteristic is known as
   (a) An attribute
   (b) A discrete variable
   (c) A continuous variable
   (d) None of above
6. Annual income of a person is
   (a) An attribute
   (b) A discrete variable
   (c) A continuous variable
   (d) a or c

7. The quickest method to collect primary data is
   (a) Personal interview
   (b) Indirect interview
   (c) Telephone interview
   (d) By observation
8. The mode of presentation of data are
   (a) Textual, tabulation and diagrammatic
   (b) Tabular, internal and external
   (c) Textual, tabular and internal
   (d) Tabular, textual and external
9. For tabulation, 'caption' is
   (a) The upper part of the table
   (b) The lower part of the table
   (c) The main part of the table
   (d) The upper part of a table that describes the column and sub-column
10. "Stub" of table is the
   (a) Left part of the table describing the columns
   (b) Right part of the table describing the columns
   (c) Right part of the table describing the rows
   (d) Left part of the table describing the rows
11. Pie-diagram is used for
   (a) Comparing different components and their relation to the table
   (b) Representing qualitative data in a circle
   (c) Representing quantitative data in circle
   (d) b or c
12. Weights are generally called –
   (a) Range
   (b) Mean
   (c) Frequencies
   (d) Mode
13. Length of a class is
   (a) The difference between the UCB and LCB of that class
   (b) The difference between the UCL and LCL of that class
   (c) a or b
   (d) Both a and b
14. Following is the class and frequency

Class	**Frequency**
0-10	5
10-20	8
20-30	15
30-40	6
40-50	4

For the class 20-30, cumulative frequency is

(a) 20

(b) 13

(c) 15

(d) 28

15. (Class frequency) / (Width of the class) is defined as

(a) Frequency density

(b) Frequency distribution

(c) Both

(d) None

16. Mode of a distribution can be obtained form

(a) Histogram

(b) Less than type Ogives

(c) More than types Ogives

(d) Frequency polygon

17. An area diagram is

(a) Histogram

(b) Frequency polygon

(c) Ogive

(d) None

18. Median of distribution can be obtained from

(a) Less than type Ogives

(b) Point of Intersection of Less than and greater than Ogives

(c) Both a and b

(d) None of these

19. Most of the commonly used frequency curves are

(a) Mixed

(b) Inverted J-shaped

(c) U-shaped

(d) Bell-shaped

20. The distribution of profits of a company follows

(a) J-shaped

(b) U-shaed

(c) Bell-shaped frequency curve

(d) Any of these

21. Cost of sugar in a month under the heads Raw-materials, Labour, direct production and others were 12, 20, 35 and 23 units respectively. What is the difference between the central angles for the largest and smallest components of the cost of sugar?

(a) 72°

(b) 48°

(c) 56°

(d) 92°

22. The number of accidents for seven days in a locality are given below:

C	0	1	2	3	4	5	6
Frequency	15	19	22	31	9	3	2

What is the number of cases when 3 or less accident occurred?

(a) 56

(b) 6

(c) 68

(d) 87

23. The following data relate to the marks of a group of students:

Marks	**No.of Students**
Below 10	15
Below 20	38
Below 30	65
Below 40	84
Below 50	100

How many students got marks more than 30?

(a) 65

(b) 50

(c) 35

(d) 43

24. Find the number of observations between 250 and 300 from the following data

Value (Greater than)	200	250	300	350
Frequency	56	38	15	0

(a) 56

(b) 23

(c) 15

(d) 8

25. Refer following table: Frequency distribution of weights of 16 students

Weight in Kg. (Class interval)	No. of students (Frequency
44-48	4
49-53	5
54-58	7
Total	**16**

Find class mark for the first class interval

(a) 4

(b) 46

(c) 44

(d) 48

(i) Find width of class interval for the second class interval

(a) 4

(b) 5

(c) 46

(d) 44-48

(ii) Find Frequency density of the second class interval

(a) 0.80

(b) 0.90

(c) 1.00

(d) 1.10

(iii) Find Relative frequency for the second class interval

(a) 1/11

(b) 5/4

(c) 5/16

(d) 1/4

(iv) Find relative frequency for the third class interval

(a) 7/16

(b) 7/4

(c) 16/7

(d) None of the above

Answers:

1	d	2	d	3	c	4	a	5	b	6	b	7	c	8	a
9	d	10	d	11	a	12	c	13	a	14	d	15	a	16	a
17	a	18	c	19	d	20	c	21	d	22	d	23	c	24	b
25 (i)	b	(ii)	b	(iii)	c	(iv)	c	(v)	a						

Study Note - 4

MEASSURES OF CENTRAL TENDENCY AND DISPERSION

4.1 MEASURES OF CENTRAL TENDENCY

In the previous chapters, data collection and presentation of data were discussed. Even after the data have been classified and tabulated one often finds too much details for many uses that may be made of the information available. We, therefore, frequently need further analysis of the tabulated data. One of the powerful tools of analysis is to calculate a single average value that represents the entire mass of data. The word average is very commonly used in day-to-day conversation. For example, we often talk of average work, average income, average age of employees, etc. an 'Average' thus is a single value which is considered as the most representative or typical value for a given set of data. Such a value is neither the smallest nor the largest value, but is a number whose value is somewhere in the middle of the group. For this reason an average is frequently referred to as a measure of central tendency or central value. Measures of central tendency show the tendency of some central tendency show the tendency of some central value around which data tends to cluster.

Objectives of Averaging:

There are two main objectives of the study of averages:

(i) **To get one single value that describes the characteristic of the entire data**. Measures of central value, by condensing the mass of data in one single value, enable us to get an idea of the entire data. Thus one value can represent thousands, lakhs and even millions of values. For example, it is impossible to remember the individual incomes of millions of earning people of India and even if one could do it there is hardly any use. But if the average income is obtained, we get one single value that represents the entire population. Such a figure would throw light on the standard of living of an average Indian.

(ii) **To facilitate comparison**. Measures of central value, by reducing the mass of data in one single figure, enable comparisons to be made Comparison can be made either at a point of time or over a period of time. For example, the figure of average sales for December may be compared with the sales figures of previous months or with the sales figure of another competitive firm.

Characteristics of a Good Average:

Since an average is a single value representing a group of values, it is desirable that such a value satisfies the following properties:

(i) **It should be easy to understand**. Since statistical methods are designed to simplify complexity, it is desirable that an average be such that can be readily understood; otherwise, its use is bound to be very limited.

(ii) **It should be simple to compute**. Not only an average should be easy to understand but also it should be simple to compute so that it can be used widely. However, though case of computation is desirable, it should not be sought at the expense of other advantages, i.e., if in the interest of greater accuracy, use of a more difficult average is desirable one should prefer that.

(iii) **It should be based on all the observations**. The average should depend upon each and every observation so that if any of the observation is dropped average itself is altered.

(iv) **It should be rigidly defined**. An average should be properly defined so that it has one and only one interpretation. It should preferably be defined by an algebraic formula so that if different people compute the average from the same figures they all get the same answer (barring arithmetical mistakes).

(v) **It should be capable of further algebraic treatment**. We should prefer to have an average that could be used for further statistical computations. For example, if we are given separately the figures of average income and number of employees of two or more factories we should be able to compute the combined average.

(vi) **It should have sampling stability**. We should prefer to get a value which has what the statisticians call 'sampling stability'. This means that if we pick 10 different groups of college students, and compute the average of each group, we should expect to get approximately the same values. It does not mean, however, that there can be no difference in the value of different samples. There may be some difference but those averages in which this difference, technically called sampling fluctuation, is less are considered better than those in which this difference is more.

(vii) **It should not be unduly affected by the presence of extreme values**. Although each and every observations should influence it unduly. If one or two very small or very large observations unduly affect the average, i.e., either increase its value or reduce its value, the average cannot be really typical of the entire set of data. In other words, extremes may distort the average and reduce its usefulness.

The following are the important measures of central tendency which are generally used in business:

A. Arithmetic mean.

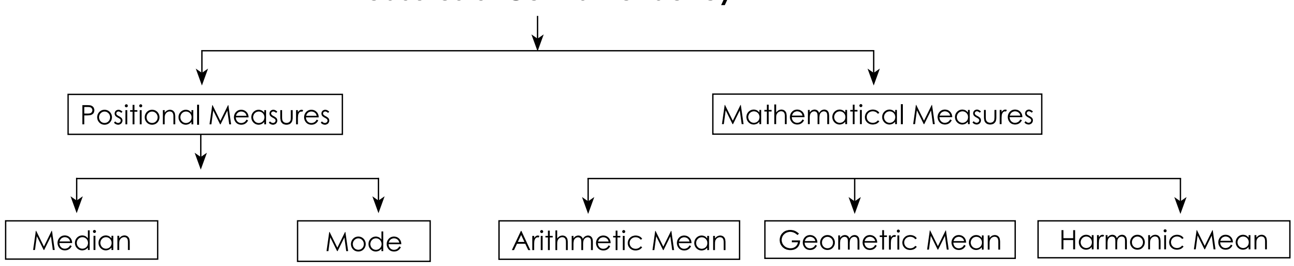
B. Median

C. Mode

D. Geometric mean, and

E. Harmonic mean

Measures of Central Tendency



I. ARITHMETIC MEAN: A.M is denoted by $\overline{X}$. It is a mathematical measurement. It is calculated by different methods of the following.

(a) Individual Series: -

(i) Direct Method: - $\overline{X} = \dfrac{\Sigma x}{N}$

(ii) Short Cut Method (or) Indirect Method: $\overline{X} = A + \frac{\Sigma dx}{N}$

(iii) Step Deviation method: $\overline{X} = A + \frac{\Sigma dx^1}{N} \times i$

Where 'N' is No. of terms in the given series.

Σx is Sum of terms, A is the assumed mean.

dx is the deviation of items from assumed mean i.e. dx = x – A

i is the common factor, $dx^1 = \frac{dx}{i}$

(b) Discrete Series:

(i) Direct Method: - $\overline{X} = \frac{\Sigma fx}{N \text{ or } \Sigma f}$

(ii) Short Cut Method (or) Indirect Method: $\overline{X} = A + \frac{\Sigma fdx}{N \text{ or } \Sigma f}$

(iii) Step Deviation method: $\overline{X} = A + \frac{\Sigma fdx^1}{N \text{ or } \Sigma f} \times i$

Where f is frequency, N = $\sum f$ = Total frequency

A is assumed mean, dx = x – A, $dx^1 = \frac{dx}{i}$

(c) Continuous Series:

(i) Direct Method: - $\overline{X} = \frac{\Sigma fm}{\Sigma f}$

(ii) Short Cut Method (or) Indirect Method: $\overline{X} = A + \frac{\Sigma fdx}{\Sigma f = N}$

(iii) Step Deviation method: $\overline{X} = A + \frac{\Sigma fdx}{N} \times i$

Where A is assumed mean, m is the mid value of the class interval

f is the frequency, N = $\sum f$ = Total frequency, i is the common factor

$dx^1 = \frac{dx}{i}$, dx = m – A

(d) To Calculate Combined $\overline{X}$:

$$\overline{X}_{123........n} = \frac{N_1\overline{X}_1 + N_2\overline{X}_2 + \cdots\cdots\cdots + N_n\overline{X}_n}{N_1 + N_2 + \cdots\cdots\cdots + N_n}$$

$\overline{X}_{123........n}$ = Combined mean of the groups = $\overline{X}$

$\overline{X}_1$ = A.M. of first group, $\overline{X}_2$ = A.M. of second group $\overline{X}_n$ = A.M. of n^{th} group,

N_1 = No. of terms in the first group N_2 = No. of terms in the second group

N_n = No. of terms n^{th} group

Note: In case $N_1 = N_2 = N_3 = \ldots\ldots\ldots\ldots\ldots N_n$ then

$$\overline{X} = \frac{\overline{X}_1 + \overline{X}_2 + \cdots\cdots\cdots + \overline{X}_n}{n}$$

(e) TO DETERMINE CORRECT $\overline{X}$, WHEN SOME TERMS ARE INCLUDED WRONGLY:

Incorrect total of n terms = $n\overline{X}$

Correct Total = Incorrect total + Correct terms – In correct terms

$$\therefore \text{Correct } \overline{X} = \frac{\text{Correct Total}}{\text{No. of Terms}}$$

(f) WEIGHTED ARITHMETIC MEAN: $\overline{X}_w = \frac{\Sigma wx}{\Sigma w}$

Where x is variable, W is assigned weight, $\overline{X}_w$ is weighted A.M.

Examples:

1. Find Mean for the following figures.

30	41	47	54	23	34	37	51	53	47

Solution:

Adding all the terms and using the formula

$\overline{X} = \frac{\Sigma x}{N}$, Here N = 10. And $\sum X = 417$ n = number of observations = 10

$= \frac{417}{10} = \mathbf{41.7}$

2. Calculate A.M. from the following data:

Marks obtained:	4	8	12	16	20
No. of students	6	12	18	15	9

Solution:

Marks X:	**No. of students**	
	F	fX
4	6	24
8	12	96
12	18	216
16	15	240
20	9	180
	N = 60	$\sum fx = 756$

As $\overline{X} = \frac{\Sigma fx}{N} = \overline{X} = \frac{756}{60} = \mathbf{12.6}$

3. Use direct method to find $\overline{X}$

Income	10-20	20-30	30-40	40-50	50-60	60-70
No. of Persons	4	7	16	20	15	8

Solution:

Income X	Mid Value m	No. of Persons f	fm
10-20	15	4	60
20-30	25	7	175
30-40	35	16	560
40-50	45	20	900
50-60	55	15	825
60-70	65	8	520

As $\overline{X} = \frac{\Sigma fx}{N} = \overline{X} = \frac{3040}{70} = \mathbf{43.43}$

4. $\overline{X}$ of 20 terms was found to be 35. But afterwards it was detected that two terms 42 and 34 were misread as 46 and 39 respectively. Find correct X.

Solution:

$\overline{X}$ of 20 terms = 35

(incorrect) Total of 20 terms = 35 x 20 = 700

Correct Total = 700 + 42 + 34 – 46 – 39 = 691

$\therefore$ Correct $\overline{X} = \frac{691}{20} = \mathbf{34.55}$

5. The mean of wages in factory A of 100 workers is ₹ 720 per week. The mean wages of 30 female workers in the factory was ₹650 per week. Find out average wage of male workers in the factory.

Solution:

$N = 100,\ N_1 = 30,\ \overline{X}_1 = 650,\ \overline{X}_2 = ?$

$N_1 + N_2 = N\ 30 + N_2 = 100 \qquad N_2 = 70$

$\overline{X}_{12} = 720$

$$\overline{X}_{12} = \frac{N_1\overline{X}_1 + N_2\overline{X}_2}{N_1 + N_2}$$

$$720 = \frac{30 \times 650 + 70\overline{x}}{100}$$

$72000 = 19500 + 70\ \overline{X}_2$

$70\ \overline{X}_2 = 52500$

$$\overline{X}_2 = \frac{52500}{70} = \mathbf{750}$$

Properties of Arithmetic Mean: The important properties of arithmetic mean are given below:

(i) The sum of the deviations of the terms from the Actual mean is always zero.

(ii) The sum of the squared deviations of the items from arithmetic mean is minimum i.e. less than the sum of the squared deviations of the items from any other value.

(iii) If we have arithmetic mean and the number of items of two or more than two groups, we can calculate the combined average of groups.

(iv) If the terms of a series are increased, decreased, multiplied or divided by some constant, the mean also increases, decreases, multiplied or are divided by the same constant.

(v) The standard error of the arithmetic mean is less than that of any other measure of central tendency.

II. **GEOMETRIC MEAN (g)**: The geometric mean is obtained by multiplying the values of the items together and then taking it to its root corresponding to the number of items. It is denoted by 'g'.

i.e., $g = \sqrt[n]{x_1 \cdot x_2 \cdot x_3 \cdots\cdots x_n}$

(a) Individual Series: g = Anti log $\left(\frac{\Sigma \log n}{n}\right)$

Where g is G.M, x is items, 'n' is No. of terms

(b) Discrete Series: g = Anti log $\left(\frac{\Sigma f \log x}{N}\right)$, Where N = total frequency

(c) Continuous Series: g = Anti log $\left(\frac{\Sigma f \log m}{N}\right)$, Where m is mid value of the C.I. $(N = \Sigma f)$

(d) Weighted Geometric mean: g = Anti log $\left(\frac{\Sigma W \log x}{\Sigma W}\right)$, Where W is weights

(e) Combined Geometric Mean: g = Anti log $\left(\frac{n_1 \log g_1 + n_2 \log g_2}{n_1 + n_2}\right)$

- The Geometric mean is relative value and is dependent on all items
- The geometric mean is never larger than the arithmetic mean. It is rare that it may be equal to the arithmetic mean.
- The Geometric mean of the products of corresponding items in two series is equal to product of their geometric mean.

III. **HARMONIC MEAN**: Harmonic mean of a given series is the reciprocal of the arithmetic average of the reciprocal of the values of its various items.

(a) **Individual Series**: H.M. = $\frac{n}{\sum_{i=1}^{n} \frac{1}{x_i}}$

where n is No. of items and x_i is $x_1, x_2, x_{3.......} x_n$

(b) **Discrete Series**: H.M. = $\frac{N}{\sum \frac{f}{x}}$

Where N = $\sum$ f = Total frequency, f is frequency

(c) **Continuous Series**: H.M = $\frac{N}{\sum \frac{f}{m}}$, Where m is mid value of the C.I, N = $\sum f$

(d) **Weighted H.M**: $H.M_w = \frac{\Sigma W}{\Sigma \frac{w}{x}}$, Where W is weights

(e) **Let x, y are two numbers then**: A.M = $\frac{x+y}{2}$, G.M = $\sqrt{xy}$, H.M = $\frac{2xy}{x+y}$

(f) **Relationship among the Average**: In any distribution where the original items differ in size, then either the values of A.M > G.M > H.M (or)

H.M < G.M < A.M in case all items are identical then A.M. = G.M = H.M

(g) **Combined Harmonic Mean**: $\left[\frac{N_1 + N_2}{\frac{N_1}{H.M_1} + \frac{N_2}{H.M_2}}\right]$

IV. MEDIAN AND OTHER POSITIONAL MEASURES: Median is denoted by M. It is a positional measurement. Median is dividing a series in two equal parts. i.e., the middle most items are called median.

(a) Individual Series: The terms are arranged in ascending (or) descending order.

(1) When number of terms is odd then

Median (M) = size of $\left(\frac{N+1}{2}\right)$ th item

Where M is Median, N is No. of terms in the given series.

(2) When number of terms is even, then Median (M) =

mean of = $\left(\frac{N}{2}\right)$ and $\left(\frac{N}{2}+1\right)$th terms = $\frac{\left[\left(\frac{N}{2}\right)+\left(\frac{N}{2}+1\right)\right]}{2}$ the terms.

(b) Discrete Series:

Median (M) = Size of $\left(\frac{N+1}{2}\right)$ th item

Where N = $\sum f$ = Total frequency

(c) Continuous Series:

Median (M) = $L_1 + \frac{N_1 - c.f}{f} \times c$

Where L_1 is lower limit of median class interval

c.f is the value in column just above $N_1 = \frac{N}{2}$

f is the frequency of median class, c is the class interval of median class

(d) Quartiles, Quintiles, Octiles, Deciles and Percentiles:

(1) Individual Series: First arrange the items in Ascending order

First Quartile (or) Lower Quartile Q_1 = size of $\left(\frac{N+1}{4}\right)^{th}$ term

Third Quartile (or) Upper Quartile Q_3 = size of 3 $\left(\frac{N+1}{4}\right)^{th}$ term

n^{th} quintile $q_n = \frac{n(N+1)}{5}^{th}$ term

n^{th} Octile $o_n = \frac{n(N+1)}{8}^{th}$ term

n^{th} Decile $D_n = \frac{n(N+1)}{10}^{th}$ term

n^{th} Percentile $p_n = \frac{n(N+1)}{100}^{th}$ term

Where N is total No. of items in the given series

In case of quintile n = 1, 2, 3, 4

Octile n = 1, 2, 3, 4, 5, 6, 7

Decile n = 1, 2, 3, 4, 5, 6, 7, 8, 9

Percentile n=1, 2, 3,………… 99

(ii) **Discrete Series**: First find out cumulative frequency column.

First Quartile (or) Lower Quartile Q_1 = size of $\left(\frac{N+1}{4}\right)^{th}$ item

Third Quartile (or) Upper Quartile Q_3 = size of 3 $\left(\frac{N+1}{4}\right)^{th}$ item

n^{th} quintile q_n = Size of $\frac{n(N+1)}{5}^{th}$ item

n^{th} Octile O_n = Size of $\frac{n(N+1)}{8}^{th}$ item

n^{th} Decile D_n = Size of $\frac{n(N+1)}{10}^{th}$ item

n^{th} Percentile P_n = Size of $\frac{n(N+1)}{100}^{th}$ item

Where $N = \sum f$ = Total frequency

(iii) **Continuous Series**: i.e. = $L_1 + \frac{N_1 - c.f}{f} \times c$

Where $N_1 = \frac{N}{2}$ for Median

$= \frac{N}{4}$ for first Quartile (Q_1)

$= 3\frac{N}{4}$ for third quartile (Q_3)

$= \frac{nN}{5}$ for n^{th} quintile (q_n)

$= \frac{nN}{8}$ for n^{th} Octile (O_n)

$= \frac{nN}{10}$ for n^{th} Deciles (D_n)

$= \frac{nN}{100}$ for n^{th} Percentile (P_n)

6. Find Median from following data:

17	19	21	13	16	18	24	22	20

Solution:

Arranging the terms in ascending order

13	16	17	18	19	20	21	22	24

Total number of terms = 9 or n = 9

Now $\frac{n+1}{2} = \frac{9+1}{2}$ 5

Median = 5^{th} term = 19

7. Compute Median for following data:

X:	10	20	30	40	50	60	70
F:	4	7	21	34	25	12	3

Solution:

X	**F**	**C_f**
10	4	=4
20	7	(4+7)=11
30	21	(11+21)=32
40	34	(32+34)=66
50	25	(66+25)=91
60	12	(91+12)=103
70	3	(103+3)=106
	N =106	

Now M = Size of $\left[\frac{N+1}{2}\right]^{th}$ term

= size of $\left[\frac{106+1}{2}\right]^{th}$ term

= Size of 53.5^{th} term

Thus Median = **40**

8. Calculate Median from following data. Case of unequal class-internals.

Class Intervals	4-8	8-20	20-28	28-40	40-60	60-72
Frequency	7	12	42	56	39	22

Solution:

X	f	C_f
4-8	7	7
8-20	12	19
20-28	42	61
28-40	56	117
40-60	39	156
60-72	22	178
	N = 178	

$N_1 = \frac{178}{2} = 89$, Cf = 61, f = 56, L = 28, i = 12,

$M = L + \frac{N_1 - c.f}{f} \times i = 28 + \frac{89-61}{56} \times 12 = \mathbf{34}$

V. MODE: It is denoted by 'Z'. Mode may be defined as the value that occurs most frequently in a statistical distribution.

(i) **Individual Series**: The terms are arranged in any order, Ascending or Descending. If each term of the series is occurring once, then thee is no mode, otherwise the value that occurs maximum times are known as Mode.

(ii) **Discrete Series**: Here the mode is known by Inspection Method only. Here that variable is the mode, where the frequency is highest. For such a distribution we have to prepare (a) Grouping Table (b) Analysis Table.

(iii) **Continuous Series**:

$$\text{Mode } (Z) = L_1 + \frac{f_1 - f_0}{2f_1 - f_0 - f_2} \times C \text{ (or) } Z = L_1 + \frac{D_1}{D_1 + D_2} \times C$$

Where L_1 is the lower limit of modal class interval

f_1 is the frequency corresponding to modal class interval

f_0 is the frequency preceding Modal class interval

f_2 is the frequency succeeding Modal class interval

C is the length of Modal class interval

$D_1 = f_1 - f_0$, $D_2 = f_1 - f_2$

Note: Class intervals must be exclusive, equal, in ascending order, not cumulative.

If Modal value lies in any other interval than with highest frequency, then

$$\text{Mode (Z)} = L_1 + \frac{f_2}{f_0 + f_2} \times C$$

9. Find Mode from the following data.

12	14	16	18	26	16	20	16	11	12	16	16	20	24

Solution:

Arrange above data in ascending order

11	12	12	14	15	16	16	16	16	18	20	20	24	26

Here we get 16 four times, 12 and 20 two times each and other terms once only. Thus Z = 16.

10. Find Mode from the following data

X:	5	10	15	20	25	30	35	40	45
F:	1	3	4	9	11	12	3	2	1

Solution:

(Note: - Since we can't make use of inspection method as the frequencies are not most concentrated about highest frequency 12. Thus we will have to proceed for the tables.

Grouping Table:

X	f(I)	II	III	IV	V	VI
5	1					
10	3	4				
15	4				16	
20	9	13	20			24
25	11			32		
30	12	23	15		26	
35	3					17
40	2	5	4	7		
45	2					

Analysis Table:

Column X	I	II	III	IV	V	VI	Total
5							-
10							-
15						X	1
20			X	X		X	3
25		X	X	X	X	X	5
30	X	X		X	X		4
35					X		1
40							-
45							-

Here 25 has occurred maximum times (5), thus Modal Value is 25

11. Calculate Mode for the following data

C.I	0-10	10-20	20-30	30-40	40-50	50-60	60-70
F:	4	13	21	44	33	22	7

Solution:

Grouping Table

C.I	f(I)	II	III	IV	V	VI
0-10	4					
10-20	13		34	38		
20-30	21 f_0	65			78	
30-40	44 f_1		77			98
40-50	33 f_2			99		
50-60	22	55			62	
60-70	7		29			

Analysis Table

C.I	I	II	III	IV	V	Total
0-10						-
10-20					X	1
20-30		X			X	2
30-40	X	X	X	X	X	5
40-50			X	X		2
50-60				X		1
60-70						-

Thus Modal Interval is 30-40

L = 30

$D_1 = 44 - 21 = 23$; $D_2 = 44 - 33 = 11$

$i = 10$

As $Z = L + \frac{D_1}{D_1 + D_2} \times i$ or, $Z = L + \frac{Z_1 + Z_0}{2Z_1 - Z_0 - Z_2} \times i$, ($Z_1 = 44$, $Z_0 = 21$, $Z_2 = 33$)

$$Z = 30 + \frac{23}{23+11} \times 10 = 30 + \frac{230}{34} = 30 + 6.76 = \mathbf{36.76}$$

(iv) Mean, Median and Mode – Their Relation: $Z = 3M - 2\overline{X}$

Where Z is Mode, M is Median, $\overline{X}$ is Mean

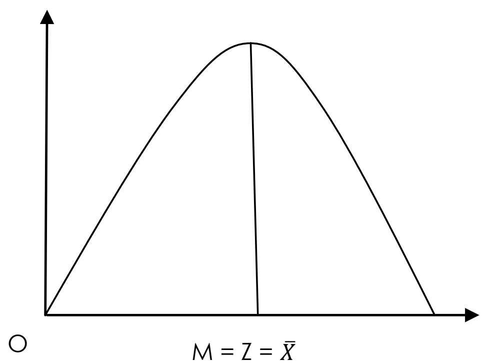
This formula was expressed by Karl Pearson. It is called Bi-mode.

(v) Incase of Symmetrical Series, the mean, median and mode coincide. i.e., $Z = M = \overline{X}$

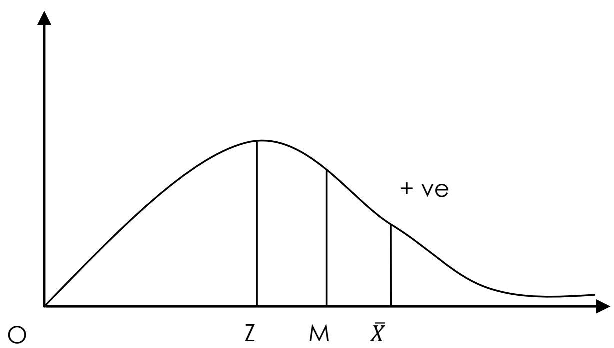
(vi) Incase of Positive Skewed: If the tail is towards right, then it is called positive skewness. It means $Z < M < \overline{X}$

(vii) Negative Skewed: When the tail of a distribution is towards left, then the skew ness is negative i.e. $Z > M > \overline{X}$

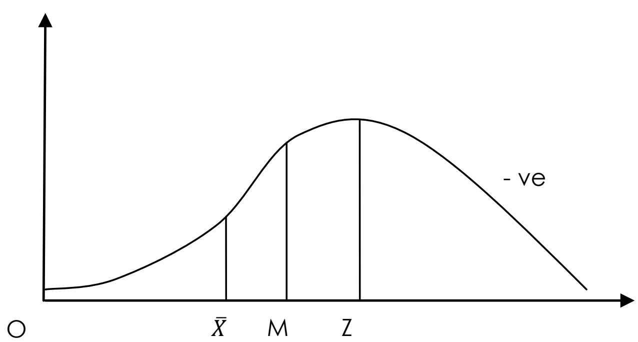
(viii) It should be noted that in both '+'ve and '-'ve skewed distributions that median lies in between the mode and the mean.



Symmetrical



Positive Skewed



Negative Skewed

Uses of various Averages: The use or application of a particular average depends upon the purpose of the investigation. Some of the cases of different averages are as follows:

(a) **Arithmetic Mean**: Arithmetic mean is considered an ideal average. It is frequently used in all the aspects of life. It possesses many mathematical properties and due to this it is of immense utility in further statistical analysis. In economic analysis arithmetic mean is used extensively to calculate average production, average wage, average cost, per capita income, exports, imports, consumption, prices etc. When different items of a series have different relative importance, then weighted arithmetic mean is used.

(b) **Geometric Mean**: Use of Geometric mean is important in a series having items of wide dispersion. It is used in the construction of Index Number. The averages of proportions, percentages and compound rates are computed by geometric mean. The growth of population is measured in it as population increases in geometric progression.

(c) **Harmonic Mean**: Harmonic mean is applied in the problems where small items must get more relative importance than the large ones. It is useful in cases where time, speed, values given in quantities, rate and prices are involved. But in practice, it has little applicability.

(d) **Median and Partition values**: Median and partition values are positional measures of central tendency. These are mainly used in the qualitative cases like honesty, intelligence, ability etc. In the distributions which are positively skewed, median is a more suitable average. These are also suitable for the problems of distribution of income, wealth, investment etc.

(e) **Mode**: Mode is also positional average. Its applicability to daily problems is increasing. Mode is used to calculate the 'modal size of a collar', modal size of shoes,' or 'modal size of ready-made garments' etc. It is also used in the sciences of Biology, Meteorology, Business and Industry.

MULTIPLE CHOICE QUESTIONS

Exercise - I

1. Measures of central tendency are called averages of the ____ order.
   (a) 1^{st}
   (b) 2^{nd}
   (c) 3^{rd}
   (d) None
2. A measure of central tendency tries to estimate the
   (a) Central value
   (b) Lower value
   (c) Upper value
   (d) None
3. The most commonly used measure of central tendency is:
   (a) Mode
   (b) Median
   (c) Mean
   (d) None
4. The algebraic sum of deviations of a set of observations from their AM is
   (a) Negative
   (b) Positive
   (c) Zero
   (d) None of these
5. If there are 3 observations 15, 20, 25 then the sum of deviation of the observations from their AM is
   (a) 0
   (b) 5
   (c) -5
   (d) None of these
6. The sum of the squares of the deviations of the variable is ___ when taken about AM
   (a) Maximum
   (b) Zero
   (c) Minimum
   (d) None
7. Pooled mean is also called
   (a) Mean
   (b) Geometric mean
   (c) Grouped mean
   (d) None

8. The mean of first five prime numbers is
   (a) 3
   (b) 3.6
   (c) 7
   (d) 5.6
9. Mean of 25, 32, 43, 53, 62, 59, 48, 31, 24, 33 is
   (a) 44
   (b) 43
   (c) 42
   (d) 41
10. The AM of 1, 3, 5, 6, x, 10 is 6. The value of x is
   (a) 10
   (b) 11
   (c) 12
   (d) None
11. The mean height of 8 students is 152 cm. Two more students of heights 143 cm and 156 cm join the group. New mean height is equal to
   (a) 153
   (b) 152.5
   (c) 151.5
   (d) 151
12. If each item is reduced by 15. AM is
   (a) Reduced by 15
   (b) Increased by 15
   (c) Reduced by 10
   (d) None
13. Find the AM for the following distribution:

Class Interval	350-369	370-389	390-409	410-429	430-449	450-469	470-489
F	23	38	58	82	65	31	11

   (a) 416.71
   (b) 520.13
   (c) 432.62
   (d) 225.71
14. Following is an incomplete distribution having modal mark as 44

Marks	0.20	20-40	40-60	60-80	80-100
No. of students	5	18	?	12	5

What would be the mean marks?

(a) 45

(b) 46

(c) 47

(d) 48

15. Given that the mean height of a group of students is 67.45 inches. Find out the missing frequencies for the following incomplete distribution of height of 100 students.

CI	60-62	63-65	66-68	69-71	72-74
F	5	18	?	?	8

(a) 35, 16

(b) 52, 12

(c) 42, 27

(d) 12, 9

16. The average marks scored by 50 students in a class were calculated to be 38. Later it was found, that marks of two students were wrongly copied as 34 and 23 instead of 43 and 32. Find correct average marks.

(a) 37.36

(b) 39.00

(c) 38.36

(d) None of these

17. If a variable assumes the values 1, 2, 3…. 5 with frequencies as 1, 2, 3,….5 then what is the AM?

(a) 11/3

(b) 5

(c) 4

(d) 4.50

18. If $y = 5x - 20$ & $\bar{x} = 30$ then the value of $\bar{y}$ is

(a) 130

(b) 140

(c) 30

(d) None

19. The mean salary for a group of 40 female workers is 5200 per month and that for a group of 60 male workers is 6800 per month. What is the combined mean salary?

(a) 6500

(b) 6200

(c) 6160

(d) 6100

20. If there are two groups containing 30 and 20 observations and having 50 and 60 as arithmetic means, then the combined arithmetic mean is
    (a) 55
    (b) 56
    (c) 54
    (d) 52
21. The average salary of a group of unskilled workers is ₹10,000 and for a group of skilled workers is ₹15,000. Combines salary is ₹12,000. What is the percent of skilled workers?
    (a) 40%
    (b) 50^
    (c) 60%
    (d) None of these
22. ____ is useful in averaging ratios, rates and percentages
    (a) AM
    (b) GM
    (c) HM
    (d) None
23. ____ is used when rate of growth or decline required.
    (a) Mode
    (b) AM
    (c) GM
    (d) None
24. What is the GM for the numbers 8, 24, and 40?
    (a) 24
    (b) 12
    (c) $8\sqrt[3]{15}$
    (d) 10
25. The greater of the two numbers whose arithmetic mean is 34 and the geometric mean is 16
    (a) 4
    (b) 256
    (c) 68
    (d) 64
26. If the AM and GM for two numbers are 6.50 and 6 respectively then the two numbers are
    (a) 6 and 7
    (b) 9 and 4
    (c) 10 and 3
    (d) 8 and 5

27. The relationship between AM, GM & HM is
   (a) GM = (AM) X (HM)
   (b) $(GM)^2$ = (AM) X (HM)
   (c) GM = $(AM \times HM)^2$
   (d) $GM^2 = (AM)^2 \times (HM)^2$
28. ____ & _____ are called ratio averages
   (a) HM & GM
   (b) HM & AM
   (c) AM & GM
   (d) None
29. ____ is used for calculation of speed and velocity
   (a) GM
   (b) AM
   (c) HM
   (d) None is used
30. What is the HM of 1, ½, 1/3, …. 1/n?
   (a) n
   (b) 2n
   (c) $\frac{2}{(n+1)}$
   (d) $\frac{n(n+1)}{2}$
31. An aeroplane flies from A to B at the rate of 500 km/hour and comes back from B to A at the rate of 700 km/hour. The average speed of the aeroplane is
   (a) 60 km per hour
   (b) 583.33 km per hour
   (c) 100 $\sqrt{35}$ km per hour
   (d) 620 km per hour
32. In ___ the distribution has open – end classes
   (a) Median
   (b) Mean
   (c) Standard deviation
   (d) None
33. ____ always lies in between the arithmetic mean and mode.
   (a) GM
   (b) HM
   (c) Median
   (d) None

34. Median divides the total number of observations into parts
    (a) 3
    (b) 4
    (c) 5
    (d) 2
35. 50% of actual values will be below & 50% of will be above ___
    (a) Mode
    (b) Median
    (c) Mean
    (d) None
36. In Ogive, obscissa corresponding to ordinate N/2 is
    (a) Median
    (b) 1^{st} quartile
    (c) 3^{rd} quartile
    (d) None
37. The second quartile is known as
    (a) Median
    (b) Lower quartile
    (c) Upper quartile
    (d) None
38. Which of the following relationship is true in a symmetrical distribution?
    (a) Median – Q_1 = Q_3 – Median
    (b) Median – Q_1 > Q_3 – Median
    (c) Median – Q_1 < Q_3 – Median
    (d) Median – Q_1 ≠ Q_3 – Median
39. What is the median for the following observations 5, 8, 6, 9, 11, 4
    (a) 6
    (b) 7
    (c) 8
    (d) None of these
40. For the values of a variable 3, 1, 5, 2, 6, 8, 4 the median is
    (a) 3
    (b) 5
    (c) 4
    (d) None

41. If the median of 5, 9, 11, 3, 4, x, 8 is 6, the value of x is equal to
   (a) 6
   (b) 5
   (c) 4
   (d) 3
42. If the difference between mean and mode is 63, the diference between mean and median is
   (a) 189
   (b) 21
   (c) 31.5
   (d) 48.5
43. If the relationship between x and y is given by 4x - 6y = 13 and if the median of x is 16. Find median of y.
   (a) 7.50
   (b) 8.00
   (c) 8.50
   (d) None of these
44. The variables x and y are related by 5x + 6y = 70 and median of x is 8. What is the median of y?
   (a) 4
   (b) 4.5
   (c) 6
   (d) 5
45. Quartiles can be determined graphically using
   (a) Histogram
   (b) Frequency polygon
   (c) Ogive
   (d) Pie chart
46. In Ogive, 1^{st} quartile is obscissa corresponding to ordinate ___
   (a) N/2
   (b) N/4
   (c) 3N/2
   (d) 3N/4
47. First quartile is ___
   (a) Lower quartile
   (b) Upper quartile
   (c) Middle quartile
   (d) Highest value

48. Third quartile is ___
    (a) Lower quartile
    (b) Upper quartile
    (c) Middle quartile
    (d) Highest value
49. In Bowley's formula how many quartiles are used.
    (a) 3
    (b) 2
    (c) 1
    (d) None
50. Find Q1 for the following observations 14, 16, 13, 15, 20, 18, 19, 22
    (a) 14
    (b) 14.25
    (c) 15
    (d) 15.25
51. What is the value of the first quartile for observations 15, 18, 10, 20, 23, 28, 12, 16?
    (a) 17
    (b) 16
    (c) 12.75
    (d) 12
52. If the first quartile is 104 and quartile deviation is 18, the third quartile will be
    (a) 140
    (b) 116
    (c) 20
    (d) 0
53. The third decile for the numbers 15, 10, 20, 25, 18, 11, 9, 12 is
    (a) 13
    (b) 10.70
    (c) 15
    (d) 75
54. Find D6 for the following observations. 7, 9, 5, 4, 10, 15, 14, 18, 6, 20
    (a) 11.40
    (b) 12.40
    (c) 13.40
    (d) 13.80

55. 25^{th} percentile is equal to

(a) 1^{st} quartile

(b) 25^{th} quartile

(c) 24^{th} quartile

(d) None

56. The abscissa of the maximum frequency in the frequency curve is the

(a) Mean

(b) Median

(c) Mode

(d) None

57. For a moderately skewed distribution, which of he following relationship holds?

(a) Mean – mode = 3 (Mean – Median)

(b) Median – Mode = 3 (Mean-Median)

(c) Mean – Median = 3 (Mean –Mode)

(d) Mean – Median =3 (Median – Mode)

58. Modal value of 0, 3, 5, 6, 7, 9, 12, 0, 2 is

(a) 6

(b) 0

(c) 3

(d) 5

59. The mode for the no.s 5, 8, 6, 4, 10, 15, 18, 10 is

(a) 18

(b) 10

(c) 14

(d) None of these

60. What is the modal value for the numbers 4, 3, 8, 15, 4, 3, 6, 3, 15, 3, 4.

(a) 3

(b) 4

(c) 15

(d) None of these

61. The mode for the no.s 7, 7, 7, 9, 10, 11, 11, 11, 12 is

(a) 11

(b) 12

(c) 7

(d) 7 & 11

62. Compute the Modal group

Marks	60-62	63-65	66-68	69-71	72-74
No. of students	15	118	142	127	18

(a) 66-68

(b) 69-71

(c) 63-65

(d) None

63. If x and y are related by x-y-10 = 0 and mode of x is known to be 23, then the mode of y is

(a) 20

(b) 13

(c) 3

(d) 23

64. For a moderately skewed distribution of marks in statistics for a group of 100 students, the mean mark and median mark were found to be 50 and 40. What is the modal mark?

(a) 15

(b) 20

(c) 25

(d) 30

65. Which of the following measure(s) satisfies (satisfy) a linear relationship between two variables?

(a) Mean

(b) Median

(c) Mode

(d) All of these

66. An ideal measure of central tendency is

(a) Moving average

(b) Median

(c) Harmonic Mean

(d) Arithmetic mean

67. Mathematical average is called

(a) Arithmetic mean

(b) Geometric mean

(c) Mode

(d) None of these

68. Sum of deviations of the items is zero from

(a) Mean

(b) Median

(c) Mode

(d) Geometric mean

69. Which of the following is based on the formula: $\overline{X} = A + \frac{\Sigma fd}{N}$

(a) Median

(b) Mode

(c) Arithmetic mean

(d) Harmonic mean

70. Which of the following is based on the following formula: $\overline{X} = A + \frac{\Sigma fd^{1}}{N} \times C$

(a) Step Deviation method

(b) Direct Method

(c) Short-cut Method

(d) None of these

71. The positional average is

(a) Harmonic mean

(b) Geometric mean

(c) Median

(d) Weighted arithmetic mean

72. For dealing with qualitative data the best average is:

(a) Median

(b) Mode

(c) Geometric mean

(d) Arithmetic mean

73. Intelligence Quotient (I.Q.) can be measured by using

(a) Quartile

(b) Median

(c) Mode

(d) Mean

74. Which partition value divides the series into two equal parts.

(a) P_{10}

(b) P_{5}

(c) P_{50}

(d) P_{90}

75. The partition value which divides the series into 10 equal parts

(a) Quartile

(b) Percentile

(c) Median

(d) Decile

76. The most ill-defined average is

(a) Median

(b) Mean

(c) Quartiles

(d) Mode

77. Which of the following is the most unstable averages

(a) Mean

(b) Median

(c) Mode

(d) G.M.

78. One of the following methods of calculating mode is:

(a) Mode = 3 Median – 2 Mean

(b) Mode = 2 Median – 3 Mean

(c) Mode = 3 Median + 2 Mean

(d) Mode = 2 Median – 2 Mean

79. A distribution with two modes is called:

(a) Unimode

(b) Bimodal

(c) Multimodal

(d) None of these

80. Mode can be located graphically with the help of

(a) Frequency rectangles

(b) Ogive curves

(c) Frequency curve

(d) Frequency polygon

81. For studying index numbers the best average is:

(a) Geometric mean

(b) Harmonic mean

(c) Arithmetic mean

(d) None of these

82. Which is the most suitable average to give equal importance to the equal rate of change

(a) Harmonic mean

(b) Geometric mean

(c) Arithmetic mean

(d) None of these

83. When compared to arithmetic mean the value of geometric mean is
    (a) Greater than arithmetic mean
    (b) Equal to arithmetic mean
    (c) Less than arithmetic mean
    (d) None of these

84. Geometric mean is
    (a) Calculated average
    (b) Locational average
    (c) Positional average
    (d) None of these

85. Geometric mean is termed as
    (a) Moving average
    (b) Ratio average
    (c) Progressive average
    (d) None of these

86. Harmonic mean is a
    (a) Positional average
    (b) Moving average
    (c) Calculated average
    (d) None of these

87. The average based on reciprocals to the numbers is
    (a) Arithmetic mean
    (b) Geometric mean
    (c) Harmonic mean
    (d) Mode

88. The average dealing with the problems of time and distance is termed as
    (a) Geometric mean
    (b) Harmonic Mean
    (c) Median
    (d) Arithmetic mean

89. Harmonic mean is used for calculating
    (a) Average growth rates of variables
    (b) Average speed of journey
    (c) Average rate of increase in net worth of a company
    (d) All the above 1 to 3.

90. In a moderately skew distribution
    (a) $\overline{X}$ > G.M. > H.M.
    (b) $\overline{X}$ < G.M. > H.M.
    (c) $\overline{X}$ > G.M. < H.M.
    (d) $\overline{X}$ < G.M. < H.M.

State the following statements are true or false

(1) The arithmetic mean is always the best measure of central tendency

(2) The sum of individual observations from mean is zero

(3) In a moderately skewed distribution A.M. < G.M. < H.M

(4) The addition of a constant value to each of the values of a series increases the average by the same value of the constant.

(5) Average alone is enough to throw light on the main characteristics of a statistical series.

(6) Median is a mathematical average

(7) The value of median and mode can be determined graphically

(8) Combined median can be calculated as in case of arithmetic mean

(9) Percentile divides the series in ten equal parts

(10) The value of median is affected more by sampling fluctuations than mean.

(11) Mode is the value that has maximum frequency

(12) Mode can be located graphically

(13) In a positively skewed distribution mode > mean

(14) Mode is a mathematical average

(15) A distribution with more than two modes is called multimodal

(16) Geometric mean is a positional measure of central tendency

(17) Geometric mean is more suitable for dealing with problems of rates and speed

(18) Geometric mean can be computed in case of open end series

(19) Combined G.M. of two or more series can be calculated

(20) The geometric mean is the nth root of the product of n items in a given distribution.

(21) Harmonic mean is used when data are given in terms of rates

(22) In a moderately asymmetrical distribution A.M. < G.M. < H.M.

(23) Harmonic mean is a positional average

(24) Harmonic mean is based on all the items in a series

(25) Harmonic mean is always less then geometric mean.

Short Answer Type Questions

1. What do you mean by "Measures of Central Tendency"?

2. What are different types of averages?

3. If some AMs are given of different series, how can we find combined $\overline{X}$ for all of those? Also provide formula.

4. What do you mean by Weighted Arithmetic Mean? How to calculate it? Give formula also.

5. Define –

   (a) Median (b) Quartiles (c) Lower Quartile (d) Upper quartile (e) Decile (f) Percentile

6. How can we calculate (i) Median (ii) Q_1 (iii) Q_3 (iv) Decile (v) Percentile in Individual or Discrete Series? Also write steps to calculate.

7. How to calculate the above said measures in a continuous series? Write steps to calculate.

8. What amendments are to be made for

   (i) Inclusive Series (ii) Cumulative series (iii) Open End series (iv) Unequal Interval Series.

9. What are Positional Value or Partition Measures? What are its various types?

10. Define mode

11. What is the Empirical relation between Mean, Mode and Median? Who invented it?

12. On what bases we Empirical relation derived. Derive the formula

13. When is empirical formula used to calculate Mode? How is it used?

14. Give the precise definition of G.M. along with the formula of its calculation

15. G.M. is a calculated average. Explain.

16. Give three merits of G.M.

17. Give three demerits of G.M.

18. Give three important properties of G.M.

19. Which formula is used to show the rise in prices or the increase in population?

20. Give meaning of H.M.

21. Give three merits of H.M.

22. Give three important limitations of H.M.

23. Prove AM > GM > HM

24. Under what circumstances A.M. = G.M. = H.M.

25. Why H.M. is indeterminate when any item of the given data is zero or negative?

Exercise Problems

1. Calculate X for the following data by direct, short cut and step deviation methods.

70	65	55	75	80	85	65	70	95

2. Calculate $\overline{X}$ for the following series by Direct, Short-cut and step-deviation Methods.

X:	5	10	15	20	25	30	35	40
F:	6	17	28	34	18	11	9	7

3. A class has 50 students with average weight of 45 kgs. Out of these there are 30 girls with average weight of 42.5 kgs. Find average wt. of boys.

4. A class of 40 students has an average of 56 marks in Math exam. But later on it was found that terms 48, 54 and 67 were misread as 68, 45 and 87. Find correct mean.

5. $\overline{X}$ for 20 items was 36, But two terms were taken as 47 and 56 instead of 67 and 65. Find correct mean.

6. Average of 10 terms is 6, Find new average if each term is (i) multiplied (ii) divided by 2 (iii) 3 is added to each term (iv) 4 is subtracted from each term.

7. Find $\overline{X}$ for the following series. (inclusive Series)

Class Interval:	3-5	6-8	9-11	12-14	15-17	18-20
Frequency:	3	7	16	34	17	3

8. Equal Intervals:

Class Interval:	Less than 50	50-75	75-100	100-125	125-150	More than 150
Frequency:	21	47	67	89	55	21

9. Mean weight of students in a class is 48 kg. If men weight of girl students is 40 kg. and that of boys is 60 kg. Find (a) % age of boys and girls (b) If there are total of 75 students, find number of each.

10. Mean of a series with 50 terms to 80; But afterwards it was noted that three terms 63, 47 and 88 were misread as 36,74 and 63. Find correct

11. From a data sheet, we get following 9 terms

13	17	11	19	9	14	21	23	16

But we are told that other six terms are missing

What will be the value of

M and Q_3, if first six last terms are missing

Q_1, and M, if last six terms are missing

Q_1, M and Q_3 if first three and last three terms are missing.

12. Median of a series is 80, but four terms 38, 66, 93 and 96 were misread as 83, 88, 39 and 69 find correct Median.

13. Calculate M, Q_1, Q_3, D_1 and P_{82} for the following data

21	13	17	11	19	9	16	23	14

14. Calculate M, Q_1, Q_3, D_6 and P_{91} for following data

X:	8	12	20	25	30	40
F:	9	16	28	46	20	10

15. Calculate Median Q_1, Q_3, D_4 and P_{86} for following data

X:	0-4	4-8	8-12	12-16	16-20	20-24	24-28	28-32
F:	4	17	36	90	123	110	66	14

16. Find M, Q_1, Q_3 for the following figures

X:	4-7	8-11	12-15	16-19	20-23	24-27	28-31
F:	7	11	23	47	36	29	17

17. Calculate D_1, M_1, Q_3 and P_{95} for the following data (Unequal Intervals)

Class Interval:	0-20	20-30	30-80	80-120	120-180	180-200
Frequency:	7	19	38	78	45	13

18. Find the value of missing frequencies if Q_1 = 36 and N = 840

X:	20-25	25-30	30-35	35-40	40-45	45-50	50-55	55-60
F:	40	50	?	150	180	130	?	60

19. Find out median and quartiles for the following data:

Marks	0-10	10-20	20-30	30-40	40-50	50-60	60-70	70-80	80-90	90-100
No. of Stu	11	18	25	28	30	33	22	15	12	10

20. Find the median, lower quartile, 7th decile of the Frequency distribution given below:
No. of students

8	12	20	32	30	28	12	4

21. Calculate P_{22} and D_8

X:	20-29	30-39	40-49	50-59	60-69	70-79
F:	2	19	20	21	15	13

22. In the batch of 15 students, 5 students failed. The marks of 10 students who passed were 9, 6, 7, 8, 8, 9, 6, 5, 4, 7 what were the median marks of all 15 students?

23. Compute Median for the following data

X:	10	20	30	40	50	60	70
F:	4	7	21	34	25	12	3

24. Given

(a) $\bar{X}$ = 42 .5, M = 41.0; Find Mode

(b) $\bar{X}$ = 160, M = 150; Find Mode

25. Given

(a) $\bar{X}$ = 56.2 , Z = 55; Find M

(b) Z= 72.0, M = 70; Find

26. Evaluate Z for the following distributions

11	7	9	14	12	21	18	14	17	21	14	15	18	10

27. Evaluate Z for following distributions

X:	3	7	10	14	22	30
F:	1	3	9	17	14	5

28. (a) Find Z for the following data

X:	10	20	30	40	50	60	70	80
F:	4	9	30	48	51	24	12	2

(Note: We find here Bimodal series. Thus calculate Z = 3 M- 2$\overline{X}$)

(b) Find out the mode of the following series:

Size (x):	5	6	7	8	9	10	11	12	13	14	15	16	17	18	19
Frequencies (f):	48	52	56	60	63	57	55	50	52	41	57	63	52	48	40

29. Find Z for following data. (Cumulative Series)

X	Less than 100	90	80	70	60	50	40	30	20
F	500	467	404	295	177	102	47	21	6

30. Find Z for the following data (Unequal Intervals, Type 1)

X:	0-5	5-10	10-20	20-40	40-50	50-60	60-80	80-95	95-100
F:	3	4	9	25	19	23	36	4	7

31. Find Z for the following data (Bimodal Series)

X	0-5	5-10	10-15	15-20	20-25	25-30	30-35	35-40
F	2	3	10	16	17	8	4	1

32. Find Z for following data $\left[Z = L + \frac{f_1 - f_0}{2f_1 - f_0 - f_2} \times i \text{ fails}\right]$

X	0-20	20-40	40-60	60-80	80-100	100-120	120-140	140-160
F	14	26	33	36	39	18	6	2

33. M and Z of the following given series in 67 and 68. If N = 115; Find the values of missing frequencies.

X	0-20	20-40	40-60	60-80	80-100	100-120	120-140
F	2	8	30	?	?	?	2

34. If X = 61 and Z = 63.2 find M.

35. Find Mode from the following data

12	14	16	18	26	16	20	16	11	12	16	15	20	24

36. Find Mode for following data

X:	4	7	11	16	25
F:	3	9	14	21	13

37. Calculate Mode for the following data

X:	0-10	10-20	20-30	30-40	40-50	50-60	60-70
F:	4	13	21	44	33	22	7

38. Calculate Geometric mean from the following data

X: (Variable)	10	52	110	120	50	80	60	37

39. From the following data calculate Geometric mean

X	135	79	38	286	59	176	7

40. Compute Geometric mean

X	10	15	18	20	25
F	2	3	5	6	4

41. The following distribution relates to marks in Economics of 60 students

Marks:	0-10	10-20	20-30	30-40	40-50	50-60
Students	3	8	15	20	10	4

Compute G.M.

42. Find w_1 given the geometric mean = 15.3

X	8.2	24.8	17	30
W	5	3	4	w_1

43. Calculate geometric mean from the following data:

X	10	110	135	120	50	59	60	7

44. Find the geometric men from the following data:

X	2	3	5	6	4
Frequencies (f)	10	15	18	12	7

45. Find the geometric mean from the following data:

x:	10-20	20-30	30-40	40-50	50-60
F	5	10	15	7	4

46. Calculate Harmonic mean from the following data

X	3834	382	63	0.8	0.4	0.03	0.009	0.0005

47. Calculate Harmonic mean from the following data

x:	10	20	25	40	50
F	20	30	50	15	5

48. Calculate Harmonic mean from the following data

x:	0-10	10-20	20-30	30-40	40-50
F	4	6	10	7	3

49. An aeroplane covers the four sides of a square at varying speeds of 500, 1000, 1500, 2000 km per hour respectively. What in the average speed of the plane around the square.

50. An aeroplane flies along the four sides of a square at speeds of 100, 200, 300 and 400 miles per hour respectively. What is the average speed of the plane in its flight around the square?

51. A man travels from Delhi to Bombay at an average speed of 30 m.p.h and returns along the same route at an average speed of 60 mph. Find the average speed for the total trip.

52. Find the value of Harmonic Mean: 5, 10, 12, 15, 22, 35, 50, 72.

53. Find the harmonic mean from the following data:

x:	10	20	40	60	120
F	1	3	6	5	4

54. Calculate Harmonic mean from the following data:

x:	0-10	10-20	20-30	30-40	40-50
F	2	7	13	5	3

55. If AM and GM of two values are 10 and 8 respectively, find those two values.

56. If for two numbers, the mean is 25 and the Harmonic mean is 9, what is the geometric mean?

57. If AM of two numbers is 17 and GM is 15, find HM.

MCQ Answers:

1	a
2	a
3	c
4	c
5	a
6	c
7	c
8	d
9	d
10	b
11	c
12	a
13	a
14	d
15	c
16	c
17	a
18	a
19	c
20	c

21	a
22	b
23	c
24	c
25	d
26	b
27	b
28	a
29	c
30	c
31	b
32	a
33	c
34	d
35	b
36	a
37	a
38	a
39	b
40	c
41	a
42	b
43	c
44	d
45	c
46	b
47	a
48	b
49	a
50	b
51	c
52	a
53	b
54	b
55	a
56	c
57	a
58	b
59	b
60	a

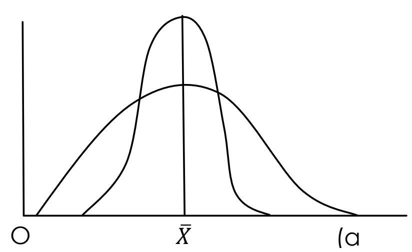
61	d
62	a
63	b
64	b
65	d
66	a
67	a
68	a
69	c
70	a
71	c
72	a
73	b
74	c
75	d
76	d
77	c
78	a
79	b
80	a
81	a
82	b
83	c
84	a
85	b
86	c
87	c
88	b
89	b
90	a

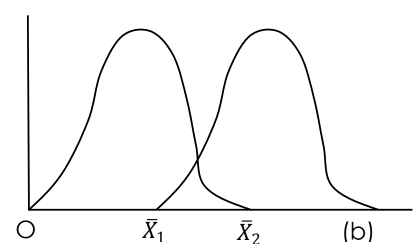
4.2 MEASURES OF DISPERSION

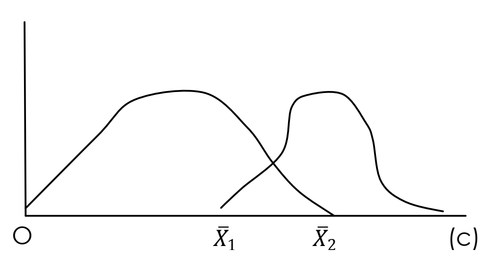
The various measures of central tendency discussed in the previous chapter give us one single value that represents the entire data. But the average alone cannot adequately describe a set of observations, unless all the observations are alike. It is necessary to describe the variability or dispersion of the observations. Also in two or more distributions the average value may be the same but still there can be wide disparities in the formation of the distributions. Measures of variation help us in studying the important characteristic of a distribution, i.e., the extent to which the observations vary from one another from some average value. The significance of the measure of variation can best be appreciated from the following example:

	Factory A wages (₹)	Factory B Wages (₹)	Factory C Wages (₹)
	2300	2310	2380
	2300	2300	2210
	2300	2304	2220
	2300	2306	2200
	2300	2280	2490
Total:	11,500	11,500	11,500
$\overline{X}$	2,300	2,300	2,300

The above data pertains to five workers each in three different factories. Since the average wage is the same in all factories, one is likely to conclude that the factories are alike in their wage structure, but a close examination shall revel that the wage distribution in the three factories differs widely from one another. In factory A, each and every worker is perfectly represented by the arithmetic mean, i.e., average wage or, in other words, none of the workers of factory A deviates from the arithmetic mean and hence there is no variation. In factory B, only one worker is perfectly represented by the arithmetic mean, the other workers vary from the mean but the variation is very small as compared to the workers of factory C. In factory C, the mean does not represent the workers as the individual wage figures differ widely from the mean. Thus we find there is no variation in the wages of workers in factory A, there is very little variation in factory B but the wages of workers of factory C differ most widely. For the student of social sciences, the mean wage is not so important as to know how these wages are distributed. Are there a large number receiving the mean wage or are there a few with enormous wages and millions with wages far below the mean? The following three diagrams represent frequency distribution with some of the characteristics we wish to emphasize:







The two curves in diagram (a) represent two distributions with the same mean $\overline{X}$, but with different variations. The two curves in (b) represent two distributions with the same variations but with unequal means, $\overline{X}_1$ and $\overline{X}_2$, Finally, (c) represent two distributions with unequal means and unequal variations.

The measures of central tendency are therefore, insufficient. They must be supported and supplemented with other measures. In this chapter, we shall be especially concerned with the measures of variation (or spread, or dispersion). Measure of variation is designed to state the extent to which the individual measures differ on an average from the mean. In measuring variation we shall be interested in the amount of the variation or its degree but not in the direction. For example, a measure of 6 centimeters below the mean has just as much variation as a measure of 6 centimeters above the mean.

Significance of Measuring Variation

Measures of variation are needed for four basic purposes:

(i) To determine the reliability of an average;

(ii) To serve as a basic for the control of the variability

(iii) To compare two or more series with regard to their variability; and

(iv) To facilitate the use of other statistical measures

A brief explanation of these points is given below:

(i) Measures of variation point out as to how far an average is representative of the entire data. When variation is small, the average is a typical value in the sense tht it closely represents the individual value and it sis reliable in the sense that it closely represents the individual value and it is reliable in the sense that it is good estimate of the average in the corresponding universe. On the other hand, when variation is large, the average is not so typical, and unless the sample is very large, the average may be quite unreliable.

(ii) Another purpose of measuring variation is to determine nature and cause of variation in order to control the variation itself. It matters of health, variation in body temperature, pulse beat and blood pressure are the basic guides to diagnosis. Prescribed treatment is designed to control their variation. In industrial production, efficient operation requires control of quality variation, the causes of which are sought through inspection and quality control programmes. Thus measurement of variation is basic to the control of cause of variation. In engineering problems, measures of variation are often specially important, in social sciences, a special problem requiring the measurement of variability is the measurement of "inequality" of the distribution of income and wealth, etc.

(iii) Measures of variation enable comparison to be made of two or more series with regard to their variability. The study of variation may also be looked upon as a means of determining uniformity or consistency. A high degree of variation would mean little uniformity or consistency whereas a low degree of variation would mean greater uniformity or consistency.

(iv) Many powerful analytical tools in statistics such as correlation analysis, the testing of hypothesis, the analysis of fluctuations, techniques of production control cost control, etc, are based on measures of variation of one kind or another.

Properties of a Good Measure of Variation: A good measure of variation should possess, as far as possible, the following properties:

(i) It should be simple to understand

(ii) It should be easy to compute

(iii) It should be rigidly defined.

(iv) It should be based on each and every observation of the distribution.

(v) It should be amendable to further algebraic treatment

(vi) It should have sampling stability

(vii) It should not be unduly affected by extreme observations.

DISPERSION: "Dispersion is a measure of variation of the items.

Methods of Measuring DISPERSION:

Types of Measures of Dispersion:

(A) **Absolute and Relative Measures**: Absolute measures of Dispersion are expressed in same units in which original data is presented but these measures cannot be used to compare the variations between the two series.

Relative measures are not expressed in units but it is a pure number.

It is the ratio of absolute dispersion to an appropriate average such as co-efficient of Standard Deviation or Co-efficient of Mean Deviation.

(B) **Methods of Measuring Dispersion**: Following methods are used to calculate dispersion.

(1) **Algebraic methods**:

(i) Methods of limits: (a) The Range (b) The Inter-quartile Range (c) The Percentile Range

(ii) Methods of moments: a) the first moment of dispersion or mean deviation (b) The second moment of dispersion or standard deviation.

(2) **Graphic Method**: Lorenz curve.

RANGE: It is the simplest of the values of Dispersion. It is merely the difference between the largest and smallest term. Symbolically; R = L - S (or) Range = Largest term – Smallest term and Coefficient of Range = $\frac{L-S}{L+S}$. It is also known as Ratio of Range or Co-efficient of scatteredness.

If the averages of the two distributions are close to each other, comparisons of the ranges show that the distribution with the smaller range has less dispersion. The average of that distribution is more typical of the group.

Methods of limits:

(i) **Range**: The difference between the largest and smallest term. It is denoted by 'R'.

Individual Series, Discrete Series and continuous series

Range = Largest term – smallest term

i.e., R = L-S and Coefficient of Range $\frac{L-S}{L+S}$

Note: Here Range is absolute measure and Co-efficient of Range is Relative measure.

Examples:

1. Find Range and coefficient of Range for following data

3	7	21	24	37	40	45

Solution:

Here L = 45 and S = 3

As Range = L-S

$\therefore$ Range = 45-3 = 42

And Coefficient of Range = $\frac{L-S}{L+S} = \frac{45-3}{45+3} = \frac{42}{48} = 0.875$

(Note Here range is absolute measure and Co-efficient of Range is Relative measure).

2. Find Range and Coefficient of Range for following data

X	5	10	15	20	25	30	35	40
f	4	7	21	47	53	24	12	6

Solution:

Going through the variables S = 5; and L = 40

$\therefore$ Range = 40-5 = 35 (R = L-S)

And Coefficient of Range $= \frac{40-5}{40+5}$ C.R $= \frac{L-S}{L+S}$

$= \frac{35}{45} = 0.778$ (Approx.)

(Note Here range is absolute measure and Co-efficient of Range is Relative measure).

3. Calculate Range and Coefficient of Range for following data

X	20-30	30-40	40-50	50-60	60-70	70-80
F	4	9	16	21	13	6

Solution:

Here L = 80 and S = 20

$\therefore$ Range = L-S = 80 – 20 = 60

And Coefficient of Range = $\frac{L-S}{L+S} = \frac{80-20}{80+20} = \frac{60}{100} = 0.6$

(ii) **Inter Quartile Range** = $Q_3 - Q_1$

Semi Inter Quartile Range = $\frac{Q_3 - Q_1}{2}$ (or) **Quartile Deviation**

Where Q_1 is lower Quartile, Q_3 is upper Quartile

Coefficient of Quartile Deviation = $\frac{Q_3 - Q_1}{Q_3 + Q_1}$

(iii) **Percentile and Deciles Range**: **Percentile Range** = $P_{90} - P_{10}$

Deciles Range = $D_9 - D_1$

Where P_{10} is 10^{th} percentile and P_{90} is 90^{th} percentile.

D_1 is first Deciles and D_9 is 9^{th} Deciles.

4. From the following data compute inter quartile range, quartile Deviation and Coefficient of Quartile Deviation.

24	7	11	9	17	3	20	14	4	22	27

Solution:

Arranging the series in Ascending Order

3	4	7	9	11	14	17	20	22	24	27

N = 11

Q_1 = Size of $\left[\frac{11+1}{4}\right]^{th}$ term

= Size of 3rd term = 7

Q_3 = Size of $\frac{3(11+1)}{4}^{th}$ term

= Size of 9^{th} term

= 22

$\therefore$ Inter Quartile Range = $Q_3 - Q_1$ = 22 – 7 = 15

Semi – Inter Quartile Range or Quartile Deviation (Q.D) = $\frac{Q_3 - Q_1}{2} = \frac{15}{2}$ = 7.5

Coefficient of Q.D = $\frac{Q_3 - Q_1}{Q_3 + Q_1} = \frac{22-7}{22+7} = \frac{15}{19}$ = 0.52 (App.)

5. Compute Inter quartile Range, Coefficient of Quartile Deviation and Percentile Range for following data.

X	4	8	12	16	20	24	28	32
f	4	9	17	40	53	37	24	16

Solution:

X	f	C.f
4	4	4
8	9	13
12	17	30
16	40	70
20	53	123
24	37	160
28	24	184
32	16	200
	N =200	

Q_1 = Size of $\left[\frac{200+1}{4}\right]^{th}$ term

= Size of 50.25th term

$\therefore Q_1$ = 16

Q_3 = Size of $\frac{3(200+1)}{4}^{th}$ term

= Size of 150.75^{th} term

$\therefore Q_3$ = 24

Inter Quartile Range = $Q_3 - Q_1$ = 24 – 16 = 8

Coefficient of Q.D = $\frac{Q_3 - Q_1}{Q_3 + Q_1} = \frac{24-16}{24+16} = \frac{8}{40}$ = 0.2

To Compute Percentile Range : -

P_{90} = Size of $\frac{90\,(200+1)}{100}$ th term

= Size of 180.9th term = 28

P_{10} = Size of $\frac{10(200+1)}{100}$ th term

= Size of 20.1th term = 12

And Percentile Range = $P_{90} - P_{10}$ = 28 – 12 = 16

6. From the following data find quartile deviation and its coefficient.

X	10-15	15-20	20-25	25-30	30-35	35-40	40-45	45-50
f	8	10	12	15	10	7	8	5

Solution:

X	f	C.f
10-15	8	8
15-20	10	18
20-25	12	30
25-30	15	45
30-35	10	55
35-40	7	62
40-45	8	70
45-50	5	75

Now N_1 for Q_1 = $\frac{N}{4}$ th item

= $\frac{75}{4}$ th item = 18.75th item

18.75th item lies in (20-25)

Where L = 20, f = 12

N_1 = 18.75 Cf = 18; i=5

Interpolating for Q_1

$$Q_1 = L + \frac{N_1 - Cf}{f} \times i$$

$$= 20 + \frac{18.75 - 18}{12} \times 5 = 20 + \frac{3.75}{12} = 23.125$$

Similarly N_1 for Q_3 = $\frac{3N}{4}$ th item

$N_1 = \frac{3 \times 74}{4}$ th item = 56.25th Item

56.25th Item lies in (35-40)

Where L = 35, I = 5, C.f = 55, f = 7, N_1 = 56.25

Putting value in the interpolation formula

Where,

L = Lower limit of Median Class

Cf = Preceeding Cumulative frequency of median class

f = Corresponding frequency of median class.

i = Width of median class.

$Q_3 = l + \frac{N_1 - C.f}{f} \times i$

$= 35 + \frac{56.25 - 55}{7} \times 5$

$= 35 + \frac{6.25}{7} = 35.893$

Quartile Deviation (Inter Quartile Range)

$= Q_3 - Q_1$

Q.D = 35.893 – 23.125

= 12.768

Coefficient of Q.D = $\frac{Q_3 - Q_1}{Q_3 + Q_1} = \frac{35.893 - 23.125}{35.893 + 23.125} = \frac{12.768}{59.018} = 0.216$

Methods of moments:

(i) **Mean Deviation (or) Average Deviation**: It is denoted by M.D.

Mean Deviation (M.D) = $\frac{\Sigma f|D|}{N}$

Where f is frequency, N is total frequency $|D| = |X - A|$

Coefficient of Mean Deviation:

(a) If deviations are taken from arithmetic mean $(\bar{X})$, Coefficient of M.D = $\frac{M.D}{\bar{X}}$; where $\bar{X}$ is mean

(b) If deviations are taken from median, (M), Coefficient of M.D = $\frac{M.D}{M}$; where M is median

(c) If deviations are taken from mode, (Z), Coefficient of M.D = $\frac{M.D}{Z}$; where Z is mode.

Individual Series:

Mean Deviation (M.D) = $\frac{\Sigma|D|}{n}$, Where $|D| = |x - \bar{X}|$, $\bar{X}$ = mean

$= |x - M|$, M = Median

$= |x - Z|$, Z = Mode

n = No. of terms in the given series.

Coefficient of Mean Deviation = $\frac{M.D}{\bar{X} \text{ or } M \text{ or } Z}$

Discrete Series:

Mean Deviation (M.D) = $\frac{\Sigma f|D|}{N}$, Where N = $\sum$ f = Total frequency

$|D| = |x - \overline{X}| = |x - M| = |x - Z|$

Continuous Series: Mean Deviation (M.D.) = $\frac{\Sigma f|D|}{N}$,

Where $|D| = |m - \overline{X}| = |m - M| = |m - Z|$, m is mid value of the class interval.

7. Compute M.D. and Coefficient of M.D. form mean and median for following series.

3	7	12	14	15	18	22

Solution:

(1) $\overline{X} = \frac{\Sigma x}{N}$

$\sum X = 3 + 7 + 12 + 14 + 15 + 18 + 22 = 91$

n = 7

$\therefore \overline{X} = \frac{91}{7} = 13$

X	dy = $\lvert X - \overline{X} \rvert$
3	10
7	6
12	1
14	1
15	2
18	5
22	9
	$\sum$ dy = 34

Now $\sum$ dy = 34

N = 7

And M.D. = $\frac{\Sigma dy}{n} = \frac{34}{7} = 4.86$

$\therefore$ M.D. FROM Mean = 4.86

And Coefficient of Mean Deviation from Mean = $\frac{\text{M.D. from Mean}}{\text{Mean}}$

$= \frac{4.86}{13} = 0.374$

(2) n = 7; $\therefore$ Median is 4th term = 14.

X	dy = $\lvert X - M \rvert$
3	11
7	7
12	2
14	0
15	1
18	4
22	8
	$\sum$ dy = 33

Now $\sum$ dy = 33, n = 7

And M.D. = $\frac{\Sigma dy}{n} = \frac{33}{7}$ = 4.71

about median

And Coefficient of Mean Deviation from Median = $\frac{\text{M.D. from Median}}{\text{Median}}$

= $\frac{4.71}{14}$ = 0.336

8. Compute M.D. from $\overline{X}$ and M for given series

X	5	10	15	20	25	30
f	3	4	8	12	7	2

Solution:

X	f	fx	C_f	dy = $\lvert X - \overline{X} \rvert$	fdy	dy = $\lvert X - M \rvert$	fdy
5	3	15	3	13.06	39.18	15	45
10	4	40	7	8.06	32.24	10	40
15	8	120	15	3.06	24.48	5	40
20	12	240	27	1.94	23.28	0	0
25	7	175	34	6.94	48.58	5	35
30	2	60	36	11.94	23.88	10	20
	N = 36	$\sum$ f x = 650			191.64		180

$\overline{X} = \frac{\Sigma fx}{N} = \frac{650}{36}$ = 18.06 (Approx.)

Median will in $\frac{N+1}{2}^{th}$ term = 18.5th term

M.D about median = $\frac{180}{36}$.   Hence Median = 20

M.D. from $\overline{X} = \frac{\Sigma fdy}{N} = \frac{191.64}{36}$ = 5.323

9. Compute Coefficient of M.D. from $\overline{X}$, M and Z, for following series where N = 100

Class Intervals	0-10	10-20	20-30	30-40	40-50
Frequency	6	28	51	11	4

Solution:

A = 25; i = 10; N = 100

C.I	X	f		dx	fdx	dy = $\lvert X - \overline{X} \rvert$	fdy	fX	
0-10	5	6	34	-2	-12	17.9	107.4	30	450
10-20	15	28		-1	-28	7.9	221.2	420	
20-30	25	51		0	0	2.1	107.1	1275	
30-40	35	11	66	1	11	12.1	133.1	385	1840
40-50	45	4		2	8	22.1	88.4	180	
		N = 100			-21		657.2		

As $\overline{X} = A + \frac{\Sigma fdx}{N} \times i$

$\overline{X} = 25 + \frac{-21}{100} \times 10$

$= 25 - \frac{21}{10} = 25 - 2.1 = 22.9$

M.D. = $\frac{\Sigma fdx}{N} = \frac{657.2}{100} = 6.572$

about $\bar{x}$

Standard Deviation: Standard Deviation is also called the Root – Mean Square Deviation, as it is the square root of the mean of the squared deviations from the actual mean. It is denoted by S.D (or) 'σ', deviations can be written with dx.

Coefficient of standard Deviation = $\frac{S.D}{\bar{X}}$, Where $\bar{X}$ is mean

(a) **Individual Series**: Deviations can be taken from **Actual mean**. S.D. = $\sqrt{\frac{\Sigma dx^2}{n}}$

Where S.D is standard deviation, n is no. of terms. dx is $(x - \bar{X})$, $\bar{X}$ is mean.

Deviations can be taken from **Assumed Mean** S.D. = $\sqrt{\frac{\Sigma dx^2}{n} - \left(\frac{\Sigma dx}{n}\right)^2}$

Where S.D. is standard deviation, n is no. of terms. dx is (x-A)

A is assumed mean.

(b) **Discrete series**: Deviations can be taken from **Actual Mean**. S.D. = $\sqrt{\frac{\Sigma fdx^2}{N}}$

Where f is frequency, N is total frequency, dx is $(x - \bar{X})$, $\bar{X}$ mean

Deviations can be taken from **Assumed mean** S.D. = $\sqrt{\frac{\Sigma fdx^2}{N} - \left(\frac{\Sigma fdx}{N}\right)^2}$

Where N is total frequency, dx = x – A, A is assumed mean

Continuous Series: From **Actual mean** S.D. = $\sqrt{\frac{\Sigma fdx^2}{N}}$, Where dx = m - $\bar{X}$ and from

Assumed mean S.D. = $\sqrt{\frac{\Sigma fdx^2}{N} - \left(\frac{\Sigma fdx}{N}\right)^2}$, where dx = m-A.

m is mid value of the class interval.

Terms are included wrongly: We know S.D. = $\sqrt{\frac{\Sigma x^2}{n} - \left(\frac{\Sigma n}{n}\right)^2}$

To correct the value of S.D. when some wrong terms are included we firstly find correct Mean

i.e. $\frac{\Sigma x}{n}$, Total of terms $(\bar{X})$

Correct total of terms $\left(\Sigma X^1\right) = \Sigma X$ + Correct terms – Incorrect terms

Correct mean = $\frac{\Sigma x^1}{n}$

We also find value of incorrect Σx^2, by substituting the given data.

Correct S.D. = $\sqrt{\frac{\Sigma x^2}{n} - \left(\frac{\Sigma x}{n}\right)^2}$

10. Calculate S.D. by Direct and Short-cut Methods

25	27	31	32	35

Solution:

Direct Method; n = 5

X	$dx = (X - \overline{X})$	dx^2
25	-5	25
27	-3	9
31	1	1
32	2	4
35	5	25
$\sum X = 150$	$\sum dx = 0$	$\sum dx^2 = 64$

$$\overline{x} = \frac{\Sigma x}{n} = \frac{150}{5} = 30$$

$$\text{S.D.} = \sqrt{\frac{\Sigma dx^2}{n}} = \sqrt{\frac{64}{5}} = \sqrt{12.8} = 3.578$$

Short-cut Method:

X	$dx = \lvert X - A \rvert$	dx^2
25	-6	36
27	-4	16
31	0	0
32	1	1
35	4	16
$\sum X = 150$	$\sum dx = -5$	$\sum dx^2 = 69$

Let Assumed Mean (A) = 31

$$\text{S.D.} = \sqrt{\frac{\Sigma dx^2}{n} - \left(\frac{\Sigma dx}{n}\right)^2}$$

$$= \sqrt{\frac{69}{5} - \left(\frac{-5}{5}\right)^2}$$

$$= \sqrt{\frac{69}{5} - 1} = \sqrt{\frac{64}{5}}$$

$$= \sqrt{12.8} = 3.578$$

$$\text{Coefficient of S.D.} = \frac{\text{S.D.}}{\overline{X}} = \frac{3.578}{30} = 0.119$$

11. Calculate Standard Deviation and Coefficient of S.D. for following data

X	2	4	6	8	10	12	15
f	5	15	20	25	25	20	8

Solution:

Direct Method

X	f	fx	dx = (X – X)	dx^2	fdx^2
2	5	10	-6	36	180
4	15	60	-4	16	240
6	20	120	-2	4	80
8	25	200	0	0	0
10	25	250	2	4	100
12	20	240	4	16	320
	N =110	∑ fx = 880			∑ fdx^2 = 920

$$= \overline{X} = \frac{\Sigma fx}{N} = \frac{880}{110} = 8$$

$$\Sigma fdx^2 = 920$$

$$\text{S.D.} = \sqrt{\frac{\Sigma fdx^2}{N}} = \sqrt{\frac{920}{110}} = \sqrt{8.364} = 2.892$$

Short-cut Method:

Let Assumed Mean (A) = 7

X	f	dx = (X–A)	fdx	fdx^2
2	5	-5	-25	125
4	15	-3	-45	135
6	20	-1	-20	20
8	25	1	25	25
10	25	3	75	225
12	20	5	100	500
	N = 110		∑ fdx = 110	∑ fdx^2 = 1030

$\sum fdx = 110$, $\sum fdx^2 = 1030$

$$\text{S.D.} = \sqrt{\frac{\Sigma fdx^2}{N} - \left(\frac{\Sigma fdx}{N}\right)^2}$$

$$= \sqrt{\frac{1030}{110} - \left(\frac{110}{110}\right)^2} = \sqrt{\frac{103}{11} - 1}$$

$$= \sqrt{\frac{103-11}{11}} = \sqrt{\frac{92}{11}}$$

$$= \sqrt{8.364} = 2.892 \text{ (Approx.)}$$

12. From the following data find out the standard deviation

X	10-20	20-30	30-40	40-50	50-60	60-70
f	10	12	15	20	14	24

Solution:

X	f	M.V.(X)	A= 45 dx	C= 10 dx´	dx^2	fdx´	fdx'^2
10-20	10	15	-30	-3	9	-30	90
20-30	12	25	-20	-2	4	-24	48
30-40	15	35	-10	-1	1	-15	15
40-50	20	45	0	0	0	0	0
50-60	14	55	10	1	1	14	14
60-70	24	65	20	2	4	48	96
	95					-7	263

By using step deviation method

$$\sigma = \sqrt{\frac{\Sigma fdx'^2}{N} - \left(\frac{\Sigma fdx'}{N}\right)^2} \times C$$

$$\sigma = \sqrt{\frac{263}{95} - \left(\frac{-7}{95}\right)^2} \times 10$$

$$= \sqrt{2.768 - 0.005} \times 10$$

$$= \sqrt{2.763} \times 10$$

= 1.662 × 10 = 16.62

13. Calculate combined Mean and S.D.

	Series-A	Series-B	Series-C
N	200	250	300
$\overline{X}$	25	10	15
S.D.	3	4	5

Solution:

Here $N_1 = 200,\quad N_2 = 250,\quad N_3 = 300$

$\overline{X}_1 = 25,\quad \overline{X}_2 = 10,\quad \overline{X}_3 = 15$

$\sigma_1 = 3,\quad \sigma_2 = 4,\quad \sigma_3 = 5$

As $$\overline{X} = \frac{N_1\overline{X}_1 + N_2\overline{X}_2 + N_3\overline{X}_3}{N_1 + N_2 + N_3}$$

$\therefore$ $$\overline{X} = \frac{200 \times 25 + 250 \times 10 + 300 \times 15}{200 + 250 + 300}$$

$$= \frac{5000 \times 2500 + 4500}{750} = \frac{12000}{750} = 16$$

$\therefore \quad d_1 = \overline{X} - \overline{X}_1 = 16 - 25 = -9$

$d_2 = \overline{X} - \overline{X}_2 = 16 - 10 = 6$

$d_3 = \overline{X} - \overline{X}_3 = 16 - 15 = 1$ And

As Combined S.D.

$$= \sqrt{\frac{N_1\sigma_1^2 + N_2\sigma_2^2 + N_3\sigma_3^2 + N_1 d_1^2 + N_2 d_2^2 + N_3 d_3^2}{N_1 + N_2 + N_3}}$$

$$\therefore \text{ Combined S.D} = \sqrt{\frac{200(3)^2 + 250(4)^2 + 300(5)^2 + 200(-9)^2 + 250(6)^2 + 300(1)^2}{200 + 250 + 300}}$$

$$= \sqrt{\frac{50(36 + 80 + 150 + 325 + 180 + 6)}{750}}$$

$$= \sqrt{\frac{776}{15}} = \sqrt{51.73} = 7.2 \text{ (Approx.)}$$

14. (a) Calculate M.D. and Q.D. if S.D = 60
    (b) Calculate S.D. and Q.D. if M.D. = 60
    (c) Calculate S.D. and M.D. if Q.D. = 60; if the series are moderately asymmetrical.

Solution:

(a) Given; S.D. = 60

M.D. = $\frac{4}{5}$ (S.D.) = $\frac{4}{5} \times 60 = 48$

Q.D. = $\frac{2}{3}$ (S.D.) = $\frac{2}{3} \times 60 = 40$

(b) Given M.D. = 60

As M.D. = $\frac{4}{5}$ (S.D.)

$\therefore$ S.D. = $\frac{5}{4}$ (M.D.) = $\frac{5}{4} \times 60 = 75$

Q.D. = $\frac{5}{6}$ (M.D.) = $\frac{5}{6} \times 60 = 50$ (OR) Q.D. = $\frac{2}{3}$ (S.D.) = $\frac{2}{3} \times 75 = 50$.

(c) Given S.D. = 60

S.D. = $\frac{3}{2}$ (Q.D.) = $\frac{3}{2} \times 60 = 90$

M.D. = $\frac{4}{5}$ (S.D.) = $\frac{4}{5} \times 90 = 72$ (or) M.D. = $\frac{6}{5}$ (Q.D.) = $\frac{6}{5} \times 60 = 72$

Properties of Standard Deviations:

1. S.D. is independent of change of origin.

   For example S.D. of 3, 4, 10, 19 will be same as that of 5, 6, 12, 21. 2 has been added to each term, but S.D. is same.

2. S.D. is not independent of scale.

   For example S.D. of 2, 4, 7, 9 is half of the 4, 8, 14, 18 as terms are half of the latter series.

3. Sum of squares taken from the Arithmetic Mean is always minimum i.e. $\Sigma(X-\bar{X})^2$ is always least.

4. S.D. of first n natural number is $\sqrt{\frac{1}{12}(n^2-1)}$

(iii) **Variance**: Variance is the Square of standard deviation. It is denoted 'σ^2'

S.D. (or) $\sigma = \sqrt{\text{variance}}$

(iv) **Combined Standard Deviation**:

$$\sigma_{123...n} = \sqrt{\frac{(N_1\sigma_1^2 + N_2\sigma_2^2 + ---) + (N_1 d_1^2 + N_2 d_2^2 + ---)}{(N_1 + N_2 + N_3 + ---)}}$$

Where $\sigma_{123...n}$ = the combined S.D. of the all series.

N_1, N_2 --- are the No. of terms in the different series.

σ_1, σ_2 --- is the S.D of different series.

d_1, d_2 ---- is the difference between combined mean and mean of 1st, 2nd ---- series respectively.

$d_1 = \bar{x} - \bar{x}_1 \quad d_2 = \bar{x} - \bar{x}_2 \quad d_3 = \bar{x} - \bar{x}_3$ -------

Note: If $N_1 = N_2 = N_3$ ---- i.e. No. of terms in each series is equal then

$$\sigma = \sqrt{\frac{(\sigma_1^2 + \sigma_2^2 + ----) + (d_1^2 + d_2^2 + ----)}{n}},$$ n is number of different series.

(v) Coefficient of variation (or) Coefficient of variability:

Coefficient of variation (or) C.V = $\frac{\sigma}{\bar{x}} \times 100$, Where σ is S.D $\bar{X}$ is A.M.

(vi) Relation between various measures of Dispersion: -

Q.D = $\frac{2}{3}$ S.D and M.D = $\frac{4}{5}$ S.D

Incase of QD, MD and SD, the relation is

(1) $\bar{X} \pm$ QD covers 50% of the total terms.

(2) $\bar{X} \pm$ MD covers 57.51% of the total terms.

(3) $\bar{X} \pm$ SD covers 68.27% of the total terms.

(4) $\bar{X} \pm$ 2SD covers 95.45% of the total terms.

(5) $\bar{X} \pm$ 3SD covers 99.73% of the total terms.

Conclusion: If we discuss about the various measures of dispersion we come to Range first. But Range is not a stable measure and has many drawbacks such as fluctuations of sampling. Thus can't be a good measure of dispersion.

Quartile Deviation: It is certainly better than Range. But here all the terms are not taken into account. It also suffers from sampling instability. It can be effectively used when class intervals are open; and to calculate M.D and S.D. these limits have to be assumed.

Mean Deviation: It is not capable of further Algebraic Treatment, although it takes into account all the terms but still, if the extreme values are big it will distort the result. Moreover, it ignores ± signs; therefore can't be called a good measure of dispersion.

Standard deviation: It is known to be the best measure of dispersion despite its some drawbacks such as extreme terms influence and assumption in case of open end intervals. Among its merits it is further capable of Algebraic Treatment, not affected by fluctuations of sampling, rigidly defined and stable.

Although extreme terms make this measure less effective, still ignoring this negative aspect, we can easily detect that it is still the best measure of Dispersion.

4.3 MEASURES OF SKEWNESS

I. Absolute measures of skewness:

Absolute skewness SK = $\overline{X}$ – Mode

Absolute skewness SK = $\overline{X}$ – Median

Absolute skewness SK = Median – Mode

II. Relative measures of skewness:

(a) The Karl Pearson's Coefficient of skewness

(b) The Bowley's Coefficient of skewness

(c) The Kelly's Coefficient of skewness

(d) Measure of skewness Based on Moments.

(a) **Karl Pearson's Coefficient of skewness**: It is based on the difference between the mean and Mode

$$S_{kp} = \frac{\overline{X} - Z}{\sigma}$$

Where S_{kp} = Karl Pearson's Coefficient of skewness,

$\overline{X}$ is mean, Z is mode, σ is S.D,

When Mode is ill-defined $S_{kp} = \frac{3(\overline{X} - M)}{\sigma}$ Where M is Median.

The result obtained with the help of this formula can vary between ± 3 only theoretically, but in practice, it rarely exceeds ± 1.

(b) **Bowley's Coefficient of skewness**: It is based on Quartiles

$$S_B = \frac{(Q_3 - M) - (M - Q_1)}{(Q_3 - M) + (M - Q_1)} = \frac{Q_3 + Q_1 - 2M}{Q_3 - Q_1}$$

Where Q_3 is lower Quartile, S_B is Bowley's Coefficient

Q_1 is upper Quartile, M is Median.

This measure is called the Quartile measure of skewness and values of the Coefficient, They obtained vary between ± 1.

(c) **Kelly's Coefficient of skewness**: Kelly used Deciles and percentiles to cover the entire data

$$S_K = \frac{D_1 + D_9 - 2M}{D_9 - D_1} = \frac{P_{10} + P_{90} - 2M}{P_{90} - P_{10}}$$

Where S_K is Kelly's Coefficient of skewness

D_1 is 1^{st} decile, D_9 is 9^{th} decile, M is median,

P_{10} is 10^{th} percentile, P_{90} is 90^{th} percentile

This method is not popular in practice and generally Karl Pearson's method is applied.

The results obtained by these formulae will generally lie between + 1 and – 1.

MULTIPLE CHOICE QUESTIONS

1. Which of the following is a relative measure?
   (a) Dispersion
   (b) Average
   (c) Coefficient
   (d) None of the above
2. Range of a the given distribution is
   (a) Equal to standard deviation
   (b) Greater than standard deviation
   (c) Less than standard deviation
   (d) None of these
3. To measure the deviations of an open end series the suitable measure of dispersion is
   (a) Quartile deviation
   (b) Mean Deviation
   (c) Standard Deviation
   (d) None of these
4. The measure of dispersion most influenced by the extreme values of distribution is
   (a) Standard Deviation
   (b) Mean Deviation
   (c) Range
   (d) Quartile deviation
5. The quartile deviation includes
   (a) Last 50%
   (b) Middle 50%
   (c) First 50%
   (d) None of these
6. Sum of deviations from median ignoring a ± signs is
   (a) The highest
   (b) Zero
   (c) The least
   (d) None of the above
7. The measure of variation that is least influenced by the extreme items is
   (a) Quartile deviation
   (b) Mean deviation
   (c) Range
   (d) Standard deviation

8. For which measure of dispersion ± signs are ignored:
   (a) Standard deviation
   (b) Quartile Deviation
   (c) Range
   (d) Mean deviation
9. For calculating mean deviation generally deviations are taken from:
   (a) Mean
   (b) Mode
   (c) G.M.
   (d) Median
10. For asymmetrical distribution
   (a) M.D. = $\frac{4}{5}\sigma$
   (b) $\frac{2}{3}\sigma$
   (c) –
   (d) None of these
11. In a normal distribution $\overline{X} \pm 3\sigma$ covers.
   (a) 99.73% items
   (b) 99.25% items
   (c) 99.9% items
   (d) None of these
12. Graphic method of calculating dispersion is
   (a) Lorenz curve
   (b) Mean Deviation
   (c) Quartile
   (d) Range
13. Coefficient of variation is calculated by the formula
   (a) $\frac{\sigma}{\overline{X}} \times 100$
   (b) $\frac{\overline{X}}{\sigma} \times 100$
   (c) $\frac{\sigma}{\overline{X}}$
   (d) $\frac{\sigma}{\overline{X}} \times 1000$
14. The formula $\sqrt{\frac{\Sigma fd^2}{N} - \left(\frac{\Sigma fd}{N}\right)^2} \times C$ is used to calculate the value of
   (a) Standard deviation
   (b) Variance

(c) Coefficient of variation

(d) None of these

15. For a normal distribution

(a) Q.D. = 2 over 3 σ

(b) Q.D. = 1 over 3 σ

(c) Q.D. = 4 over 5 σ

(d) None of these

16. Given mean and mode, which of the following is needed to find out Karl Pearson's coefficient of skewness.

(a) Coefficient of variation

(b) Standard deviation

(c) Median

(d) Nothing else is required

17. If the coefficient of skewness is zero the distribution is

(a) Normal

(b) A Symmetrical

(c) J-shaped

(d) None of these

18. Karl Pearson's coefficient of skewness is

(a) Less than Bowley's coeff. of skewness

(b) More than Bowley's coeff.of skewness

(c) Not related to Bowley's coefficient of skewness

(d) None of these

19. If coefficient of skewness is negative then

(a) $Q_3 + Q_1 > 2Q_2$

(b) $Q_3 + Q_1 = 2Q$

(c) $Q_3 + Q_1 < 2Q_2$

(d) None of these

20. When a distribution is negatively skewed then

(a) Mode < median > mean

(b) Mode < Median < Mean

(c) Mode > median < mean

(d) None of these

21. Measures of dispersion are called averages of the ____ order

(a) 1^{st}

(b) 2^{nd}

(c) 3rd

(d) None

22. Difference between the maximum & minimum value of a given data is called ____
   (a) Width
   (b) Size
   (c) Range
   (d) Class
23. Which measures of dispersion is the quickest to compute?
   (a) Standard deviation
   (b) Quartile deviation
   (c) Mean deviation
   (d) Range
24. For any two numbers range is always
   (a) Twice the SD
   (b) Half the SD
   (c) Square the SD
   (d) None of these
25. As the sample size increases, range tends to
   (a) Decrease
   (b) Increase
   (c) Same
   (d) None
26. If each item is reduced by 10, the range is
   (a) Increased by 10
   (b) Decreased by 10
   (c) Unchanged
   (d) None
27. For the observations 4, 2, 6, 3, 9, 5, 11, 70, 10 range is
   (a) 70
   (b) 2
   (c) 68
   (d) 11
28. Following are the marks of 10 students: 82, 79, 56, 70, 85, 95, 55, 72, 70, 66. Find coefficient of range
   (a) 25.66
   (b) 26.67
   (c) 27.66
   (d) 28.67

29. Following are the wages of 8 workers expressed in rupees: 82, 96, 52, 75, 70, 65, 50, 70. Find the range and its coefficient.

(a) 52, 85.25

(b) 70, 35.27

(c) 39, 33.52

(d) 46, 31.51

30. If R_x and R_y denote ranges of x and y respectively where x and y are related by 3x + 2y+10=0, what would be the relation between x and y?

(a) $R_x = R_y$

(b) $2R_x = 3R_y$

(c) $3R_x = 2R_y$

(d) $R_x = 2R_y$

31. If the range of x is 2, what would be the range of 3x + 50?

(a) 2

(b) 6

(c) -6

(d) 44

32. Quartile deviation is called ____. Hence it is_____ the Inter Quartile Range.

(a) Semi Inter quartile range, Half

(b) Quartile range, Equal

(c) Both a and b

(d) None

33. Quartile deviation is based on the

(a) Highest 50%

(b) Lowest 25%

(c) Highest 25%

(d) Middle 50% of the observations.

34. The lower & upper quartiles are used to define

(a) Standard deviation

(b) Quartile deviation

(c) Both

(d) None

35. Q.D. of the data 1, 3, 5, 7, 9, 11, 13, 15 is

(a) 4.50

(b) 8

(c) 5

(d) None of these

36. The quartiles of a variable are 45, 52, and 65 respectively. Its quartile deviation is
    (a) 10
    (b) 20
    (c) 25
    (d) 8.30
37. The quartile deviation of the daily wages (in ₹) of 7 persons given below: 12, 7, 15, 10, 17, 19, 25 is
    (a) 14.5
    (b) 5
    (c) 9
    (d) 4.5
38. $(Q_3 - Q_1) / (Q_3 + Q_1)$ is
    (a) Coefficient of Quartile deviation
    (b) Coefficient of Mean deviation
    (c) Coefficient of Standard deviation
    (d) None
39. Coefficient of quartile deviation is equal to
    (a) Quartile deviation x 100/median
    (b) Quartile deviation x 100/mean
    (c) Quartile deviation x 100/mode
    (d) None
40. If Median = 5, quartile Deviation = 2.5 then the coefficient of Quartile deviation is –
    (a) 20
    (b) 50
    (c) 125
    (d) 5
41. If median = 12, $Q_1 = 6$, $Q_3 = 22$ then the coefficient of quartile deviation is
    (a) 33.33
    (b) 60
    (c) 66.67
    (d) 70
42. 25% of the items of a data are less than 35 and 25% of the items are more than 75. Q.D. of the data is
    (a) 55
    (b) 20
    (c) 35
    (d) 75

43. 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

(a) 16

(b) 14

(c) 10

(d) 9

44. If the quartile deviation of x is 8 and 3x + 6y = 20, then the quartile deviation of y is

(a) -4

(b) 3

(c) 5

(d) 4

45. If the mean deviation of a normal variable is 16, what is its quartile deviation?

(a) 10.00

(b) 13.50

(c) 15.00

(d) 12.05

46. What is the value of Mean deviation about mean for the following numbers? 5, 8, 6, 3, 4

(a) 5.20

(b) 7.20

(c) 1.44

(d) 2.23

47. the coefficient of mean deviation about mean for the first 9 natural numbers is

(a) 200/9

(b) 80

(c) 400/9

(d) 50

48. If the relation between x and y is 5y-3x = 10 and the mean deviation about mean for x is 12, then the Mean deviation of y about mean is

(a) 7.20

(b) 6.80

(c) 20

(d) 18.80

49. The most commonly used measure of dispersion is

(a) Range

(b) Standard deviation

(c) Coefficient of variation

(d) Quartile deviation

50. "Root –mean square deviation from Mean" is
    (a) Standard deviation
    (b) Quartile deviation
    (c) Both
    (d) None
51. Most useful among all measures of dispersion is
    (a) Standard deviation
    (b) Quartile deviation
    (c) Mean deviation
    (d) None
52. When all the values are equal then variance and standard deviation would be
    (a) 2
    (b) -1
    (c) 1
    (d) 0
53. The standard deviation is always taken from
    (a) Median
    (b) Mode
    (c) Mean
    (d) None
54. The square of Standard deviation is known as
    (a) Variance
    (b) Standard deviation
    (c) Mean deviation
    (d) None
55. For a moderately skewed distribution, quartile deviation and the standard deviation are related by
    (a) S.D. = (2/3) Q.D.
    (b) S.D. = (3/2) Q.D.
    (c) S.D. = (3/4) Q.D.
    (d) S.D. = (4/3) Q.D.
56. Coefficient of Standard deviation is equal to
    (a) Standard deviation / AM
    (b) AM/ Standard deviation
    (c) Standard deviation / GM
    (d) None

57. The standard deviation of 10, 16, 10, 16, 10, 10, 16, 16 is

(a) 4

(b) 6

(c) 3

(d) 0

58. The standard deviation of first n natural numbers is

(a) [n(n+1) (2n+1)] /6

(b) $(n^2 - 1) / 12$

(c) $\sqrt{\frac{n^2-1}{12}}$

(d) n/2

59. If the Standard deviation of x is 3, what is the variance of (5-2x)?

(a) 36

(b) 6

(c) 1

(d) 9

60. $\overline{x} = 50$, $\sigma = 25$; the C.V. is

(a) 200%

(b) 25%

(c) 50%

(d) 100%

61. If mean = 10, Standard Deviation = 1.3 then coefficient of variation is –

(a) 10

(b) 31

(c) 13

(d) 20

62. What is the coefficient of variation of the following numbers: 53, 52, 61, 60, 64

(a) 8.09

(b) 18.08

(c) 20.23

(d) 20.45

63. Coefficient of variation of two series are 60% and 80% respectively. Their standard deviation are 20 and 16 respectively what are their A.M.

(a) 15 and 20

(b) 33.3. and 20

(c) 33.3. and 15

(d) 12 and 16

64. ____ of a set of observations is defined to be their sum, divided by the no. of observations.
   (a) HM
   (b) GM
   (c) AM
   (d) None

65. The mean weight of a group of 10 items is 29 and that of anther group of n items is 3. The mean of combined group of 10+n items is found to be 30. The value of n is
   (a) 2
   (b) 4
   (c) 10
   (d) 12

66. There are 5 bags of wheat weighting on an average 102 kgs and another 8 bags weighing 98 kgs on an average. What is combined mean of 13 bags?
   (a) 109.54
   (b) 99.54
   (c) 95.54
   (d) None of these

67. 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 ₹6,800 per month. What is the combined salary?
   (a) ₹6,160
   (b) ₹5,283
   (c) ₹6,000
   (d) 4,528

68. The A.M. between two numbers is 34 and their G.M. is 16 the numbers are
   (a) 4, 64
   (b) 4, 32
   (c) 32, 64
   (d) None of the these

69. If the AM and HM for two numbers are 5 and 3.2 respectively then ghe GM will be
   (a) 16.00
   (b) 4.10
   (c) 4.05
   (d) 4.00

70. The harmonic mean for the numbers 2, 3, 5 is
   (a) 2.00
   (b) 3.33
   (c) 2.90
   (d) $\sqrt[-3]{30}$

71. The median of the no.s 11, 10, 12, 13, 9 is

(a) 12.5

(b) 12

(c) 10.5

(d) 11

72. The variables x and y are related by 2x + 3y = 6 and median of x is 2. What is the median of y?

(a) 1/3

(b) 2/3

(c) 1

(d) None of these

73. In Ogive, abscissa corresponding to ordinate ___ is k^{th} percentile

(a) kN/10

(b) kN/100

(c) kN/50

(d) None

74. 10^{th} percentile is equal to

(a) 1^{st} decile

(b) 10^{th} decile

(c) 9^{th} decile

(d) None

75. 90^{th} percentile is equal to

(a) 9^{th} percentile

(b) 90^{th} decile

(c) 9^{th} decile

(d) None

76. If y = 4+3x and mode of x is 25, what is the mode of y?

(a) 75

(b) 25

(c) 79

(d) 80

77. If y = 5+7x and mode of x is 4, what is the mode of y?

(a) 28

(b) 33

(c) 4

(d) 43

78. The range of 15, 12, 20, 9, 17, 20 is
   (a) 5
   (b) 12
   (c) 13
   (d) 11
79. For the observations 6, 4, 1, 6, 5, 10, 4, 8, range is
   (a) 10
   (b) 9
   (c) 8
   (d) None
80. What is the coefficient of range for the following wages of 8 workers? ₹80, ₹65, ₹90, ₹60, ₹75, ₹70, ₹72, ₹85
   (a) ₹30
   (b) ₹20
   (c) ₹30
   (d) ₹20
81. What is the coefficient of range for the following distribution?

CI	10-19	20-29	30-39	40-49	50-59
F	11	25	16	7	3

   (a) 22
   (b) 50
   (c) 72.46
   (d) 75.82
82. If the relationship between x and y is given by 4x + 5y = 10 and the range of x is 15, what would be the range of y?
   (a) 10
   (b) 11
   (c) 12
   (d) 13
83. 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?
   (a) ₹15
   (b) ₹10
   (c) ₹19
   (d) ₹23

84. Quartile deviation for th data 1, 3, 4, 5, 6, 6, 10 is

(a) 3

(b) 1

(c) 6

(d) 1.5

85. 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.

(a) 9.23, 16.35

(b) 34.2, 19.68

(c) 9.89, 14.65

(d) 10.50, 18.42

86. If the first and third quartiles are 22.16 and 56.36, then the quartile deviation is:

(a) 17.1

(b) 34.2

(c) 51.3

(d) 43.2

87. What is the value of Mean deviation about mean for the following observations: 50, 60, 50, 50, 60, 60, 60, 50, 50, 60, 60, 60, 50.

(a) 5

(b) 7

(c) 35

(d) 10

88. The marks obtained by 10 students in an examination were as follows:

60, 65, 68, 70, 75, 73, 80, 70, 83, 86. Find mean deviation about the mean

(a) 5.3

(b) 5.4

(c) 5.5

(d) 5.6

89. What is the Mean deviation about median for the following data?

X	3	5	7	9	11	13	15
F	2	8	9	16	14	7	4

(a) 2.50

(b) 2.46

(c) 2.43

(d) 2.37

90. Compute the mean deviation about the arithmetic mean for the following data:

X	1	3	5	7	9
F	5	8	9	2	1

Also find the coefficient of the mean deviation about the AM.

(a) 1.72, 44.33

(b) 2.25, 38.63

(c) 2.36, 45.82

(d) 1.29, 47.65

91. If x and y are related as $4x + 3y + 11 = 0$ and mean deviation of x is 2.70. What is mean deviation of y?

(a) 7.20

(b) 14.40

(c) 3.60

(d) None of these

92. If two variables x and y are related by $2x + 3y - 7 = 0$ and the mean and Mean deviation about mean of x are 1 and 0.3 respectively, then the coefficient of mean deviation of y about mean is

(a) -5

(b) 12

(c) 50

(d) 4

93. What is the Standard deviation of 5, 5, 9, 9, 9, 9, 10 5, 10, 10?

(a) $\sqrt{14}$

(b) $\sqrt{42}$

(c) 4.50

(d) 8

94. If x and y are related by $y = 2x + 5$ and the Standard deviation and AM of x are known to be 5 and 10 respectively, then the coefficient of variation is

(a) 25

(b) 30

(c) 40

(d) 20

95. Find the standard deviation and the coefficient of variation for the following numbers: 5, 8, 9, 2, 6

(a) 2.45, 40.83

(b) 1.96, 39.25

(c) 2.45, 38.29

(d) 2.05, 33.24

96. If Arithmetic Mean and coefficient of variation of x are 5 and 20 respectively. What is the variance of (15-2x)?

(a) 16

(b) 2

(c) 4

(d) 32

State the following statements are true or false

(1) Absolute measure of dispersion are most suitable for comparison

(2) Range is the best measure of dispersion

(3) Quartile deviation is most suitable in case of open-end series

(4) Relative measures of dispersion make deviations in similar units comparable

(5) Range is the value of difference between mode and median.

(6) Mean deviation is the least when deviations are taken from median.

(7) Average deviation is mostly used in statistical work

(8) Coefficient of mean deviation is the ratio of M.D. to the average from which the deviations are taken.

(9) Mean deviation can never be negative

(10) M.D. = $\frac{4}{5}\sigma$

(11) There is no difference between coefficient of variation and variance

(12) In a normal distribution S.D.> M.D.> Q.D.

(13) Standard deviation = $\frac{\text{Coefficient of variation}}{\text{Mean}} \times 100$

(14) Combined standard deviation of the group can be calculated by combining the standard deviations of various sub groups.

(15) Coefficient of variation = $\frac{\text{Standard variation}}{\text{Mean}}$

(16) Bowley's measure of skewness is based on quartiles

(17) Skewness cannot be calculated in a distribution where mode is ill-defined

(18) Quartiles are equi-distant in a skewness distribution

(19) In a normal distribution the extreme deciles are equidistant from median

(20) Median can never be equal to mean in a skewed distribution

Short Answer Type Questions

1. Define Variation or Dispersion or Scatteredness.
2. Explain Absolute and Relative measures of Dispersion.
3. Name various methods of measuring Dispersion
4. Define Range. Is it positional measure? How?
5. What is Coefficient of Range? Also narrate formula
6. Define Inter-Quartile range.
7. What is the relation between Q_1, M and Q_3?
8. What do you mean by 'Percentile range'?
9. Define –
   (a) Mean Deviation from $\overline{X}$
   (b) Mean Deviation from M.
   (c) Mean Deviation from Z.
10. Provide steps to calculate Co-efficient of Average Deviation for individual series.
11. Name three co-efficient for Mean Deviations.
12. Narrate the formula to determine Mean Deviation directly.
13. When should we use the formula for direct calculation of M.D.
14. Narrate the steps to calculate co-efficient of Mean Deviation for discrete and Continuous Series.
15. Define standard Deviation and its co-efficient
16. Define variance
17. Describe the formula to Compute Combined Standard Deviation of two or more series when their $\overline{X}$ and S.D.'s are given
18. What formulae can be used when
   (a) $N_1 = N_2 = N_3$ .....
   (b) $\overline{X}_1 = \overline{X}_2 = \overline{X}_3$ (c) when both $N_1 = N_2 = N_3$ ..... and $\overline{X}_1 = \overline{X}_2 = \overline{X}_3 \ldots \ldots$
19. Define Coefficient of Variation or Coefficient of Variability
20. Define Relation between M.D., S.D. and Q.D.
21. Define Lorenz Curve. When is it used?
22. (a) What is S.D. of first n natural numbers
   (b) Explain the method of calculating standard deviation
23. Definition of Skewness
24. Absolute measures of skewness
25. Relative Measures of Skewness
26. Difference between relative and absolute measures of skewness
27. Difference between Dispersion and skewness
28. Skewness Vs. coefficient of Skewness
29. Positive Skewness Vs. Negative Skewness
30. Tests of Skewness

EXERCISE

1. Compute Range and Coefficient of Range for following data:

(a)

40	9	17	21	36	14	3	4	11	24	5	27	39	7	31

(b)

X:	1	4	7	10	15	20	30
F:	3	4	7	15	21	16	7

(c)

X:	10-15	15-20	20-25	25-30	30-35	35-40	40-45
F:	6	17	34	41	54	29	13

2. Computer Inter-Quartile range, Quartile Deviation an Coefficient of Quartile Deviation for following data:

(a)

19	9	15	21	4	3	23	11	7	26	14

(b)

X:	3	6	9	12	15	18	21
F:	4	9	14	21	28	15	8

3. Computer Inter-Quartile Range, Semi-inter-Quartile range, Coefficient of Q.D and Percentile Range for following data.

X:	0-4	4-8	8-12	12-16	16-20	20-24	24-28	28-32
F:	4	9	23	55	62	30	12	5

4. Compute Coefficient of Q.D. for following data:

X:	Less than 500	450	400	350	300	250	200	150	100
F:	150	146	130	93	47	26	15	7	3

5. Find Range and Coefficient of Range for following data:

14	18	19	21	23	25	27	29	33	34	37

6. Find Range and Coefficient of Range for following data:

X:	5	10	15	20	25	30	35	40
F:	4	7	21	47	53	24	12	6

7. Calculate Range and Coefficient of Range for following data:

X:	20-30	30-40	40-50	50-60	60-70	70-80
F:	4	9	16	21	13	6

8. From the following data compute inter quartile range, quartile Deviation and Coefficient of Quartile Deviation.

24	7	11	9	17	3	20	14	4	22	27

9. Compute Inter Quartile Range, Coefficient of Quartile Deviation and Percentile Range for following data:

X:	4	8	12	16	20	24	28	32
F:	4	9	17	40	53	37	24	16

10. From the following table showing incomes of persons, calculate the semi-inter quartile range and the co-efficient of quartile Deviation.

Income	0-10	10-20	20-30	30-40	40-50	50-60	60-70
No. of Persons	22	240	350	410	220	110	35

11. From the following data find quartile deviation and its coefficient.

X:	10-15	15-20	20-25	25-30	30-35	35-40	40-45	45-50
F:	8	10	12	15	10	7	8	5

12. Calculate M.D. from $\overline{X}$, M and Z for following data:

3	2	4	7	3	8

Also calculate its co-efficient.

13. Compare the Variability of two series with median as base.

A:	3484	4572	4124	3682	5624	4388	3680	4308
B:	487	508	620	382	405	266	186	218

14. Compute M.D. and its coefficient from $\overline{X}$, M and Z for following data:

X:	5	10	15	20	25	30	35	40
F:	16	32	36	44	28	18	12	14

15. Compute M.D. and its co-efficients from $\overline{X}$, M and Z.

X:	100-120	120-140	140-160	160-180	180-200
F:	4	6	8	10	5

Also use Direct Method to verify M.D.s.

16. Compute M.D. and Coefficient of M.D. form mean and median for following series

3	7	12	14	15	18	22

17. Compute M.D. from and M for given series.

X:	5	10	15	20	25	30
F:	3	4	8	12	7	2

18. Compute coefficient of M.D. from $\overline{X}$, M and Z, for following series where N = 100

Class Intervals	0-10	10-20	20-30	30-40	40-50
Frequency:	6	28	51	11	4

19. Calculate S.D. by Direct and Short-cut Methods

25	27	31	32	35

20. Calculate Standard Deviation and Coefficient of S.D. for following data:

X:	2	4	6	8	10	12	15
F:	5	15	20	25	25	20	8

21. (a) Compute standard deviation for following data by (1) Taking Actual Mean (2) Taking Assumed Mean (3) Step Deviation Method.

X	0-10	10-20	20-30	30-40	40-50	50-60	60-70
F	1	4	17	45	26	5	2

(b) From the following data find out the standard deviation

X	10-20	20-30	30-40	40-50	50-60	60-70
F	10	12	15	20	14	24

22. Calculate Combined Mean and S.D.

	Section A	Section B	Section C
N	200	250	300
$\overline{X}$	25	10	15
S.D	3	4	5

23. Find missing figures from following data

	Group A	Group B	Group C	Combined
N	50	?	90	200
$\overline{X}$	113	?	115	116
S.D	6	7	?	7.746

24. Calculate S.D by Direct and Short-cut Method:

22	57	44	53	47	53

25. Calculate S.D. Coefficient of S.D and Coefficient of variation by short cut method

7	19	12	14	16	13	17	18

26. Calculate Combined S.D. of the following series

	Section A	Section B	Section C
S.D	2	3	4
$\overline{X}$	10	20	30
N	8	10	12

27. In a Moderately Asymmetrical Distribution Compute
    (a) M.D. and S.D. Given Q.D. = 50
    (b) S.D. and Q.D. Given M.D. = 50
    (c) M.D. and Q.D. Given S.D. = 50
28. Calculate Sum and Sum of Squares; if $\overline{X}$ of 20 terms is 30 and S.D is 3.
29. A.M. of two no.s is 10 and S.D. is 2; find nos.
30. Calculate S.D. for first 10 natural nos.

31. Calculate S.D. for 39, 40, 41, 42, 43, 44, 45, 46, 47, 48

32. Calculate S.D. for 5, 10, 15, 20, 25…. 100.

33. Compare the variability

Series	$\overline{X}$	S.D
A	800	120
B	900	135

34. Compute $\overline{X}$, S.D., Coefficient of S.D. and coefficient of Variation for following data using 1) Actual Mean 2) Assumed Mean and 3) Step deviation methods

X	15	25	35	45	55	65
F	2	4	8	20	12	4

35. Compute $\overline{X}$, S.D., Coefficient of variation for given data

X	0-10	10-20	20-30	30-40	40-50
F	5	15	30	65	80

36. Calculate S.D. and C.V using Step-Deviation method:

X	30	40	50	60	70	80	90	100
F	5	5	8	2	10	7	2	1

37. Marks of two students in various subjects are as follows. Who is better by (1) Average (2) by consistency?

	Sub-I	Sub-II	Sub-III	Sub-IV	Sub-V	Sub-VI
A	16	18	19	22	20	18
B	7	21	48	3	44	16

38. Wages of workers of two shifts in a factory is given below 1) Which shift worker earn better 2) Which shift workers are consistent in their earnings

Wages	50-100	100-150	150-200	200-250	250-300	300-350
Shift-A	3	8	24	63	102	50
Shift-B	32	44	62	41	37	34

39. $\overline{X}$ and S.D of a series of 80 terms is 50 and 5. But later on it was found that a term 17 was wrongly included as 71. Find correct $\overline{X}$ and S.D.

40. In an office C.V of wages of male and female workers is 55% and 70%. S.D. are 22 and 15.40. Compute combined $\overline{X}$ and σ of all the workers if there are 80% male workers.

41. Mean of 40 terms is 25 nd S.D. is 4. Find sum and sum of Squares of all the terms.

42. What will happen to S.D. and variance of a series if each term is a) multiplied by 3 b) divided by 3 c) increased by 3 d) decreased by 3. Given S.D = 2.

43. Three series with equal terms have means 6,7,8 and S.D.s 1,2,3; Find combined mean and Standard Deviation.

44. Three series with equal terms and equal Mean have S.D.'s 6, 7, 8; Find combined S.D.

45. Calculate the coefficient of skewness from the following data

X:	210	240	260	290	245	255	288	272	263	277
	1	2	5	10	3	4	9	7	6	8

46. Calculate Karl Pearson's Coefficient of skewness from the following data

Income (in ₹)	5	10	15	20	25	30	35	40
No. of workers	26	29	40	35	26	18	14	12

47. The following figure relate to wage payments in two firms A and B.

	Firm A	Firm B
Mean	75	80
Median	72	70
Mode	67	62
Quartiles	62 and 78	65 and 85
Standard Deviation	13	19

Compare the features of two distributions.

48. Find the first and the second coefficient of skewness of frequency distribution with:

Mean = ₹ 79.76, Median = ₹ 79.06

Mode = ₹ 77.50, Standard deviation = 15.60

49. Calculate coefficient of quartile deviation and Bowley's coefficient of skewness from the following data:

Profit (₹ Lakhs)	Below 10	10-20	20-30	30-40	40-50	Above 50
No. of Cos.	5	12	20	16	5	2

50. From the following data, compute quartiles and find the coefficient of skewness:

Income (₹)	No. of Persons	Income (₹)	No. of Persons
Below 200	25	600-800	75
200-400	40	800-1000	20
400-600	80	Above 1,000	16

51. Explain negative and positive skewness. Calculate quartile deviation and coefficient of skewness from the following:

Median: 18.8 inches, Q_1 = 14.6 inches, Q_3 = 25.2 inches.

52. From the following data, compute quartile deviation and coefficient of skewness:

Size	4-8	8-12	12-16	16-20	20-40	24-28	28-32	32-36	36-40
Frequency	6	10	18	30	15	12	10	6	2

53. From the following data Calculate Karl Perarson's coefficient of skewness.

Marks	1	4	4	5	6

54. Find out Karl Perarson's Coefficient by Skewness from the following data:

X:	4.5	5.5	6.5	7.5	8.5	9.5	10.5	11.5
F:	35	40	48	100	125	87	43	22

55. Calculate Karl Pearson's Coefficient of Skewness from the following data:

Age in years	10-20	20-30	30-40	40-50	50-60
No. of Persons	18	20	30	22	10

56. From the following marks obtained by 7 students in a tutorial group, calculate Bowley's Coefficient of Skewness

Roll No.	1	2	3	4	5	6	7
Marks	1	5	12	22	17	9	4

57. Consider the following distributions:

	Distribution A	Distribution B
Mean	100	90
Median	90	80
Standard Deviation	10	10

(i) Distribution A has the same degree of the variation as distribution B.

(ii) Both distributions have the same coefficient of skewness. True/False? Give reasons.

58. For a moderately skewed distribution, arithmetic mean = 160, mode = 157 and standard deviation = 50, Find (i) coefficient of variation, (ii) Pearsonian coefficient of skewness and (iii) Median.

59. Given Mean = 50, C.V = 40%, Karl Pearson's Coefficient of Skewness = – 0.4. Find standard deviation and Mode.

60. In a frequency distribution the coefficient of skewness based upon the quartile is 0.6. If the sum of the upper and lower quartiles is 100 and the median is 38, find the value of upper quartile.

61. The sum of 20 observations is 300 and its sum of square is 5,000 and median is 15. Find its coefficient of skewness and coefficient of variation.

62. Find the coefficient of skewness if Difference between two quartiles = 8.

Sum of two quartiles = 22, Median = 10.5

MCQ Answers:

1	c
2	b
3	a
4	c
5	b
6	c
7	b
8	d
9	d
10	a
11	a
12	a
13	a
14	a
15	a
16	b

17	a
18	c
19	a
20	a
21	b
22	c
23	d
24	a
25	b
26	c
27	c
28	b
29	d
30	c
31	b
32	a
33	d
34	b
35	a
36	a
37	d
38	a
39	a
40	b
41	c
42	b
43	d
44	d
45	b
46	c
47	c
48	a
49	b
50	a
51	a
52	d
53	c
54	a
55	b
56	a
57	c
58	c
59	a
60	c

61	c
62	a
63	b
64	c
65	b
66	b
67	a
68	a
69	d
70	c
71	d
72	b
73	a
74	a
75	c
76	c
77	b
78	d
79	b
80	d
81	c
82	c
83	b
84	d
85	d
86	a
87	a
88	d
89	d
90	a
91	b
92	b
93	b
94	c
95	a
96	c

Study Note - 5

CORRELATION AND REGRESSION

This Study Note includes

5.1 CORRELATION

Introduction:

Managers, generally, assess the nature and degree of relationship between variables. A marketing manager likes to know the degree of relationship between advertising expenditure of the volume of sales. He expects a positive relationship between the two variables.

The manager is interested in finding out whether there is any relationship between two or more variables and if it is true, he would like to assess the strength of the relationship or the degree of relationship. Correlation coefficient helps the managers to know the degree of relationship between variables. They can know the degree of relationship by working out correlation between customer satisfaction and profitability, motivation of workers and productivity, percentage defectives and cost, etc.

In case of central tendency, dispersion and skewness, we study problems relating to one variable. But in the real world we have problems pertaining to two or more than two variables. We know there is correlation between price and quantity demanded, price and quantity supplied, age of husband and wife, income and consumption etc. The extent of relationship between any two variables can be measured with the help of correlation. The measure of correlation is called correlation index or correlation coefficient. It gives one figure which shows the degree and direction of correlation. It means the coefficient of correlation helps us in determining the closeness of the relationship between two or more than two variables.

Definition: According to A.M. Tuttle, "Correlation is an analysis of the co-variation between two or more variables".

Uses or significance of correlation analysis:

The use or utility of the study of correlation is clear from the following points:

(i) The correlation coefficient helps us in measuring the extent of relationship between two or more than two variables. The degree and extent of the relationship between two variables is, of course, one of the most important problems in statistics.

(ii) It is through correlation that we can predict about the future. For instance. If there are good monsoons, we can expect better food supply and hence can expect fall in price of food grains and other products.

(iii) If the value of a variable is given, we can know the value of another variable. It is, of course, done with the help of regression analysis.

(iv) Correlation contributes to economic behavior. It helps us in knowing the important variables on which other depend.

(v) The technique of ratio of variation and regression analysis depends totally on the findings of coefficient of correlation.

(vi) In the field of commerce and industry, the technique of correlation coefficient helps to make estimates like sales, price or costs.

(vii) The predictions made on the basis of correlation analysis are considered to be nearer to reality and hence reliable. In the words. Of Tippett, "the effect of correlation is to reduce the range of uncertainty of our prediction".

But a noteworthy point is that if the concept of correlation is not employed with proper care, it can lead to misleading conclusions. That means we should use this statistical tool with great care.

Types of correlation: There are three important types of correlation which are discussed as under:

Positive and Negative Correlation: Correlation between two variables, positive or negative depends on the direction in which the variables move.

(a) **Positive or direct correlation**: If the two variables move in the same direction i.e. with an increase in one variable, the other variable also increases or with a fall in one variable, the other variable also falls, the correlation is said to be positive. For example, price and supply are positively related. It means if price goes up, the supply goes up and vice versa. It can be shown with the help of "arrows".

$$\uparrow P \quad \uparrow S \qquad \downarrow P \quad \downarrow S$$

Price	50	60	70	80	90
Supply	100	120	140	160	180

This numerical illustration shows that as price increases the supply also increases and vice versa.

(b) **Negative or Inverse Correlation**: If two variables move in opposite direction i.e. with the increase in one variable, the other variable falls or with the fall in one variable, the other variable rises, the correlation is said to be negative or inverse. For example, the law of demand shows inverse relation between price and demand. It can be shown with the help of arrows:

$$\uparrow P \quad \downarrow D \qquad \downarrow P \quad \uparrow D$$

Price	50	60	70	80	90
Demand	180	160	140	120	100

This table shows that with the increase in price, the demand falls and vice versa.

Simple and Multiple Correlations:

(a) **Simple Correlation**: When there are only two variables and the relationship is studied between those two variables, it is case of simple correlation. Relationships between height and weight, price and demand or income and consumption etc. are examples of simple correlation.

(b) **Multiple Correlation**: When there are more than two variables and we study the relationship between one variable and all the other variables taken together then it is a case of multiple correlation. Suppose there are three variables 1, 2, 3 we can study the multiple correlation between A and B & C taken together or between B and A & C together etc. It can be denoted as $R_{1.23}$ or $R_{2.13}$ or $R_{3.12}$

Partial and Total Correlation:

(a) **Partial Correlation**: When there are more than two variables and the relationship between any two of the variables is studied assuming other variables as constant it is a case of partial correlation.

This, in fact, is an extension of multiple correlation. Suppose we study the relationship between rainfall and crop, without taking into consideration the effects of other inputs like fertilizers, seeds and pesticides etc., this technique will be known as partial correlation. Symbolically if x, y, z are the three variables then partial correlation between x and y excluding z will be given by $r_{xy.z,}$ $r_{xz.y}$ or $r_{yz.x}$

(b) **Total Correlation**: When the correlation between the variables under study taken together at a time, is worked out, it is called total correlation.

The main point worth consideration is that this line will indicate positive relation if 'a' is positive and in case 'a' is negative the correlation will also be negative. In such type of correlation the value of the coefficient of correlation is always + 1 or -1 depending on the sign of 'a' in the equation of y = ax + b. Correlation will be +1 if 'a' is +ve and -1 if 'a' is –ve.

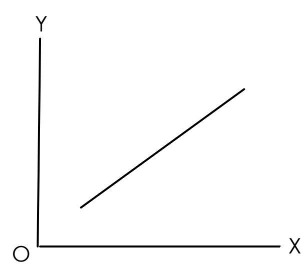
Linear and Non-Linear Correlation:

(a) **Linear Correlation**: The correlation between two variables will be linear if corresponding to a unit change in one variable, there is a constant change in the other variable over the entire range of the values. For instance, we consider the following data:

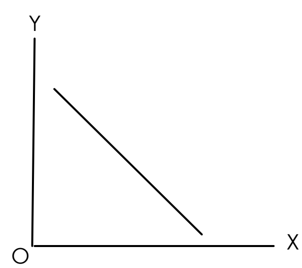
X	1	2	3	4	5
Y	2	4	6	8	10

It shows that for a unit change in value of x, there is a constant change i.e. 2 in the corresponding values of y. Mathematically, it can be shown as Y = 2x + 0 = 4. In general, two variables are linearly related if there exists a relationship of the form Y = a+bX between them.

In the above equation, 'a' is intercept, whereas b is the slope. If we plot the values of the two variables, we will get a straight line. The main point worth consideration is that this line will indicate positive relation if 'a' is positive and in case 'a' is negative then correlation will also be negative. In such types of correlation the value of the coefficient of correlation is always +1 or -1 depending on the sign of 'a' in the equation of y = ax + b. Correlation will be +1 if 'a' is +ve and -1 if 'a' is –ve. This type of relation does not exist in economics and other social sciences. This type of relation can exist only in physical sciences. However, it has great theoretical importance in economics and other social sciences.



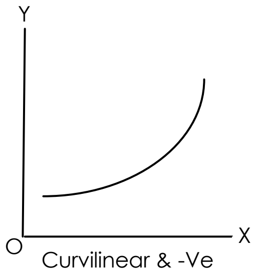
POSITIVE AND LINEAR CORRELATION



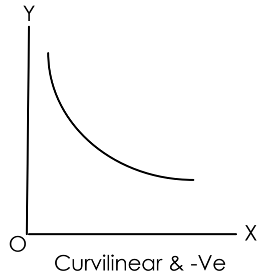
NEGATIVE AND LINEAR CORRELATION

(b) **Non-linear Correlation**: The relationship between two variables will be non-linear or curvi-linear, if corresponding to a unit change in one variable, the other variables change at a different rate. If such data is plotted, we do not get a straight line but a curve type figure. It means the slope of the plotted curve is not constant.

Mathematically the relationship between x and y will never be of the form y = ax +b, but in the form of $y = ax^2 + bx + c$ or $y = a.b^x$ etc. When the values are plotted on the graph paper, we will get the graphs as curve rather than the straight lines, as shown



Curvilinear & -Ve
CORRELATION



Curvilinear & -Ve
CORRELATION

Such types of correlations are found very commonly in the fields of economics and the other social sciences. As such these are very important in the study of social sciences.

Degree and interpretation of correlation coefficient:

(a) According to Karl Pearson, the coefficient of correlation lies between two limits i.e., ±1. It implies if there is perfect positive relationship between two variables, the value of correlation would be +1. On the contrary, if there is perfect negative relationship between two variables, the value of the correlation will be -1. It means r lies between +1 and -1. Within these limits the value of correlation is interpreted as:

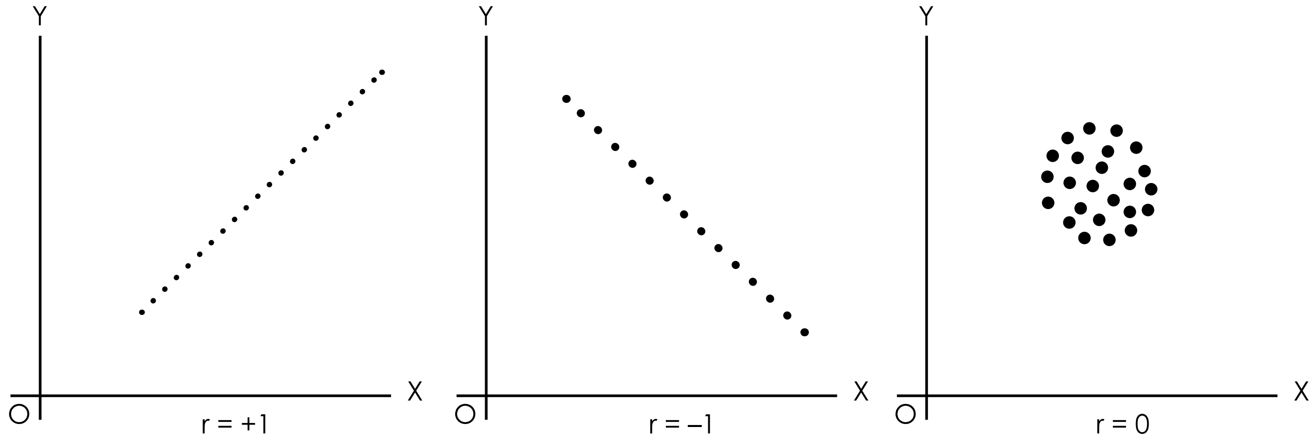
(i)	When r = +1	Perfect positive correlation
(ii)	When r > + 0.75 but < +1	High degree of positive correlation
(iii)	When r > + 0.5 but < + 0.75	Moderate degree of positive correlation
(iv)	When r > + 0 but < + 0.5	Low degree of positive correlation
(v)	When r = 0	No correlation at all
(vi)	When r < - 0.75 but > - 1	High degree of negative correlation
(vii)	When r < - 0.5 but > - 0.75	Moderate degree of negative correlation
(viii)	When r < 0 but > -0.5	Low degree of negative correlation
(ix)	When r = -1	Perfect negative correlation

Methods of Studying Correlation:

In correlation analysis, we are required to know the relationship between variables and the extent of that relationship. There are two methods which visualize the relationship between the two variables i.e. 1) the Scatter Diagram Method and 2) Graphic Method. These are based on graphs and diagrams. Then there are mathematical methods in which we include, 3) Karl Pearson's coefficient of correlation, 4) Rank Correlation Method, 5) Concurrent Deviation Method, and 6) Method of Least Squares.

Scatter Diagram Method: This method is also known as dot diagram, datagram or scattergram. Scattergram is one of the simplest method of diagrammatic representation of a bivariate (two variables) distribution. It provides the simplest tool of determining the correlation between two variables. The term scatter refers to the dispersion or spread of the dots on the graph. We should keep the following points in mind while interpreting the correlation between two variables through scatter diagram:

(i) If the points plotted are very close to each other, it shows high correlation, otherwise poor correlation is expected



(ii) If the points on the diagram show upward or downward trend, then we say there is correlation. But in case no trend is shown by the points, it shows that the variables are uncorrelated.

(iii) If there is upward trend from left to right, the correlation is positive, that means the values of two variables move in the same direction. On the other hand, if the points show a downward trend from left to right, the correlation in negative as the values of the two variables move in opposite direction.

(iv) The correlation would be perfect or equal to one if all the points lie on a straight line starting from left bottom and going up towards the right top. On the other hand, the correlation would be perfect and negative if all the points lie on a straight line starting from top left to fall to right bottom.

Graphic Method:

This method is also known as correlogram or simple graph method. It is the simplest method to determine the presence of correlation between the two variables. If the values of the two variables are plotted on the graph paper we will get two curves, one for X variable and another for Y variable individually. By looking at the direction and closeness of the two curves, we can know whether the variables are related or not. If these two curves, we can know whether the variables are related or not. If these two curves move in the same direction, correlation is said to be positive. On the other hand, if the curves are moving in the opposite directions, correlation is said to be negative. This cn be made clear with the help of an example.

Example: From the following data ascertain whether there is correlation between X and Y variables:

X	80	84	112	106	174	176	128
Y	88	96	108	112	136	156	112

The above graph shows that the variables X and Y are quite closely related. The thick line shows the values of X variable whereas dotted line shows the values of Y variable.

Algebraic or Mathematical Methods: Some of the methods of calculation of correlation coefficient are based on algebraic or mathematical treatment. The value of the coefficient of correlation by these formulae too remains between ±1. Following are the main mathematical methods –

1. Karl Pearson's / Covariance method
2. Rank Correlation method
3. Concurrent Deviation method

Karl Pearson's Method: Karl Pearson, a reputed statistician, in 1890, has constructed a well set formula based on mathematical treatment for determining the coefficient of correlation. The formula is named after his name as Karl Pearson's formula and is popularly known as 'Karl Pearson's coefficient of correlation'. It is also named as, 'Product moment coefficient'.

Example 1. Calculate the co-efficient of correlation from the following data:

X	9	8	7	6	5	4	3	2	1
Y	15	16	14	13	11	12	10	8	9

Solution: Calculation of co-efficient of correlation

X	$x = (X - \overline{X})$	x^2	Y	$y = (Y - \overline{Y})$	y^2	xy
9	4	16	15	3	9	12
8	3	9	16	4	16	12
7	2	4	14	2	4	4
6	1	1	13	1	1	1
5	0	0	11	-1	1	0
4	-1	1	12	0	0	0
3	-2	4	10	-2	4	4
2	-3	9	8	-4	16	12
1	-4	16	9	-3	9	12
$\sum X = 45$	$\sum x = 0$	$\sum x^2 = 60$	$\sum Y = 108$	$\sum y \quad 0$	$\sum y^2 = 60$	$\sum xy = 57$

$$\overline{X} = \frac{\Sigma X}{N} = \frac{45}{9} = 5$$

$$\overline{Y} = \frac{\Sigma Y}{N} = \frac{108}{9} = 12$$

$$\sum x^2 = 60, \sum y^2 = 60, \sum xy = 57$$

$$r = \frac{\Sigma xy}{\sqrt{\Sigma x^2 \times \Sigma y^2}} = \frac{57}{\sqrt{60 \times 60}} = \frac{57}{60} = +0.95$$

Example 2. From the following data compute the co-efficient of correlation between X and Y:

	X Series	**Y Series**
No. of items	15	15
Arithmetic Mean	25	18
Square of deviations from mean	136	138

Summation of product of deviations of X and Y series from their respective Arithmetic Mean is 122.

Solution:

Denoting deviations of X and Y from the arithmetic means by x and y respectively the given data are

$$\sum x^2 = 136, \sum y^2 = 138, \sum xy = 122$$

We apply Karl Pearson's methd

$$r = \frac{\Sigma xy}{\sqrt{\Sigma x^2} \times \sqrt{\Sigma y^2}} = \frac{122}{\sqrt{136 \times 138}} = \frac{122}{137} = +0.89$$

$$x = X - \overline{X}$$
$$y = Y - \overline{Y}$$

Example 3. Calculate the co-efficient of correlation between x and y series from the following data:

X	1	2	3	4	5
Y	3	4	6	7	10

Solution:

X	Y	x^2	y^2	xy
1	3	1	9	3
2	4	4	16	8
3	6	9	36	18
4	7	16	49	28
5	10	25	100	50
$\sum X = 15$	$\sum y = 30$	$\sum X^2 = 55$	$\sum y^2 = 210$	$\sum xy = 107$

$$r = \frac{N\sum xy - \sum x \sum y}{\sqrt{N\sum x^2 - \left(\sum x\right)^2}\sqrt{N\sum y^2 - \left(\sum y\right)^2}}$$

Substituting values, we set

$$r = \frac{5(107) - 15 \times 30}{\sqrt{5(55) - (15)^2}\sqrt{5(210) - (30)^2}}$$

$$= \frac{535 - 450}{\sqrt{275 - 225}\sqrt{1050 - 900}} = \frac{85}{\sqrt{50}\sqrt{150}} = \frac{85}{7.07 \times 12.25} = \mathbf{0.98}$$

Thus there is near perfect positive correlation between two series.

Example 4. The following table gives the soil temperature and the germination time at various places. Calculate the co-efficient of correlation and interpret the value

Temperature	57	42	40	38	42	45	42	44	40	46	44	43
Germination Time	10	26	30	41	29	27	27	19	18	19	31	29

Take 44 and 26 as assumed means

Solution:

We assume temperature as x and germination time as Y.

X	(X-44) dx	dx^2	Y	(Y-26) dy	dy^2	dxdy
57	13	169	10	-16	256	-208
42	-2	4	26	0	0	0
40	-4	16	30	+4	16	-16
38	-6	36	41	+15	225	-90
42	-2	4	29	+3	9	-6
45	+1	1	27	+1	1	+1
42	-2	4	27	+1	1	-2
44	0	0	19	-7	49	0
40	-4	16	18	-8	64	+32
46	+3	4	19	-7	49	-14
44	0	0	31	+5	25	0
43	-1	1	29	+3	9	-3
N = 12	$\sum dx = -5$	$\sum dx^2 = 255$		$\sum dy = -6$	$\sum dy^2 = 704$	$\sum dxdy = -306$

$$r = \frac{\Sigma dxdy - \frac{\Sigma dx \Sigma dy}{N}}{\sqrt{\Sigma dx^2 - \frac{(\Sigma dx)^2}{N}}\sqrt{\Sigma dy^2 - \frac{(\Sigma dy)^2}{N}}}$$

$$= \frac{-306 - \frac{(-5)^2 - (-6)}{12}}{\sqrt{255 - \frac{(-5)^2}{12}}\sqrt{704 - \frac{(-6)^2}{12}}}$$

$$= \frac{-306 - 2.5}{\sqrt{255 - 2.1}\sqrt{704 - 3}} = \frac{-308.5}{\sqrt{252.9}\sqrt{701}}$$

Example 5. Given

No. of pairs of observations In X and Y series = 8

X series Arithmetic average = 74.5

X series assumed average =69

X series standard deviation =13.07

Y series arithmetic average = 125.5

Y series assumed average = 112

Y series standard deviation= 15.85

Summating of products of corresponding deviations of X and Y series = 2176

Calculate coefficient of correlation.

Solution:

$$r = \frac{\Sigma xy - N(\bar{X} - A_x)(\bar{Y} - A_y)}{N\sigma_x\sigma_y}$$

Substitute the values in the formula

$$= \frac{2176 - 8(74.5 - 69)(125.5 - 112)}{8(13.07 \times 15.85)}$$

$$= \frac{2176 - 594}{16573.276} = \frac{1582}{1657.276} = +\mathbf{0.955}$$

Properties of Karl Pearson's coefficient:

The following are the main properties of Karl Pearson's coefficient of correlation.

1. In case correlation is present, then coefficient of correlation would lie between ±1. If correlation is absent, then it is denoted by zero or $-1 \leq r \leq +1$.
2. Coefficient of correlation is based on a suitable measure of variation as it takes into account all items of the variable.
3. Coefficient of correlation measures both the direction as well as degree of change
4. If there is accidental correlation, in that case the coefficient of correlation might lead to fallacious conclusions. It is known as non-sense or spurious correlation.
5. The coefficient of correlation does not prove causation but it is simply a measure of co-variation. It is because variations in X and Y series may be due to
   (i) some common cause,

(ii) some mutual dependence,

(iii) some change and

(iv) some causation of the subject to be relative

6. It is independent of changes of scale and origin of the variables X and Y.

7. Coefficient of correlation is a geometric mean of two regression coefficients. Symbolically $r = \pm\sqrt{b_{xy} \times b_{yx}}$.

8. Coefficient of correlation is independent of the unit of measurement

9. Coefficient of correlation works both ways i.e., $r_{xy} = r_{yx}$

10. If the value of x and y are linearly relate with each other i.e., if we have the relation between x and y as y = ax + b, the correlation coefficient between x and y will be +1 and if the relation between x and y is as y = -ax + b, then 'r' will be -1, 'a' being a negative constant.

Coefficient of Correlation and Probable Error:

Probable error is an instrument which measures the reliability and dependability of the value of 'r', the Karl Pearson's coefficient of correlation. The probable error of coefficient of correlation helps in interpreting its value. The coefficient of correlation is generally computed from samples, which are subject to errors of sampling. From the interpretation point of view Probable Error is very useful.

According to Horace Secrist. "The probable error of 'r' is an amount which if added to and subtracted from the average correlation coefficient produces amounts within which the chances are even that a coefficient of correlation from a series selected at random will fall".

According to Wheldon, "probable error defines the limit above and below the size of the coefficient determined within which there is an – equal chance that the coefficient of correlation similarly calculated from other samples will fall"

Karl Person's Probable error is calculated by using the following formula

P.E. = $0.6745 \dfrac{1-r^2}{\sqrt{N}}$, Where Standard Error = $\dfrac{1-r^2}{\sqrt{n}}$.

Where P.E. = Probable Error

r = Co-efficient of correlation

N = Number of pairs of observations

This method is used in interpreting whether 'r' is significant or not. If 'r' is more than P.E. then there is correlation and it is significant if 'r' is more than six times P.E.

Utility or Importance of Probable Error:

Following are the main points of utility of probable error:

(1) **Determination of Limits**: The Probable error of coefficient of correlation determines the two limits (r ± P.E.) within which, coefficients of correlation of randomly selected samples from the same universe will fall

(2) **Interpretation of 'r'**

The interpretation of r based on the probable error is as follows:

(i) If 'r' is less than the probable error (r < P.E.), then there is no evidence of correlation in two variables i.e., correlation is insignificant.

(ii) If 'r' is greater than probable error (r > P.E.), then the correlation is significant. If r is not more than six times the probable error ($r \ngtr 6$ P.E.), then it is not significant.

(iii) If the probable error is small, correlation is definitely existing where 'r' is above 0.5

(iv) If r > 0.3 and probable error is relatively small, 'r' is not treated as marked

(v) In other cases noting can be calculated with certainty

Example 6. Coefficient of correlation between advertising expenditure and sales for 9 items was observed as + 0.69. Find out probable error of correlation coefficient and comment on the significance of r.

Solution:

$$\text{P.E.} = 0.6745\ \frac{1-r^2}{\sqrt{N}} = 0.6745\frac{1-(0.69)^2}{\sqrt{9}} = 0.6745\ \frac{1-0.476}{3} = 0.6745\frac{0.524}{3} = 0.118$$

P.E. = 6 x 0.118 = **0.71**

So, we can say that r is less than 6 times the value of probable error. Hence the value of r is not significant.

Example 7. To study the correlation between the weights and heights of the students of a college, a sample of 100 is taken from the universe. The sample study gives the coefficient of correlation between two variables as 9. Within what limits does it hold good for the universe.

Solution:

Here r = 0.9, N = 100

$$\text{SE of r} = \frac{1-r^2}{\sqrt{N}} = \frac{1-(0.9)^2}{\sqrt{100}} = \frac{1-81}{10} = \mathbf{0.019}$$

$\therefore$ Lower Limit = r – 3 (SE) = 0.9 – 0.057 = **0.843**

$\therefore$ Upper Limit = 4 + 3 (SE) = 0.9 + 0.057 = **0.957**

Rank Correlation (Spearman's Method)

In 1904, Prof. Charles Edward Spearman had devised a method of computing coefficient of correlation. It is based on the ranking of different items in the variable. This method is useful where actual item values are not given, simply their ranks in the series are known. Thus it is a good measure in cases where abstract quantity of one group is correlated with that of the other group. In Spearman's coefficient of correlation, we take the differences in ranks, squaring them and finding out the aggregate of the squared differences. Symbolically

$$R = 1-\frac{6\Sigma D^2}{N(N^2-1)} = 1-\frac{6\Sigma D^2}{N^3-N}$$

Here r_k = Coefficient of rank correlation

D = Rank differences

N = Number of Pairs

In rank coefficient of correlation, we shall study three different cases

Case I When ranks are not given

Case II When ranks are given

Case III When ranks are equal

The value of rank correlation falls between ±1.

Case I: When ranks are not given

Example 8. Compute Rank correlation from the following table

X	415	434	420	430	424	428
Y	330	332	328	331	327	325

Solution:

X	R_1	Y	R_2	$(R_1 - R_2) = D$	D^2
415	6	330	3	3	9
434	1	332	1	0	0
420	5	328	4	1	1
430	2	331	2	0	0
424	4	327	5	-1	1
428	3	325	6	-3	9
					$\sum D^2 = 20$

$$r_k = 1\frac{6\Sigma D^2}{N\,(N^2-1)}$$

$$= 1 - \frac{1(20)}{6\,(6^2-1)} = 1 - \frac{120}{210} = \frac{210-120}{210} = \frac{90}{210} = \frac{3}{7} = \mathbf{0.429}$$

Case 2: When ranks are given:

Example 9. The ranks of students in Hindi (R_1) and Economics (R_2) are given

Hindi (R_1)	6	1	5	2	4	3
Economics (R_2)	3	1	4	2	5	6

Solution:

R_1	R_2	$(R_1 - R_2) = D$	D^2
6	3	3	9
1	1	0	0
5	4	1	1
2	2	0	0
4	5	-1	1
3	6	-3	9
			$\sum D^2 = 20$

$$R = 1 - \frac{6\Sigma D^2}{N^3 - N}$$

$$= 1 - \frac{1(20)}{6\,(6^2-1)} = 1 - \frac{120}{210} = \frac{210-120}{210} = \frac{90}{210} = \frac{3}{7} = \mathbf{0.429}$$

Case – III. When ranks are equal (Tied or repeated Ranks)

In certain cases, we may find equal ranks in case of two or more than two values. In that case, each individual item is given an average rank. If two individuals are ranked equal at third place, they are each given the rank $\frac{3+4}{2}$ = 3.5, while if three are ranked at third place then $\frac{3+4+5}{3}$ = 4 would be the common rank for third, fourth and fifth place.

When equal ranks are given to some entries, some adjustment in the above formula has to be made in calculating the rank correlation, the adjustment is made by adding ½ $(m^3 - m)$ to the value of ΣD^2. Here 'm' stands for the number of items which have common rank. In case, there are more than one such group of items with common rank, the value is added as many times as the number of such groups The formula in that case is written as

$$r_k = 1 - \frac{6\left\{\Sigma D^2 + \frac{1}{12}(m^3 - m) + \cdots\right\}}{N^3 - N}$$

Example 10.

Eight students have obtained the following marks in Accountancy and economics. Calculate the Rank Co-efficient of correlation.

Accountancy (X)	25	30	38	22	50	70	30	90
Economics (Y)	50	40	60	40	30	20	40	70

Solution:

Calculation of Rank Correlation

X	R_1	Y	R_2	$(R_1 - R_2) = D$	D^2
25	2	50	6	-4	16.00
30	3.5	40	4	-0.5	0.25
38	5	60	7	-2	4.00
22	1	40	4	-3	9.00
50	6	30	2	+4	16.00
70	7	20	1	+6	36.00
30	6.5	40	4	-0.5	0.25
90	8	70	8	0	0.00
					$\Sigma D^2 = 81.5$

$$r_k = 1 - \frac{6\left\{\Sigma D^2 + \frac{1}{12}(m^3 - m) + \frac{1}{12}(m^3 - m)\right\}}{N^3 - N}$$

$\Sigma D^2 = 81.5 \, N = 8$

As item 30 is repeated 2 times in X series so m = 2

In series Y the item 40 is repeated 3 times, so m = 3

$$\therefore R = r_k = 1 - \frac{6\left\{81.5 + \frac{1}{12}(2^3 - 2) + \frac{1}{12}(3^3 - 3)\right\}}{8^3 - 8}$$

$$= 1 - \frac{6(81.5 + 0.5 + 2)}{504} = 1 - \frac{6(84)}{504} = 1 - \frac{504}{504} = 1 - 1 = 0$$

MULTIPLE CHOICE QUESTIONS

1. Bivariate Data are the data collected for
   (a) Two variables
   (b) More than two variables
   (c) Two variables simultaneously
   (d) Two variables at different points of time
2. Simple correlation is called
   (a) Linear correlation
   (b) Nonlinear correlation
   (c) Both
   (d) None
3. If high values of one tend to low values of the other, they are said to be
   (a) Negatively correlated
   (b) Inversely correlated
   (c) Both
   (d) None
4. The correlation is said to be positive
   (a) When the values of two variables move in the same direction
   (b) When the values of two variables move in the opposite direction
   (c) When the values of two variables would not change
   (d) None of these
5. Age of Applicants for life insurance and the premium of Insurance – correlations are
   (a) Positive
   (b) Negative
   (c) Zero
   (d) None
6. The correlation between sale of cold drinks and day temperature is __
   (a) Zero
   (b) Positive
   (c) Negative
   (d) None of these
7. In case "The ages of husbands and wives' correlation is
   (a) Positive
   (b) Negative
   (c) Zero
   (d) None

8. Whatever may be the value of r, positive or negative, its square will be
   (a) Negative only
   (b) Positive only
   (c) Zero only
   (d) None only
9. The value of correlation coefficient lies between
   (a) -1 and +1
   (b) -1 and 0
   (c) 0 and 1
   (d) None
10. What are the limits of the correlation coefficient?
   (a) No limit
   (b) -1 and 1, including the limits
   (c) 0 and 1, including the limits
   (d) -1 and 0, including the limits
11. If $y = a + bx$, then what is the coefficient of correlation between x and y?
   (a) 1
   (b) -1
   (c) 1 or -1 according as $b > 0$ or $b < 0$
   (d) None of these
12. Scatter diagram helps us to
   (a) Find the nature correlation between two variables
   (b) Compute the extent of correlation between two variables
   (c) Obtain the mathematical relationship between the two variables
   (d) Both a and c
13. Scatter diagram is considered for measuring
   (a) Linear relationship between two variables
   (b) Curvilinear relationship between two variables
   (c) Neither a or b
   (d) Both a and b
14. If the plotted points in a scatter diagram lie from upper left to lower right, then the correlation is
   (a) Positive
   (b) Zero
   (c) Negative
   (d) None

15. If the coefficient of correlation between x and y is 0.28, co-variance between x and y is 7.6 and the variance of x is 9, then the S.D. of y series is;

(a) 9.8

(b) 10.1

(c) 9.05

(d) 10.05

16. If for two variable X and Y, the covariance, variance of X and variance of Y are 40, 16 and 256 respectively, what is the value of the correlation coefficient?

(a) 0.01

(b) 0.625

(c) 0.4

(d) 0.5

17. If the relationship between two variables x and y is given by 2x+ 3y+4 = 0, then the value of the correlation coefficient between x and y is

(a) 0

(b) 1

(c) -1

(d) Negative

18. If Cov (u.v) = 3, $\sigma_u^2 = 4.5$, $\sigma_y^2 = 5.5$, then r(u,v) is:

(a) 0.347

(b) 0.603

(c) 0.07

(d) 0.121

19. What is the coefficient of correlation from the following data?

X	1	2	3	4	5
Y	8	6	7	5	5

(a) 0.75

(b) - 0.75

(c) - 0.85

(d) 0.82

20. Maximum value of Rank Correlation coefficient is

(a) -1

(b) +1

(c) 0

(d) None

21. The sum of the difference of rank is
    (a) 1
    (b) -1
    (c) 0
    (d) None
22. For a group of 8 students, the sum of squares of differences in ranks for Maths and Stats mark was found to be 50. What is the value of rank correlation coefficient?
    (a) 0.23
    (b) 0.40
    (c) 0.78
    (d) 0.92
23. For 10 pairs of observations, no. of concurrent deviations was found to be 4. What is the value of the coefficient of concurrent deviation?
    (a) $\sqrt{0.2}$
    (b) $\sqrt{0.2}$
    (c) 1/3
    (d) -1/3
24. The coefficient of concurrent deviation for p pairs of observations was found to be $\frac{1}{\sqrt{3}}$. If the number of concurrent deviations was found to be 6, then the value of p is
    (a) 10
    (b) 9
    (c) 8
    (d) None of these

EXERCISE-I:

1. From the following table calculate the coefficient of correlation by Karl Pearson's method and also fill the gap.

X	8	4	12	6	10
Y	11	13	?	10	9

Arithmetic means of X and Y series are 8 and 10 respectively.

2. Given Number of pairs of observations of X and Y = 9

X Series Arithmetic average = 70.5,

X Series Standard deviation = 9.07,

X series Assumed average = 65.0,

Y Series Assumed average = 108.0

Y Series Arithmetic Average = 121.5

Y Series standard deviation = 11.85

Summation of products of corresponding deviation of X and Y series = 1451

Calculate coefficient of correlation.

3. Find the rank correlation for the following distribution

Marks in Economics	48	60	72	62	56	40	39	52	30
Marks in Accountancy	62	78	65	70	38	54	60	32	31

4. Calculate correlation coefficient from the following results:

$N = 10, \sum x = 100\ \sum y = 150, \sum(X-10)^2 = 180, \sum(X-15)^2 = 215\ \sum(X-10)\ (Y-15) = 60$

5. If the covariance between X and Y variables is 10 and the variance of X and Y are respectively 16 and 9, find the coefficient of correlation.

6. From the data given below, find the number of items.

$r = 0.5 \qquad \sum xy = 120 \qquad \sum X^2 = 90 \qquad \sigma_y = 8$

(Where x and y are deviations from arithmetic average).

7. From the following data compute the coefficient of correlation between X and Y:

	X-Series	Y-Series
Arithmetic Mean	15	28
Sum of Squares of deviations from mean	144	225

Summation of products of deviations of X and Y series from their respective means = 20

8. In a question on correlation the value of r is 0.64 and its PE=0.1312. What was the value of N?

9. From the marks obtained by 8 students in Accountancy and Statistics, compute rank coefficient of correlation.

Marks in Accountancy	60	15	20	28	12	40	80	20
Marks in Statistics	10	40	30	50	30	20	60	30

10. The coefficient of rank correlation between marks in Quantitative Mathematics and Economics obtained by a certain group of students is 7/11. The sum of the squares of differences in ranks is 60. What is the number of students in the group?

11. The coefficient of rank correlation of the marks obtained by 10 students in statistics and accountancy was found to be 0.5. It was later discovered that the difference in ranks in the two subjects obtained by one of the students was wrongly taken as 3 instead of 7. Find the correct coefficient of rank correlation.

12. Find rank correlation coefficient from the following data

X	10	12	15	22	28	30	45	60	72
Y	32	35	42	48	52	30	65	68	70

13. Find correlation coefficient for data given below:

Age (in years)	42	36	48	43	55	52	38
Blood Pressure	132	120	140	133	142	148	122

14. Calculate Karl Pearson's co-efficients of correlation between expenditure and advertisement and sales from the data given below:

Advertising expenses (1000) ₹	39	78	65	62	90	82	75	25	98	36
Sales (Lakh ₹)	47	84	53	58	86	62	68	60	91	51

The following table gives the marks obtained in economics and statistics by certain students in an examination. Calculate the value of Coefficient of Correlation.

Marks in Economics	44	42	58	56	79	76	66
Marks in Statistics	56	49	53	59	65	78	58

EXERCISE-II

1. Find the coefficient of correlation and probable error from the data:

X	67	69	71	75	85	93	87	73
Y	95	80	87	80	79	75	80	85

2. Calculate Karl Pearson's coefficient of correlation between Imports and exports:

Imports ₹ (billion)	42	44	58	55	89	98	66
Exports ₹(billions)	56	49	53	58	65	76	58

3. From the following data compute the coefficient of correlation between X and Y.

	X Series	Y Series
Arithmetic mean	25	18
Square of deviations form A.M.	136	138

Summation of products of deviations of X and Y series from their respective means = 122. Number of pairs of value is 15.

4. Calculate Rank Correlation Coefficient for the following:

Marks in English	29	28	17	15	20	26	27	25	34	19
Marks in Maths	31	32	25	29	42	15	43	32	20	40

5. Ten competitors in a beauty contest are ranked by 3 judges in the following order:

Judge A:	1	6	5	10	3	2	4	9	7	8
Judge B:	3	5	8	4	7	10	2	1	6	9
Judge C:	6	4	9	8	1	2	3	10	5	7

Use Rank correlation coefficient to determine which pair of judges has the nearest approach to common taste in beauty.

6. Calculate the coefficient of correlation by any suitable method:

A	115	168	170	127	118	129	135	140
B	2	6	8	11	1	1	4	3

Also calculate probable error

7. Calculate Karl Pearson's coefficient of correlation from the following data:

Values of X	100	110	115	116	120	125	130	135
Values of Y	18	18	17	16	16	15	13	10

8. Calculate coefficient of correlation and its probable error from the given data

Variable of X	42	44	58	55	89	98	66
Variable of Y	56	49	53	58	65	76	58

9. From the data given below, compute the correlation coefficient by the methods of concurrent deviations.

Year	1996	1997	1998	1999	2000	2001	2002
Supply	150	154	160	172	160	165	180
Price	200	180	170	160	190	180	172

10. Making use of the data summarized below calculate the coefficient of correlation (r_{12})

Case	A	B	C	D	E	F	G	H
X_1	10	6	9	10	12	13	11	9
X_2	9	4	6	9	11	13	18	4

The coefficient of correlation between the variables X and Y is 0.64, their covariance is 16. The variance of X is 9. Find the standard deviation of Y.

12. The following table gives the scores obtained by 11 students in English and Tamil translation. Find the rank correlation coefficient.

Scores in English	40	46	54	60	70	80	82	85	85	90	95
Scores in Tamil	45	45	50	43.	40	75	55	72	65	42	70

13. Calculate the coefficient of concurrent deviation from the following:

X	60	55	50	55	30	70	40	35	80	80	75
Y	65	40	35	75	63	80	35	20	80	60	60

14. Calculate Pearson's coefficient of correlation from the following data using 44 and 26 as the origin of X and Y respectively.

X	43	44	46	40	44	42	45	42	38	40	42	57
Y	29	31	19	18	19	27	27	29	41	30	26	10

MCQ Answers:

1	c	2	a	3	c	4	c	5	a	6	b	7	a	8	b	9	a	10	b
11	c	12	a	13	d	14	c	15	c	16	b	17	c	18	b	19	c	20	b
21	c	22	b	23	d	24	a												

5.1 REGRESSION

Introduction:

The analysis of coefficient of correlation between two variables, examines the extent and degree of correlation ship between two variables, i.e., it measures the closeness with which two or more variables co-vary in a given period of study. Similarly, we can estimate or predict the value of a variable given the value of another variable on the basis of functional relationship between them. The statistical technique of estimating or predicting the unknown value of a dependent variable from the known value of an independent variable is called regression analysis. Sir Francis Galton introduced the concept of 'regression' for the first time in 1877 where he studies the case of one thousand fathers and sons and concluded that the tall fathers tend to have tall sons and short fathers have short sons, but the average height of the sons of a group of tall fathers is less than that of the fathers and the average height of the sons of a group of short fathers is greater than that of the fathers. The line showing this tendency to go back was called by Galton a "Regression Line". The modern statisticians use the term 'estimating line' instead of regression line as this concept is more classificatory now.

Sales depend on promotional expense. It is possible to predict sales for a given promotion expense. Regression is more useful for business planning and forecasting.

It economics it is the basic tool for estimating the relationship among economic variables that constitute the essence of economic theory. If we know the two variables, price (X) and demand (Y) are closely related; in that case we can find the most probable value of Y for a given value of X.

Definition:

According to Morris M. Blair, "Regression is the measure of the average relationship between two or more variables in terms of the original units of the data."

Objectives of Regression analysis:

Regression analysis does the following

1. Explain the variations in the dependent variable as a result of using a number of independent variables.
2. Describe the nature of relationship in a precise manner by way of regression equation.
3. It is used in prediction and forecasting problems
4. It helps in removing unwanted factors.

Classification of regression analysis: The regression analysis can be classified on the following bases: -

(i) Change in Proportion; and

(ii) Number of variables

Basis of Change in Proportions: On the basis of proportions the regression can be classified into the following categories: -

1. Linear regression and
2. Non-linear regression

1. **Linear regression Analysis Model**: When dependent variable moves in a fixed proportion of the unit movement of independent variable, it is called a linear regression. Linear regression, when plotted on a graph paper, forms a straight line. Mathematically the relation between X and Y variables can be expressed by a simple linear regression equation as under-

$$\mathbf{y_i = a + bx_i + e_i}$$

Where a and b are known as regression parameters, e_i denotes residual terms, x_i presents value of independent variable and y_i is the value of dependent variable. 'a' express the intercept of the regression line of y on x. i.e. value of dependent variable say y, when the value of independent variable, that is 'x' is zero. Again 'b' denotes the slope of regression line of y on x. Again e_i denotes the combined effect of all other variables, (not taken in the model) on y. this equation is known as classical simple linear regression model.

(2) **Non-linear Regression Analysis**: Contrary to the linear regression model, in non-linear regression, the value of dependent variable say 'y' does not change by a constant absolute amount for unit change in the value of the independent variable, say 'x'. If the data are plotted on a graph, it would form a curve, rather than a straight line. This is also called curvi-linear regression.

On the basis of Number of Variables: On the basis of number of variables regression analysis can be classified as under:

1. Simple regression
2. Partial Regression
3. Multiple Regression

1. **Simple regression**: When only two variables are studied to find the regression relationships, it is known as simple regression analysis. Of these variables, one is treated as an independent variable while the other as dependent one. Functional relationship between price and demand may be noted as an example of simple regression.
2. **Partial Regression**: When more than two variables are studied in a functional relationship but the relationship of only two variables is analysed at a time, keeping other variables as constant, such a regression analysis is called partial regression.
3. **Multiple Regression**: When more than two variables are studied and their relationships are simultaneously worked out, it is a case of multiple regression. Study of the growth in the production of wheat in relation to fertilizers, hybrid seeds, irrigation etc., is an example of multiple regression.

REGRESSION LINES:

A regression line is a graphic technique to show the functional relationship between the two variables X and Y. i.e, dependent and independent variables. It is a line which shows average relationship between two variables X and Y. Thus, this is a line of average. This is also called an estimating line as it gives the average estimated value of dependent variable (Y) for any given value of independent variable (X).

According to Galton, "The regression lines show the average relationship between two variables".

METHODS OF DRAWING REGRESSION LINES:

The regression lines can be drawn by two methods as given below: -

1. Free Hand Curve Method
2. The method of Least Squares

1. **Free Hand Curve Method**: This method is also known as the method of Scatter Diagram. This is a very simple method of constructing regression lines. At the same time it is a crude and very rough and rarely used method of drawing regression lines. In this method, the value of paired observations of the variable are plotted on the graph paper. It takes the shape of a scattered diagram scattered over the graphic range of X axis and Y axis. The independent variable is taken on the vertical axis. A straight line is drawn through the scattered points on the graph that it confirms the following requisites: -

   (a) It is at the maximum possible nearer to all the points on the graph

(b) It is at the equi-distance of all the points on either sides of the line

(c) It passes through the centre of scattered points.

Example: From the following data relating to X and Y, draw a regression line of Y on X.

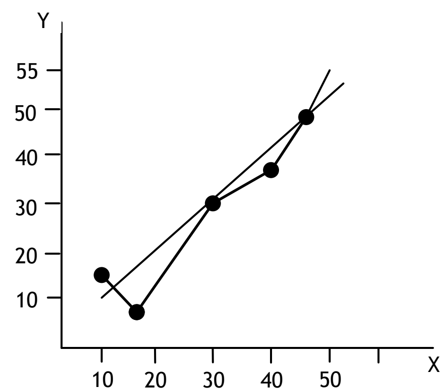
X	10	16	24	36	48
Y	20	12	32	40	55

Solution:

To draw a regression line of Y on X, X is taken on 'X' axis and Y on Y axis and a line is drawn by free hand as under –

This method being most subjective is not commonly used in practice.

2. **Method of Least Squares**: The other method of drawing a line of regression is the method of least squares According to the least squares method the line should be drawn through the plotted points in such a way that the sum of the squares of the deviations of the actual Y values from the computed 'Y' values is the minimum or the least. The line which fits the points in the best manner should have $\sum(Y-Y_c)^2$ as minimum. A line fitted by this method is called the lines of best fit. One of the characteristics of the line of best fit is that deviations above the line are equal to the deviations below the line. It implies that the total of the positive and negative deviations is Zero, i.e., $\sum(Y-Y_c) = 0$. The line of best fit or the straight line goes through the overall mean of the data i.e. $\overline{X}$, $\overline{Y}$.



Methods of Calculating Regression Equations or Derivation of Regression Lines:

Following are the two methods to form the two regression equations, that equation for Y on X and, for X on Y.

1. Regression equations through normal equations
2. Regression equations through regression co-efficient.

1. **Regression Equations through Normal Equations**: The two main equations generally used in regression analysis are:

   (i) Y on X,

   (ii) X on Y

   for Y on X, the equation is $Y_c = a + bX$

   for X on Y, the equation is $X_c = a + bY$

'a' and 'b' are constant values and 'a' is called the intercept. In the case of Y on X it is an estimated value of Y when X is zero and similarly in the case of X on Y, it shows the value of X when Y is zero. 'b' represents the slope of the line, that is change per unit of an independent variable. It is also known as regression coefficient of Y on X or X on Y as the case may be and also denoted as b_{yx} for Y on X and b_{xy} for X on Y. If 'b' is having positive sign before it, regression line will be upward sloping and in case of negative sign, the line shall be sloping downwards.

Y_c or X_c are the values of Y or X computed from the relationship for a given X or Y.

Regression Equation of Y on X:

The regression equation of Y on X can be written as $Y_c = a + bX$

We can arrive at two normal equations as follows

Given $Y = a + bx$….(i)

Now summate (∑) Eq. (i)

$\sum Y = Na + b \sum X$ ………..(ii)

Now multiply the whole equation (ii) by X; we get

$\sum XY = a \sum X + b \sum X^2$ ………..(iii)

Equation (ii) and (iii) are called normal equations

Regression Equation of X on Y

The regression of X on Y is expressed as

$X_c = a + b\,Y$

For determining the values of 'a' and 'b we determine two normal equations which can be solved simultaneously.

Summate Eq. (i)

$\sum X = Na + b \sum Y$ ………..(ii)

Multiply Eq. (ii) by Y; we get

$\sum XY = a \sum Y + b \sum Y^2$ ………..(iii)

Equation (ii) and (iii) are normal equations.

Example: Given the bivariate data:

X	2	6	4	3	2	2	8	4
Y	7	2	1	1	2	3	2	6

(a) Fit the regression line of Y on X and hence predict Y, if X = 20

(b) Fit the regression line of X on Y and hence predict X, if Y = 5.

Solution:

In this case both the regression lines are needed. It is necessary to find 'a' and 'b'. Given X, Y can be estimated and vice versa.

Computation of Regression Equations:

X	Y	X^2	Y^2	XY
2	7	4	49	14
6	2	36	4	12
4	1	16	1	4
3	1	9	1	3
2	2	4	4	4
2	3	4	9	6
8	2	64	4	16
4	6	16	36	24
$\sum X = 31$	$\sum Y = 24$	$\sum X^2 = 153$	$\sum Y^2 = 108$	$\sum XY = 83$

In order to find the values of a and b, two equations are to be solved simultaneously.

$\sum Y = N a + b \sum X$

$\sum XY = a \sum X + b \sum X^2$

Substitute the values; we get

$24 = 8a + 31b$ …(i)

$83=31a +153b$ …(ii)

Multiply Eq. (i) by 31 and Eq. (ii) by 8 we get

$744=248a + 961b$ ….(iii)

$664 = 248a + 1224b$ ….(iv)

Subtract (iii) from (iv)

$248a + 1224b = 664$

$\underline{-248a - 961b = -744}$

$263b = - 80$

$b = \frac{-80}{263} = - 0.30$ approx

Substituting b = - 0.30 in Eq. (i), we get

$24 = 8a + 31\ (-030)$

$24 = 8a - 9.3$

$8a = 33.3$

$a = \frac{33.3}{8} = 4.1625$

Thus regression of Y on X is

Y = 4.1265 – 0.30 X Ans.

Similarly we can solve second regression equation with the help of the simultaneous equations. The regression equation of X on Y is

$X = a + bY$

The two normal = ns are

$\sum X = N a + b \sum Y$

$\sum XY = a \sum Y + b \sum Y^2$

Substitute the values in Eq. (i) and (ii)

$31 = 8a + 24b$ …(i)

$83 = 24a + 108b$ …(ii)

Multiply Eq. (i) by 3 and subtract (ii) from (i) we get

$93 = 24a + 72b$ …(iii)

$\underline{-83 = -24a - 108b}$ …(iv)

$10 = -36b$

$b = \frac{-10}{36} = \frac{-5}{18}$ = -0.28 approximately

Put this is Eq. (i)

$8a + (24)(-0.28) = 31$

$8a - 6.72 = 31$

$8a = 31 + 6.72$

$a = \frac{37.72}{8} = 4.715$

Thus the regression of X on Y is

X = 4.715 – 0.28 Y Ans.

(a) Now let us predict the value of Y if X = 20

Y on X regression equation is

$Y = 4.1625\ 0.30X = 4.1625 - 0.30(20) = 4.1625 - 6.00$

Y = -1.8375 Ans.

(b) Now let us predict X if Y = 5

$X = 4.715 - 0.28\ (5)$

$X = 4.715 - 1.40$

X = 3.315 Ans.

Regression Equations through Regression Coefficients:

Regression coefficient refers to the constant value multiplied to the independent variable in a given relation. Say a relation Y = a + bx, here b (the slope of the regression line) is the regression coefficient, since it is a multiple of independent variable x, Regression equations or lines can easily be arrived at by the use of regression coefficients. For this purpose, we are required to calculate mean, standard deviation and correlation coefficient of the given series. The following are the main methods to calculate regression coefficient Y on X (b_{yx}) or X on Y (b_{xy}).

1. Taking deviations from actual mean
2. Taking deviations from Assumed mean
3. Applying actual observations
4. Applying grouped data

Regression equation of X on Y. This can be written as

$$X - \overline{X} = r\frac{\sigma_x}{\sigma_y}(Y - \overline{Y})$$

$r\dfrac{\sigma_x}{\sigma_y}$ is known as the regression coefficient of X on Y. It is denoted by b_{xy}.

Regression equation of Y on X

$$Y - \overline{Y} = r\frac{\sigma_y}{\sigma_x}(X - \overline{X})$$

$r\dfrac{\sigma_y}{\sigma_x}$ is the regression coefficient of Y on X. It is denoted by b_{yx}

When deviations are taken from actual means, the regression coefficient of Y on X can be written as

$$r\frac{\sigma_y}{\sigma_x} = \frac{\Sigma xy}{\Sigma x^2} = b_{yx}$$

Example: Given the bivariate data:

X	1	5	3	2	1	2	7	3
Y	6	1	0	0	1	2	1	5

Find regression equations by taking deviations of items from the means of X and Y respectively

Solution:

Computation of Regression Equations:

	$(X - \overline{X})$				$(Y - \overline{Y})$	
X	x	X^2	Y	y	y^2	xy
1	-2	4	6	4	16	-8
5	2	4	1	-1	1	-2
3	0	0	0	-2	4	0
2	-1	1	0	-2	4	2
1	-2	4	1	-1	1	2
2	-1	1	2	0	0	0
7	4	16	1	-1	1	-4
3	0	0	5	3	9	0
$\sum X = 24$		$\sum X^2 = 30$	$\sum Y = 16$		$\sum Y^2 = 36$	$\sum XY = -10$

Regression Equation of Y on X i.e.

$y = bx$ ....(i)

$$b = \frac{\Sigma XY}{\Sigma X^2} = \frac{-10}{30}$$

$$\overline{X} = \frac{\Sigma X}{N} = \frac{24}{8} = 3$$

$$\overline{Y} = \frac{\Sigma Y}{N} = \frac{16}{8} = 2$$

Substitute the values in (i); we get

$y = \frac{-10}{30} x = -0.33x$

But   y = (Y – 2) and x = (X -3)

Y – 2 = -0.33 (X-3)

Y – 2 = -0.33 X + 0.99

Y = -0.33 X + 2.99

Y = 2.99 – 0.33 X

Regression equation of X on Y i.e.

i.e. x = by

$b = \frac{\Sigma xy}{\Sigma y^2}$

Substituting the values, we get

$x = \frac{-10}{36} y = -0.28y$

But x = (X -3) = -0.28 (Y-2)

(X-3) = -0.28(Y – 2)

X -3 = -0.28 Y + 0.56

X = 3.56 – 0.28 Y

Deviations taken from Assumed Means X and Y:

In practice we get means of fractions and for simplicity we take deviations from assumed means. When the deviations are taken from the assumed means, the procedure for finding regression equations remains the same. In case of actual means the regression equations are

$$X - \overline{X} = r\frac{\sigma_x}{\sigma_y}\left(Y - \overline{Y}\right)$$

The value of r will now be obtained as follows

$$r\frac{\sigma_x}{\sigma_y} = \frac{\Sigma dxdy - \frac{\Sigma dx \times \Sigma dy}{N}}{\Sigma dy^2 - \frac{(\Sigma dy)^2}{N}} = b_{xy}$$

dx = (X - A) and dy = (Y – A)

Properties of Regression Coefficient:

The main properties of regression coefficients are as under:

1. Both the regression coefficients b_{xy} and b_{yx} cannot be greater than unity, i.e., either both are less than unity or one of them must be less than unity. In other words, the square root of the product of two regression coefficient must be less than or equal to 1 or $\sqrt{bxy \times byx} \geq -1$ or $\sqrt{b_{xy} \times b_{yx}} \leq 1$.
2. Both the regression coefficients will have the same sign i.e.,
   (i) If b_{xy} is positive, then b_{yx} will also be positive.
   (ii) If b_{xy} is negative, then b_{yx} will also be negative.

(iii) Both b_{xy} and b_{yx} must have same signs. If both are positive, r will be positive and vice-versa.

(iv) $\sqrt{b_{yx} \times b_{xy}} \leq 1$ or $\sqrt{b_{xy} \times b_{yx}} \geq -1$.

3. Correlation Co-efficient is the geometric mean between regression co-efficients i.e., $r = \pm\sqrt{b_{xy} \times b_{yx}}$

Example: Past data on household income and expenditure reveals that

(a) The average absolute increase in income in relation to increase in expenditure is ₹1.5 crore and

(b) The average absolute increase in expenditure in relation to increase in income in ₹50 crore

Find the coefficient of correlation between household income and expenditure.

Solution:

Let the Increase in Income by = x

And the increase in expenditure = y.

The value of regression coefficient $b_{xy} = 1.5$

Similarly the regression coefficient

$b_{yx} = 0.5$

To find r $r = \sqrt{b_{xy} \times b_{yx}} = \sqrt{1.5 \times 0.5}$

$= \sqrt{0.75} = 0.866$

r = 0.866 **Ans.**

4. The arithmetic mean of b_{yx} and b_{xy} is greater than or equal to Coefficient of Correlation i.e.

$$\frac{b_{yx} + b_{xy}}{2} \geq r$$

5. Since $b_{yx} = r\,\frac{\sigma_y}{\sigma_x}$, we can find any of these four values, given the other three.

6. If $6_x = 6_y$, then the coefficient of Correlation (r) is equal to regression Coefficient i.e. $r = b_{yx} = b_{xy}$

7. If r = 0 then b_{xy} and b_{yx} both are zero

8. If $b_{xy} = b_{yx}$ then it is also equal to Coefficient of Correlation. It means $b_{xy} = b_{yx} = r$

9. Regression Coefficients are independent of change of origin but not of scale

Some Important Relations:

1. Relation between regression coefficients and standard deviations

(a) The regression coefficient of Y on X is equal to $b_{yx} = r\frac{\sigma_y}{\sigma_x}$

(b) The regression coefficient of X on Y is equal to $b_{xy} = r\frac{\sigma_x}{\sigma_y}$

2. Relation between Regression coefficient and co-efficient of Determination (r^2):

Coefficient of Determination is the product of both the coefficients of regression i.e.,

$\therefore r = \pm\sqrt{b_{xy} \times b_{yx}} \rightarrow r^2 = b_{xy} \times b_{yx}$

3. In case of a perfect correlation b_{yx} should be the reciprocal of b_{xy}

For a perfect correlation r = 1 but $r^2 = b_{yx} \times b_{xy}$

$\therefore (1)^2 = b_{yx} \times b_{xy}$

$\therefore \dfrac{1}{b_{xy}} = b_{yx}$

4. The point of intersection of the Two Regression Lines: Both the regression lines intersect at a point where the values of X and Y would be the mean values i.e., X and Y. Thus if we solve the two regression equations as simultaneous equations the values of X ad Y so obtained and respectively the values of $\overline{X}$ and $\overline{Y}$.

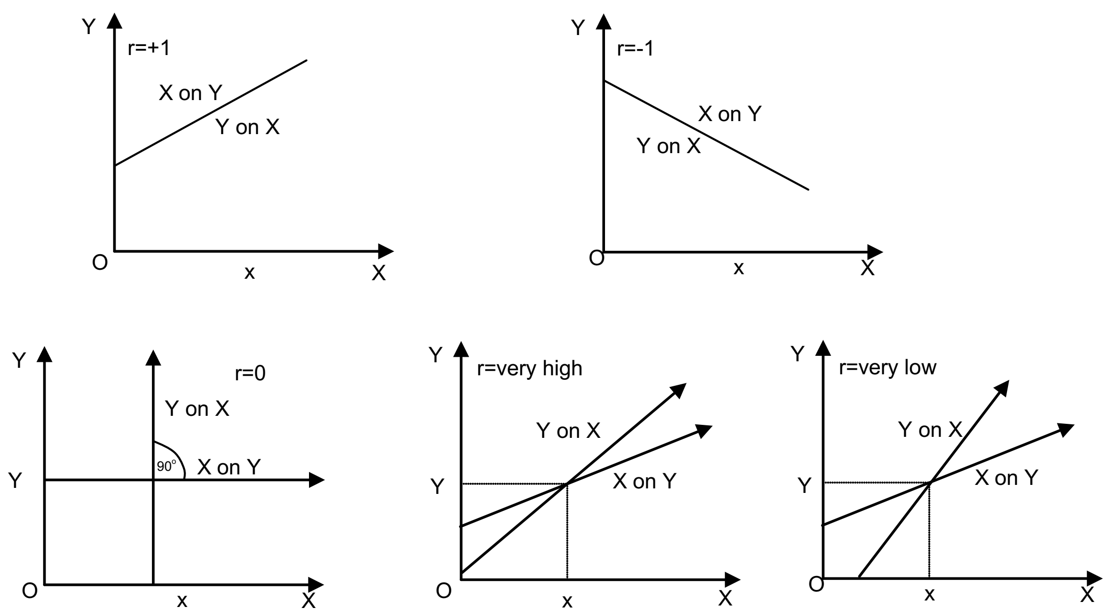
Correlation and Regression:

Correlation and Regression analysis are two important statistical tools to study the functional relationship between the variables. Coefficient of correlation is a measure of degree of covariance between x and y whereas the aim of regression analysis is to study the nature of relationship between the variables, so that we may know the value of one variable on the basis of another.

However, we can study the relation and difference between the correlation and regression as under.

Relation between Correlation and Regression:

The regression lines are used to predict the value of dependent variable given the different values of independent variable. These two regression lines show the average relationship between two variables. If there is perfect correlation (positive or negative, i.e., r ±1) both the lines will coincide, as shown in figure 1 and 2 In this case there will be one regression line. In case r = 0, i.e., both the variables are independent, both the lines will cut each other at right angle i.e., parallel to OX and OY, as shown in figure 3. Higher the value of r, closer will be the regression lines to each other, as shown in figure 4. On the other hand if the value



of r is low, the gap between regression lines will be wider, as shown in the figure. It is important to note that the point, where both the regression lines intersect each other, gives us the mean values of X and Y Variables, i.e., $\overline{X}$ and $\overline{Y}$.

Difference between Correlation and Regression

The difference between correlation and regression can be analysed as under –

1. **Nature of Relationship**: Correlation analysis tests the closeness of the variables, whereas regression analysis measures the extent of change in dependent (y) variable due to change in the independent (x) variable i.e., both the nature and extent of functional relationship between the variables is studied.

2. **Relationship**: In regression analysis, the causal relationship in variables moving in the same or opposite direction is studied while in correlation analysis, the study is made by taking into consideration the cause and effect relationship between the variables. i.e. only closeness of the variables is studied.

3. **Mutual dependence of variables**: Correlation studies the mutual dependence of variables but in regression analysis the functional relationship showing dependence of one variable upon other is analysed.

4. **Spurious Correlation**: In correlation chance or spurious correlation between two variables, having no practical importance, may be observed, but there is no chance of existence of such type of relation in regression analysis.

5. **Mathematical treatment**: Correlation is having no scope of further mathematical treatment whereas regression can be used for further treatment.

6. **Origin & Scale**: Regression analysis is independent of change of origin but not of scale, whereas correlation is both independent of change in origin and scale.

7. **Relative and Absolute Measures**: Regression analysis is an absolute measure showing the change in the value of y or x for unit change in the value of x or y. whereas correlation coefficient is a relative measure of linear relationship between x and y and is independent of the measurement. It is a number which lies between ±1.

8. **Applicability**: Correlation has very limited scope of application. It is delimited to the linear relationship between two variables, but the scope of applicability of regression analysis is very wide. It can be covered under linear as well as non-linear relationship between the variables.

9. **Differentiation in variables**: In correlation both variables are considered at par for study purposes, whereas in regression analysis variables are differentiated as dependent and independent variables.

10. **Symmetrical or Asymmetrical formation:** Correlation is symmetrical in formation, i.e. $r_{xy} = r_{yx}$. The measurement of the co-variability of two variables is symmetrical in formation. It means 'r' is a both way relationship of x on y or y on x, but in regression, the approach is one way of analysis treating one variable as dependent and the other as an independent, thus making the analysis treating one variable as dependent and the other as an independent, thus making the analysis Asymmetrical, that is $b_{yx} \neq b_{xy}$.

MULTIPLE CHOICE QUESTIONS

1. The line x = a + by represents the regression equation of
   (a) Y on x
   (b) X on y
   (c) Both of above
   (d) None
2. The y = a + bx represents the regression equation of
   (a) y on x
   (b) x on y
   (c) Both
   (d) None
3. The regression coefficients are zero if r is equal to
   (a) 2
   (b) -1
   (c) 1
   (d) 0
4. b_{yx} is called regression coefficient of
   (a) X on Y
   (b) Y on X
   (c) Both
   (d) None
5. b_{xy} is called regression coefficient of
   (a) X on Y
   (b) Y on X
   (c) Both
   (d) None
6. The slope of the regression line of y on x is
   (a) b_{yx}
   (b) b_{xy}
   (c) b_{xx}
   (d) b_{yy}
7. Since the correlation coefficient r is the ____ of the two regression coefficients b_{yx} and b_{xy}
   (a) A.M.
   (b) G.M.
   (c) H.M.
   (d) None

8. r, b_{xy}, b_{yx} all have ___ sign.
   (a) Different
   (b) Same
   (c) Both
   (d) None
9. The two lines of regression meet at:
   (a) $(\overline{X}, \overline{Y})$
   (b) (σ_x, σ_y)
   (c) (σ_x^2, σ_y^2)
   (d) (x, y)
10. Two regression lines coincide when
   (a) r = 0
   (b) r = 2
   (c) r = +1 or -1
   (d) None
11. If x and y satisfy the relationship y = -5 + 7x, the value of r is
   (a) 0
   (b) -1
   (c) +1
   (d) None
12. For the regression equation of Y on X, 2x + 3y + 50 = 0. The value of b_{yx} is
   (a) 2/3
   (b) -2/3
   (c) -3/2
   (d) None
13. In the line $y = 19 - \frac{5x}{2}$, b_{yx} is equal to
   (a) 19/2
   (b) 5/2
   (c) -5/2
   (d) None
14. In the equation $x = \frac{35}{8} - \frac{2y}{5}$, b_{xy} is equal to
   (a) -2/5
   (b) 35/8
   (c) 2/5
   (d) 5/2

15. If two regression lines are: x + 3y = 7 and 2x + 5y = 12 then $\bar{x}$ and $\bar{y}$ are respectively.

(a) 2, 1

(b) 1,2

(c) 2,3

(d) 2,4

16. Out of the two lines of regression given by x + 2y = 4 and 2x + 3y – 5 = 0, the regression line of x on y is:

(a) 2x+3y-5 = 0

(b) x+2y = 4

(c) x+2y = 0

(d) The given lines can't be regression lines.

17. If the coefficient of correlation between two variables is -0.2, then the coefficient of determination is

(a) 0.4

(b) 0.02

(c) 0.04

(d) 0.16

18. If the coefficient of correlation between two variables is -0.9, then the coefficient of determination is

(a) 0.9

(b) 0.81

(c) 0.1

(d) 0.19

19. If r = 0.6 then the coefficient of non-determination is

(a) 0.4

(b) -0.6

(c) 0.36

(d) 0.64

20. Find the coefficient of correlation when its probable error is 0.2 and the number of pairs of item is 9.

(a) 0.505

(b) 0.332

(c) 0.414

(d) 0.316

21. When one regression coefficient is positive, the other would be

(a) Negative

(b) Positive

(c) Zero

(d) None of them

22. If b_{yx} and b_{xy} are negative, r is
    (a) Positive
    (b) Negative
    (c) Zero
    (d) None
23. Where r is zero the regression lines cut each other making
    (a) An angle of 45°
    (b) 60°
    (c) 30°
    (d) 90° i.e., parallel to OX and OY
24. The line $y = 13 - \frac{3x}{2}$ is the regression equation of
    (a) y on x
    (b) x on y
    (c) both
    (d) none
25. $x = \frac{31}{6} - \frac{y}{6}$ is the regression equation of
    (a) y on x
    (b) x on y
    (c) both
    (d) none
26. If $2x + 5y - 9 = 0$ and $3x - y - 5 = 0$ are two regression equation, then find the value of mean of x and men of y.
    (a) 1, 2
    (b) 2, 2
    (c) 2, 1
    (d) 1, 1
27. For the variables x and y, the regression equations are given as $7x - 3y - 18 = 0$ and $4x - y - 11 = 0$. After finding the arithmetic means of x and y, compute the correlation coefficient between x and y.
    (a) 0.5642
    (b) 0.7638
    (c) 0.2135
    (d) 0.9841
28. Given the following data:

Variable	X	Y
Mean:	80	8
Variance:	4	9

Coefficient of correlation = 0.6

What is the most likely value of y when x = 90?

(a) 90

(b) 103

(c) 104

(d) 107

EXERCISE - I

1. Given the bivariate data

X	2	6	4	3	2	3	8	4
Y	7	2	1	1	2	3	2	6

Obtain regression equations taking deviations from 5 in case of X and 4 in case of Y.

2. Past 10 years data on Rainfall and output of wheat in a certain village offered the following results.

Av. Wheat output	25 Qtl
Av. Rainfall	20 Cms
Variance of wheat output	3 Qtl
Variance of Rainfall	5 Cms
r	0.65

Find the most likely wheat output per acre when rainfall is 35 cms.

3. In a partially destroyed laboratory record relating to correlation data, the following results are legible

σ_x^2 = 9, Regression equations 8X – 10 Y + 66 = 0

40X – 18X = 214.

What were (a) the mean values of X and Y, (b) σ_y, c) the co-efficient of correlation between X and Y?

4. If x = 0.85y and y = 0.89x

σ_x = 3, calculate σ_y and r.

5. If two regression coefficients are 0.8 and 1.2 then what would be the value of coefficient of correlation?

6. If two regression coefficients b_{xy} = 0.87 and b_{yx} = 0.49, find 'r'.

7. Given $\overline{X}$ = 40 σ_x = 10 σ_y = 1.5

r_{xy} = 0.9, $\overline{Y}$ = 6

Estimate the value of X if the value of Y is 10.

8. Given

	X Series	Y Series
Mean	18	100
Standard Deviation	14	20

Coefficient of correlation between X and Y is +0.8. Find out

(a) The most probable value of Y if X is 70 and most probable value of X if Y is 90.

(b) If the regression coefficients are 0.8 and 0.6, what would be the value of the coefficient of correlation?

9. Given that the means of X and Y are 65 and 67, their standard deviations are 2.5 and 3.5 respectively and the coefficient of correlation between them is 0.8.

(i) Write down the regression lines

(ii) Obtain the best estimate of X when Y = 70

(iii) Using the estimated value of X as the given value of X, estimate the corresponding value of Y.

10. The correlation coefficient between the variables X and Y is r = 0.60. If $\sigma_x = 1.50$, $\sigma_y = 2.00$, $\overline{X} = 10$, $\overline{Y} = 20$, find the equations of the regression lines (i) Y on X (ii) X on Y.

11. Find out σ_y and r from the following data: 3x = y, 4y = 3x and $\sigma_x = 2$.

12. Given that the regression equation of Y on X and X on Y are respectively Y = X and 4X-Y = 3. Find the correlation coefficient between X and Y.

13. From the following data calculate (i) coefficient of correlation (ii) Standard deviation of Y.

X = 0.854 Y; Y = 0.89X; $\sigma_x = 3$

14. If the two lines of regression are

4X – 5Y + 30 = 0 and 20X – 9Y-107 = 0

Which of these is the line of regression of X on Y. Find r and σ_y when $\sigma_x = 3$.

15. From the following regression equations, calculate $\overline{X}$, $\overline{Y}$ and r

20x – 9Y = 107

4X – 5Y = -33

EXERCISE - II

1. The following table gives the age of car of a certain make and annual maintenance costs. Obtain the regression equations for cost related to age, Estimate maintenance cost of a car whose age is 10 Years.

Age of cars (x) in years	2	4	6	8
Maintenance Cost (in hundred of ₹ Y)	10	20	25	30

2. Find two regression equations from the following data:

X	10	25	34	42	37	35	36	45
Y	56	64	63	58	73	75	82	77

3. Obtain the regression equations for the following:

X	15	27	27	30	34	38	46
Y	120	140	150	170	180	200	250

4. Given the following. Variance X = 36.

12X -85 Y + 99 = 0, 6X + 27 Y=321

Calculate

(i) The average values of X and Y

(ii) SD of Y

(iii) Regression coefficient.

5. Obtain the line of regression of Y on X for the following data:

Age (Yrs) X	66	38	56	42	72	36	63	47	55	45
Blood Pressure Y	145	124	147	125	160	118	149	128	150	124

Estimate the blood pressure of a man whose age is 50 years.

6. In trying to evaluate the effectiveness in its advertising compaign. A firm compiled the following information.

Year	1998	1999	2000	2001	2002	2003	2004	2005
Adv. Expenditure ('000) ₹	12	15	15	23	24	38	42	48
Sales (Lakhs ₹)	5.0	5.6	5.8	7.0	7.2	8.8	9.2	9.5

Calculate the regression equation of sales on advertising expenditure. Estimate the probable sales when advertisement expenditure is ₹60 thousand.

7. Find out the likely output corresponding to a rainfall of 40 cms from the following data

	Rainfall (in cms)	Output (in quaintals)
Average	30	50
S.D	5	10

r = 0.8

8. Given the following data, calculate:

(i) The probable value of Y when X = 12

(ii) The probable value of X when Y = 30

	X	Y
Mean	27.6	14.8
S.D.	40	20
r = 0.8		

9. Given the following values estimate the yield of wheat when the rainfall is 15.5 cm

	Mean	S.D
Yield of wheat (kg/unit area)	10.7	8.1
(Annual Rainfall (cm)	20.4	5.0

Coefficient of correlation between yield and rainfall = + 0.52

10. Two random variables have the least square regression lines with equations 3x + 2y - 26 = 0 and 6x + y – 31 = 0. Find the mean value and the correlation coefficient between X and Y.

11. Estimate (a) the shall of advertising expenditure of ₹100 lakhs and (b) the advertisement expenditure for sales of ₹47 crores from the data given below:

Sales (₹ crores)	14	16	18	20	24	30	32
Adv. Exp. (₹ lakhs)	52	62	65	70	76	80	78

Given the means of two variables, X and Y are 68 and 150, their standard deviations are 2.5 and 20 and correlation between. Then is 0.6, write down the regression equations of X on Y.

13. The following data show the years of services (x) and average salary per month (y) with respect to 7 persons working in a college. Fit a strength line regression even out of it on x.

S.No.	1	2	3	4	5	6	7
x (years)	1	3	7	6	2	4	5
y (₹'000)	5	7	7	8	6	4	3

MCQ Answers:

1	b
2	a
3	d
4	b
5	a
6	a
7	b
8	b
9	a
10	c
11	c
12	b
13	c
14	a
15	b
16	a
17	c
18	b
19	d
20	b
21	b
22	b
23	d
24	e
25	b
26	c
27	b
28	d

Study Note - 6

PROBABILITY

This Study Note includes

6.1 GENERAL CONCEPT

The concept of probability is difficult to define in precise terms. In ordinary language, the word probable means likely or chance. The probability theory is an important branch of mathematics. Generally the word, probability, is used to denote the happening of a certain event, and the likelihood of the. occurrence of that event, based on past experiences. By looking at the clear sky, one will say that there will not be any rain today. On the other hand, by looking at the cloudy sky or overcast sky, one will say that there will be rain today. In the earlier sentence, we aim that there will not be rain and in the latter we expect rain. On the other hand a mathematician says that the probability of rain is 0 in the first case and that the probability of rain is 1 in the second case. In between 0 and 1, there are fractions denoting the chance of the event occurring.

If a coin is tossed, the coin falls dawn. The coin has two sides ; head and tail. On tossing a coin, the coin may fall down either with the head up or tail up. A coin, on reaching the ground, will not stand on its edge or rather, we assume ; so the probability of the coin coming down is 1. The probability of the head coming up is 50% and the tail coming up is 50% ; in other words we can say the probability of the head or the tail coming up is $\frac{1}{2}$, $\frac{1}{2}$ ince 'head' and 'tail' share equal chances. The probability that it will come down head or tail is unity.

A brief, History

The first development of theory of probability came from the gamblers. Historically, this theory originated in the 17th century. Prior to this, an Indian Mathematician, Jerome Carbon (1501 - 1579) was the first man to write a book title "Book on Games of Chance" in 1663. Through this, the gamblers could minimise their risks and safeguard from cheatings. Chevalier de mere, a notable French gambler became attached to the problems of gambling and Approach the French mathematicians Balise, Bascal (1623—1662) and Pietre/de F ,mat (1601 — 1665). These two mathematicians developed the theory of probability. Subsequently, many authors like James Bernoulli (1654-1705), Laplace (1749-1827), De moivre (1667-1754), Thomas Bayes (1702—1761) etc., were inspired to develop the theory of Probability.

Today the theory of probability has been developed to a great extent. It is applied in all the disciplines. It is extensively used in business, economic problems, etc.

- Meaning

It is difficult to give a clear or generally accepted meaning of probability. In our day-.to-day life, we come across sentences like, for example :

1. "When a coin is tossed, it must fall down".
2. "It is impossible to live without oxygen".
3. "Probably, I win the match".

When we look the above sentences, there is *certainty* in the first example ; *impossibility* in the second example and uncertainty in the third example. The expectation in these examples are certainty, impossibility and uncertainty. In the first example, the certainty is there and mathematicians say that the probability of coming down the coin is 1 (one). In the second example, mathematicians say that the probability of living without oxygen is 0 (zero). In the third example, the happenings depend upon chance or likelihood. Agai >, the event is not certain or, in other words, there is uncertainty abou the happening of the event. In ordinary language, the word probability means uncertainty about happenings. In mathematics or statistics, a numerical measure of uncertainty is provided by the important branch of Statistics, — called theory of probability. Thus we can say, that the theory of probability describes "certainty by 1 (one), "impossibility' by 0 (zero) and uncertainties by the co-efficients which lie between 0 and 1.

To-day, the probability theory has been developed to great extent and there is not even a single discipline where probability theory is not used. It is extensively used in Commerce and Economics. The theory of probability provides a numerical measure of the element of uncertainty.

For example, have a look at some of the business situations, characterised by uncertainty:

A businessman having a choice of investing in two different projects, each having different projects, each having different initial investment. The decision has to be taken on the choice, the outcome of which is contingent upon the level of demand, - INVESTMENT PROBLEM.

When a new product is developed, the problem is to decide whether or not to introduce the product in addition to the existing Product-Mix.

The decision-maker may not be sure about acceptability of the product – problem of introducting a new product.

PROBABILITY

When the probability of an event is absolutely certain, the probability is said to be unity or 1. Every person is certain to die one day hence the probability is equal to unity. Mathematicians will say that the probability of man's death is one. Similarly the probability of man surviving without blood is 0. Thus a probability can be 1 and 0. In between 1 and 0, there are fractions (decimals) denoting the probabilities of all sort of events occurring. Similarly, if we toss a coin, the probability of head up or tail up is unity because it will not stand on its edge. The probability of head coming up is $\frac{1}{2}$ and tail coming up is $\frac{1}{2}$.

6.2 SOME USEFUL TERMS

Before discussing the theory of probability, let us have an understanding of the following terms :

Random Experiment or Trial :

If an experiment or trial can be repeated under the same conditions, any number of times and it is possible to count the total number of outcomes, but individual result i.e. individual outcome is not predictable. Suppose we toss a coin. It is not possible to predict exactly the outcomes. The outcome may be either head up or tail up. Thus an action or an operation which can produce any result or outcome is called a random experiment or a trial.

Example :

1. Tossing a coin is an experiment or trial. When you toss, it falls head up or trail up.
2. When you roll a die, it is an experiment. A die Is a solid cube ; the six faces are marked numerically 1, 2,3,4,5,6 or with dots to denote the numbers. The outcome of the roll may be any particular side of the cube (1 to 6) coming up. This .is a case of chance, a trial.

Event :

Any possible outcome of a random experiment is called an event. Performing an experiment is called trial and outcomes are termed as events.

An event whose occurrence is inevitable when a certain random experiment is performed, is called a sure event or certain event. At the same time, an event can never occur when a certain random experiment is performed is called an impossible event. The events may be simple or composite. An event is called simple if it corresponds to a single possible outcome. For example, in rolling a die, the chance of getting 2 is a simple event. Further in tossing a die, chance of getting event numbers (1, 3, 5) are compound event.

Example :

1. Tossing a coin is a random experiment or trial and getting a head or a trial is an event;
2. Drawing a ball from an urn containing red and white balls is a trial and getting a red or white ball is an event.

Sample space

The set or aggregate of all possible outcomes is known as sample space. For example, when we roll a die, the possible outcomes are 1, 2, 3, 4, 5, and 6 ; one and only one face come upwards. Thus, all the outcomes—1, 2, 3, 4, 5 and 6 are sample space. And each possible outcome or element in a sample space called sample point.

Mutually exclusive events or cases :

Two events are said to be mutually exclusive if the occurrence of one of them excludes the possibility of the occurrence of the other in a single observation. The occurrence of one event prevents the occurrence of the other event. As such, mutually exclusive events are those events, the occurrence of which prevents the possibility of the other to occur. All simple events are mutually exclusive. Thus, if a coin is tossed, either the head can be up or tail can be up; but both cannot be up at the same time.

Similarly, in one throw of a die, an even and odd number cannot come up at the same time. Thus two or more events are considered mutually exclusive if the events cannot occur together.

Equally likely events :

The outcomes are said to be equally likely when one does not occur more often than the others.

That is, two or more events are said to be equally likely if the chance of their happening is equal. Thus, in a throw of a die the coming up of 1, 2, 3, 4, 5 and 6 is equally likely. For example, head and tail are equally likely events in tossing an unbiased coin.

Exhaustive events

The total number of possible outcomes of a random experiment is called exhaustive events. The group of events is exhaustive, as there is no other possible outcome. Thus tossing a coin, the possible outcome are head or tail ; exhaustive events are two. Similarly throwing a die, the outcomes are 1, 2, 3, 4, 5 and 6. In case of two coins, the possible number of outcomes are 4 *i.e.* (2^2), *i.e., HH, HT TH* and TT. In case of 3 coins, the possible outcomes are 2^3=8 and so on. Thus, in a throw of *n*" coin, the exhaustive number of case is 2^n.

Independent Events

A set of events is said to be independent, if the occurrence of any one of them does not, in any way, affect the Occurrence of any other in the set. For instance, when we toss a coin twice, the result of the second toss will in no way be affected by the result of the first toss.

Dependent Events

Two events are said to be dependent, if the occurrence or non-occurrence of one event in any trial affects the probability of the other subsequent trials. If the occurrence of one event affects the happening of the other events, then they are said to be dependent events. For example, the probability of drawing a king from a pack of 52 cards is 4/52, ; the card is not put back ; then the probability of drawing a king again is 3/51. Thus the outcome of the first event affects the outcome of the second event and they are dependent. But if the card is put back, then the probability of drawing a king is 4/52 and is an independent event.

Simple and Compound Events

When a single event take place, the probability of its happening or not happening is known as simple event.

When two or more events take place simultaneously, their occurrence is known as compound event (compound probability) ; for instance, throwing a die.

Complementary Events :

The complement of an events, means non-occurrence *of A* and is denoted by $\bar{A}$. $\bar{A}$ contains those points of the sample space which do not belong to *A.* For instance let there be two events *A* and *B. A* is called the complementary event of *B* and vice verse, if *A* and *B* are mutually exclusive and exhaustive.

Favourable Cases

The number of outcomes which result in the happening of a desired event are called favourable cases to the event. For example, in drawing a card from a pack of cards, the cases favourable to "getting a diamond" are 13 and to "getting an ace of spade" is only one. Take another example, in a single throw of a dice the number of favourable cases of getting an odd number are three -1,3 and 5.

6.3 MEASUREMENT OF PROBABILITY

The origin and development of the theory of probability dates back to the seventeenth century. Ordinarily speaking the probability of an event denotes the likelihood of its happening. A value of the probability is a number ranges between 0 and 1. Different schools of thought have defined the term probability differently. The various schools of thought which have defined probability are discussed briefly.

Classical Approach (Priori Probability)

The classical approach is the oldest method of measuring probabilities and has its origin in gambling games. According to this approach, the probability is the ratio of favourable events to the total number of equally likely events. If we toss a coin we are certain that the head or tail will come up. The probability of the coin coming down is 1, of the head coming up is $\frac{1}{2}$ and of the tail coming up is $\frac{1}{2}$.

It is customary to describe the probability of one event as 7" (success) and of the other event as '*q*' (failure) as there is no third event.

$$P = -\frac{\text{Number of favourable cases}}{\text{Total number of equally likely cases}}$$

If an event can occur in 'a' ways and fail to occur in '*b*' ways and these are equally to occur, then the probability of the event occurring, $\frac{a}{a+b}$ is denoted by *P*. Such probabilities are also known as unitary or

theoretical or mathematical probability. *P* is the probability of the event happening and *q* is the probability of its not happening.

$$P = \frac{a}{a+b} \text{ and } q = \frac{b}{a+b}$$

$$\text{Hence } P + q = \frac{a}{(a+b)} + \frac{b}{(a+b)} = \frac{a+b}{a+b} = 1$$

Therefore

$$P + q = 1. 1 - q. 1 - p = q.$$

Probabilities can be expressed either as ratio, fraction or percentage, such as - or 0.5 or 50%

Tossing of a coin and throwing of a die of this type.

Limitations of Classical Approach:

1. This definition is confined to the problems of games of chance only and cannot explain the problem other than the games of chance.
2. We cannot apply this method, when the total number of cases cannot be calculated.
3. When the outcomes of a random experiment are not equally likely, this method cannot be applied.
4. It is difficult to subdivide the possible outcome of experiment into mutually exclusive, exhaustive and equally likely in most cases.

Example 1:

What is the chance of getting a king in a draw from a pack of 52 cards?

Solution :

The total number of cases that can happen

= 52 (52 cards are there).

Total number of kings are 4 ; hence favourable cases=4 Therefore probability of drawing a king $= \frac{4}{52} = \frac{1}{13}$.

Example 2:

Two coins are tossed simultaneously. What is the probability of getting a head and a tail ?

Solution :

The possible combinations of the two coins turning up with head (*H*) or tail *(T)* are *HH, HT, TH, TT*. The favourable ways are two out of these four possible ways and all these are equally likely to happen.

Hence the probability of getting a head and a tail is $\frac{2}{4} = \frac{1}{2}$.

Example 3 :

One card is drawn at random from a well-shuffled pack of 52 cards. What is the probability that it will be *(a)* a diamond *(b)* a queen ?

Solution :

(a) There are 13 diamond cards in a pack of 52 cards. The number of ways in which a card can be drawn from that pack is 52. The number favourable to the event happening is 13.

Hence probability of drawing a diamond

$$\frac{13}{52}=\frac{1}{4}$$

(b) There are 4 queens in the pack ; and so the number of ways favourable to the event = 4

The probability = $\frac{4}{52}=\frac{1}{13}$

Example 4 :

Two cards are drawn from a pack of cards at random. What is the probability that it will be (a) a diamond and a heart *(b)* a king and a queen (c) two kings ?

Solution :

(a) The number of ways of drawing 2 cards from out of 52 cards

$$= {}^{52}C_2 = \frac{52\times 51}{1\times 2} = 26\times 51$$

The number of ways of drawing a diamond and a heart

=13 x 13

The required probability

$$=\frac{13\times 13}{26\times 51}=\frac{13}{102}$$

(b) The number of ways of drawing a king and a queen = 4 × 4

The required probability

$$=\frac{4\times 4}{26\times 51}=\frac{6}{663}$$

(c) Two kings can be drawn out of 4 kings in ${}^4C_2 = \frac{\underline{|4}}{\underline{|4-2}\,.\,\underline{|2}} = \frac{\underline{|4}}{\underline{|2}\,.\,\underline{|2}} = \frac{\cancel{4}\times 3\times 2\times 1}{\cancel{2}\times 1\times \cancel{2}\times 1} = 6$ ways

The probability of drawing 2 kings

$$=\frac{6}{26\times 51}=\frac{1}{221}$$

Example 5 :

A bag contains 7 red, 12 white and 4 green balls. What is the probability that :

(a) 3 balls drawn are all white and

(b) 3 balls drawn are one of each colour ?

Solution :

(a) Total number of balls

= 7 +12 + 4 = 23

Number of possible ways of drawing 3 out of 12 white

$= {}^{12}C_3$

Total number of possible ways of drawing 3 out of 23 balls

$= {}^{23}C_3$

Therefore, **probability** of drawing 3 white balls $= \frac{{}^{12}C_3}{{}^{23}C_3} = \frac{220}{1771} = 0.1242$

(b) Number of possible ways of drawing 1 out of red — ${}^{7}C_1$ Number of possible ways of drawing 1 out of 12 white = ${}^{12}C_1$ Number of possible ways of drawing 1 out of 4 green - ${}^{4}C_1$

Therefore the probability of drawing balls of different colours

$$= \frac{{}^{7}C_1 \times {}^{12}C_1 \times {}^{4}C_1}{{}^{23}C_3} = \frac{7 \times 12 \times 4}{1771}$$

= 0.1897

Relative Frequency Theory of probability :

Classical approach is useful for solving problems involving game 2 of chances—throwing dice, coins, etc. but if applied to other types of problems it does not provide answers. For instance, if a man jumps from a height of 300 feet, the probability of his survival will, not be 50%, since survival and death are not equally alike.

Similarly, the prices of shares of a Joint Stock Company have three alternatives i.e. the prices may remain constant or prices may go up or prices may go down. Thus, the classical approach fails to answe questions of these type.

If we toss a coin 20 times, the classical probability suggests that we; should have heads ten times. But in practice it may not be so. These empirical approach suggests, that if a coin is tossed a large number or time, say, 1,000 times, we can expect 50% heads and 50% tails. Vor Miscs explained, "If the experiment be repeated a large number of times under essentially identical conditions, the limiting value of the ratio of the number of times the event *A* happens to the total, number of trials.of the experiments as the number of trials increases indefinitely, is called the probability of the occurrence of *A*".

Thus, $P(A) = \lim_{n \to \infty} \frac{m}{n}$

The happening of an event is determined on the basis of past experience or on the basis of relative frequency of success in the past. For instance, if a machine produces 10% unacceptable articles of the total output. On the basis of such experience or experiments, we may arrive at that (i) the relative frequency obtained on the basis of past experience can be shown to come very close to the classical probability. For example, as said earlier, a coin is tossed for 6 times, we may not get exactly 3 heads and 3 tails. But, the coin is tossed for larger number of times, say 10,000 times, we can expect heads and tails very close to 50% (ii) There are certain laws, according to which the 'occurrence' or 'non-occurrence of the events take place. Posterior probabilities, also called Empirical Probabilities are based on experiences of the past and on experiments conducted. Thus, relative frequency can be termed as a measure of probability and it is calculated on the basis of empirical or statistical findings. For instance if a machine produces 100 articles in the past, 2 particles were found to be defective, then the probability of the defective articles is 2/100 or 2%.

Limitations of Relative Frequency Theory of Probability:

1. The experimental conditions may not remain essentially homogeneous and identical in a large number of repetitions of the experiment.
2. The relative frequency —, may not attain a unique value no matter however large N may be.
3. Probability **P(A) defined** can never be obtained in practice. We can **only attempt** at a close estimate of P(A) by making N sufficiently large.

Personalistic View of Probability :

The personalistic theory of probability is also known as subjective, theory of probability. This theory is commonly used in business decision making! Here, the decisions reflects the personality of the decision maker. Persons may arrive at different probability assignment because of differences in value or experience etc. That is the personality of the decision maker is reflected in a final decision. For instance, a student would top the list in B.Com. Examination this year. A subjectivist would assign a weight between zero and one to this event according to his belief for its possible occurrence.

Axiomatic Approach of Probability :

This approach was introduced by a Russian Mathematician A.N. Kolmogorov in 1933. This approach is mathematical in character and is based on set theory. No precise definition has been given but some concepts are laid down and certain properties or postulates commonly known as axioms are defined. The probability calculations are based on the axioms. The axiomatic probability includes the concept of both classical and empirical definitions of probability. The probability of an event ranges from 0 to 1. The probability of the entire space is one and for mutually exclusive events (disjoint) the probability of the happening of either *A* or *B* denoted by *P (A*UB) (read as *A* union *B)* shall be

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

For events of simultaneous occurrence or the probabilities of both *A* and *B* happening denoted by $P(A \cap B)$ read as probability *A* intersection *B* shall be :

$P(A \cap B)=P(A)P(B)$, *when A & B are Independent.*

Example 6:

An urn contains 8 white and 3 red balls. If two balls are drawn at random, find the probability that *(a)* both are white, *(b)* both are red and (c) one is of each colour.

Solution :

Total number of balls in the urn = 8 + 3 =11

Two balls can be drawn out of 11 balls in $^{11}C_2$ ways.

Exhaustive number of cases = $^{11}C_2 = \frac{11 \times 10}{2} = 55 \cdot$

(a) Two white balls to be drawn out of 8 white, can be done in $^8C_2 = \frac{8 \times 7}{2} = 28$ ways.

 The probability that both are white $= \frac{28}{55}$

(b) Two red balls to be drawn out of 3 red balls can be done in $^3C_2$ = 3 ways.

 Hence, the probability that both are red $= \frac{3}{55}$

(c) The number of favourable cases for drawing one white ball and one red ball is

$^8C_1 x ^3C_1 = 8 \times 3 = 24$.

Therefore, the probability (one red and one white) $= \frac{24}{55}$

Example 7 :

Two cards are drawn from a pack of cards at random. What is the probability that it will be *(a)* a diamond and a heart, *(b)* a king and a queen (c) two kings ?

Solution :

(a) The number of ways of drawing 2 cards from out of 52 cards

$= {}^{52}C_2 = \frac{52 \times 51}{1 \times 2} = 26 \times 51$

The number of ways of drawing a diamond and a heart

= 13 x13

The required probability

$= \frac{13 \times 13}{26 \times 51} = \frac{13}{102}$

(b) The number of ways of drawing a king and a queen

= 4 X 4

The required probability

$= \frac{4 \times 4}{26 \times 51} = \frac{8}{663}$

(c) Two kings can be drawn out of 4 kings in $^4C_2 = 6$ ways.

The probability of drawing 2 kings

$= \frac{6}{26 \times 51} = \frac{1}{221}$

Example 8 :

Two dice are thrown. Find the probability that :

(a) the total of the numbers on the dice is 8,

(b) the first die shows 6,

(c) the total of the numbers on the dice is greater than 8,

(d) the total of the numbers on the dice is 13,

(e) both the dice show the same number,

(f) the sum of the numbers shown by the dice is less than 5,

(g) the sum of the numbers shown by the dice is exactly 6.

Solution :

When the dice are thrown, the possible combinations are $6^2 = 36$, which are listed below :

			First die			
Second die	1	2	3	4	5	6
1	(1,1)	(1,2)	(1,3)	(1,4)	(1,5)	(1,6)
2	(2, 1)	(2, 2)	(2, 3)	(2, 4)	(2, 5)	(2, 6)
3	(3,1)	(3, 2)	(3, 3)	(3,4)	(3, 5)	(3, 6)
4	(4,1)	(4, 2)	(4, 3)	(4, 4)	(4, 5)	(4,6)
5	(5,1)	(5, 2)	(5,3)	(5, 4)	(5, 6)	(5,7)
6	(6, 1)	(6, 2)	(6, 3)	(6, 4)	(6, 5)	(6, 6)

(a) The cases which have a total of 8 are

(2, 6), (3, 5), (4, 4), (5, 3), (6, 2) = 5

Therefore, the probability $= \frac{5}{36}$

(b) The first die shows 6 in the following cases (6, 1), (6, 2), (6, 3), (6, 4), (6, 5), (6, 6) = 6

Therefore, the probability $= \frac{6}{36} = \frac{1}{6}$

(c) The cases which give a total of more than 8 are

(3, 6) (4, 5), (4, 6), (5, 4), (5, 5), (5, 6), (6, 3), (6, 4), (6, 5), (6, 6) = 10.

Therefore, the probability of getting more than 8

$= \frac{10}{36} = \frac{5}{18}$

(d) The probability of getting 13 or more than 13 is impossible.

Therefore, the probability = 0

(e) The favourable cases are :

(1, 1), (2, 2), (3, 3), (4, 4), (5, 5), (6, 6) ,.

Therefore the probability $= \frac{6}{36} = \frac{1}{6}$

(J) The cases which have less than 5 are 6 and they are (1, 1), (1, 2), (1, 3), (2, 1), (2, 2) and (3, 1).

Therefore, the probability $= \frac{6}{36} = \frac{1}{6}$

(g) The cases which have exactly 6 are :5

(5, 1), (4, 2),(3, 3), (2, 4), (1, 5)

Therefore, the probability is $= \frac{5}{36}$

Example 9 :

Tickets are numbered from 1 to 100. They are well shuffled and a ticket is drawn at random. What is the probability that the drawn ticket has :

(a) an even number,

(b) a number 5 or a multiple of 5,

(c) a number which is greater than 75,

(d) a number which is a square ?

Solution :

(a) The total number of exhaustive, mutually exclusive and equal cases is 100. There are 50 even numbered tickets.

Therefore, favourable cases to the event is 50.

Therefore, the probability $=\frac{50}{100}=\frac{1}{2}$

(b) Suppose A denotes the number of happenings that the drawn ticket has a number 5 or a multiple of 5. These are 20 cases i. e., 5, 10, 15, 20,...100.

Therefore, $P(A)=\frac{20}{100}=\frac{1}{5}$

(c) There are 25 cases, which have a number greater than 75. Say A will denote it.

Therefore, $P(A)=\frac{25}{100}=\frac{1}{4}$

(d) There are 10 favourable cases which give squares between 1 and 100 i.e., 1, 4, 9, 16, 25, 36, 49, 64, 81, 100.

Therefore, $P(A)=\frac{10}{100}=\frac{1}{10}$

Example 10 :

Four cards are drawn from a pack of 52 cards without replacement. What is the probability that they are all of different suits ?

Solution :

The required probability would be :

$$1\times\frac{39}{51}\times\frac{26}{50}\times\frac{13}{49}=\frac{2,197}{20,825}$$

6.4 THEOREMS OF PROBABILITY

We have studied what probability is and how it can be measured. We dealt with simple problems. Now we shall consider some of the laws of probability to tackle complex situation. There are two important theorems, viz., (1) the Addition Theorem and (2) the Multiplication Theorem.

Addition Theorem :

The simplest and most important **rule used in** the **calculation is** the **addition rules, it** states, "If two events are mutually **exclusive, then** the probability of the occurrence of either *A* or *B* is the **sum** of **the** probabilities of *A* and *B.* Thus,

$P(A \text{ or } B) = P(A)+P(B)$

Example 11 :

A bag **contains** 4 white, 3 black and 5 red balls. What is the probability of getting a white or a red ball at random in a single draw ?

Solution :

The probability of getting a white ball = $\frac{4}{12}$

The probability of getting a red ball = $\frac{5}{12}$

The probability of a white or a red $=\frac{4}{12}+\frac{5}{12}=\frac{9}{12}$

or $\frac{9}{12}\times 100=75\%$

When events are not mutually exclusive *i.*

The addition theorem studied above is not applicable when the events are not mutually exclusive. In such cases where the events are not mutually exclusive, the probability is :

$P(A \text{ or } B) = P(A) + P(B) - P(A \text{ and } B)$

Example 12 :

Two students A" and V work independently On a problem. The probability that A'will solve it is 3/4 and the probability that *Y* will solve it is 2/3. What is the probability that the problem will be solved ?

Solution :

$P(A \text{ or } B) = P(A)+P(B)-P(A \text{ and } B)$

The probability that A will solve the problem is = 3/4

The probability that *Y* will solve the problem is—2/3

The events are not mutually exclusive as both of them may solve the problem.

Therefore, the probability $=\frac{3}{4}+\frac{2}{3}-\left(\frac{3}{4}\times\frac{2}{3}\right)$

$$=\frac{17}{12}-\frac{6}{12}=\frac{11}{12}$$

Alternatively:

The probability that X will solve it and Y fail to solve it = 3/4 x 1/3 = 3/12

$\therefore$ Probability that the problem will be solved $=\frac{2}{3}+\frac{3}{12}=\frac{11}{12}$

Alternatively

The probability that X will fail to solve and will Y solve it

= 1/4 x 2/3 = 2/12

$\therefore$ Probability that the problem will be solved $=\frac{3}{4}+\frac{2}{12}=\frac{9+2}{12}=\frac{11}{12}$

Alternatively :

The probability that neither X nor Y will solve it $\frac{1}{4}\times\frac{1}{3}=\frac{1}{12}$

Hence, the probability that the problem will be solved

$=1-\frac{1}{12}=\frac{11}{12}$

Multiplication

When it is desired to estimate the chances of the happening of successive events, the separate probabilities of these successive events are multiplied. If two events A and B are independent, then the probability that both will occur is equal to the product of the respective probabilities. We find the probability of the happening of two or more events in succession.

Symbolically :

$$P(A \text{ and } B)=P(A)xP(B)$$

Example 13 :

In two tosses of a fair coin, what are the chances of head in both ?

Solution :

Probability of head in first toss = 1/2

Probability of head in the second toss = 1/2

Probability of head in both tosses = 1/2x1/2 -1/4

Example 14 :

The probability that X and Y will be alive ten years hence is 0.5 and 0.8 respectively. What is the probability that both of them will be alive ten years hence ?

Solution :

Probability of X being alive ten years hence = 0.5

Probability of Y being alive ten years hence = 0.8

Probability of X and Y both being alive ten years hence =.5x.8=0.4

When events are dependent :

If the events are dependent, the probability is conditional. Two events *A* and *B* are dependent ; *B* occurs only when *A* is known to have occurred.

P (B | A) means the probability of B given that A has occurred.

$$P(B \mid A) = \frac{P(AB)}{P(A)};\ P\left(\frac{A}{B}\right) = \frac{P(AB)}{P(B)}$$

Example 15 :

A man want to marry a girl haying qualities: White complexion the probability of getting such girl is 1 in 20. Handsome dowry - the probabihty of getting is 1 in 50. Westernised style - the probability is 1 in 100.

Find out the probability of his getting married to such a girl, who has all the three qualities.

Solution :

The probability of a girl with white complexion $= \frac{1}{20}$ or 0.05. The probability of a girl with handsome dowry or 0.02. The probabihty of a girl with westernised style $= \frac{1}{100}$ or 0.01. Since the events are independent, the probabihty of simultaneous occurrence of all three qualities =

$$\frac{1}{20} \times \frac{1}{50} \times \frac{1}{100} = 0.05 \times 0.02 \times 0.01 = 0.00001$$

Example 16 :

A university has to select an examiner from a list of 50 persons, 20 of them women and 30 men, 10 of them knowing Hindi and 40 not. 15 of them being teachers and the remaining 35 not. What is the probability of the University selecting a Hindi-knowing women teacher ?

Solution :

Probability of selecting a women $= \frac{20}{50}$

Probability of selecting a teacher $= \frac{15}{50}$

Probability of selecting a Hindi-knowing candidate $= \frac{10}{50}$

Since the events are independent the probabihty of the University selecting a Hindi-knowing woman teacher is :

$$\frac{20}{50} \times \frac{15}{50} \times \frac{10}{50} = \frac{3}{125} \text{ or } 0.024.$$

Example 17 :

A ball is drawn at random from a box containing 6 red balls, 4 white

balls and 5 blue balls. Determine the probability that it is :

(i) Red (ii) white, (iii) Blue, (iv) Not Red and (v) Red or White.

Solution :

$$P = \frac{\text{No. of favourable cases}}{\text{Total No. of equally likely cases}}$$

(i) Probability of Red $= \frac{6}{15}$ or 0.40

(ii) Probability of white $= \frac{4}{15}$ or 0.267

(iii) Probability of Blue $= \frac{5}{15}$ A or 0.333

(iv) Probability of not Red $= \frac{9}{15}$ or 0.60

(v) Probability of Red and White $= \frac{10}{15}$ or 0.667

6.5 BAYES' THEOREM

This theorem is associated with the name of Reverend **Thomas Bayes**. It is also known as the inverse probability. Probabilities can be revised when new information pertaining to a random experiment is obtained. One of the important applications of the conditional probability is in the computation of unknown probabilities, on.the basis of the information supplied by the experiment or past records. That is, the applications of the results of probability theory involves estimating unknown probabilities and making decisions on the basis of **new** sample information. This concept is referred to as Bayes' Theorem. Quite often the businessman has the extra information on a particular event, either through a personal belief or from the past history of the events. Revision of probability arises from a need to make better use of experimental information. Probabilities assigned on the basis of personal experience, before observing the outcomes of the experiment are called prior probabilities. For example, probabilities assigned to past sales records, to past number of defectives produced by a machine, are examples of prior probabilities. When the probabilities are revised with the use of Bayes' rule, they are called posterior probabilities. Bayes' theorem is useful in solving practical business problems in the light of additional information. Thus popularity of the theorem has been mainly because of its usefulness in revising a set of old probability (Prior Probability) in the light of additional information made available and to derive a set of new probabilily (i.e. Posterior Probability)

Bayes' Theorem : An event A can occurre only if on one of the mutually exclusive and exhaustive set of events B_1, B_2, ... B_n occurs. Suppse that the unconditional probabilities

$$P(B_1), P(B_2), ... P(B_n)$$

and the conditional probabilities

$$P(A/B_1), P(A/B_2), ... P(A/B_n)$$

are known. Then the conditional probability $P(B_i/A)$ of a specific event B_i, when A is stated to have actually accquared, is given by

$$P(B_i / A) = \frac{P(B_i).P(A / B_i)}{\sum_{i=i}^{n} P(B_i) . P(A / B_i)}$$

This is known as Bayes' Theorem.

The following example illustrate the application of Baye's Theorem.

The above calculation can be verified as follows :

If 1,000 scooters were produced by the two plants in a particular week, the number of scooters produced by Plant I & Plant II are respectively :

I,000 x 80% = 800 scooters

1,000 x 20% = 200 scooters

The number of standard quality scooters produced by Plant I :

800 x 85/100 = 680 scooters

The number of standard quality scooters produced by Plant II :

200 x 65/100 = 130 Scooters.

The probability that a standard quality scooter was produced by Plant I is :

$$=\frac{680}{680+130}=\frac{680}{810}=\frac{68}{81}$$

The probability that a standard quality scooter was produced by Plant II is :

$$=\frac{130}{680+130}=\frac{130}{810}=\frac{13}{81}$$

The same process *i.e.* revision can be repeated if more information is made available. Thus it is a good theorem in improving the quality of probability in decision making under uncertainty.

$$\frac{\text{Event}}{B \cap A_1}\ \frac{\text{Pr obability}}{0.80 \times 0.85}=0.68$$

$$B \cap A_1\ 0.20 \times 0.65=0.13$$

Similarly

$$\therefore P(B \cap A_1)=0.20 \times 0.65=0.13$$

Example 18 :

You note that your officer is happy on 60% of your calls, so you assign a probability of his being happy on your visit as 0.6 or 6/10. You have noticed also that if he is happy, he accedes to your request with a probability of 0.4 or 4/10 whereas if he is not happy, he acedes to the request with a probability of 0.1 or D or $\frac{1}{10}$. You call one day, and he accedes to your request. What is the probability of his being happy ?

Solution :

Let- *H* be the Hypothesis that the officer is happy and the Hypothesis that the officer is not happy

$$P(H)=\frac{6}{10} \qquad P(H)=\frac{4}{10}$$

Let *A* be the event that he accedes to request

$$P(A/H)=\frac{4}{10},\ P(A/\overline{H})=\frac{1}{10}$$

To find P(H/A), according to Baye's Theorem,

$$P(H/A) = \frac{P(H) \times P(A/H)}{P(H) \times P(A/H) + P(\bar{H}) \times P\left(\frac{A}{\bar{H}}\right)} = \frac{\frac{6}{10} \times \frac{4}{10}}{\frac{6}{10} \cdot \frac{4}{10} + \frac{4}{10} \cdot \frac{1}{10}}$$

$$= \frac{\frac{24}{100}}{\frac{24}{100} + \frac{4}{100}} = \frac{24}{28} = \frac{6}{7} = 0.857$$

Example 19 :

A company has two plants to manufacture scooters. Plant I manufactures 80% of the scooters and plant II manufactures 20%. At Plant I, 85 out of 100 scooters are rated standard quality or better. At Plant II, only 65 out of 100 scooters are rated standard quality or better. What is the probability that the scooter selected at random came from Plant I if it is known that the scooter is of standard quality ?

What is the probability that the scooter came from Plant II if it is known that the scooter is of standard quality.

Solution :

Let A_1 be the event of drawing a scooter produced by Plant I and A_2 be the event of drawing a scooter produced by Plant II. *B* be the event of drawing a standard quality scooter produced by either Plant I or Plant II

Then, from the first information :

$$P(A_1) = \frac{80}{100} = 80\% = 0.80$$

$$P(A_2) = \frac{20}{100} = 20\% = 0.2$$

From the additional information :

$$P(B \mid A_1) = \frac{85}{100} = 85\%; \quad P(B \mid A_2) = 65\%$$

The required values are computed in the following table :

Event	**Prior**	**Conditional**	**Joint**	**Posterior Probability**
	Probability(2)	Probabilitv(3)	Probability (4)	(Revised)(5) (=4^P(fl))
A_1	0.80	0.85	0.68	$\frac{0.68}{0.81} = \frac{68}{81}$
A_2	0.20	0.65	0.13	$\frac{0.13}{0.81} = \frac{13}{81}$
	1		P(B) = 0.81	1

From the first information we may say that the standard scooter is drawn from Plant I since $P(A_1) = 80\%$ which is greater than $P(A_2) = 20\%$,

From the additional information i.e. at Plant I, 85 out of 100 and 65 out of 100 are rated standard quality, we can give better answer, Thus we may conclude that the standard quality of scooter is more likely drawn from the output by Plant I.

Example 20 :

Box I contains three defective and seven non-defective balls, and Box II contains one defective and nine non-defective balls. We select a box at random and then draw one ball at random from the box.

(a) What is the probability of drawing a non-defective ball ?

(b) What is the probability of drawing a defective ball ?

(c) What is the probability that box I was chosen, given a defective ball is drawn ?

Solution :

$P(B_1)$ or Probability that Box I is chosen $= \frac{1}{2}$ $P(B_1)$ or

Probability that Box I is chosen $= \frac{1}{2}$.

$P(B_2)$ or Probability that Box II is chosen $= \frac{1}{2}$

P(D) - Probability that a defective Ball is drawn P(ND) = Probability that a non-defective Ball is drawn

Joint Probability

$\frac{1}{2}\times\frac{3}{10}=\frac{3}{20}$ $\quad$ $\frac{1}{2}\times\frac{1}{10}=\frac{1}{20}$

$\frac{1}{2}\times\frac{7}{10}=\frac{7}{20}$ $\quad$ $\frac{1}{2}\times\frac{9}{10}=\frac{9}{20}$

(a) *P(ND)=P* (Box I and non-defective) +P(Box II non-defective)

$$=\left(\frac{1}{2}\times\frac{7}{10}\right)+\left(\frac{1}{2}\times\frac{9}{10}\right)=\frac{16}{20}$$

(b) *P(D)* = P (Box I and defective) +/P (Box II and defective)

$$=\left(\frac{1}{2}\times\frac{3}{10}\right)+\left(\frac{1}{2}\times\frac{1}{10}\right)=\frac{4}{20}$$

(c) Bayes' Theorem :

$$P(B_1 / D)=\frac{P(B_1 \text{ and} D)}{P(D)}=\frac{3/20}{4/20}=\frac{3}{4}$$

$P(B_1)$ and $P(B_2)$ are called prior probabilities and *P(B/D)* and $P(B_2/D)$ are called posterior probabilities. The above information is summarised in the following table :

Event	Prior Probability	Conditional Probability	Joint Probability	Posterior Probability
B_1	$\frac{1}{2}$	3/10	3/20	3/4
B_2	$\frac{1}{2}$	1/10	1/20	1/4
	1		4/10	1

6.6 ODDS

We must know the concept of odds. The word odd is frequently used in statistics. Odds relate the chances in favour of an event to the chances against it. For instance, the odds are 2 : 1 that *A* will get a job, means that there are 2 chances that he will get the job and 1 chance against his getting the job. This can also be converted into probability as getting the job = 2/3. Therefore, if the odds are *a* : *b* in favour of an events, then *P(A) - a/(a+b)*. Further, it may be noted that the odds are *a : b* in favour of an event is the same as to say that the odds are *b : a* against the event.

If the probability of an event is *p,* then the odds in favour of its occurrence are P to (*1-p*) and the odds against its occurrence are *l-p to p*.

Example 21 :

Suppose it is 11 to 5 against a person who is now 38 years of age living till he is 73 and 5 to 3 against *B* who is 43 Living till he is 78, find the chance that at least one of these persons will be alive 35 years hence.

Solution :

The probability that *A* will die within 35 years $=\frac{1}{16}$

The probability that *B* will die within 35 years $=\frac{5}{8}$

The probability that both of them will die within 35 years

$$=\frac{11}{16}\times\frac{5}{8}=\frac{55}{128}$$

The probability that both of them will not die i.e.

atleast one of them will be alive

$$=1-\frac{55}{128}=\frac{73}{128}$$

Example 22 :

Two cards are drawn at random from a well-shuffled pack of 52 cards. What is the probability that :

(a) both are aces,

(b) both are red,

(c) at least one is an ace ?

Solution :

(a) Let *A* indicate the event of drawing 2 aces.

$$P\left(\frac{A}{A}\right)=P(B)\times P\left(\frac{A}{A}\right)$$

P(A) : drawing of an ace first

$P\left(\frac{A}{A}\right)$: conditional probability of an ace at the second draw, given that the first was an ace.

Therefore,

$$P\left(\frac{A}{A}\right)=\frac{4}{52}P\left(\frac{A}{A}\right)=\frac{3}{51}$$

$$P\left(\frac{A}{A}\right)=\frac{4}{52}\times\frac{3}{51}=\frac{3}{51}=\frac{12}{2652}=\frac{1}{221}$$

(b) Let R indicate the event of drawing 2 red cards

$$P\left(\frac{R}{R}\right)=P(R)\times P\left(\frac{R}{R}\right)$$

$$=\frac{26}{52}\times\frac{25}{51}=\frac{650}{2652}=\frac{25}{102}$$

(c) Let *E* indicate the event of drawing an ace. Then the probability that at least an ace is drawn is denoted by *P(E)*. Probability of not drawing an ace :

$$P\left(\frac{E}{E}\right)=P(E)\times P\left(\frac{E}{E}\right)$$

$$=\frac{48}{52}\times\frac{47}{51}=\frac{2256}{2652}=\left(\frac{188}{221}\right)$$

Therefor, probability of drawing at least on ace

$$=1-\frac{188}{221}=\frac{33}{221}$$

Example 23 :

The odds in favour of a certain event are 2 to 5 and the odds against another event independent of the former are 5 to 6. Find the chance that one at least of the events will happen.

Solution :

The chance that the 1st event happens and the 2nd one does not happen

$$=\frac{2}{7}\times\frac{5}{11}=\frac{10}{77}$$

The chance that the 1st event does not happen and the 2nd happens.

$$=\frac{5}{7}\times\frac{6}{11}=\frac{30}{77}$$

The chance that both the events happen

$$=\frac{2}{7}\times\frac{6}{11}=\frac{12}{77}$$

The chance that one at least of the events will happen,

$$=\frac{10}{77}+\frac{30}{77}+\frac{12}{77}=\frac{52}{77}$$

Alternatively :

The first event does not happen

$$=1-\frac{2}{7}=\frac{5}{7}$$

The second event does **not** happen

$$=1-\frac{6}{11}=\frac{5}{11}$$

The chance that both do not happen

$$=\frac{5}{7}\times\frac{5}{11}=\frac{25}{77}$$

$$=1-\frac{25}{77}=\frac{52}{77}$$

The chance that one at least will happen

Example 24 :

What is the chance that a leap year, selected at random will contain 53 Sundays ?

Solution:

As a leap year consist of 366 days it contains 52 complete weeks and two more days.

The two consecutive days make the following combinations :

(a) Monday and Tuesday

(b) Tuesday and Wednesday

(c) Wednesday and Thursday

(d) Thursday and Friday

(e) Friday and Saturday

(f) Saturday and Sunday, and

(g) Sunday and Monday

If *(f)* or (g) occur, then the year consists of 53 Sundays.

Therefore the number of favourable cases = 2

Total number of cases = 7

The probability = $\frac{2}{7}$

Example 25 :

Find the chance of throwing more than 15 in one throw with 3 dice.

Solution :

Total number of cases = 6 x 6 X 6 = 216

Throwing more than 15 means getting 16, 17 or 18.

Possible ways of throwing 16 are (6, 6, 4), (6, 5, 5) (6, 4, 6), (5, 5, 6) (5, 6, 5) **and** (4, 6, 6).

Number of favourable cases = 6.

The probability of getting 16 with three dice = $\frac{6}{216}$

Possible ways of throwing 17 are (6, 6, 5), (6. 5, 6), ar 6, 6

3 The probability of throwing 17 with three dice = $\frac{6}{216}$

There is only one way of throwing 18 with three dice namely (6,6,6).

The probability of throwing 18 with three dice = $\frac{6}{216}$

The three cases are mutually exclusive.

Therefore the probability of throwing more than 15.

$$=\frac{6}{216}+\frac{3}{216}+\frac{1}{216}=\frac{10}{216}=\frac{5}{108}$$

Example 26 :

A problem in statistics is given to three students *A, B, C* whose chances of **solving** it are 1/2, 1/3, 1/4 respectively. What is the probability **that the** problem will be solved ?

Solution:

The probability that A fails to soive the problem $=1-\frac{1}{2}=\frac{1}{2}$

The probability that *B* fails to solve the problem $=1-\frac{1}{3}=\frac{2}{3}$

The probability that C fails to solve the problem $=1-\frac{1}{4}=\frac{3}{4}$

The probability that the problem is not solved by *A, B* and C $=\frac{1}{2}\times\frac{2}{3}\times\frac{3}{4}=\frac{1}{4}$

Therefore, the probability that the problem is solved $=1-\frac{1}{4}=\frac{3}{4}$

Example 27 :

Assuming that half the population is vegetarian so that the chance of an individual being a vegetation is $\frac{1}{2}$ and assuming that 100 investigators can take a sample of 10 individuals to see whether they are vegetarian, how many investigators would you expect to report that three people or less were vegetarians?

Solution :

We have p = $\frac{1}{2}$, q =1-$\frac{1}{2}$=, n = 10, N =100

Number of investigators getting 3 or less vegetarians *i.e.* 0,1, 2, 3 Vegetarians

$$=100\left(\frac{1}{2}\right)^{10}+100\times\ {}^{10}C_1\left(\frac{1}{2}\right)^{1}+\left(\frac{1}{2}\right)^{9}+100\times{}^{10}C_2\left(\frac{1}{2}\right)^{2}\left(\frac{1}{2}\right)^{8}$$

$$+100\times{}^{10}C_3\left(\frac{1}{2}\right)^{3}\left(\frac{1}{2}\right)^{7}$$

$$=100\times\left(\frac{1}{2}\right)^{10}+100\times10\times\left(\frac{1}{2}\right)^{10}+100\times45\times\left(\frac{1}{2}\right)^{10}$$

$$+100\times120\times\left(\frac{1}{2}\right)^{10}$$

$$=100\times\left(\frac{1}{2}\right)^{10}(1+10+45+120)$$

$$=100\times\frac{1}{1024}\times176=17(\text{approximately})$$

Example 28 :

An ordinary die is tossed twice and the difference between the number of spots turned up is noted. Find the probability of a difference of 3.

Solution :

The sample space consists of 36 values.

The event space has the following 6 cases : (1, 4), (2, 5), (3, 6), (4, 1), (5, 2), (6, 3)

The required probability = $\frac{6}{36}$

Example 29 :

From a pack of 52 cards, two cards are drawn at random ; find the chance that one is a knave and the other a queen.

Solution :

Sample space = ${}^{52}C_2$

Event space = ${}^{4}C_1 \times {}^{4}C_1$

(as there are 4 queens and 4 knaves in the pack)

Required Probability $= \frac{{}^{4}C_1 \times {}^{4}C_1}{{}^{52}C_2} = \frac{8}{663}$

Example 30 :

A bag contains 7 red balls and 5 white balls. 4 balls are drawn at random. What is the probability that (i) all of them are red ; (ii) two of them are red and two white ?

Solution :

(i) Favourable cases ${}^{7}C_4$, Exhaustive cases ${}^{12}C_4$

Probability $= \frac{{}^{7}C_4}{{}^{12}C_4} = \frac{105}{495} = \frac{7}{33}$

(ii) Favourable cases = ${}^{7}C_2 \times {}^{5}C_2$

Exhaustive cases = ${}^{12}C_4$

Probability $= \frac{{}^{7}C_2 \times {}^{5}C_2}{{}^{12}C_4} = \frac{12 \times 10}{495} = \frac{14}{33}$

Example 31 :

A petrol pump proprietor sells on an average ₹ 80,000 worth of petrol on rainy days and an average of ₹ 95,000 on clear days. Statistics from the Metereological Department show that the probability is 0.76 for clear weather and 0.24 for rainy weather on coming Monday. Find the expected value of petrol sale on coming Monday.

Solution :

X_1 = ₹ 80,000; $P_1 = 0.24$

X_2 = ₹ *95,000* $P_2 = 6.76$

The required probability = $P_1 X_1 + P_2 X_2$

= 0.24 X 80,000 + 0.76 X 95,000

= 19,200 + 72,200 = ₹ 91,400.

The expected value of petrol sale on coming Monday = ₹ 91,400

Example 32 :

A bag contains 6 white and 9 black balls. Two drawings of 4 balls are made such that (a) the Balls are replaced before the second trial (b) the balls are not replaced before the second trial. Find the probability that the first drawing will give 4 white and the second 4 black balls in each case.

Solution :

(a) When the balls are replaced before the second trial the number of ways in which 4 balls may be drawn is $^{15}C_4$

The number of ways in which 4 white balls may be drawn = $^{6}C_4$

The number of ways in which 4 black balls may be drawn = $^{9}C_4$

Therefore, the probability of drawing 4 white balls at first trial

$$= \frac{^{6}C_4}{^{15}C_4} = \frac{1}{91}$$

The Second trial of drawing 4 black balls.

$$= \frac{^{9}C_4}{^{15}C_4} = \frac{9 \times 8 \times 7 \times 6}{4!} \times \frac{4!}{15 \times 14 \times 13 \times 12} = \frac{6}{65}$$

Therefore the chance of the Compound event $= \frac{1}{91} \times \frac{6}{65} = \frac{6}{5915}$

(b) When the balls are not replaced :

At the first trial, 4 balls may be drawn in $^{15}C_4$ ways and 4 white balls may be drawn in $^{6}C_4$ ways.

Therefore the chance of 4 white balls at first trial $= \frac{^{6}C_4}{^{15}C_4} = \frac{1}{91}$ (as above)

When 4 white balls have been drawn and removed, the bag contains 2 white and 9 black balls.

Therefore at the second trial, 4 balls may be drawn in $^{9}C_4$ ways and 4 black balls maybe drawn in $^{9}C_4$ ways So, the chance of 4 black balls at the second trial

$$= \frac{^{9}C_4}{^{11}C_4}$$

$$= \frac{9 \times 8 \times 7 \times 6}{4!} \times \frac{4!}{11 \times 10 \times 9 \times 8} = \frac{21}{55}$$

Therefore the chance of the compound event $= \frac{1}{91} \times \frac{21}{55} = \frac{3}{715}$

Example 33 :

A salesman is known to sell a product in 3 out of 5 attempts while another salesman is 2 out of 5 attempts. Find the probability that (i) No sale will be effected when they both try to sell the product and (ii) Either of them will succeed in selling the product.

Solution :

Let the two salesmen be *A* and *B.*

P (A) = The probability that the salesman A is able to sell the

product = $\frac{3}{5}$

P (B) = The probability that the salesman B is able to sell the product $\frac{2}{5}$

(i) probability that no sale will be effected $=\left(1-\frac{3}{5}\right)\left(1-\frac{2}{5}\right)=\frac{6}{25}$

(ii) probability that either of them will succeed in selling the product

$$=\frac{3}{5}+\frac{2}{5}-\frac{3}{5}\times\frac{2}{5}=\frac{19}{25}$$

Example 34 :

A class consists of 100 students, 25 of them are girls and 75 boys, 20 of them are rich and remaining poor, 40 of them are fair complexioned. What is the probability of selecting a fair complexsioned rich girl ?

Solution:

Probability of selecting a fair complexioned student $=\frac{40}{100}=\frac{2}{5}$

Probability of selecting a rich student $=\frac{20}{100}=\frac{1}{5}$

Probability of selecting a girl $=\frac{25}{100}=\frac{1}{4}$

Since the events are independent, by multiplication rule of probability, the

probability of selecting a fair complexioned rich girl $=\frac{2}{5}\times\frac{1}{5}\times\frac{1}{4}=\frac{2}{100}=0.02$

Example 35 :

Three groups of workers contain 3 men and one woman, 2 man and 2 women, and 1 man and 3 woman respectively. One worker is selected at random from each group. What is the probability that the group selected consists of 1 man and 2 woman ?

Solution :

There are three possibilities :

(i) Man is selected from the first group and women from second and third groups; or

(ii) Man is selected from the second groups and women from first and third groups; or

(iii) Man is selected from the third groups and women from first and second groups.

∴ the probability of selecting a group of one man & two woman

$$=\left(\frac{3}{4}\times\frac{2}{4}\times\frac{3}{4}\right)+\left(\frac{2}{4}\times\frac{1}{4}\times\frac{3}{4}\right)+\left(\frac{1}{4}\times\frac{1}{4}\times\frac{2}{4}\right)$$

$$=\frac{18}{64}+\frac{6}{64}+\frac{2}{64}=\frac{13}{32}$$

6.7 SOME IMPORTANT TERMS AND CONCEPTS

1. **Random Experiment or Trial:**

   An experiment is characterized by the property that its observations under a given set of circumstances do not always lead to the same observed outcome but rather to different outcomes which follows a sort of statistical regularity. It is also called a Trial.

   For example: tossing a coin, or throwing a dice.

2. **Sample Space:**

   A set of all possible outcomes from an experiment is called a sample space. Let us toss a coin, the result is either head or tail, Let I denote head and 0 denote tail, Mark the point 0, 1 on a straight line. These points are called sample points or event points. For a given experiment there are different possible outcomes and hence different sample points. The collection of all such sample points is a Sample Space.

3. **Discrete Sample Space**:

   A sample space whose elements are finite or infinite but countable is called a discrete sample space.

   For example, if we toss a coin as many times as we require for turning up one head, then the sequence of points $S_1 = (1)$, $S_2 = (0, 1)$, $S_3 = (1, 0, 0)$, $S_4 = (0,0,0,1)$ etc., is a discrete sample space.

4. **Continuous Sample Space**:

   Sample spaces whose elements are infinite and uncountable or assume all the values on a real line R or on an interval of R is called a continuous sample space. In this case the sample points build up a continuum, and the sample space is said to be continuous.

   For example, all the points on a line or all points on a plane is a sample space.

5. **Event**:

   A sub collection of a number of sample points under a definite rule or law is called an event.

   For example: let us take a dice. Let us faces 1, 2, 3, 4, 5, 6 be represented by E_1, E_2, E_3, E_4, E_5, E_6 respectively. Then all the E_1's are sample points Let E be the event of getting an even number on the dice. Obviously, $E = (E_2, E_4, E_6)$, which is a subset of the set $(E_1, E_2, E_3, E_4, E_5, E_6)$

6. **Null Event**:

   An event having no sample point is called a null event and is denoted by Φ .

7. **Simple Event**:

   An event consisting of only one sample point of a sample space is called a simple event.

   For example, let a dice be rolled once and A be the event that face number 5 is turned up then A is a simple event.

8. **Compound Events**:

   When an event is decomposable into a number of simple events, then it is called a compound event.

   For example: the sum of the two numbers shown by the upper faces of the two dice is seven in the simultaneous throw of the two unbiased dice, is a compound event as it can be decomposable.

9. **Exhaustive Cases or Events**:

   It is the total number of all the possible outcomes of an experiment.

   For example: When we throw a dice, then any one of the six faces (1, 2, 3, 4, 5, 6) may turn up and, therefore, there are six possible outcomes. Hence there are six exhaustive cases of events in throwing a dice.

10. Mutually Exclusive Events:

If in an experiment the occurrence of an event precludes or prevents or rules out the happening of all other events in the same experiment, then these events are said to be mutually exclusive events

For example: in tossing a coin, the events head and tail are mutually exclusive, because if the outcome is head, then the possibility of getting a tail in the same trial is ruled out.

11. Equally likely Events:

Events are said to be equally likely if there is no reason to expect any one in preference to other.

For example: in throwing a dice, all the six faces (1,2,3,4,5,6) are equally likely to occur.

12. Collectively Exhaustive Events:

The total number of events in a population exhausts the population. So they are known as collectively exhaustive events.

13. Equally Probable Events:

If in an experiment all possible outcomes have equal chances of occurrence, then such events are said to be equally probable events.

For example: in throwing a coin, the events head and tail have equal chances of occurrence, therefore, they are equally probable events.

14. Favourable Cases:

The cases which ensure the occurrence of an event are said to be favourable to the event.

15. Independent and Dependent Events:

When the experiment are conducted in such a way that the occurrence of an event in one trial does not have any effect on the occurrence of this or other events at a subsequent experiment, then the events are said to be independent. In other words, two or more events are said to be independent if the happening of any one does not depend on the happening of the other. Events which are not independent are called dependent events.

16. Classical Definition of Probability:

If an experiment has n mutually exclusive, equally likely an exhaustive cases, out of which m are favourable to the happening of the event A, then the probability of the happening of A is denoted by P (A) and is defined as:

$$P(A) = \frac{m}{n} = \frac{\text{No. of cases favourable to A}}{\text{Total (Exhaustive) number of cases}}$$

Note 1: Probability of an event which is certain to occur is 1 and the probability of an impossible event is zero.

17. The probability of occurrence of any event lies between 0 and 1, both inclusive

Addition Theorem or Theorem on total Probability

Statement: If a events are mutually exclusive, then the probability of happening of any one of them is equal to the sum of the probabilities of the happening of the separate events, i.e., in other words, if, E_1, E_2, E_3….E_n be n events and $P(E_1)$, $P(E_2)$,….. $P(E_n)$, be their respective probabilities then
$P(E_1 + E_{2+} E_3 + …. + E_n) = P(E_1) + P(E_2) +…..+ P(E_n)$.

18. Multiplicative Theorem or Theorem on Compound Probability

Simple and compound events: a single event is called a simple event. When two or more than simple events occur in connection with each other, then their simultaneous occurrence is called a compound event. If A and B are two simple events. The simultaneous occurrence of A and B is called a compound event and is denoted by AB. or $A \cap B$

Conditional probability: The probability of the happening of an event B. when it is known that A has already happened is called the conditional probability of B and is denoted by P(B/A)

i.e., P(B/A) $\Rightarrow$ conditional probability of B given that A has already occurred.

Similarly, P(A/B) $\Rightarrow$ conditional probability of A given that B has already happened.

Mutually Independent Events: An event A is said to be independent of the event B

if P(A/B) = P(A), i.e., the probability of the happening of A is independent of the happening of B.

Theorem on Compound Probability:

Statement: The probability of the simultaneous occurrence of the two events A and B is equal to the probability of one of the events multiplied by the conditional probability of other given the occurrence of the first, i.e.,

P(AB) = P(A).P(B/A) = P(B).P(A/B)

Cor.1. If the events A and B are statistically independent then P(B/A) = P(B) and P(A/B) = P(A).

$\therefore$ P(AB) = P(A) P(B/A) = P(A) × P(B). $\quad |\because\ P(B/A) = P(B)|$

19. Complementary Events:

The event 'A occurs' and the event 'A does not occur' are called complementary events. The "event A does not occur" is denoted by A^c or $\overline{}$ and is read as complementary of A. It is important to note that $P(A) + P(A^c) = 1 \Rightarrow P(A^c) = 1 - P(A)$.

20. STATEMENT OF BAYEES' THEOREM:

Theorem: Let $E_1, E_2, \ldots E_n$ be the set of n mutually exclusive and exhaustive event whose union is the random sample space S, of an experiment. If A be any arbitrary event of the sample space of the above experiment with $P(A) \neq 0$, then the probability of the event E_1, when the event A has actually occurred is given by $P(E_1/A)$,

where $P(E_1/A) = \dfrac{P(A \cap E_1)}{P(A \cap E_1) + P(A \cap E_1) + P(A \cap E_n)}$

We know that $P(A \cap E_1) = P(E_1)\,P(A/E_1)$

$$\therefore \qquad P(E_1/A) = \frac{P(E_1)\,P(A/E_1)}{\Sigma\, P(E_1)\,P(A/E_1)}$$

MULTIPLE CHOICE QUESTIONS

1. Initially, probability was a branch of
   (a) Physics
   (b) Statistics
   (c) Mathematics
   (d) Economics
2. If events are mutually exclusive, then –
   (a) Their probabilities are less than ne
   (b) Their probabilities sum to one
   (c) Both events cannot occur at the same time
   (d) Both of them contain every possible outcome of an experiment
3. If the events A and B are mutually exclusive then $P(A \cap B)$ is equal to-
   (a) 1
   (b) P(A) P(B)
   (c) 0
   (d) P(A)+P(B) - P(AUB)
4. If for two events A and B, P(AUB) = 1, then A and B are
   (a) Mutually exclusive events
   (b) Equally likely events
   (c) Exhaustive events
   (d) Dependent events
5. If P(A) = 1, then the event A is known as
   (a) Symmetric event
   (b) Dependent event
   (c) Improbable event
   (d) Sure event
6. If P(A) = 1, then the event A is known as
   (a) Symmetric event
   (b) Dependent event
   (c) Improbable event
   (d) Sure event
7. If an unbiased coin is tossed once, then the two events head and tall are
   (a) Mutually exclusive
   (b) Exhaustive
   (c) Equally likely
   (d) All these

8. The probability of an event can assume any value between
   (a) -1 and 1
   (b) 0 and 1
   (c) -1 and 0
   (d) None of these
9. ____ of all probabilities is equal to 1.
   (a) Sum
   (b) Difference
   (c) Product
   (d) None of the above
10. Sum of probability of an event A and its complement is _
   (a) 1
   (b) 0
   (c) ½
   (d) -1/2
11. If A is an event and A^c its complementary event then
   (a) $P(A) = P(A^c) - 1$
   (b) $P(A^c) = 1 - P(A)$
   (c) $P(A) = 1 + P(A^c)$
   (d) None
12. If for two events A and B, $P \neq P(A \cap B) \neq P(A)\ P(B)$, then the two events A and B are
   (a) Independent
   (b) Dependent
   (c) Not equally likely
   (d) Not exhaustive
13. Probability of occurrence of atleast one of the events A and B is denoted by
   (a) P(AB)
   (b) P(A+B)
   (c) P(A/B)
   (d) None
14. Probability of occurrence of A as well as B is denoted by
   (a) P(AB)
   (b) P(A+B)
   (c) P(A/B)
   (d) None of these

15. When the no. of cases favourable to the event A= 0 then P(A) is equal to
    (a) 1
    (b) 0
    (c) ½
    (d) None
16. If P(A) = 5/11 then probability of complement of A is equal to –
    (a) 6/11
    (b) 5/11
    (c) 1
    (d) 5/6
17. If P(A) = 7/8 then $P(A^c)$ is equal to
    (a) 1
    (b) 0
    (c) 7/8
    (d) 1/8
18. All possible outcomes of a random experiment forms the
    (a) Events
    (b) Sample
    (c) Both
    (d) None
19. S = {1, 2, 3, 4, 5, 6} is the ____ when a die is tossed.
    (a) event
    (b) sample
    (c) set
    (d) sample space
20. The value of P(S) were S is the sample space is
    (a) -1
    (b) 0
    (c) 1
    (d) None
21. The probability space in tossing two coins is
    (a) {(H, H), (H, T), (T, H)}
    (b) {(H, T), (T, H), (T, T)}
    (c) {(H, H), (H, T), (T, H), (T, T)}
    (d) None

22. Which set of function define a probability space on S {A, B, C}?

(a) P(A) = P(B) = 0, P(C) = 1

(b) P(A) = 1/3, P(B) = 0, P(C) = 2/3

(c) Both ((a) & ((b)

(d) Neither ((a) nor ((b)

23. Which of the following set of function define a probability space on S = {A, B, C}

(a) P(A) = 1/3, P(B) = 1/2, P(C) = ¼

(b) P(A) = 1/3, P(B) = 0, P(C) = 2/3

(c) Both a and b

(d) Neither a nor b

24. Let P be a probability function on S = {x1, x2, x3}if P(X1) = ¼ and P(X3) = 1/3 then P(X2) is equal to

(a) 5/12

(b) 7/12

(c) ¾

(d) none of these

25. The probability of occurrence of atleast one of the 2 events A and B (which may not be mutually exclusive) is given by

(a) P(A+B)=P(A)-P(B)

(b) P(A+B)=P(A)+P(B)-P(AB)

(c) P(A+B)=P(A)-P(B)+P(AB)

(d) P(A+B)=P(A)+P(B)+P(AB)

26. If P(A) = ½, P(B) = 3/5 and the events A& B are independent then P $(A \cap B)$ is-

(a) 7/10

(b) 3/10

(c) 5/10

(d) 9/10

27. If P$(A \cap B)$ = 0.60 and P(AUB) = 0.70 for two events A and B, then P(A) + P(B) is

(a) 1.30

(b) 0.90

(c) 1.00

(d) 0.75

28. If for two independent events A and B, P(AUB) = 2/3 and P(A) = 2/5, what is P(B)?

(a) 4/15

(b) 4/9

(c) 5/9

(d) 7/15

29. A and B are two events such that P(A) = ½, P(B) =1/4 and $P(A \cap B)$ = 1/5. Find P(AUB).
    (a) 4/5
    (b) 11/20
    (c) 3/5
    (d) None of these
30. A, B and C are three mutually exclusive and exhaustive events such that P(A) = 2 P(B) = 3 P(C). What is P(B)?
    (a) 6/11
    (b) 6/22
    (c) 1/6
    (d) 1/3
31. If A, B and C are three mutually exclusive and exhaustive events, then P(A) + P(B)+P(C) equals to
    (a) 1/3
    (b) 1
    (c) -1
    (d) Any value between 0 and 1
32. If P(A) = 6/9 then the odds against the event A is
    (a) 3/9
    (b) 6/3
    (c) 3/6
    (d) 3/15
33. If p:q are the odds in favour of an event, then the probability of that event is
    (a) p/q
    (b) $\frac{p}{p+q}$
    (c) $\frac{q}{p+q}$
    (d) None of these
34. The odds in favour of one student passing a test are 3:7. The odds against another student passing at are 3:5. The probability that both pass is
    (a) 7/16
    (b) 21/80
    (c) 9/80
    (d) 3/16
35. The odds in favour of one student passing a test are 3:7. The odds against another student passing at are 3:5. The probability that both fall is
    (a) 7/16
    (b) 21/80
    (c) 9/80
    (d) 3/16

36. A bag contains 20 discs numbered 1 to 20. A disc is drawn from the bag. The probability that the number on it is a multiple of 3 is

(a) 5/10

(b) 2/5

(c) 1/5

(d) 3/10

37. A number of selected from the set S = {1, 2, 3, 4..25}. The probability, that it would be divisible by 4 or 7, is

(a) 0.26

(b) 0.46

(c) 0.36

(d) None of these

38. A bag contains 30 balls numbered from 1 to 30. One ball is drawn at random. The probability that the number of the drawn ball will be multiple of 3 or 7 is

(a) 7/15

(b) 13/30

(c) ½

(d) None of these

39. If an unbiased coin is tossed twice, the probability of obtaining at least one tail is

(a) 0.25

(b) 0.50

(c) 0.75

(d) 1.00

40. Two unbiased coins are tossed. The probability of obtaining one head and one tail is

(a) ¼

(b) 2/4

(c) ¾

(d) None

41. Three coins are tossed together. The probability of getting exactly two heads is

(a) 5/8

(b) 3/8

(c) 1/8

(d) None

42. 4 coins are tossed. The probability that there are 2 heads is

(a) ½

(b) 3/8

(c) 1/8

(d) none of these

43. Probability of throwing an even number with an ordinary six faced die is
    (a) ½
    (b) 1
    (c) 0
    (d) -1/2
44. If two unbiased dice are rolled, what is the probability of getting sum of the points neither 6 nor 9?
    (a) 0.25
    (b) 0.50
    (c) 0.75
    (d) 0.80
45. Two dice are thrown together. The probability that 'the event the difference of no.s. shown is 2' is
    (a) 2/9
    (b) 5/9
    (c) 4/9
    (d) 7/9
46. Two dice are thrown together. The probability of the event that the sum of no.s. Shown is greater than 5 is
    (a) 13/18
    (b) 15/18
    (c) 1
    (d) None
47. A card is drawn from a well-shuffled pack of playing cards. The probability that it is a king is
    (a) 1/13
    (b) ¼
    (c) 4/13
    (d) None
48. If a card is drawn at random form a pack of 52 cards, what is the chance of getting a Spade or an ace?
    (a) 4/13
    (b) 5/13
    (c) 0.25
    (d) 0.20
49. The probability of drawing a black ball from a bag containing 8 white and 5 black balls –
    (a) 5/13
    (b) 8/13
    (c) ½
    (d) 1

50. A bag contains 10 red and 10 green balls. A ball is drawn from it. The probability that it will be green is

(a) 1/10

(b) 1/3

(c) ½

(d) None of these

51. The probability that A speaks truth is 4/5, while the probability for B is ¾. The probability that they contract each when asked to speak on a fact is

(a) $\frac{3}{20}$

(b) $\frac{1}{5}$

(c) $\frac{7}{20}$

(d) $\frac{4}{5}$

52. A problem in probability was given to three CA students A, B and C whose chances of solving it are 1/3, 1/5 and ½ respectively. What is the probability that the problem would be solved?

(a) 4/15

(b) 7/8

(c) 8/15

(d) 11/15

53. If the overall percentage of success in an exam is 60, what is the probability that out of a group of 4 students, at least one has passed?

(a) 0.6525

(b) 0.9744

(c) 0.8704

(d) 0.0256

54. If the probability of a house A winning a race is 1/6 and the probability of a horse B winning the same race is ¼, what is the probability that one of the horses will win

(a) 5/12

(b) 7/12

(c) 1/12

(d) None

55. A man can kill a bird once in five shots. The probabilities that a bird is not killed is

(a) 4/5

(b) 1/5

(c) 3/5

(d) 2/5

56. What is the probability that a leap year selected at random would contain 53 Saturdays?
    (a) 1/7
    (b) 2/7
    (c) 1/12
    (d) ¼
57. What is the probability that 3 children selected at random would have different birthdays?
    (a) $\frac{364 \times 363 \times 362}{(365)^3}$
    (b) $\frac{6 \times 5 \times 4}{(7)^3}$
    (c) $\frac{1}{365}$
    (d) $\frac{1}{7} \times 3$
58. Baye's Theorem is useful in
    (a) Revising probability estimates
    (b) Computing conditional probabilities
    (c) Computing Sequential probabilities
    (d) None of these
59. Let A & B are two events with P(A) = 2/3, P(B) = ¼ and $P(A \cap B)$ = 1/12 then P(A/B) is equal to
    (a) 2/3
    (b) 1/3
    (c) 1/8
    (d) 7/8
60. In connection with a random experiment, it is found that P(A) = 2/3, P(B) = 3/5 and P(AUB) = 5/6. Find P(B/A).
    (a) 13/18
    (b) ½
    (c) 13/20
    (d) 5/18
61. If $P(A \cap B) = 0.10$ and $P(\overline{B}) = 0.80$, then P(A/B) is
    (a) 0.25
    (b) 0.40
    (c) 0.50
    (d) 0.75
62. A and B are two events such that P(A) = 1/3, P(B)=1/4, P(A+B)=1/2 then P(B/A) is equal to
    (a) ¼
    (b) $\frac{1}{3}$
    (c) ½
    (d) none of these

EXERCISE - I

1. What is the probability of getting an even number in a single throw with a dice?

2. What is the probability of getting tail in a throw of a coin?

3. A bag contains 6 white balls, 9 black balls. What is the probability of drawing a black ball?

4. What is the probability that if a card is drawn at random from an ordinary pack of cards. It is i) a red card ii) a club, iii) one of the court cards (Jack or Queen or King).

5. What is the probability of throwing a number greater than 3 with an ordinary dice?

6. What is the probability that a leap year, selected at random, will have 33 Sundays?

7. What is the probability of getting a total of more than 10 in a single throw with two-dice?

8. A card a drawn from an ordinary pack of playing cards and a person bets that it is a spade or an ace. What are the odds against this winning this bet?

9. A dice is rolled. What is the probability that a number 1 or 6 may appear on the upper face?

10. If the probability of the horse. A winning the race is 1/3 and the probability of the horse 8 winning the same race is 1/6. What is the probability that one of the horses will win the race?

11. If the coins are tossed. What is the probability that all will show a head?

12. A card is drawn from a pack of 32 cards and then a second card is drawn. What is the probability that both the cards drawn are queen?

13. A bag contains 5 while and 3 black balls. Two balls are drawn at random one after the other without replacement. Find the probability of drawing a black ball in the first second attempt respectively.

14. Find the chance of throwing at least one up in a single throw with three dices.

15. 4 coins are tossed. Find the probability that at least one head turns up.

16. A problem in Mathematics is given to three students. Dayanad, Ramesh and Naresh whose chances of solving it are $\frac{1}{2},\frac{1}{3},\frac{1}{4}$ respectively. What is the probability that the problem will be solved?

17. A man and his wife appear for an interview for two posts. The probability of the husbands selection is $\frac{1}{7}$ and that of the wife's selection is $\frac{1}{5}$. What is the probability that only one of them will be selected?

18. A salesman has a 60% chance of making a sale chance of making a sale to each customer the behavior of successive customers is independent. If two customers A and B enter, what is the probability that the sales man will make a sale to A or B?

19. Two students X and Y work independently on a problem. The probability that X will solve it is (3/4) and the probability that Y will solve it is (2/3). What is the probability that the problem will be solved?

20. The probability that a student passes a Physics test is (2/3) and the probability that he passes both Physics and English test is (14/45). The probability that he passes at least one test is (4/5). What is the probability that the student passes the English test?

21. Find the probability of the event A: a) if the odds in favour are 3:2; b) if the odd against it are 1:4.

22. A problem in statistics is given to two students A and B. The odds in favour of A solving the problem are 6 to 9 and against B solving the problem are 12 to 10. If A and B attempt, find the probability of the problem being solved.

23. The odds against a certain event are 5 to 3 and the odds in favour of another event, independent of the former are 7 to 5. Find the chance that at least one of the events will happen.

24. Suppose that it is 11 to 5 against to person who is now 38 years of age living all he is 73 and 5 to 3 against B now 43 living till he is 78 years. Find the chance that at least one of those persons will be alive 35 years hence.

25. The chances of living a person who is now 35 years old till he is 75 are 8:6 and of living another person now 40 years old till he is 80 are 4:5. Find the probability that at least one of these persons would die before completing 40 years hence.

26. From a pack of 52 cards, two are drawn at random. Find the chance that one is a king and the other is a queen.

27. Two cards are drawn from a well shuffled pack of playing cards. Determine the probability that both are aces.

28. Four persons are chosen at random from a group continuing 3 men, 2 women and 4 children. Show that the chance that exactly two of them will be children is 10/21

29. In a factory, there are 6 skilled workers and 4 unskilled workers. What is the probability that (a) worker selected a skilled worker, (b) the two workers selected are skilled?

30. Four cards are drawn from a full pack of cards. Find the probability that two are spades and two are hearts.

31. A committee of three is to be chosen from a group consisting of 5 men and 5 women. If the selection is made at random, find the probability that (a) all three are men, (b) two are men.

32. A bag contains 5 white and 8 red balls. Two drawings of 3 balls are made such that a) balls are replaced before the second trial and b the balls are not replaced before the second trial. Find the probability that the first drawing will give 3 while and the second 3 red with in each case.

33. A speaks truth in 75% and B in 80% of the cases. In what percentage of cases are they likely to contradict each other narrating the same incident?

34. Two dice are tossed. What is the probability that total is divisible by 3 or 4?

35. The probability that a person A who is now 25 years old, lives for another 30 years is 2/5, and the probability that the person B who is now 45 years old lives for another 30 years is 7/16, Find the probability that at least one of these persons will be alive. #0 years hence.

36. Let A and B be events with P(A) = 1/3, P(B) = 1/4 , $P(A \cap B) = 1/12$.

    Find (i) P(A/B), (ii) P(B/A), (iii) $P(B/A^c)$, (iv) $P(A \cap B^c)$

37. Given that P(A) = 3/8, P(B) = 5/8 and P(AUB)= ¾, Find $P(A \cap B)$ and P(B/A). Show whether A and B are independent.

38. The odds against a student X solving a business statistics problem are 8 to 6 and odds in favour of the student Y solving the problem are 14 to 16.

    (a) What is the chance that the problem will be solved if they both try independent of each other?

    (b) What is the probability that none of them is able to solve the problem?

39. The probability that A can solve the problem is 4/5, B can solve it is 2/3, and C can solve it is 3/7. If all of them try independently, find the probability that the problem will be solved.

40. The probability that A can solve a problem is 2/3 and that B can solve is 3/4. If both of them attempt the problem, what is the probability that the problem get solved?

EXERCISE - II

41. If P(A)=1/4, P(B) = 1/2, P(AUB) = 5/8, then P $(A \cap B)$ is:

42. A and B are events and P(A) = 0.4, P(AUB) = 0.7. If A and B are independent, the P(B) is:

43. The key for a door is in a bunch of 10 keys. A man attempts to open the door by trying keys at random discarding the wrong key. The probability that the door is opened in the fifth trial is:

44. Two dice are rolled simultaneously. The probability that the sum of the two numbers on the dice is a prime number is:

45. The probability that a number selected at random from the set of numbers {1,2,3,….,100} is a cube is:

46. The probability of choosing at random a number that is divisible by 6 or 8 from among 1 to 90 is:

47. Two unbiased six faced dice are thrown. The probability that the sum of the numbers on faces of them is a prime number greater than 5 is:

48. Two dice are thrown at a time and the sum of the numbers on them is 6. The probability of getting the number 4 on anyone of the dice is

49. A single letter is selected at random from the word PROBABILITY. The probability that it is a vowel is

50. The probabilities of two events A and B are 0.25 and 0.40 respectively. The probability that both A and B occur is 0.15. The probability that neither A nor B occurs is:

51. In a competition A, B and C are participating. The probability that A wins is twice that of B. the probability that B wins is twice that of C. The probability that A loses is

52. A bag contains 5 black balls, 4 white balls and 3 red balls. If a ball is selected at random, the probability that it is red or black is:

53. One die and a coin are tossed simultaneously. The probability of getting 5 on the top of the die and tail on the coin is:

54. The probability of getting qualified in IIT JEE and AIEEE by a student are respectively 1/5 and 3/5. The probability that the student gets qualified for one of these tests is:

55. If P(AUB) = 0.8 P$(A \cap B)$ = 0.3, then P($\overline{A}$) + P($\overline{B}$)=

56. A bag X contains 2 white and 3 black balls and another bag Y contains 4 white and 2 black balls. One bag a selected at random and a ball is drawn from it. Then the probability for the ball chosen be while is to

57. The probability that A speaks truth is 4/5, while its probability for B is ¾, The probability that they contradict each other when asked to speak on a fact is:

58. A coin and a six faced die, both unbiased, are thrown simultaneously. The probability of getting a head on the coin and an odd number on the die is:

59. Seven balls are drawn simultaneously from a bag containing 5 white and 6 green balls. The probability of drawing 3 while and 4 green balls is

60. The probability of happening an event A in one trial is 0.4. The probability that the event A happens at least once in three independent trials is

61. A bag contains tickets numbered from 1 to 20. Two tickets are drawn. The probability that both the numbers are prime is

62. A card is drawn from a pack of 100 cards numbered 1 to 100. The probability of drawing a number which is a square is:

63. If three dice are thrown simultaneously, then the probability of getting a score of 5 is

64. One card is drawn from a pack of 52 cards. The probability that it is a king or diamond is

65. Four persons are chooses at random from a group of 3 men, 2 women and 4 children. The chance that exactly 3 of them are children is:

66. A fair coin is tossed three times. Find the probability of getting at most one head and two consecutive heads.

67. Given two mutually exclusive events A and B. such that P(A) = 0.45 and P(B) = 0.35, then P(A or B)=

68. A box contains 25 tickets numbered 1, 2, ...25. If two tickets are drawn at random, then the probability that the product of their numbers is even is:

69. The letters of the word 'SOCIETY' are arranged in a row. What is the chance that the word formed begins with S and end in Y?

70. Three numbers are chosen from 1 to 30. The probability that they are consecutive is:

EXERCISE - III

71. Two dice are thrown simultaneously. The probability of getting a pair of aces is:

72. A and B are events, such that P(AUB) = ¾, P $(A \cap B)$ = ¼, P $(\overline{A})$ = 2/3, then $P(\overline{A} \cap B)$ is:

73. In a single throw of two dice, the probability of getting a total of 7 to 9 is

74. The probability of having at least one tail in 4 throws with a coin is:

75. 3 mangos and 3 apples are in a box, If 2 fruits are chosen at random, the probability that one is a mango and the other is an apple is:

76. There are four letters and four envelops bearing addresses at random. The probability that the letters are placed in correct envelops is:

77. From a well shuffled pack of playing cards, two cards are drawn one by one without replacement. The probability that both are aces is:

78. A pack of cards contains 4 aces, 4 kings, 4 queens and 4 jacks. Two cards are drawn at random. The probability that at least one of them is an ace is:

79. If A and B are events, such that P(A)= 0.42, P(B) = 0.48 and $P(A \cap B) = 0.16$, the P(A+B) is equal to

80. Three identical dice are rolled. The probability that the same number will appear on each of them is

81. A problem is given to three persons and their chances of solving it are 1/3, 1/5, 1/6 respectively. The probability that none will solve it is:

82. From a group of 5 boys and 3 girls, three persons are chosen at random. Find the probability that there are more girls than boys.

83. For a biased dice, the probabilities for the different faces to turn up are

Face	1	2	3	4	5	6
P	0.10	0.32	0.21	0.15	0.05	0.17

The dice is tossed and you are told the either face 1 or face 2 has turned up, then the probability that it is face 1 is

84. The chance that the vowels are separated in an arrangement of the letters of the word 'HORROR' is:

85. The problem in mathematics is given to 3students whose chances of solving individually are ½, 1/3 and ¼, The probability that the problem will be solved at least by one is

86. An urn contains 9 balls two of which are red, three blue and four black. Three balls are drawn at random. The probability that they are of same colour is

87. If A and B are two incidences and P(A) = 3/8, P(B) = ½ , $P(A \cap B)$ = ¼, then the value of PA` U B`) is:

88. An unbiased dice is rolled four times. The probability that the minimum number on any toss is not less than 3 is:

89. A cricket club has 15 members of which only 5 can bowl. If the names of 15 members are put into a box and 11 names are drawn at random, then the probability of obtaining 11 member team containing exactly three bowlers is:

90. The probability for a randomly chosen month to have its 10^{th} day as Sunday is

91. In a non-leap year, the probability of getting 53 Sundays or 53 Tuesdays or 53 Thursdays is:

92. If A and B are two events, such that P(A) > 0 and P(B) ≠ 1, then $P\left(\overline{A}/\overline{B}\right)$ is equal to

93. The odds against an event A are 5:2 and odds to favour of another independent event B are 6:5. The chances that neither A nor B occurs is

94. A bag contains 5 brown and 4 white socks. A man pulls out 2 socks. The probability that they are of the same colour is:

95. A purse contains 4 copper coins and 3 silver coins, the second purse contains 6 copper coins and 2 silver coins. A coin is taken out from any purse, the probability that it is a copper coin is:

96. A sample space has only three points a_1, a_2, a_3. If a_1 is likely to occur as a_2 and a_3 is twice as likely to occur as a_3 then $P(a_3)$ is

97. From a pack of cards, two are drawn, the first being replaced before the second is drawn. The chance that the first is a diamond and the second is a king is:

98. If $P(A \cap B) = 0.15$, $P(B^c) = 0.10$, then $P(A/B)$ is

99. If there are 3 children in a family, then probability that there is one girl in the family is

100. Two events A and B have probabilities 0.25 and 0.50 respectively. The probability that both A and B occur simultaneously is 0.14. Then the probability that neither A nor B occur is:

MCQ Answers:

1	c
2	c
3	c
4	c
5	d
6	c
7	d
8	b
9	a
10	a
11	b
12	b
13	b
14	a

15	b
16	a
17	d
18	b
19	d
20	c
21	c
22	c
23	b
24	a
25	b
26	b
27	a
28	b
29	b
30	b
31	b
32	a
33	b
34	d
35	b
36	d
37	c
38	b
39	c

40	b
41	b
42	b
43	a
44	c
45	a
46	a
47	a
48	a
49	a
50	c
51	c
52	d
53	b
54	d
55	a
56	b
57	a
58	b
59	b
60	c
61	c
62	a