Ratio, Proportion and Logarithm

(I) Ratios:-

* A ratio is a comparison of 2 similar quantities using the division operation.

* A ratio is usually expressed in the form of a fraction (a/b) or (a:b) and is always unit free .

* It is always expressed in the simplest form of a:b.

* In a ratio, the 2 quantities are of similar nature and with same units. However, if the quantities are dissimilar, then the ratio is known as rate.

* However, in many scenarios, rate and ratio are interchangeably used. (Ratio is a pure fraction)

* In a ratio a:b, the terms a&b are known as antecedent (first term) and consequent (second term) respectively.

(i) CONTINUED RATIO:-

When more than 2 quantities are compared, such a ratio is known as continued ratio. (ex. a:b:c)

(ex. a:b:c:d)

(ii) COMPOUND RATIO:-

When a ratio is obtained by multiplying 2 or more ratios then such ratio is known as compound ratio.

e.x. Compound ratio of 1:2 & 3:4 = $\frac{1}{2}$ X $\frac{3}{4}$ = $\frac{3}{8}$ = $\frac{3}{8}$

* When a quantity is divided in a ratio, ex. divide K in the ratio a:b, then,

1st step = K X $\frac{a}{a+b}$

2nd step = K X $\frac{b}{a+b}$

e.x. Total student = 100(k)

Girl =40(a) } not given

Boys =60(b)

Ratio of Girls to Boys = 40:60 = 2:3 (a:b)

Then to find number of boys and girls,

Girls = K X $\frac{a}{a+b}$ = 100 X $\frac{2}{2+3}$ = 40

Boys = K X $\frac{b}{a+b}$ = 100 X $\frac{3}{2+3}$ = 60

- When a quantity is increased or decreased in the ratio a:b, then
  New quantity = Old quantity X $\frac{b}{a}$
  Ex. what is the value when q is increased in the ratio 3:4?
  It means if old value = 3x then new value is 4x
  New quantity = q X $\frac{4}{3}$ = 16

Ex. what is the value when 8 is decreased in the ratio 4:1 ?

It means if old value = 4x, then new value is 1x.

New quantity = 8 X $\frac{1}{4}$ = 2

$\therefore \frac{old\ quantity}{New\ quantity} = \frac{a}{b}$

- Duplicate and Triplicate ratio :-
  Duplicate ratio of $\frac{a}{b} = \frac{a^2}{b^2}$
  Triplicate value of $\frac{a}{b} = \frac{a^3}{b^3}$
- Sub duplicate and sub triplicate ratio :-
  Sub duplicate ratio of $\frac{a}{b} = \frac{\sqrt{a}}{\sqrt{b}}$
  Sub triplicate ratio of $\frac{a}{b} = \frac{\sqrt[3]{}}{\sqrt[3]{}}$
- Inverse ratio = $\frac{a}{b} = \frac{b}{a}$

(II) Proportion:-

1) An equality between 2 ratios is called proportion i.e. if a:b = c:d, then we say, a,b,c,d are in proportion.

2) All the values need not be of similar kind. However, the two terms of the first ratio must be similar to each other and so must be the two terms of 2nd ratio each other and so must be the two terms of 2nd ratio.

3) a, b, c, d are the 1st, 2nd,3rd,4th terms of the proportion.

4) a, d are known as extremes, b, c are known as means.

a:b = c:d

∴ $\frac{a}{b} = \frac{c}{d}$

∴ ad = bc

∴ Product of extremes of = product of means

(Cross product rule)

<u>Continued Proportion:-</u>

If a:b = b:c,

Then a,b,c are said to be in continued proportion and b is known as mean proportional to a&c.

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

∴ b^2 = ac

∴ b = $\sqrt{ac}$

∴ Mean proportion is always geometric mean of extrems

*<u>Properties of Proportion:-</u>

i) If $\frac{a}{b} = \frac{c}{d}$

∴ $\frac{b}{a} = \frac{d}{c}$ -------------- Invertendo

ii) If $\frac{a}{b} = \frac{c}{d}$

∴ $\frac{a}{c} = \frac{b}{a}$ -------------- Alternando

iii) If $\frac{a}{b} = \frac{c}{d}$

then $\frac{a+b}{b} = \frac{c+d}{d}$ ----------------- Compondendo –(1)

iv) If $\frac{a}{b} = \frac{c}{d}$

then $\frac{a-b}{b} = \frac{c-d}{d}$ ---------------- Dividendo –(2)

v) If $\frac{a}{b} = \frac{c}{d}$

then $\frac{a+b}{a-b} = \frac{c+d}{c-d}$ ---------------- Compondendo / dividendo $\frac{(1)}{(2)}$

vi) If $\frac{a}{b} = \frac{c}{d} = \frac{e}{f} = \frac{g}{h}$ ------- =K

then $\frac{a+c+e+g+\_\_\_\_}{b+c+f+h+\_\_\_\_} = k$ = each original ratio

----Addendo/or Theorem an equal ratios

vii) If $\frac{a}{b} = \frac{c}{d} = \frac{e}{f}$ = ------ K

then, $\frac{a-c-e.......}{b-d-f.......}$ =K = each ratio

----- subtragendo

Viii) $\frac{a-b}{a+b} = \frac{c-d}{c+d}$ ------- dividendo / compondendo $\frac{(2)}{(1)}$

*<u>Indices:-</u>

(1) 2+2+2+2 ------- 20 times = 2*20

a+a+a+a -------- n times = a*n

2*2*2*2 ------- 20 times = 2^{20}

a*a*a*a ------- n times = a^n

(2) Properties of Indices

i) $a^m * a^n = a^{m+n}$

ii) $\frac{a^m}{a^n} = a^{m-n}$

iii) $a^0 = 1$

iv) $(a^m)^n = a^{mn}$

v) $(ab)^n = a^n \times b^n$

vi) $a^{-m} = \frac{1}{a^m}$

vii) $\sqrt{a} = a^{\frac{1}{2}}$

$\sqrt[3]{a} = a^{\frac{1}{3}}$

$\sqrt[m]{a} = a^{\frac{1}{m}}$

viii) $a^x = a^y$, then x=y

ix) $a^x = b^x$, then a= $|b|$

$(2^2)^3$	$\neq$	2^{2^3}
↓		↓
$2^2 \times 2^2 \times 2^2$		$2^{(2^3)}$
2^{2+2+2}		2^8
$2^{2\times3}$		
2^6		

Important concept

(Point to remember)

∗ Proof of (If a + b + c=0, then $a^3 + b^3 + c^3 = 3abc$)

a + b + c=0

a + b = -c

$(a + b)^3 = (-c)^3$

$a^3 + b^3 + 3ab\,(a + b) = -c^3$

$a^3 + b^3 + c^3 + 3ab\,(-c) = 0$

$a^3 + b^3 + c^3 = +3abc$

(III) Logarithms:-

1) Logarithm is just another way of writing an indices equation

Ex.

i) Exponential form:- 2^3=8

Logarithmic form:- log_{2^8}=3

2 raised to what is 8?

ii) Exponential form:- $64^{1/2}$

Logarithmic form:- log_{64^8}=$^1/_2$

64 raised to what is 8

iii) Exponential form:- 12^x= 1728

Logarithmic form:- log_{12}1728=8

Properties of log

1) Log a^1=0
2) Log a^a=1
3) Log x^a+ logx^b= logx^{ab}
4) Log x^a- logx^b= log$x^{a/b}$
5) Log a^m= m log_{x^a}
6) $log_{x^{m^a}}$= $\frac{1}{m}$ log_{x^a}
7) Change of base theorem

log_{b^a} = b^a = $\frac{log\ x^a}{log\ x^b}$ x ⟶ any random value

8) Log c^a X Log b^c = Log b^a
9) Log c^a X Log b^c X Log d^b = Log d^a
10) Log b^a = $\frac{1}{Log\ a^b}$

$$\left(\frac{1}{Log\ a^b} = \frac{1}{\frac{Log\ b}{Log\ a}} = \frac{Log\ a}{Log\ b} = Log\ b^a \right)$$

11) $a^{Log\ a^b}$ = b
12) $Log\ b^a$ X $Log\ a^b$ = 1
13) $log\ b^a$ X $Log\ c^b$ X $Log\ c^a$ = 1
14) If a^3 + b^3 =0
15) $a^{Log\ b}$ = $b^{Log\ a}$

Logarithms can be of 2 types:-

1) Common Logarithms:-
   - Log taken with the base as 10.
   - The Log values of different numbers in the Log table are values of Log with base 10.
2) Natural Logarithms (denoted by Ln)
   - Log taken with the base as e (exponential numbers, 2.71828)

*Find Log values:-

For finding the Log values of any number with base 10, use the following steps:-

Case I:- If the given number is greater than 1 :-

(I)

i) Characteristic = No. of digits (before decimal) – 1

ii) Ignore the decimal point in given number

iii) Using first 4 digits of the given no. find Mantissa from the Log table.

iv) Log value of the given no. = char . mandissa

Decimal point

(II) Using calculator :-

i) Take the given number on calc

ii) Press √ sign 13 times.

iii) Subtract 1

iv) Multiply by 3558

Case II:- For number less than 1 (Log of –ve was don't exist)

(I) i) Characteristics = no. of zero immediately after decimal +1.

However, we have to write characteristic with bar sign (negative value)

ii) Ignore the decimal of the given no.

iii) Find mandissa using first 4 digits from the Log table-

this mantissa always has a positive sign

iv)∴ Log value = -characteristics + mantissa

*<u>Finding antilog values.</u>

Case 1:- for numbers greater than 0 (+ve values)

(I) Closing Log table:-

- (i) The given number is of 2 parts
  - (a) Characteristic = Part before the decimal
  - (b) Mantissa = Part after the decimal
- (ii) Using mantissa find first 4 digits from the antilog table.
- (iii) No. of digits before decimal = characteristic +1
- (iv) Place the decimal after the above no. of digits.

(II) Using Calculator :-

- (i) Take the given number on calculator.
- (ii) Divide by 3557
- (iii) Add 1
- (iv) Press 'X=' 13 times

Case 2:-

(I)

- i) Convert the negative number in the bar format (- characteristic + mantissa)
- ii) Using mantissa, find first four digits from antilog table
- iii) Number of zero (after decimal and before the digits = characteristic)
- iv) Place the above number of zeros after the decimal but before the digits.

Equations

(I) <u>Linear equations in 1 variable (highest power:1)</u>

- Only one solutions
- Solve mostly by option- wise

(II) <u>Linear equations in 2 or 3 variables (highest power: -)</u>

- Simultaneous equation
- Method

i) Method of elimination

ii) Cramer's rule

- If ax + by= c and dx + cy= f, then

i) If $\frac{a}{d}$= $\frac{b}{e}$= $\frac{c}{f}$ then infinite solutions (.....same equation)

ii) If $\frac{a}{d}$= $\frac{b}{e}$ $\neq \frac{c}{f}$, no solution (.....parallel lines)

iii) If $\frac{a}{d} \neq \frac{b}{e}$, one unique solution

(III) <u>Quadratic Equations</u>

These equations are usually in 1 variable where the highest power of the variable is 2

the general form of Q.E. is: - ax^2 + bx + c= 0

To find the solution of such an equation: - 2 methods:-

i) <u>Method of factorization</u>

In this method, we find 2 factors of given equation by splitting the middle term.

The product to be equal to 0, one of the factors must be equal to 0.

ii) <u>Formula method</u>

If factorization is not possible, any quadratic equation can be solved for x by using the below formula,

If ax^2 + bx + c= 0,

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

i.e. $x = \frac{-b+\sqrt{b^2-4ac}}{2a}$ Or $x = \frac{-b-\sqrt{b^2-4ac}}{2a}$

∗ Properties of Quadratic equations & their roots (solution).

i) If the roots of the quadratic equation are denoted by $\propto$ & β, then:

$\propto + \beta = -b + \frac{-b+\sqrt{b^2-4ac}}{2a} + \frac{-b-\sqrt{b^2-4ac}}{2a}$

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

a) ∴ $\propto + \beta = \frac{-b}{a}$

b) $\propto \beta = \frac{c}{a}$

ii) Nature of Roots

The nature of the roots of a quadratic equations are determined by value of $b^2 - 4ac$ (also known as discriminant / determinant) denoted by Δ.

$\Delta = b^2 - 4ac$	Nature of roots
1) Δ=0	Real, rational equal
2) Δ<0	Imaginary, unequal
3) Δ>0 prfect square	Real , rational, unequal
4) Δ>0 but not -11-	Real , irrational, unequal

iii) If the 2 roots of the equation are known and we have to construct the equation then.

Equation⇒ x^2 - (sum of roots) x + (product of roots) = 0

⇒ $x^2 - (\propto + \beta)x + (\propto \beta) = 0$

(IV) Cubic equations:-

- For solving cubic equations we have to find one of solutions by trial & error.
- We may use the following rule for finding the 1st solution:-

For a cubic equation, $ax^3 + bx^2 + cx + d = 0$

1) If a + c= b + d , then (x +1) is a factor of the equation ∴ x= -1 is a solution of the equation.

2) If a + c= -(b + d), then (x-1) is a factor of the equation ∴ x= 1 is a solution of the equation
3) If a + c= (b + d)=0 then (x+1), (x-1) both are factors & ∴ x=1 & n= -1 are 2 out of the 3 solutions of the equation.

- The remaining factors can be found using syntactic division.

LINEAR INEQUALITIES

An equation is equality between 2 mathematics expressions.
An equation where the highest power of the variables is 1 is known as linear equations.

The solution of linear equations is usually a single point.
An inequality is when 1 mathematical expression is greater than or less than another mathematical expression.
The solution for in equations is a range of values for the variables.

Linear inequalities are inequality where the highest power of the variable is 1.

I) <u>Linear inequality in single variable.</u>

a) (1) $x + 3 > 8$ (2) $x - 8 < 10$ (3) $3x > 15$ (4) $\frac{x}{3} > 10$

$x > 5$ $x < 10$ $x > 5$ $x > 30$

(5) $-x < -8$ (6) $-x < 8$ (7) $x + 8 < 3$

$x > 8$ $x > -8$ $x < -5$

(Sign reverse) (Sign reverse)

b) Graphs

i) $x >, 3$

ii) $x > 3$

c) Representation of answer:-

i) $x > 3$

$\Rightarrow 3 < x < \infty$

$\Rightarrow (3,\infty)$

ii) $x \geq 3$

$\Rightarrow 3 \leq x \leq \infty$

$\Rightarrow [3,\infty)$

iii) $x < -5$

$\Rightarrow -\infty < x < -5$

$\Rightarrow (-\infty, -5)$

iv) $x \leq 5$

$\Rightarrow -\infty < x \leq 5$

$\Rightarrow (-\infty, 5]$

v) $x > 2$ & $x > 5$

$\Rightarrow x > 5$

vi) $x > 2$ & $x < 5$

$\Rightarrow 2 < x < 5$

vii) $x \geq 2$ & $x \leq 5$

$\Rightarrow 2 \leq x \leq 5$

$\Rightarrow [2,5]$

viii) $x < 2$ & $x < -1$

$\Rightarrow x < -1$

$\Rightarrow (-\infty, -1)$

II) <u>Linear Inequality in 2 variables:-</u>

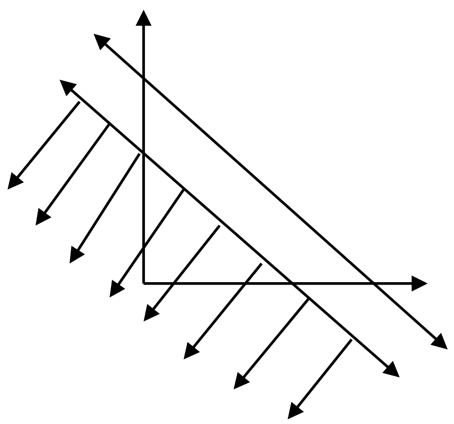
For these inequalities, we need graphs to represent the range of values of the variables.

a) Examples:

i) $2x + 3y \leq 6$

$3x + 4y \leq 12$

Inequalities	corresponding Equation	points	region
$2x + 3y \leq 6$	$2x + 3y = 6$	(0, 2), (3, 0)	towards origin
$3x + 4y \leq 12$	$3x + 4y = 12$	(0, 3), (4, 0)	towards origin

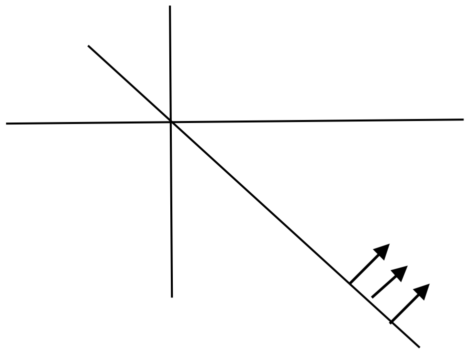


Rule: If a line passes through the first quadrant, then the area which is towards origin is less than area & area which is away from origin is more than area.

i) When the line does not pass through the first quadrant.

$3x + 5y \geq 0$

$3x + 5y = 0$ (0,0), (5,-3)



Substitute say (1, 0) in LHS.

$3 + 0 > 0$

$\therefore$ (1,0) is greater than side

Take a random point and substitute in LHS. Check with the questions if we want less than/ greater than region

Number System

- Real
  - Rational (where $\frac{p}{q}$, p, q ∈ N)
    - Whole nos. (0, 1, 2, 3...)
    - Natural nos. (1, 2, 3 ...)
    - Integers (-2,-1, 0, 1, 2...)
  - Irrational ($\sqrt{3}, \sqrt{5}, \sqrt{8}, \sqrt{10}$...)
- Imaginary (Root of negative numbers)

∗ To find a^n on calculate:-

i) Take a on calc.
ii) Press '√ ' sign 12 times
iii) Subtract 1
iv) Multiply by n
v) Add 1
vi) Press ' x = ' 12 times

∗ To find log x on calculator:-

i) Take x on calc.
ii) Press '√ ' sign 13 times
iii) Subtract 1
iv) Multiply by 3558

∗ To find anti log x on calculator:-

i) Take x on calc.
ii) Divide by 3557
iii) Add 1
iv) Press ' x= ' 13 times

Time Value of Money

Rent is a cost for using someone else's asset similarly interest is cost of using someone else fund funds.

Interest is calculated in two ways:-

1) Simple interest

   In this type of interest, the interest is always calculated on the original amount of investment a.k.a. principal

   ∴ If rate of interest is same, the interest amount for each year will be same.

   - SI for 1 year = $\frac{P\times R}{100}$
   - SI for N years= $\frac{P\times R\times N}{100}$

The balance amount / amount due after N years is given by

$$A = P + SI$$

$$\therefore A = P + \frac{PRN}{100}$$

$$\therefore A = P\left(1+\frac{RN}{100}\right)$$

2) Compound interest

   In compound interest, interest is always calculated on the balance outstanding at the start of the year.

   The amount due after n years of investment is given by:-

   $A= P\left(1+\frac{r}{100}\right)^{n}$ OR

   $A= P\ (1+i)^{n}$....... where $i = \frac{R}{100}$

   C.I = A-P

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

   $\therefore$ $C.I = P\ ((1+i)^{n} - 1)$

∗ Rule of doubling: - (Rule of 69)

The number of years required for an amount to double itself at r% p.a C.I, is given by

$N = \frac{69}{r} + 0.35$ (approx.) or $\left(=\frac{72}{r}=\frac{70}{r}\right)$

Depreciation:-

s.v = o.c $(1-d)^n$

s.v = scrap value

o.c = original cost

d = rate depreciation

n = no. of years

∗ Annuity:-

When a series of payment or receipt of an equal amount is made at regular interval of time, such a series of payment/ receipts is known as annuity.

∗ Types of annuity:-

i) Annuity regular
   When annuity is paid/ at the end of each period, such an annuity is called received annuity regular.
ii) Annuity immediate/ Annuity due
   When annuity is paid/received at the start of each period, such an annuity is called annuity immediate.

∗ Present value of an Annuity:-

A single payment or receipt made today corresponding a series of future receipts or payment. This single payment or receipt is the present value of the annuity.

∗ Present value (PV) of an annuity regular:-

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

Where A is annuity amount, i is rate of interest

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

∗ P.V of annuity immediate:-

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

OR

* <u>Future Value of an Annuity:-</u>

A single payment or receipt after n year's equivalent to a series of future payments/receipts of equal amounts is known as future value of annuity.

FV of annuity regular = A $\left[\frac{(1+i)^n-1}{i}\right]$

FV of annuity immediate = A$\left[\frac{(1+i)^n-1}{i}\right] \times (1+i)$

Relation between F.V and P.V.

P.V$(1+i)^n$ = F.V

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

* <u>Perpetuity:-</u> (indefinite term)

An annuity which is perpetual in nature is perpetuity. The future value of a perpetuity cannot be calculated however using so of G.P.,

P.V of perpetuity= $\frac{A}{i}$

Where A is annuity amount and i is R/100

P.V of growing perpetuity = $\frac{A}{i-9}$ (where g is growth rate of interest)

* <u>Valuation of Bond:-</u>

A bond is a type debt security (loan).

In bonds, the company is required to pay the interest annually to the investor for a fixed term.

Note: - The interest is always calculated at the face value of the bond

The company is also required to redeem the principal amount after the term of the loan. This redemption may be done at par or at premium or at a discount of the face value.

Value of bond = Present value of all future cash inflows from the bond

(From POV of investor)

* Effective rate of interest:-

i) $\in = (1+i)^n - 1$

$\in$ = effective rate of interest in decimal

i = rate of interest per conversion period in decimal

n = no. of conversion periods.

ii) $i = (\in +1)^{1/n} - 1$ → To find nominal rate of interest

$r(p-a) = i \times no.of\ conversion\ period$

1) P.V of Annuity regular:-

P.V= $\frac{A}{(1+i)^1} + \frac{A}{(1+i)^2} + \frac{A}{(1+i)^3} + \cdots \frac{A}{(1+i)^n}$

Calc: - $[A \div (1+i)(\text{=M+"...n times)MRC}]$

2) P.V of annuity immediate:-

P.V= $\left(\frac{A}{(1+i)^1} + \frac{A}{(1+i)^2} + \frac{A}{(1+i)^3} + \cdots \frac{A}{(1+i)^n}\right) \times (1+i)$

Calc: - $[A \div (1+i)(\text{"=M+"...n times)MRC}\times(1+\text{i})]$

3) F.V of annuity regular:-

F.V = $\frac{A(1+i)^1 + A(1+i)^2 + A(1+i)^3 + \cdots . A(1+i)^n}{(1+i)}$

Calc:- $[(1+i) \times A(\text{"=M+"...n times) MRC} \div (1+\text{i})]$

4) F.V of annuity immediate:-

F.V= $A(1+i)^1 + A(1+i)^2 + A(1+i)^3 + \cdots . A(1+i)^n$

Calc:- $[(1+i) \times A(\text{"=M+"...n times) MRC}]$

* F.V of annuity immediate (alternative):-

P.V= $A(1+i)^1 + A(1+i)^2 + A(1+i)^3 + \cdots A(1+i)^{n-1} + A$

Calc:- $[(1+i) \times A(\text{"=M+"...n-1 times) AM+MRC}]$

* P.V of annuity immediate (alternative):-

P.V = $\left(\frac{A}{(1+i)^1} + \frac{A}{(1+i)^2} + \cdots + \frac{A}{(1+i)^{n-1}} + A\right)$

Calc:-[A÷(1+i) ('=M+"...n-1 times) AM+MRC]

Permutation and Combination

(Basic Concepts)

(No. of different ways of doing things)

∗ Factorial Notation:-

- Denoted by ! or ∟

e.x. n ! ....

∟n ......n factorial

The product of all natural nos. upto a certain number is factorial value of that number.

Ex. 5 ! = 5× 4 ×3×2×1

8 ! = 8×7×6×5×4×3×2×1

8 ! = 8×7×6×5 !

8 ! = 8×7 !

- List of values of factorial of 1st 10 natural nos.

1 ! = 1

2 ! = 2

3 ! = 6

4 ! = 24

5 ! = 120

6 ! = 720

7 ! = 5040

8 ! = 40320

9 ! = 362880

10 ! = 3628800

0 != 1

n ! = n(n-1) (n-2) (n-3) × ....×3×2×1.......(n terms)

∗ Permutation (order is important)

When n things are to be arranged in r places we write this as n permutation r given below:-

$$nP_r = \frac{n!}{(n-r)!}$$

Properties:-

i) $nP_n = n!$
ii) $nP_1 = n$
iii) $nP_{n-1} = n!$
iv) $nP_0 = 1$

Note:-

i) Unless otherwise mentioned, assume that use have to solve permutation question without repetition. (Repetition of thing/letters/digits)

Always together, never together, No. 2 together

(I) Always Together:-
(1) From a group of the things which are to be always together.
(2) Arrange this group with the remaining things.
(3) Multiply by the no. of ways in which things within the group can be arranged among themselves.

(II) Never Together:-
(1) Find total ways of arranging all things (without any restriction).
(2) Find answer of always together.
(3) Never together = (step 1) – (step 2)

(III) No. 2 together:-
(1) Arrange the things on which there is no restriction among themselves.
(2) Mark the places in between and at the start & end of the above letters (already arranged).
(3) Arrange letters on which there is restriction in the above marked places.

Note:-

If the restriction is on 2 things, never together = no.2 together

*If Repetition of things is allowed

1) Don't use permutation.

2) Solve for each place differently (place-wise).

* When n things are to be arranged of which n1 things are of same kind (are same, ex- A, A), n2 things are also same, -----

Then the no. of ways of arranging all things n

$$\text{Things} = \frac{n!}{n1! \times n2! \times \text{-----}}$$

(When repetition is in available things)

*Circular Permutation:-

When n things are to be arranged in a circle, the no. of ways of arranging them is given by:-

$$\frac{nPn}{n} = \frac{n!}{n} = \frac{\cancel{n}(n-1)!}{\cancel{n}} = (n-1)!$$

However, when n flower or n needs are to be arranged in circle to form a garland necklace, etc. No. of dif. garlands, etc. That can be formed is given by:-

$$\frac{(n-1)!}{2}$$

*When n person are to be arranged at a round table, such that no person has same 2 neighbours, then no. of ways of arrangement is given by :-

$$\frac{(n-1)!}{2}$$

*Combination: - (order is not important)

The no. of ways in which r things can be selected out of n different things is given by

n Cr

i) $$nCr = \frac{nPr}{r!}$$

ii) $$n\,C_r = \frac{n!}{(n-r)!\,r!}$$

Properties:-

1) $n\,C_n = 1$
2) $n\,C_o = 1$
3) $n\,C_1 = n$
4) $n\,C_r = n\,C_{n-r}$
5) $n\,C_{n-1} = n\,C_1 = n$
6) The number of ways of selecting 1 or more things out of available n different things is given by ;-

   $nC1 + nC2 + ----- + nCn = 2n-1$

   $nCo + nC1 + nC2 + ----- nCn = 2n$ ← (when things can be zero)
7) $nCr + nCr-1 = n+1Cr$
8) The no. of ways of selecting one or more things out of available n things of which n1 are of same kind, n2 are also of same kind and so an is given by :-

   (Refer Pg. 147 NB)

   $(n1+1) * (n2+1) * (n3+1) * ----- -1$

*Application of PnC in Geometry:-

1) Number of parallelograms:-

   Set of m parallel lines intersect another

   Set of n parallel lines

   No. of parallelograms = $mC2 * nC2$

2) No. of lines / segments:-

   Case I] Available -> m non – collinear points

No. of Lines = mC2

Case II] Available -> m points of which n are collinear

Direct Method:-

m-n C2 + (m-nC1 * nC1) + 1

(Refer on pg. 148)

Indirect Method:-

mC2 – nC2 + 1

3) No. of triangles:-

Case 1] m non-collinear points

No. of triangles = mC3

Case II] m points of which n are collinear

Direct method:-

(m-n C3) + (m-n C2 * nC1) + (m-n C1 * nC2)

Indirect method:-

mC3 – nC3

4) No. of diagonals:-

An n sided polygon will contain n points (non-collinear)

Total no. of lines = n C2

(-) no. of sides = (-) n

No. of diagonals = nC2 – n

5) Points of intersection:-

i) Maximum no. of points of intersection for n different lines = nC2

ii) Maximum no. of points of intersection for n different circles.

= no. of different pairs of circle * 2 points of intersection per pair

= nC2 * 2

iii) Maximum no. of points of intersection for n different lines and m different circles.

I) 2 lines : nC2
II) 1 line & 1 circle : nC1 * mC1 = 2
III) 2 circles : mC2 * 2
= (nC2 * 1) + (nC1- * mC1 C * 2) + (mC2 * 2)

When n things are to be arranged in r place taking all at a time, use should solve this not place wise but thing wise.

1^{st} thing can be arranged in any of the r places. rC1

Similarly, for each thing, no. of arrangements = rC1

Total no. of arrangements = rC1 * rC1 * ----- n times

$$\frac{rC1}{n\ 1}$$

(n-1) Pr + r * (n-1) P (r-1) = nPr

(Not important but learn it)

Number of side of polygon = n

Then no. of diagonals = $\frac{n(n-3)}{2}$

Sequence and Series –
Arithmetic & Geometric Progression

∗ Sequence:-

When numbers are listed in a specific order and there is a fixed rule/pattern using which we can estimate the next number in the series, then such a list of number is called a sequence.

In this chapter, we will study mainly 3 such patterns of sequence, namely;

i) Arithmetic progression (A.P)
ii) Geometric progression (G.P)
iii) Harmonic progression (H.P)

∗ Arithmetic progression:-

i) When difference between 2 consecutive numbers of a sequence is constant/ some, then such a sequence is called an A.P.

ii) The terms of this sequence are denoted by t_n (n_{th} term of the sequence)

1st term = t_1 = a (t_1 is denoted by a)

Common difference = d (dif between any 2 consecutive)

= t_n - t_n-1

iii) Terms of an AP may be written as follows:

a, a+d, a+2d, a+3d, a+4d,..........

t_1 , t_2, t_3, t_4, t_5,........

iv) Nth term of an AP is given by

tn= a+(n-1)d

v) Sum of first n terms of an AP is given by

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

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

∗ Relation between Sn and tn

sn - sn_{-1}=tn

∗ Inserting arithmetic means between 2 numbers.

If you asked to insert n arithmetic means between 2 nos. are a and b, then the sequence of such nos. formed must be in Arithmetic progression.

i.e. a,n_1, n_2, n_3,......b should be in AP.

e.x insert 5 A.M between 35 and 95

35,45,55,65,75,85,95

a= 35, t_7=95

95= 35 + (n-1)d

60=6d (∵ n=7 :- 35,95,5 AMs)

∴ d=10

Sum of first n odd numbers:-

S=n^2

∗ <u>Geometric progression (G.P)</u>

i) If in a sequence, the ratio between any 2 consecutive terms is constant / same, then such sequence is G.P.

ii) The terms of this sequence are denote by tn (n_{th} term of the sequence)

iii) 1st term= t_1=a

Common ratio= r

= ratio between any 2 consecutive terms

= $\frac{t2}{t1}$=$\frac{t3}{t2}$=$\frac{tn}{tn-1}$

iv) The terms can do written as follows:-

a, ar, ar^2, ar^3,.......,t_{50}

v) Nth term of a GP is given by:-

tn= $a.r^{n-1}$

vi) Sum of first n terms of a GP is given by:-

i) Sn= $\frac{a\times(r^n-1)}{(r-1)}$ ...... if $|r|$>1

ii) Sn= $\frac{a\times(1-r^n)}{(1-r)}$ .... if $|r|$<1

iii) Sn= n×a .... if r=1

iv) Sn=a if r= -1 and n= odd

Sn=0 if r= -1 and n= even

* While considering terms which are in GP, use the following pattern.

(I) For odd number of terms:-

(Middle term = a)

i) $\frac{a}{r^2}, \frac{a}{r}$, a, ar, ar^2 ii) $\frac{a}{r}$, a, ar

II) For even number of terms:-

i) $\frac{a}{r^3}, \frac{a}{r}$, ar, ar^3 ii) $\frac{a}{r_5}, \frac{a}{r_3}, \frac{a}{r}$,ar,$ar^3$, ar^5

(middle 2 terms =$\frac{a}{r}$ and ar)

(Common ratio = r^2

* While considering terms in an A.P, use the following pattern:-

(I) <u>For odd numbers of terms:-</u>

i) a-2d,a-d,a,a+d,a+2d   ii) a-d,a,a+d

(II) <u>For even no. of terms. (middle 2 terms = a-d, a+d)</u>

i) a-3d, a-d, a+d, a+3d   ii) a-5d, a-3d, a-d, a+d, a+3d,a+5d

* <u>Inserting n GMs between 2 terms a and b.</u>

This means the resulting sequence of numbers starting with a, then n GMs and then b should be a G.P

* <u>Sum of an infinite GP</u>

When a GP has infinite terms, then Sn of such GP is given by:-

S∞= $\frac{a}{1-r}$ if $|r|$ < 1

S∞=∞   if $|r|$ > 1

* <u>Harmonic progression:-</u>

If in a sequence, the reciprocals of the given terms are in A.P. then the given terms are in H.P

e.x. 6,3,2,........ → H.P

$\frac{1}{6}, \frac{1}{3}, \frac{1}{2}$, ...→ A.P

* If we add, subtract, multiply, divide any constant from all terms of an AP, the resulting sequence will also be an AP.

* If the whole question is in general terms, we can solve by taking any values satisfying the given conditions and then solve.

Sets, Relation & Functions

∗ Sets:-

A set is a collection of well- defined distinct objects. Each object in a set is called element of the set.

Sets are denoted by capital letters.
Elements are denoted by small letters.

The no. of element in a set is known as cardinal number of the set.

n(A)= no. of elements in set A

∗ A set can be written in mainly 2 forms;

i) Roaster form or braces form:- (listing method)
   In this form, the list of all elements separated by commas is made and written within brace brackets
ii) Set –builder form (rule method/ property method) (or Algebraic form)
   In this form, instead of making a list of all the elements, we just write a rule or property which describes all the elements of the set.

∗ Types of Sets

1) Finite set with finite no. of elements.
2) Infinite set: set with infinite no. of elements
3) Singleton set: A set with only a single element.
4) Empty set: A set with 0(zero) elements
   Empty set is denoted by:-(Null set)
   - A= { } or
   - A= Ø

   Where, A is an empty set
   Note:-
   If a set A contains an element a , this is written as 'a belongs to A', denoted as 'a ∈ A'
5) Subset:
   When all the elements of 1 set belongs to another set then the 1^{st} set is subset of the 2^{nd} set
   Ex. If all elements of set A belong to set B, then A is a subset of B, denoted by A C_ B

6) Superset

If A is a subset of B, then B is a superset of A

If A C_ B then B $\geq$ A

7) Proper subset:

If A C_ B and A$\neq$ B, then A is a proper subset of B

Note:-

For a given set with n elements,

No. of subsets= 2^n

No. of proper subsets= 2^n-1

8) Power set of A

For a given set A, a set formed using all the subsets of set a is known as power set of A.

Denoted as P (A)

Ex. A= {1,2,3}

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

9) Universal set

A set which contains all the type of objects which are being considered is called universal set.

* Mathematical operation on set:-

i) Union of 2 sets:-

When 2 sets are combined together and all the different elements which are either in the 1st set or in the 2nd set are listed to form a new set.

This new set is the union of the 2 sets.

ii) Intersection of 2 sets:-

For given 2 sets, where a set is formed using all the common elements present in both the sets, such a set is called as intersection of 2 sets.

iii) Compliment of a set:-

For a given set, a set which contains all the elements which are set which are not in the given set but are in the universal set is called compliment of a set.

Notations & keywords:-

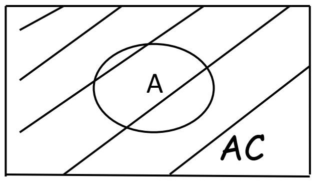
1) Union:-
   Notation A∪B, A + B
   Key words A or B, at least 1 of A or B, either A or B

2) Intersection:-
   Notation A ∩ B or AB
   Key words A and B, both A&B

3) Compliment:-
   Notation A^c or A' or $\bar{A}$
   Key words Not in A except A

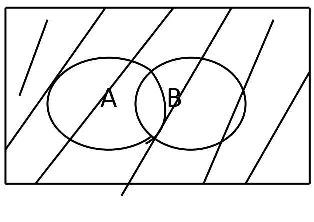
∗ Basic properties of cardinal number of above operations: (With Venn diagram)

i) N(A^c) = n(∪)- n(A)

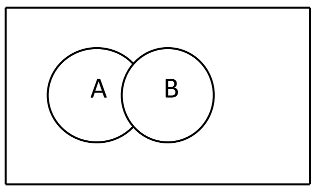


Demorgan's Law (2,3)

ii) n$(A \cup B)^c$= n($A^c \cap B^c$)
(Neither in A nov in B)

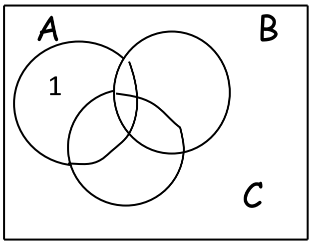


iii) n$(A \cap B)^c$= n($A^c \cup B^c$)



iv) n(A∪B)= n(A) + n(B)- n(A∩B)

v) $n(A \cup B \cup C) = n(A) + n(B) + n(C) - n(A \cap B) - n(A \cap C) - n(B \cap C) + n(A \cap B \cap C)$



A= 1+2+4+5

B= 2+3+5+6

C= 4+5+6+7

$n(A) + n(B) + n(C) - n(A \cap B) - n(A \cap C) - n(B \cap C) + n(A \cap B \cap C)$

$= [1+2+4+5] + [2+3+5+6] + [4+5+6+7] - [2+5] - [4+5] - [5+6] + [5]$

$= 1+2+3+4+5+6+7$

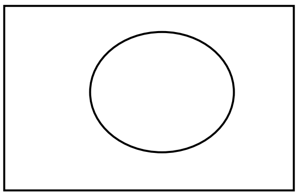
$= n[A \cup B \cup C]$

vi) If $A \subset B$, then

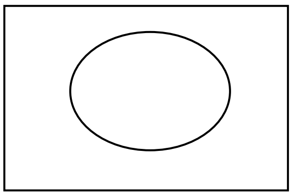
i) $A \cap B = A$

ii) $A \cup B = B$

iii) $A \cap B' = \emptyset$



$A \cap B = A$



$A \cup B = B$

vii) $A \cap B \subset A$ and $A \cap B \subset B$

$n(A \cap B) \leq A$ and $n(A \cap B) \leq B$

viii) $A \subset A \cup B$ and $B \subset A \cup B$

$n(A) \leq n(A \cup B)$ and $n(B) \leq n(A \cup B)$

ix) Equivalent set:

2 sets who have same number of elements.

Equal sets:-

2 sets who have same elements are called equal sets.

- All equal sets are equivalent sets however; a pair of equivalent sets may / may not be equal sets.

x) $A \cap A' = \emptyset$

$A \cup A'$ = universal set

$A \cap \emptyset = \emptyset$

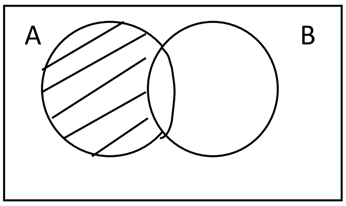
A∪ ∅= A

xi) n(A-B) = n (A) – n(A∩B)

Set A-B will have all the elements which are in set A but not in set B

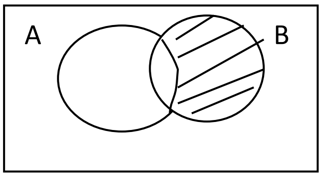
i.e. elements only in A

i.e. elements in A but not in B



(A-B)

n(B-A)= n(B) –n(B∩A



Note:-

i) LHS ⊂ RHS

Elements of LHS should be elements of RHS

ii) LHS ∈ RHS

LHS should be an element of RHS

∗ <u>Cartesian product set</u>

- Ordered pairs:- ex. (1,2), (3,4), (5,6)

When 2 elements are written in the form of pair (a, b) in a specific order and these pairs become an element of set, such pairs are called ordered pairs.

- Cartesian product sets:-

For 2 sets A&B, their product set is given by $A \times B = \{(a,b): a \in A, b \in B\}$

Ex. A= {1,2,3}

B={4,5}

A×B ={(1,4),(1,5),(2,4),(2,5),(3,4),(3,5)}

B×A= {(4,1),(4,2),(4,3),(5,1),(5,2),(5,3)}

• $A \times B \neq B \times A$ ....Not equal sets

• $n(A \times B) = n(A) \times n(B)$

• $n(B \times A) = n(B) \times n(A)$

$\therefore n(A \times B) = n(B \times A)$ ....equivalent sets

• If $A=\emptyset$ or $B=\emptyset$, $A \times B = \emptyset$

∗ <u>Relations & functions:-</u>

For 2 sets A & B, a set R is called a relation from A to B, if RC(A×B)
Similarly, if RC (B×A), then R is a relation from B to A.

* <u>Functions:-</u>

A set is said to be a functions from A to B if:-

i) It is a relation from A to B
ii) No.2 ordered pairs have same 1st component
iii) Each element of A has an image in B.
(All elements of A must be in ordered pair)

- <u>Image and pre image:-</u>

For any ordered pair (a, b) 'b' is called the image of 'a' whereas 'a' is called the pre image of 'b'.

i.e.(pre image, image)

Note:-

In a function, a pre- image cannot have more than 1 image.
However an image may have more than 1 pre- image.

4) Domain, Co-Domain & Range of function for a function $f: A \rightarrow B$
Set A= domain of function f
Set B= Co-domain of function f.
Range= Set of all elements of B which have pre-image in A.(set of 2nd components of ordered pairs of the function)

∗ <u>Types of functions:-</u>

1) <u>One- one function:- (injective f^n)</u>

A function where no 2 pre- image have same image

i.e. every 1st component has different image

i.e. 2nd component of ordered pair doesn't repeat

2) <u>Many- one function:-</u>

A function where at least 2 pre- images have same image i.e. 2nd component repeats.

3) <u>Onto functions (subjective functions)</u>

If in a functions from A to B, every element of B has a pre- image in A, then its an into function

Note:-

Bijective function:- A function which is one- one (injective) as well as onto (surjective) function is called bijective function.

* <u>Domain, co-domain, range of a relation:-</u>

For a R: A→B;

- Domain= Set of all first components of the ordered pairs of R.
- co-domain = Set B (2nd set)
- Range = set of all second components of the ordered pairs of R

Note:-

Range ⊆ co-domain

* Types of relations:-

Let S= {a,b,c,....}

R be a relation from s to s

i.e. R⊆ s× s

(1) Reflexive relation:-

R is said to be reflexive relation if it contains all the possible orderd pairs in the form of (a, a) such that $a \in s$.

(2) Symmetric Relation:-

For R to be symmetric relation, if $(a, b) \in R$, (b, a) should also belong to R.

(3) Transitive relation:-

For R to be transitive relation, if $(a, b) \in R$ & $(b, c) \in R$, then (a, c) should also belong to R.

(4) Relation which is reflexive, symmetric and transitive is equivalence.

∗ Inverse function

• A function is invertible if and only if it is a one-one and onto function (bijective functions).

• For a given function, y= f(x), this represents value of y according to the value of x.
[Value of y expressed in terms of x]

• When we invert y= f(x); we must get a function which represents value of x according to some value of y.
[Value of x expressed in terms of y]

• If y= f(x); then f(y) =x

∗ Composite Function:-

When 1 function is substituted in another function then the result is known as composite function

• fog (x) = f [g(x)]..... Substitute g(x) in f(x)

• gof (x) = g [f(x)]...... Substitute f(x) in g(x)

• fof (x) = f [f(x)]...... Substitute f(x) in f(x)

• gog (x) = g [g(x)]..... Substitute g(x) in g(x)

Note:-

If forgiven $f(x)$ and $g(x)$;

$f[g(x)] = g[f(x)] = x$;

Then $f(x)$ & $g(x)$ are inverse of each other.

$\{0, 1\}$ = set containing two elements 0&1.

$[0, 1]$= from 0 to 1 both inclusive

$(0, 1)$ = from 0 to 1 both exclusive

$(0, 1]$ = from 0 to 1 (0 exclusive 1 inclusive)

$[0, 1)$ = from 0 to 1 (0 inclusive, 1 exclusive)

Differentiation/ Derivations

For a given function f(x), derivative of f(x) means the rate of change in f(x) with respect to change in value of x.

i.e.

Derivative of f(x) = $\frac{change\ in\ f(x)}{change\ in\ x}$

Notation:

a) $\frac{d}{dx}$ ⇒ Derivative with respect to x

b) $\frac{d}{dx}$ y = $\frac{dy}{dx}$ ⇒ derivative of y w. r. t. x.

c) $\frac{d}{dx}$ f(x) = f^1(x)

Derivative is defined as the ratio of change in value of y w. r. t a very small (lose to 0) change in value of x.

∴ $\frac{dy}{dx}$= $\frac{change\ in\ f(x)}{change\ in\ x}$= $\frac{\Delta y}{\Delta x}$

Here, since the denominator, Δ in x, is very small we need to take the limiting value of the above function.

∴ $\frac{dy}{dx}$= $h \overset{lim}{\longrightarrow} 0$ f $\frac{f(x+h)- f(x)}{h}$ ------- (1)

Where,

h ⇒ change in value of x i.e. Δx= h

x= x + h

y= f(x) = f(x + h)

Δ in y = f(x + h) – f(x)

(Derivatives by definition/ derivatives by first principle) (∵ the above formula (1))

- Using concept of limits, derivatives of various types of functions are stated below:

* Standard Results:-

i) $\frac{dx}{dx}$= 1 $\neq$ 1

ii) $\frac{d}{dx}$ c = 0 $\neq$ 0

iii) $\frac{d}{dx}$ x^n= nx^{n-1} $\neq nx^{n-1}$ (algebraic functions,) (variable raised to constant)

iv) $\frac{d}{dx}$ a^x = a^x. log a (exponential function) (constant raised to variable)

v) $\frac{d}{dx}$ e^x= e^x

vi) $\frac{d}{dx}$ log x = $\frac{1}{x}$ (logarithmic function)

vii) $\frac{d}{dx}\sqrt{x}$= $\frac{1}{2\sqrt{x}}$

viii) $\frac{d}{dx}\frac{1}{\sqrt{x}}$ = $\frac{-1}{2x\sqrt{x}}$

ix) $\frac{d}{dx}\frac{1}{\sqrt{x}}$ = $\frac{-1}{2x}$

* Properties

i) $\frac{d}{dx}$ f(x) $\pm$ g(x) =$\frac{d}{dx}$ f(x) $\pm$ $\frac{d}{dx}$ g(x)

ii) $\frac{d}{dx}$ k. f (x) = k $\frac{d}{dx}$ f(x)

iii) $\frac{d}{dx}$ F(x). g(x) = f(x)= f(x) $\frac{d}{dx}$ g(x) + g(x) $\frac{d}{dx}$ f (x)

I.e. uv rule.

$\frac{d}{dx}$ u.v = u$\frac{d}{dx}$v + v$\frac{d}{dx}$u

iv) 'u/v' rule

$\frac{d}{dx}\frac{u}{v} = \frac{\left(v\frac{d}{dx}u - u\frac{d}{dx}v\right)}{v^2}$

v) <u>Function of a function (chain rule):-</u>

If y= f [g(x)],

Then let g(x) =u and y=f (u)

Diff. w.r.t.x	diff. w.r.t.u.
$\frac{d}{dx}$u = g'(x)	$\frac{dy}{du}$ = f'(u)

$\therefore \frac{dy}{dx} = \frac{dy}{du} \times \frac{du}{dx} = f'(u) \times g'(x)$

This method of solving is known as method of substitution.

A shortcut to all these steps is the chain rule where,

If y= f [g(x)]

$\frac{dy}{dx}$= = f'[g(x)] × g'(x)

vi) Variable raised to variable (use log) (NB pg. 134)

vii) <u>Parametric function:-</u>

When x and y are not directly related to each other, but both of them are related to a common variable t, then such functions are called parametric functions.

To find $\frac{dy}{dx}$, follow the below steps. (N.B pg. 135)

y= f (t), x= g (t)

$\frac{dy}{dt}$ = f'(t), $\frac{dx}{dt}$= g'(t)

du -f'

viii) <u>Implicit functions:-</u>

These are functions which are combinations of 2 variables together.

For finding $\frac{dy}{dx}$ follow the below steps.

i) Diff. both sides w.r.t.x (uv, u/v, chain rule)

ii) All the terms with dy/dx are kept on LHS whereas the other are taken to RHS.

iii) Take dy/dx common from LHS

iv) Find value of dy/dx

ix) <u>Multi –order Derivative:- (higher order derivative)</u>

Y= f(x)

$\therefore \frac{d}{dx}$ y = $\frac{d}{dx}$ f(x)

$\frac{dy}{dx}$= f'(x) …… first order derivative

$\frac{d}{dx}\frac{dy}{dx}$= $\frac{d}{dx}$ f'(x)

$\frac{d^2y}{dx^2}$ = f''(x) ……second order derivative

$\frac{d^3y}{dx^3}$ = f'''(x)…… third order derivative

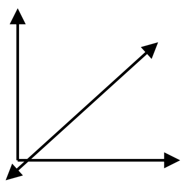
x) Some examples where taking log simplifies the function before taking derivative (pg.138)

∗ <u>Application of Derivatives:-</u>

1. <u>Slope of line/Gradient of curve:-</u>

• Slope of live = tan measurement of the angle made by the line with positive side of x axis.

Slope (m) = tanØ



• slope of line is the derivative of the equation of the line/curve.

2. <u>Marginal functions:-</u>

For any given f^n (cost, utility, profit function), Marginal f^n denotes change in the value of the f^n ÷ change in the no. of units

∴ They are ratios of change in the value of function to change in x (no. of units)
To find marginalf^n, we have to take derivative of the totalf^n.

3. <u>Maxima and Minima of functions:-</u>

a) Increasing/ decreasing function.

i) f^n is increasing f^n when x and y both increase or both decrease.

$\frac{dy}{dx} = \frac{\Delta y}{\Delta x}$

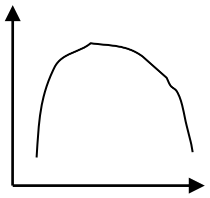
For an increasingf^n, dy/dx will be positive

ii) f^n is decreasing when x ↑ and y ↓ or x↓ and y↑.

For a decreasingf^n, dy/dx will be < 0.

b) Maxima/Minima of a f^n

(1)



In the diagram f^n is ↑ till A, and ↓ after point A.

$\frac{dy}{dx}$ Helps us to determine

Whether the function is increasing OR decreasing.

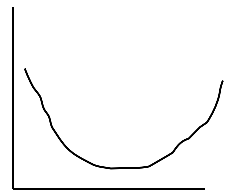
$\frac{d^2y}{dx^2}$ helps to determine the rate of increase in inc. f^n or rate of decrease in a dec. f^n.

In the above graph, $\frac{dy}{dx}$ =0 at point A.

Whereas d^2y will be negative (<0) since

And we get the point of maxima (point A) (fig.1)

(2)



In the diagram, f^n is ↓ till A and ↑ after A

However at A, it is neither inc. nor dec.

∴ $\frac{dy}{dx}$ = 0 at point A.

Further $\frac{d^3y}{dx^3}$ > 0 since f^n ↑ at ↑ rate and we get point of minima (point A).

Conclusion:-

To find minima/maxima of a function

Steps:-

1) Find dy/dx
2) Find values of x where dy/dx
3) Find $d^2y\ /dx^2$
4) Substitute values of x as per step 2 in d^2y/dx^2. If d^2y/dx^2 > 0, we get maxima at that value of x.
5) Substitute the above value of x in the original function to find minima/ maxima of f^n.

Unit I - Measures of Central Tendency

1) These measures give us a single representative value for the whole data.
They measure the central location of all the given observations.
These measures give us a value of uniformity in variability.

2) These measures are known as averages of first order

3) Data may be represented in 3 different forms
- Discrete series
- Ungrouped frequency distribution
- Grouped frequency distribution

4) There are mainly 3 measures of central tendencies:- Mean, Median and Mode

5) Mean, median, mode may be calculated using the formula in the below table

Particulars	**Discrete series**	**Ungrouped freq. distribution**	**Grouped freq. distribution**
Arithmetic mean (AM) ($\bar{x}$)	$\frac{\sum x}{n}$	$\frac{\sum fx}{\sum f}$	$\frac{\sum fm}{\sum f}$
Median (positional average)	$(\frac{n+1}{2})$ value	$(\frac{n+1}{2})$ value	$(\frac{n}{2})$ value = $L_1 + (\frac{\frac{n}{2} - cf_{\circ}}{f} \times h)$
Mode	Observation which repeats max no. of times	Observation with highest frequency	$L_1 + (\frac{f - f_{\circ}}{(f - f_{\circ}) + (f - f1)}) \times h$

- **For the above formula of median for grouped frequency**

Median class = class where the median will lie
= class with $(\frac{n}{2})^{th}$ value (class corresponding to closest cumulative frequency greater than n/2)

L_1 = Lower limit of median class
$C.f._0$ = c.f. of class above the median class
f = frequency of median class
h = height/ width of median class

- **For the above formula of mode for grouped frequency-**

Modal class = class where mode will lie
= class with highest frequency

L_1 = lower limit of modal class
f = frequency of modal class
f_0 = frequency of class above modal class
f_1 = frequency of class below modal class

Steps for Calculating Median for Grouped Frequency Distribution:
Step 1: Arrange the data in ascending order.
Step 2: Prepare c.f. (cumulative frequency less than type) column
Step 3: Calculate n/2
Step 4: Find Median Class = Class corresponding to closest cumulative frequency greater than n/2
Step 5: Find median using the above formula and the notations.

Properties of Central Tendencies

1) Arithmetic mean is dependent on all observations whereas median and mode are not dependent on all the observation.

2) Median is the least affected by the extreme values of the data.

3) For open end classification (where classes are open on either end), median is the most suitable measure of Central Tendency.

4) For study of zoology or study of shoe size, mode is the most suitable measure of Central Tendency.

5) The sum of the deviations of all the observation from its arithmetic mean is always 0 i.e. $\sum(x - \bar{x})= 0$

6) The sum of absolute deviations is minimum when the deviations of all observations are taken from their median.
i.e. $\sum |x - A|$ is minimum when A = median

7) The sum of squares of deviations of a set of observations has the smallest value , when the deviations are taken from AM.

8) Relation between mean, median and mode:-
- a) For positively skewed data (skewness towards right) –
  Mean > Median > Mode
- b) For negatively skewed data (skewness towards left) –
  Mean < Median < Mode
- c) For symmetric data (Zero Skewness) –
  Mean = Median = Mode

9) Also, for moderately asymmetrical data , an approximate relation between mean , median & mode, is given by-

Mean – Mode = 3(Mean – Median)

10) Mode is not rigidly defined. It is an ill-defined measure of central tendency. Some distribution may have multi modes, while some may not have any mode.

Mode is the most appropriate measure of central tendency for ordering particular design of a cloth showroom or for ordering shoe sizes.

Mode has no mathematical property

∴ It cannot be treated algebraically

11) Of the 3 central tendencies, mean is the most affected by sampling fluctuations, whereas median is not much affected by sampling fluctuations.
Mode is little affected by sampling fluctuations.

12) Of the 3 central tendencies, mean is the most used C.T.
However, mode is the most popular measure of C.T.

13) Effect of change of origin and scale on central tendencies:
All the measures of central tendency are affected by both, change of origin as well as scale

i.e. if	$y = ax + b$	or	$y = \frac{x-a}{b}$
then	$CT_y = (a \times CT_x) + b$	or	$CT_y = \frac{CTx - a}{b}$

All 3, mean, mode, median satisfy a linear relationship between 2 variables.

14) AM ($\bar{x}$) = $(A+ \frac{\Sigma fi\ di}{\Sigma f} \times C)$ ------ step deviation method
A= Assumed mean (AM= assumed mean + arithmetic mean deviation of terms)
$d_i = \frac{xi-A}{c}$
c= class length

15) Formula of mode is applicable even if classes are of unequal as well as equal width.

16) Moving average are the average of a series of overlapping averages, each of which is based on certain number of items within a series. M.A are used for smoothening a time series.

Arithmetic mean, Geometric mean, Harmonic mean

i) A.M (Refer table on 1st page)

ii) G.M (Geometric mean):-

a) For Discrete Series-

G.M= $(x_1 \times x_2 \times x_3 \ldots\ldots\ldots)^{1/n}$

b) For Ungrouped frequency distribution :-

G.M = $[(x_1{}^{f_1}) \times (x_2{}^{f^2}) \times (x_n{}^{f_n})]^{1/\Sigma f}$

c) Grouped frequency distribution :-

G.M = $[(m_1{}^{f_1}) \times (m_2{}^{f^2}) \times \ldots\ldots \times (m_n{}^{f_n})]^{1/\Sigma f}$

iii) H.M. (Harmonic mean):-

H.M is the reciprocal of the arithmetic mean of the reciprocals of all observations

a) Discrete series :-

H.M= $\frac{n}{\Sigma(\frac{1}{x})}$

b) Ungrouped frequency :-

H.M= $\frac{\Sigma f}{\Sigma(\frac{f}{x})}$

c) Grouped frequency distribution :-

H.M = $\frac{\Sigma f}{\Sigma(\frac{f}{m})}$

Properties of AM, GM and HM

1)	Relation bet^n AM, GM and HM:-	AM ≥ GM ≥ HM
	For any number of distinct observations,	AM > GM > HM
	If all the observations are identical,	AM = GM = HM

2) GM^2= AM × HM

∴ GM = $\sqrt{AM \times HM}$

i.e. The GM of all observations is the GM of AM and HM of all observations.

3) GM= $(x_1 \times x_2 \times x_3 \times ......x_n)^{1/n}$

$\therefore$ log GM = $\frac{1}{n}$ log $(x_1 \times x_2 \times x_3 \times ......x_n)$

$\therefore$ log GM = $\frac{1}{n}$ (logx_1 + logx_2+.......+logx_n)

$\therefore$ log GM= $\frac{\sum \log x}{n}$ ....(for discrete series)

$\therefore$ log GM = AM of log of all observations

i.e. log value of GM of all observations is equal to AM of log value of all observations.

$\therefore$ log GM = $\frac{\sum f \log x}{\sum f}$ (for group frequency)

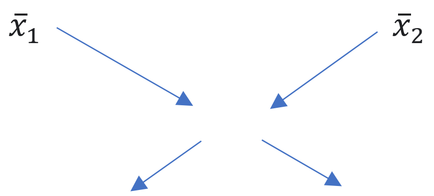
4) Combined Arithmetic mean:- (Grouped mean/ Pooled mean)

A group with n_1 observations has mean as $\bar{x}_1$ whereas another group with n_2 obs. has mean as $\bar{x}_2$. When the 2 groups are combined together, the combined mean of the whole group is given by:-

$$\bar{x}_{1\,2} = \frac{(n_1 \times \bar{x}_1)+(n_2 \times \bar{x}_2)}{n_1+n_2}$$

Shortcut for finding ratio between n_1 and n_2 (given $\bar{x}_1$, $\bar{x}_2$ *and* $\bar{x}_{1\,2}$ *are available*)

Butterfly Method-



$\bar{x}_{1\,2}$

$\bar{x}_2 - \bar{x}_{12}$ $\bar{x}_{12} - \bar{x}_1$

$$N_1 : N_2 = \bar{x}_2 - \bar{x}_{12} : \bar{x}_{12} - \bar{x}_1$$

5) Combined Harmonic Mean:-

A group with n_1 observations and H_1 as H.M is combined with another group with n_2 obs. and H_2 as H.M, then combined H.M is given by:-

$$H_{12} = \frac{n_1 + n_2}{\frac{n_1}{H_1} + \frac{n_2}{H_2}}$$

6) Use of Harmonic Mean:-

Whenever 1 thing is constant and we have to find average of the other thing which varies, (other thing is itself an average), then we'll always use harmonic mean.

Ex.1) For same distance travelled, if there are multiple speeds given, then average speed = harmonic mean of all the speeds.

Ex.2) You have a fixed amount of ₹ with you market 1 offers apples @ ₹ x_1 per dozen whereas market 2 offers apples @₹x_2 per dozen then the average price of apples of the 2 market is the harmonic mean of x_1 & x_2.

7) For calculating average of rates, ratios, population growth and percentage, we usually use GM, sometimes HM is also used.
GM is useful in construction of index number.
AM is the most stable of all measures of C.T.

8) Majority of the scenarios, arithmetic mean is the most used measures of central tendency.

9) If z = x.y (∵ x, y, z all are a set of values)
Then GM (z) = GM(x) × GM(y)

If $z = \frac{x}{y}$, then GM(z) = Gm(x) / Gm(y)

10) GM is not much affected by sampling fluctuations.

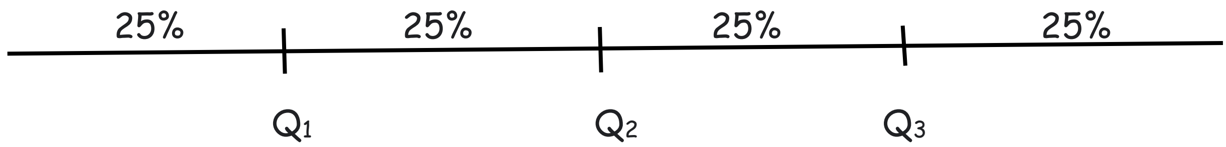
11) HM is good substitute to weighted average.

Partition Values

1) Partition values are values which divide the whole data in n equal parts.

2) There are mainly three types of Partition Values-
a) Quartiles b) Deciles c) Percentiles

3) **Quartiles:-** i) Divide the data in 4 equal parts.
ii) There are 3 quartiles- Q_1, Q_2, Q_3
iii) Each part will consist 25% of the total observations



iv) There are 25% observations less than and 75% observations more than Q_1
v) There are 50% observations less than and 50% observations more than Q_2

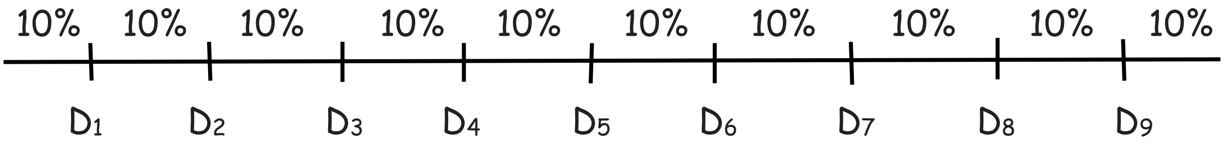
vi) There are 75% observations less than and 25% observations more than Q_3

vii) The limits for the middle 50% observations are given by Q_1 (Lower Limit) & Q_3 (Upper Limit)

Deciles:- i) Divide the data in 10 equal parts.

ii) There are 9 deciles- $D_1, D_2, D_3, \dots\dots, D_9$

iii) Each part will consist 10% of the total observations



iv) There are 20% observations less than and 80% observations more than D_2

v) There are 50% observations less than and 50% observations more than D_5

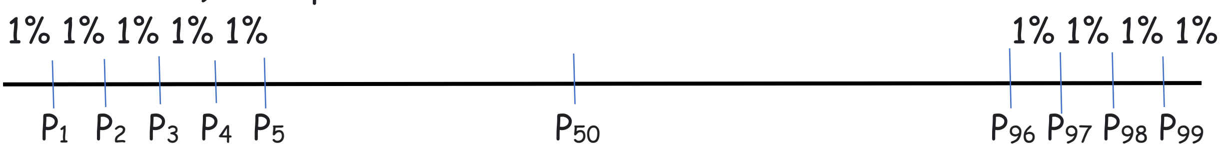
vi) There are 70% observations less than and 30% observations more than D_7

vii) The limits for the middle 40% observations are given by D_3 (Lower Limit) & D_7 (Upper Limit)

Percentiles:- i) Divide the data in 100 equal parts.

ii) There are 99 percentiles- $P_1, P_2, P_3, \dots\dots, P_{99}$

iii) Each part will consist 1% of the total observations.



iv) There are 35% observations less than and 65% observations more than P_{65}

v) There are 50% observations less than and 50% observations more than P_{50}

vi) There are 55% observations less than and 45% observations more than P_{55}

vii) The limits for the middle 30% observations are given by P_{35} (Lower Limit) & P_{65} (Upper Limit)

viii) The minimum value above which 40% observations lie is P_{60}

4) Formulae for calculating Partition Values:

Particulars	Discrete series	Ungrouped	Grouped
Quartiles (Q_k)	$\frac{k(n+1)^{th}}{4}$ value	$\frac{k(n+1)^{th}}{4}$ value	$(\frac{k(n)}{4})^{th}$ value $= L_1 + (\frac{\frac{k(n)}{4} - cf_0}{f}) \times h$

Deciles (D_k)	$\frac{k(n+1)^{th}}{10}$ value	$(\frac{k+n}{10})^{th}$ value	$(\frac{kn}{10})^{th}$ value $=L_1+ (\frac{\frac{k(n)}{10} - cf_0}{f}) \times h$
Percentiles (P_k)	$\frac{k(n+1)^{th}}{100}$ value	$\frac{k(n+1)^{th}}{100}$ value	$(\frac{kn}{100})^{th}$ value $=L_1+ (\frac{\frac{k(n)}{100} - cf_0}{f}) \times h$

Note- These formulae are very similar to formulae for calculating Median and hence the steps for calculations are also same.

5) **Note:-**

e.g : For discrete series, let's assume n=7

then Q_2 = $\frac{2(7+1)^{th}}{4}$ value

= 4^{th} value

Here, 4 is called the Frequency/Rank of Q_2 where as 4^{th} value is Q_2

6) Relation between different Partition Values:-

i) $Q_2 = D_5 = P_{50}$ = Median

ii) $Q_1 = P_{25}$, $Q_3 = P_{75}$

iii) $D_2 = P_{20}$, $D_7 = P_{70}$, etc

6) **Graphical method to calculate partition values:-**

i) To find partition values graphically, we plot the variable on the x axis whereas the cumulative frequency (less than type) on the y axis against the Upper limits of all the classes

ii) For finding Q_k, the abscissa of the point on the less than ogive (c.f. curve less than type) whose ordinate is equal to rank of Q_k is the value of $Q_k \backslash D_k \backslash P_k$

(Abscissa = x coordinate,
Ordinate = y coordinate)

Note: The abscissa of the point of intersection of the less than ogive and more than ogive is Q_2 or median or D5 or P50 (ordinate = $\frac{n}{2}$; rank of median)

7) Quartiles are used for measuring central tendency, dispersion and skewness.

8) Values of median, mode can be determined graphically and also quartiles, deciles, percentiles but not mean.

9) Quartiles are used for measuring C.T dispersion and skewness.
Quartiles are used in Bowley's formula and NOT in Person's formula of skewness.

10) Calculation of quartiles, deciles, percentiles can be obtained graphically from ogive.

Unit II - Measures of Dispersion

1) Measures of dispersion are used to calculate the amount of variation in the data. Dispersion is the amount of deviation of the obs. usually from an appropriate measure of central tendency.
Deviation may be positive, negative or zero.

2) More the variation in data, less is the consistency and higher is the risk, whereas, less variation means more consistency and low risk.

3) Measures of dispersion are known as averages of 2nd order.

4) There are 4 measures of dispersion, namely:
a) Range b) Quartile Deviation (Q.D.) c) Mean Deviation (M.D.) d) Standard Deviation (S.D.)

Absolute measures of Dispersion

Particulars	**Discrete series**	**Ungrouped freq. distribution**	**Grouped freq. Distribution**
Range	Highest - lowest value	Highest - lowest value	Upper limit of last class - lower limit of first class
Quartile deviation (QD) (semi- inter quartile range)	$\frac{Q3-Q1}{2}$	$\frac{Q3-Q1}{2}$	$\frac{Q3-Q1}{2}$
Mean deviation (mean of absolute deviations)	$\frac{\sum\lvert x-A\rvert}{n}$	$\frac{\sum f\lvert x-A\rvert}{\sum f}$	$\frac{\sum f\lvert m-A\rvert}{\sum f}$
Standard deviation (Root mean squared deviation)(δ)	$\sqrt{\frac{\sum(x-\bar{x})^2}{n}}$ Or	$\sqrt{\frac{\sum f(x-\bar{x})^2}{\sum f}}$ Or	$\sqrt{\frac{\sum f(m-\bar{x})^2}{\sum f}}$ Or

	$\sqrt{\frac{\sum x^2}{n} - \left(\frac{\sum x}{n}\right)^2}$	$\sqrt{\frac{\sum fx^2}{\sum f} - \left(\frac{\sum fx}{\sum f}\right)^2}$	$\sqrt{\frac{\sum fm^2}{\sum f} - \left(\frac{\sum fm}{\sum f}\right)^2}$

NOTE: Variance = $(\mathrm{S.D})^2$

5) The above measures are all absolute measures of dispersion i.e. they are unit dependent.

6) Standard deviation of first n natural numbers is given by:-

$\delta_x = \sqrt{\frac{n^2-1}{12}}$

7) Since absolute measures of dispersion are unit dependent, they are not comparable. To make these measures comparable, we form relative measures of dispersion (unit free)

Relative Measures of Dispersion

Co-efficient of range	= $\frac{H.V-L.V}{H.V+L.V} \times 100$
Co-efficient of Q.D.	= $\frac{Q3-Q1}{2}$ / $\frac{Q3+Q1}{2}$ = $\frac{Q3-Q1}{Q3+Q1} \times 100$ Or $\frac{QD}{Q2} \times 100$
Coefficient of M.D	= $\frac{M.D\ above\ A}{A} \times 100$
Coefficient of variation (C.V)	= $\frac{\delta x}{x} \times 100$

8) These relative measures are unit free.

9) CV is the most used relative measure of dispersion. Higher the CV, lower is the consistency whereas lower the C.V, higher is the consistency.

10) Standard deviation is the most used measure of dispersion.
For open end classification, quartile deviation is used.

11) For moderately skewed data, an approximate relation bet^n QD, MD and SD is as follows: -

QD:MD: SD= 10:12:15 (SD>MD>QD>)

12) Combined mean & combined S.D

$\bar{x}_{12} = \frac{n1\,\bar{x}1 + n2\,\bar{x}2}{n1+n2}$

$r_{12} = \sqrt{\frac{n1\delta1^2 + n2\delta2^2 + n1d1^2 + n2d2^2}{n1+n2}} = \sqrt{\frac{n1(\delta1^2 + d1^2) + n2(\delta2^2 + d2^2)}{n1+n2}}$

Where,

$d_1 = \bar{x}_1 - \bar{x}_{12}$

$d_2 = \bar{x}_2 - \bar{x}_{12}$

n_1 ,$\bar{x}_1$, δ_1= n, mean & SD of 1st groups.

n_2 ,$\bar{x}_2$, δ_2 = n, mean & SD of 2nd groups.

13) Effect of change of origin & change of scale:-

All the absolute measures of dispersion are affected by change of sale but not by change of origin.

(Note:- Theoretically measures of dispersion are always +ve)

i.e. if y = ax + b or $y = \frac{x-a}{b}$

then M.Dy = MDx × $|a|$ or MDy = $\frac{MDx}{|b|}$

14) Q.D and M.D have no mathematical properties.

15) For any 2 given numbers, SD is always half of range.

$\delta = \frac{1}{2} \times Range$

16) And also, mean deviation is always half of range

M.D = $\frac{1}{2}$ ×range

17) For any 2 numbers, MD = δ

18) QD and MD are not much affected by sampling fluctuations.

19) Variance of sum of two random variables is the sum of their variances i.e.

V(x + y) = V(x) + V(y)

20) Mean deviation is the least when deviations are taken from Median.

21) Criteria for ideal Measure of C.T & Dispersion: -

- Properly defined
- Easy to comprehend
- Simple to compute

- Based on all obs.
- Unaffected by (or least affected by) sampling fluctuations and extreme values
- Should have some desirable mathematical properties.

22) Frequencies are also called weights.
Simple averages is also called unweighted averages
In simple average, each value is considered only once while in weighted avg. each value is considered as many times as it occurs.
Simple & weighted avg. equal when all weights are equal.

23) A distribution is said to be symmetrical when the frequency rises & falls from the highest value in the equal proportion.

Probability

Experiment: A performance that produces certain result.

Events: The results (outcomes) of a random experiment are known as events.

i) Simple/ elementary

Event cannot be decomposed into further events

ii) Composite/ compound

Event that can be decomposed into 2 or more events.

iii) Mutually exclusive events

2 or more events are said to be mutually exclusive if they cannot occur simultaneously.

(i.e. A and B are mutually exclusive if:-

$n(A \cap B) = 0$

$\therefore P(A \cap B) = 0$

iv) Mutually exhaustive events

2 events are said to be mutually exhaustive if atleast 1 of the 2 events will occur

$n(A \cup B) = n(S)$ → S= Sample space

$P(A \cup B) = 1$

v) Equally likely event

2 events are said to be equally likely when the probability of their occurrence is equal.

Probability:-

For a given experiment, the probability of an event is the chance of that event occurring in the future when the experiment will be conducted.

It is the ratio of the number of ways in which event can occur to the number of ways in which the experiment can occur

It is denoted by:-

$$P(A) = \frac{\text{no. of outcomes favourable to A}}{\text{total no. of outcomes of exp.}}$$

$$= \frac{n(A)}{n(S)}$$

Independent events

2 events are said to be independent if the occurrence or non-occurrence of 1 event does not affect the occurrence or non-occurrence of the other event.

$\therefore$ $P(A \cap B) = P(A) \times P(B)$

Note:-

If A & B are independent; following are also independent:-

i) A & B'
ii) A' & B
iii) A' & B'

Properties:-

i) $P(A) = \frac{n(A)}{n(S)}$

ii) $P(A') = 1 - P(A)$ .... $[P(A') = P(A') = P(\bar{A})]$

iii) $P(A \cup B) = P(A) + P(B) - P(A \cap B)$

iv) $P(A \cup B \cup C) = P(A) + P(C) + P(A \cap B \cap C)$
- $P(A \cap B) - P(A \cap C) - P(B \cap C)$

v) If A & B are exclusive;
$P(A \cap B) = 0$

vi) If A&B are exhaustive

A(A∪B)=1

vii) If A & B are independent

P(A∩B) = P(A)× P(B)

viii) P(only A) = P(A but not B)

= P (A∩B')

= P(A-B)

= P(A) – P (A∩B)

ix) P (only B)= P(B but not A)

= P(B∩A')

= P(B – A)

= P(B) – P(A∩B)

x) If A and B are exclusive as well as exhaustive,

P(A) + P(B) = 1

xi) If A,B,C are exclusive and exhaustive,

P(A) + P(B) + P(C)= 1

xii) If A is a sure event, P(A)= 1

xiii) If A is an impossible event, P(A)=0

xiv) De Morgan's law:-

$P(A \cup B)^c = P(A^c \cap B^c)$

$P(A \cap B)' = P(A' \cup B')$

$P(A' \cap B)' = P(A \cup B')$

And so on ....

$P(A \cup B') = P(A' \cap B)^C$

= P $(only\ B)^c$

= 1- P (only B)

= 1- [P(B)- P(A∩B)]

xv) P(A+B) = P(A∪B)= P(A) + P(B) – P(A∩B)

xvi) P(A-B)= P(only A) = P(A) - P(A∩B) = P(A∪ B')

= P(A but not B)

xvii) P(A.B) = P(A∪B)

xviii) P(A/B)= conditional probability

∗ <u>Conditional probability:-</u>

• Probability of an event if a condition (another event) is met is known as conditional probability.

• Probability of event A if event B occurs is the conditional probability of event A given by:- P(A/B)

•P(A/B) means probability of A if probability of B has already occurred.

•P(B/A) =$\frac{P(A\cap B)}{P(A)}$

•P(A/B)= $\frac{P(A\cap B)}{P(B)}$

•P(A∩B)= P(A)×P(B/A)

P(A∩B)= P(B) ×P(A/B)

•P(A) × P(B/A)= P(B) × P(A/B)

•If A and B are independent,

P(A/B) = P(A)

$\therefore \frac{P(A \cap B)}{P(B)} = P(A)$

$\therefore P(A \cap B) = P(A) \times P(B)$

$\therefore P(A/B) = P(A)$

$P(B/A) = P(B)$

∗ <u>Probability distribution</u>

When the total probability of 1 is distributed in different probabilities for different outcomes of an experiment, such a table showing probabilities against different events is called probability distribution

Ex. When a coin is tossed, prob. Distribution:-

X	P
Head	$1/2$
Tail	$1/2$
	1

Ex. When a dice is rolled, prob. distribution:-

X	P
1	$1/6$
2	$1/6$
3	$1/6$
4	$1/6$
5	$1/6$
6	$1/6$

iii) When 2 cards are drawn from a deck prob. distribution.

X	P	
Both face cards	12_{C_2}/52C2	= 11/221
1 face /normal	12_{C_1} ×40c1/52c2	= 80/221
Both normal	40_{C_2}/52c2	= 130/221

* Random variable

• When a variable is used to express all the different possible outcomes of a random experiment such a variable is called random variable.

• Usually when a random variable is used, the outcomes of an experiment are mapped to different numbers.

• There are mainly 2 types of random variables:-

a) Discrete r.v.

b) Continuous r.v.

1) Discrete random variable:-

• When values which random variable can assume – jumps from 1 point to another (no smooth increase), then it is discrete r.v.

• Values of discrete r.v. can be counted or they may be count ably infinite

Ex. no. of stars in the sky,

no. of heads when coin is tossed n times

no. of student in a class.

• Probability of each individual value of r.v. exists and can be calculated.

2) Continuous random variable:-

• When values which the random variable can assume increase continuously from 1 point to another, i.e. there will be infinite values between 2 point intervals; it is continuous r.v.

• These are never counted, but measured to an approximate value.

• Probability is always calculated within a range of values of random variables and not at individual value of r.v.

• Probability at a point does not exist probability within a range exists.

Ex. exact mass/height of a person.

* Probability Functions:-

i) When probability of a random variable is defined by a function, then such a function is known as probability function.
ii) Probability function of a discrete r.v. is known as probability mass function.
iii) Probability function of a continuous r.v. is known as probability density function.

* Expected value and variance of a discrete random variable:-

i) When we calculate mean of the different values a discrete r.v. can assume, then this mean is known as expected value of the variable.
ii) Mean= Expected value ($\in (x)$)= $\sum$xp
iii) Variance of a D.R.V:-

Var (x) = $\in [x - \in (x)]^2$

= $\in (x^2)$- $[\in (x)]^2$

= $\in x^2$p- $(\sum xp)^2$

* Properties of mean & variance of a random variable

- Expected value of a constant

$\in$(k)= k

- Expected value of (x + y), (x-y)

$\in$(x+y)= $\in$(x) + $\in$(y)

$\in$(x-y)= $\in$(x) - $\in$(y)

- $\in$(ax + b) = a. $\in$(x) + b
- var (ax + b)= a^2. Var (x)
- $\in$(x.y) = $\in$(x) × $\in$(y).... if x,y are independent

* Odds in favour, odds against:-

$$\frac{\text{Favourable outcomes to } A}{\text{unfavourable outcomes to } A} \qquad \frac{\text{Unfavourable outcomes to } A}{\text{favourable outcomes to} A}$$

* • If n coins are tossed (OR if a coin is tossed n times);

n(s) = 2^n

• If n dice are thrown (OR if a dice is thrown n times);

n(S) = 6^n.

* Probability function & properties of discrete random variables & continuous random variable

i) Probability functions of D.R.V.

• f(x) which defines probability of a discrete random variable is known as probability mass function.

• $\sum f(x) = 1$

$F(x) \geq 0$ & $f(x) \leq 1$

ii) Probability f^n of C.R.V.

• f(x) which defines probability of a continuous random variable is k.a. probability density function.

• In C.R.V. probability cannot be calculated for each & every value of random variable as there are infinite values

∴ for C.R.V , we always find probability within a range of values of random variable

The probability of a continuous random variables is given by the area under the curve of the probability function.

To find area under any curve, we use integration of the equation of the curve.

For a given f(x) [Probability density function] where

a < x < b

$_c^d\int f(x)\,dx = 1$

P[c < x < d] ...... a < c , d < b

$_c^d\int f(x)\,dx = 0 \leq x \leq 1$

* Classical and statistical definition of probability

• $P(A) = \frac{n(A)}{n(S)}$ ...... classical definition /

. priori definition

$= \frac{favourable\ outcomes\ to\ event\ A}{total\ outcomes}$

In this definition, we can divide all the outcomes of the experiment in n number of mutually exclusive, exhaustive and equally likely outcomes.

$$P(A) = \frac{MA}{m} = \frac{\text{no.of mutually exc,exh.,equally likely events favourable to A}}{\text{Total no.of mutually exc.exh. and equally likely events}}$$

Limitations:-

- Applicable only when total no. of events is events is finite
- Can be used only when events are equally likely.
- Limited application –coin, dice, cards→ possible events known in advance

If no prior knowledge/ uncertainty→ def. is. not applicable

- Statistical / Relative frequency definition:-

i) Since one of the limitations of classical def is that n(s) should be finite, statistical def. applies when n(s) tends to infinity.

ii) As per statistical definition.

$$P(A) = \lim \frac{FA}{n}$$

Where FA= no. of favourable cases to event A

n= n(s)

Independence of events

- 2 events are said to be independent if

i) $P(A/B) = P(A)$

ii) $P(B/A)= P(B)$

$\therefore\ P(A\cap B)= P(A) \times P(B)$ ———— (1)

∗ Equation (1) is the necessary & sufficient condition for independence of 2 events.

- For 3 events A,B& C the 3 of them are said to be independent if:-
  - $P(A \cap B) = P(A).\ P(B)$
  - $P(A \cap C) = P(A).\ P(C)$
  - $P(B \cap C) = P(B).\ P(C)$
  - $P(A \cap B \cap C) = P(A).\ P(B).\ P(C)$

These are the 4 conditions- necessary and sufficient to conclude that A, B, C are independent.

∗ <u>Axiomatic or modern def. of probability:-</u>

When an event A is defined on the sample space (s) and the probability of an event is defined as a function of the sample space, this probability is known as axiomatic def. of probability.

It satisfies the following axioms:-

i) $P(A) \geq 0$ for every $A \subseteq S$ (subset)

ii) $P(S) = 1$

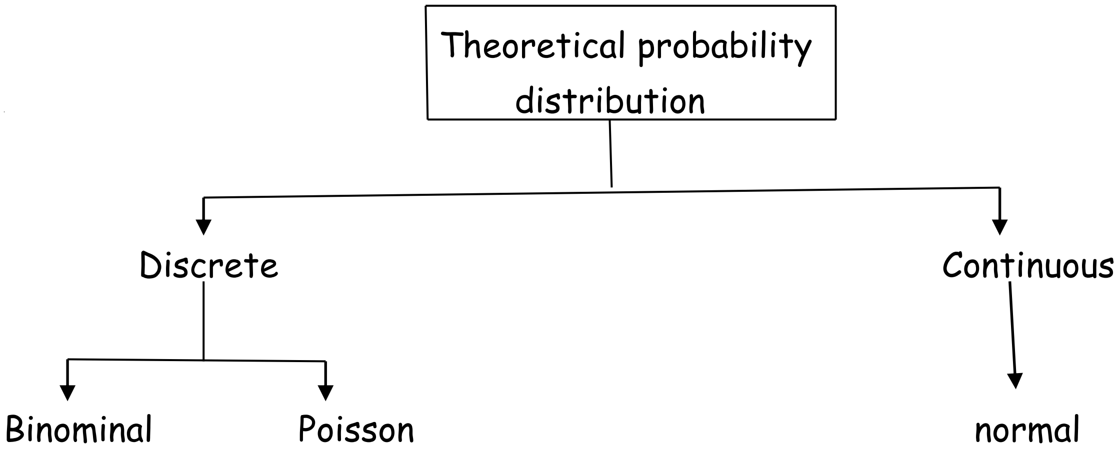
Expected value (mean) of continuous random variable:-

$\in(x) = \int xp\ dx.$

Variance of c.r.v:-

$$\in (x^2) = \int x^2 p\ dx - \left[\int xp\ dx\right]^2$$

Theoretical Distributions



i) <u>Binominal distribution :-</u>

- It is based on Bernoulli's trials.
- The characteristics of Bernoulli's trials are:-

i) Each trial results in 2 mutually exclusive and exhaustive outcomes. One of the outcome is denoted as a success & the other as failure

ii) The trials are independent of each other

iii) p= Probability of success in 1 trial
q= Probability of failure in 1 trial
∴ q= 1-p

iv) The no. of trials is a finite positive integer.

- The probability mass function of a binomial distribution is given as

$P[x= r] = {}^{n}C_{r} \,.\, P^{r} .\, q^{n-r}$

Where,

r= no. of successful trials reqd.

n= total no. of trials

p= probability of success in 1 trial

q= probability of failure in 1 trial

* Properties of binomial distribution:-

- Mean of a binomial distribution is given by

U= np

- Variance of a binomial distribution is given by

6^2= npq.

∴ 6 = $\sqrt{npq}$

- Binomial distribution is a discrete probability distribution with 2 parameters, n and p.

∴ Binomial distribution is a bi-parametric distribution.

- Binomial distribution may be uni-modal or bi-modal; depending on the value of n & p.

i) If (n + 1) p is not a perfect integer then

Mode = highest integer in (n + 1) p

ii) If (n + 1) p is a perfect integer then
Mode= (n + 1) p and (n + 1) p-1

- Variance = npq

∴ r^2 = uq

r^2 is always less than u.

- Variance will attain max value if p=q=$\frac{1}{2}$ & max value = $\frac{n}{4}$

- Values required to form complete probability distribution table are called as parameters.

For forming complete binomial distribution table we need 2 values n & p.

This is written as:-

x ~B (n,p)

:-[x follows binomial distribution with parameters n and p].

- If x ~B (n_1, p)

& y ~ B (n_2,p)

Then:- x + y ~ B (n_1 + n_2,p)

∗ Poisson distribution

i) It is an extension of binomial distribution
ii) It is a special case of binomial distribution where 'n' is a very large number and 'p' has a very small value.
iii) It is also a discrete random variable distribution.
iv) It is a uni- parametric distribution with 'm' as parameter
v) m= mean of the distribution
   m= n.p
vi) Variance (6^2) = np
   ∴ mean = variance
vii) The probability mass function for Poisson distribution is given by:-
   P[x=r] = $\frac{e^{-m} m^r}{r!}$,(e= 2.71828)
   , (m= mean = np)
viii) Mode:-
   Poisson distribution maybe uni-modal or bi-modal depending on the value of m
   i) If m is a perfect integer, then there are 2 modes, (m) and (m-1)
   ii) If m is not a perfect integer, then mode= highest integer in the value of m.
ix) If x ~ p (m_1) & y ~ p (m_2)
   Then x + y ~ p $(m_1 + m_2)$

• Skewness in Binomial & Poisson distribution

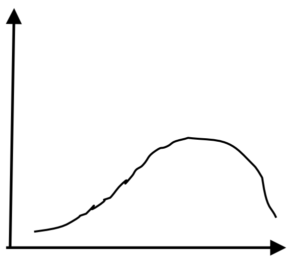
i) For binomial distribution

P[x=r] =$^nC_r\, p^r\, q^{n-r}$

1) If p > 0.5, q < 0.5

The value of probability mass function will be higher when the value of r is higher

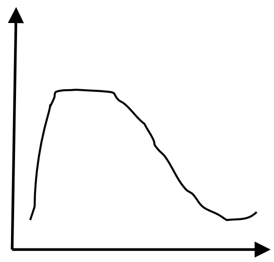
∴ Skewness of binomial distribution will be negative skewness (skewness towards left)



2) If P , 0.5; q > 0.5

The value of probability mass function will be higher when value of r is lower

∴ Skewness of binomial distribution will be positive skewness (towards right).



Conclusion

If p > 0.5 negative skewness (left)

If p < 0.5 positive skewness (right)

ii) Poisson distribution

- The value of p is very small in poisson (less than 0.5).
- Skewness is always positive (towards right)

* Normal Distribution (for continuous variables)

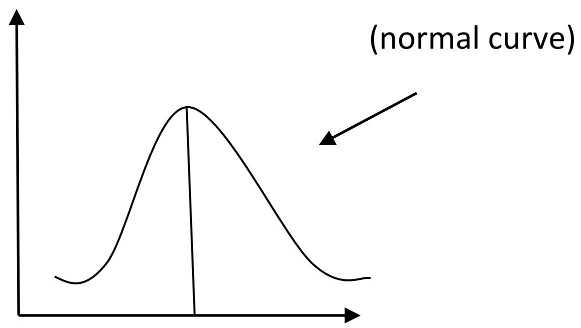
- It is also known as Gaussian distribution
- It is a continuous random variable distribution
- The probability density function of the normal distribution is given by:-

$$f(x) = \frac{1}{r\sqrt{2x}} * e^{-\frac{(x-u)^2}{2r^2}}$$ u =(mean)

$$f(x) = \frac{1}{r\sqrt{2x}} * \text{Exp}\left(-\frac{(x-u)^2}{2r^2}\right)$$

- Since normal distribution is a continuous random variable distribution, probability can never be found at a certain point.
  Probability is always calculated within a range

- For finding probability of a continuous r.v. are find area under the curve of the probability function.
- For finding area under the curve, we use definite integration.
- ∴ p[a<x<b] = $\int_a^b f(x) a^I x$
- The curve of the P.D.F. of normal distribution is a.k.a. normal curve.



- Normal distribution is a symmetric distribution.
  i) Mean = median = mode.
  ii) Skewness = 0.
- The shape of the normal curve is bell – shaped curve asymptotic to x-axis. (gets closer to x-axis on both ends, however never touches x-axis).
- Normal distribution is a bi-parametric distribution with 2 parameters {u(mean), r^2(various).}
  X ~ N (u, r^2)
- Additive property
  If X ~ N ($u_{1\,,r1^2}$) & y ~ N ($u_{2\,,r2^2}$)
  Then X + Y ~ N ($u_{1+}\,u_{2\ ,}\ r_1{}^2 + \ r2^2$)
- For finding values of probabilities of a normal distribution , a normal distribution table is used
  The normal distribution table is so formed such that u=0 and r=1.
- For using values of the normal distribution table we have to convert out normal variable (z), using change of scale such that mean of z=0 and r_2 = 1

  $$Z = \frac{x-un}{rx}$$

  This z is known as standard normal variable where,
  u_z = 0 and $r_{2=1}$
  The normal distribution table is also called z-table.

In exam, instead of giving the whole z table, certain values will be given in the question which may be given in 2 ways.

Ex. I) p[z = 0 to z = 1] = p [o < z < 1]= 0.3413

II) Ø(1) = p [-00 < Z < 1] = P [z < 1]= o.8413.

*Properties of normal distribution:-

* Since normal distribution is a symmetric distribution,

Mean = median = mode

- QD : MD : SD = 10 : 12 : 15
  (This is an approximate relation b/w measures of dispersion
- A more specific relation b/w MD & SD :-
  MD = 0.86
- Qualities of Normal Distribution
  Q1 = u – 0.675 r   Q.D. = 0.675 r
  Q3 = u + 0.675 r
  i.e. 1^{st} and 3^{rd} quartiles are 0.675r away from the mean on each side.

- For standard normal distribution,

U = 0, r = 1 , M.D. = 0.8,

Q1 = -0.675, Q3 =0.675

* Points of inflexion:-

The points of inflexion are the points where the normal curve changes its shapr from convex to concave or Concave to convex

The points of inflexion are given by :-

(u – r) and (u + r) .

- For standard normal distribution,
  Points of inflexion are -1 and +1.
- If z = 2, it means that the required value of X is 2 standard deviations on the right side of mean
  z = -1.2, means that required value of X is 1.2 standard deviations less than mean.
- Probability of x b/w l1 + r , u –r is 0.6826.
  (i.e. Probability of z b/w -1 and +1) i.e.68.26%

- Probability of X b/w u = 26 , u – 26 is 95.44%

  (I.e. Probability of z b/w -2, and +2) is 95.44%

- Probability of x b/w u + 36, u – 36 is 99.74%.

- Method usually applied for fitting a binomial distribution is known as:-

→ Method of moments.

∗ <u>Conditions for Poisson distribution:</u>

- Probability of finding success in a very small time interval (t, t + dt) is kt,

Where k (> 0) is a constant

- Probability of having more than 1 success in this time interval is very low.

- Probability of having success in this time interval is independent of t as well as earlier success.

∗ The p.m.f. of a random variable is denoted by f(x)

Whereas,

The cumulative probability is denoted by f(x)

$F(x) = P[X \leq x]$

X → Random Variable; x → read probability

Correlation and Regression

∗ Measures of central tendency and dispersion were single variable measures. Correlation and regression are two- variate measures.

Therefore in this chapter we'll be studying the movement of 2 variables together i.e. their co-movement.

∗ The variance of a variable x gives us the movement of x from its own arithmetic mean.

$$\text{Variance } ((6^2 x) = \frac{\sum(x-\bar{x})^2}{n}$$

$$= \frac{\sum(x-\bar{x})(x-\bar{x})}{n}$$

$$= \frac{\sum x^2}{n} - \left(\frac{\sum x}{n}\right)^2$$

*<u>Co-Variance:-</u>

The covariance between 2 variables x and y gives us the co-movement of the variables x & y, through which use can determine the nature of relationship between x and y.

$$\text{Co-variance } (cov_{xy}) = \frac{\sum(x-\bar{x})(y-\bar{y})}{n}$$

$$= \frac{\sum xy}{n} - \left(\frac{\sum x}{n}\right)\left(\frac{\sum y}{n}\right)$$

$$= \frac{\sum xy}{n} - (\bar{x})(\bar{y})$$

*<u>Nature of relationship:-</u>

● If x and y both move in the same direction (i.e. both increase or both decrease), then they have a positive relationship (direct relation)

● If x and y move in opposite direction (i.e. one increase and the other decreases), then they have negative relationship (inverse relation).

- If $(cov_{xy}) > 0$, positive relation between x & y.

  If $(cov_{xy}) < 0$, negative relation between x & y.

  If $(cov_{xy}) = 0$, no relation between x & y.

∗ Covariance of x & y (cov_{xy}) has no upper and lower limits (i.e. $-\infty < (cov_{xy}) < \infty$) Therefore we can only determine the nature of relationship but not the strength of relationship.

∗ Covariance is a unit dependent measure to make covariance unit free and to determine the strength of relationship, we use correlation co-efficient.

In correlation both x and y are inter- dependent

∗ Karl person's correlation co-efficient:

(Product moment correlation co-efficient)

$$r_{xy} = \frac{(cov_{xy})}{6x \times 6y} = \frac{n\sum xy - (\sum x)(\sum y)}{\sqrt{n\sum x^2 - (\sum x)^2}\sqrt{n\sum y^2 - (\sum y)^2}}$$

- $-1 < r\ xy < 1$

∗ Correlation coefficient is a unit free measure

∗ Spearman's rank correlation coefficient:-

- If we want to find correlation between 2 attributes, we usually use spearman's rank correlation coefficient.

- In spearman's rank correlation coefficient we assign ranks to the given attributes or variables in the following manner:-

1st rank is given for the observation with the highest value and so on.

Spearmen's rank correlation coefficient is given by:-

$$r = 1 - \frac{6\sum D^2}{n(n^2-1)}$$

Where, D= difference of ranks

∗ If while giving ranks , there are ties then give an average rank to all the same values and calculate spearman's rank correlation using the below formula:-

$$r = 1 - 6\left[\frac{\sum D^2 + \frac{1}{12}t_1(t1^2-1) + \frac{1}{12}t_2(t2^2-1) + \cdots}{n(n^2-1)}\right]$$

Where,

t_1, t_2,......=no. of times a rank is repeating

(Refer NB page 193 for question (9)

When

t= 2 → 0.5

t= 3 → 2

t= 4 → 5

t= 5 → 10

∗ Coefficient of concurrent deviations:-

In this method, we look at the directions of movement of the variables but not the quantum of movement

- Coefficient of concurrent deviations is calculated using the following steps

1) Assign a positive sign in front of a value of variable if it is greater than the precious value.

Similarly, assign a negative sign if the value of variable is less than the precious value.

(These signs indicate the directions of movement of x and y individually)

2) Prepare a column which shows the product of deviations (direction)

3) Calculate coefficient of concurrent deviations using the following:-

$$r= \sqrt{\frac{12(-m)}{m}}$$

Where,

c= no. of concurrent deviations

= no. of positive signs in product column

m= no. of deviations

= n-1

4) In the above formulas, r and (2c-m) will have same sign

Conclusion: - If c=m (i.e. all positive signs in product r= +1.
If c= 0 (i.e. all –ve signs in product)

r= -1

∗ <u>Effect of change of origin & change of scale:-</u>

- Covariance is affected by change of scale but not by change of origin.

If u= $\frac{x-a}{b}$ and v= $\frac{y-c}{d}$

Then, $\boxed{cov_{uv} = \frac{1}{b} \times \frac{1}{d} \times cov_{xy}}$

- Correlation co-efficient is neither affected by change of origin nor by change of scale.

u= $\frac{x-a}{b}$ v= $\frac{y-c}{d}$

$$\boxed{r_{uv} = \frac{|b||d|}{bd} \times r_{xy}}$$

i) If b and d have some sign, $r_{uv}=r_{xy}$
ii) If b and d have opposite sign , $r_{uv}= -r_{xy}$

∗ Coefficient of determination and no-determination :- (not very important – learn only from encircled formula part)

While studying a specific share in share market the variance in the market prices of that share is known as risk of that share.

i.e. risk of a share = variance (r^2)

- There are <u>2 components of this risk</u>

1) Systematic risk (external) (market)

2) Unsystematic risk (internal)

1) It is that part of the risk (that part of variance) which can be explained due to variance in the share market

It is determined using the correlation between share and the share market

Explained variance= $r^2 \times r^2$

$\therefore \boxed{r^2 = \frac{Explained\ variance}{total\ variance}}$

r^2 is known as co-efficient of determination.

2) Unsystematic risk is that part of the risk which cannot be explained due to external factors. It is due to the own performance of the company (internal factors)

- Unexplained variance = total variance – explained

$= r^2 - r^2 \times r^2$

$= r^2(1 - r^2)$

1-r^2= unexplained variance

Total variance

1-r^2 is known as co-efficient of non-determination

* Regression Analysis

- In correlation, we studied the co-movement of both variables where both of them were considered to be interdependent.

- In regression analysis, we will be studying the movement of 1 variable will be dependent on another).

- Regression analysis is divided in two parts:-

i) Regression co-efficient

ii) Regression equation

- Regression co-efficient

i) y on x (y is dependent on x):-

byx= $\frac{cov_{xy}}{r^2 x}$

$=\frac{n\sum xy - \sum x.\sum y}{n\sum x^2 - (\sum x)^2}$

'byx' gives us a ratio of change in y according to change in x

i.e. byx=$\frac{\Delta y}{\Delta x}$

ii) x on y (x is dependent on y)

bxy= $\frac{cov_{xy}}{6^2 y}$

= $\frac{n\sum xy-(\sum x)(\sum y)}{n\sum y^2-(\sum y)^2}$

'bxy' gives us a ratio of change in x according to change in y.

i.e. bxy=$\frac{\Delta x}{\Delta y}$

- Regression Equation:-

i) y on n

byx= $\frac{\Delta y}{\Delta x}$

∴ byx=$\frac{y-\bar{y}}{x-\bar{x}}$

∴ byx(x-$\bar{x}$)= y-$\bar{y}$

∴ y-$\bar{y}$ = byx(x-$\bar{x}$)........ Regression equation of y on x

ii) x on y

bxy= $\frac{\Delta x}{\Delta y}$

∴ xy= $\frac{x-\bar{x}}{y-\bar{y}}$

∴ x-$\bar{x}$ = xy(y-$\bar{y}$).......... regression equation of x on y.

Relation between r, bxy and byx

i) r^2= bxy × byx $\left[\left(\frac{cov_{xy}}{rx\, y}\right)^2 = \frac{cov_{xy}}{r^2 y} \times \frac{cov_{xy}}{6^2 x}\right]$

r= $\sqrt{bxy \times byx}$

ii) ∴ r is the geometric mean (G.M) of bxy & byx

iii) -1 <_ r <_ 1

∴ 0 <_ r^2 <_ 1

∴ 0 <_ bxy × byx <_ 1

iv) ∴ if 1 of the regression coefficient is more than 1, then the other will be less than 1.

v) $bxy = \text{r} \times \frac{rx}{ry}$ ∴ bxy= byx $\times \frac{r^2x}{r^2y}$

byx= r $\times \frac{ry}{rx}$ ∴ byx= bxy $\times \frac{r^2y}{r^2x}$

vi) The point of intersection of 2 lines is $(\bar{x}, \bar{y})$ (i.e. the 2 regression lines of x and y intersect at their arithmetic means)

vii) In a given equation of y on x , y= a+bx. coefficient of x= byx =b

In a given equation of x on y, x= a+by

Co-efficient of y= bxy = b

viii) Slope of y on x equation = byx

Slope of x on y equation = $\frac{1}{bxy}$

ix) r, bxy, byx → all 3 have same signs

∗ Effect of change of scale & origin on regression

If u= $\frac{x-a}{b}$ & v= $\frac{y-c}{d}$

∴ buv = $\frac{cov\ uv}{r^2v}$ bvu= $\frac{cov\ uv}{r^2u}$

∴ buv=$\frac{1/b}{1/d}$ × bxy bvu= $\frac{1/d}{1/b}$ × byx

(Remember here, that change of scale is $1/_b$ and resp. and not b and d)

<u>Conclusion:-</u>

Regression co-efficient are affected by change of scale but not change of origin.

2nd formula: -

. bvu= $\frac{coefficient\ of\ y}{coefficient\ of\ x}$ × byx

buv= $\frac{coefficient\ of\ x}{coefficient\ of\ y}$ × bxy

Measure	Change of Origin	scale	Effect
Central tendencies	✔	✔	If u= ax+b Ctu= a × ctx + b

Dispersion	×	✔	If u= ax + b M.Du =$\lvert a \rvert$ × MDx
Covariance	×	✔	If u= $\frac{x-a}{b}$ & v= $\frac{y-c}{d}$ cov_{uv}= $\frac{1}{bd}$ cov_{xy}
correlation	×	× (only sign can change)	If u =$\frac{x-a}{b}$ & v=$\frac{y-c}{d}$ ruv = $\frac{\lvert b \rvert \lvert d \rvert}{bd}$ × rxy
Regression	×	✔	If u= $\frac{x-a}{b}$ & v= $\frac{y-c}{d}$ buv= $\frac{1/b}{1/d}$ × bxy bvu= $\frac{1/d}{1/b}$ × byx

∗ Probable error:-

It is a method of obtaining correlation coefficient of population.

P.ϵ= 0.674 × $\frac{1-r^2}{\sqrt{N}}$ or $\frac{2}{3}$ × $\frac{1-r^2}{\sqrt{N}}$

S.ϵ= $\frac{1-r^2}{\sqrt{N}}$

∴ P.ϵ= $\frac{2}{3}$ × S.ϵ

Where,

P.ϵ= probable error of cor. Corf. S.ϵ= standard error of

r= cor.coeff. from sample obs. N= no. obs

Limit of correlation coefficient is given by P= r +_ P.ϵ ,

Where P= correlation coefficient of population

Assumption while probable errors are significant

i) If r < 6× P.ϵ , there is no evidence of correlation
ii) If value of r is more than 6 times of the probable error, then the presence of correlation coefficient is certain
iii) Since r lies between -1 & +1, probable error is never negative (∵ r^2 is always +ve)

* V.V. imp

If r= -1 or +1, the 2 regression lines co-incide with each other (become identical).

If r= 0, regression lines are perpendicular to each other

* Scatter diagram for correlation

(r=1) perfect correlation

(0.5→ 1) strong positive correlation

(0→0.5) weak positive correlation

(r= -1)

Perfect negative correlation

(-1 < r <- 0.5)

Strong negative correlation

(-0.5 < r < 0)

Weak negative correlation

(r=0) (r=0)

No correlation curvilinear correlation

Lower left to upper right→ positive

Upper left to lower right → negative

∗ The regression lines are best-fit lines which represent all pairs of (x,y) and make a linear relation between them

i) While forming regression equation y on x, the best fit line where all of the vertical distances of the points to this line are minimised

ii) While forming regression equation x on y, the best fit line is such a line where all of the horizontal distances of all points to this line are minimised

∗ Bi-variate frequency table:-

When a discrete data is given of all the pairs of (x,y), we can convert it in a bi-variate frequency table using tally marks.

If the table includes m rows and n columns then the total number of cells in bi-variate table $=m\times n$

From a bivariate frequency distribution we can obtain 2 types of uni- variate distributions.

i) Marginal distribution:-

- Marginal distribution of x-
this table shows the different classes of x and their total frequencies (i.e. ignore y completely).

- Marginal distribution of y-
This table shows different classes of y and their total frequencies (i.e. ignore x completely).

Note:-

In total, we can form 2 tables of marginal distribution from a bi-variate frequency distribution

ii) Conditional distribution:-

- Conditional distribution of x when y= a to b. This table will have classes of x and frequency of x when y= a to b. (a-b)

- Conditional distribution of y when x= a-b this table will have classes of y and frequency of y when x= a-b

Note:-

In total, we can form m+n tables of conditional distribution from a bivariate freq. distribution. (with m rows & n columns)

- Errors in case of regression equations are

→ All (positive, negative, 0)

- Difference between observed value and estimated value in regression analysis is known as

→ Error or residue

- Method applied for deriving the regression equations is known as

→ Least squares

- Method of concurrent deriving is the quickest method to find correlation between 2 variables.

- Correlation coefficient only measures linear relationship between variables.

Angle between regression lines

The acute angle between the 2 regression lines is given by

$$\theta = tan^{-1}\left|\frac{bxy.byx-1}{bxy+byx}\right|$$

$$= tan^{-1}\left|\frac{r^2-1}{bxy+byx}\right|$$

$$= tan^{-1}\left|\frac{r^2-1}{r} \times \frac{6x.ry}{r^2x+r^2y}\right|$$

- <u>Correlation coefficient for bivariate frequency distribution:-</u>

$$r= \frac{n\sum fuv-(\sum fu)(\sum fv)}{\sqrt{\sum fu^2-(\sum fu^2)}\sqrt{n\sum fv^2-(\sum fv)^2}}$$

- byx:-

y⟹ dependent/ regression /explained variable

x⟹ independent variable / predictor/ explanatory

- Correlation coefficient (r)= 0 :-

a) No correlation OR

b) No linear relationship

∗ <u>Spurious correlation :-</u> (Nonsense correlation)

- Correlation coefficient is not zero even though the 2 variables are not related.

- This is due to existence of 3rd variable which is related to the both the variables under consideration.

Ex. Production of rice and iron in India.
→ Positive correlation over a period of 20 years due to time element that affects both rice, iron.

Index Numbers

- Ratio are comparison of 2 similar things

- Index numbers are also ratios where we compare 2 prices/ quantities / values over 2 time periods i.e. we will compare price of a product/ group of products from 1 periods with another period

- The period of which the prices are compared –

Year 1)⟸ current year (jis saal ka comparison ho raha hai).
The period with which comparison is made-

Year 0)⟸ base year (jis saal ke saath comparison ho raha hai)

- Relative:-

i) Price relative – When price of a single product is compared over 2 period, its price relatives.

P.R. = $\frac{P_1}{P_0} \times \frac{P_n}{P_0} \times 100$

Where,

P_1 = price in current year

P_0 = price in base year

ii) Quantity relatives- When quantity of single products is compared over 2 periods, its quantity relative.

Q.R= $\frac{Q_1}{Q_0} \times 100 = \frac{Q_n}{Q_0} \times 100$

Where,

Q_1 = Quantity in current year

Q_0 = Quantity in base year

iii) Value relative – When value of a single product is compared over two periods, it is value relative

V.R= $\frac{V_1}{V_0} \times 100 = \frac{V_n}{V_0} \times 100$

Where,

V_1= value in current year = $p_1 \times q_1$.

v_0 = value in base year = $p_0 \times q_0$.

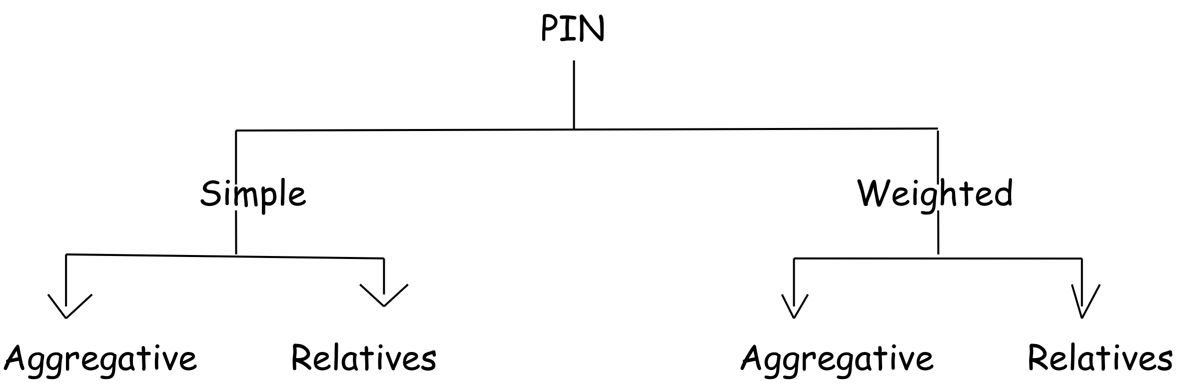
Note:-

For a single product VR= PR × QR

iv) <u>Link relatives:</u> - When comparison is made between successive years, such relatives – link relatives.

ex. $\frac{P_1}{P_0}, \frac{P_2}{P_1}, \frac{P_3}{P_2}, \frac{P_4}{P_3}$,....... are link relatives

∗ <u>Price index numbers (PIN):- (Po_1)</u>



• Price index numbers is a comparison of prices of multiple products from 1 year to another year.

• There are various methods developed to calculate PIN.

i) <u>Simple method (simple avg.):-</u>

a) <u>Simple aggregative PIN:-</u>

$$Po_1 = \frac{\sum P_1}{\sum P_0} \times 100$$

b) i) <u>PIN using price relatives by A.M.</u>

$$Po_1 = \frac{\left(\frac{P_1}{P_0} \times 100\right)}{n}$$

ii) <u>PIN using price relatives by G.M.</u>

$$Po_1 = \left[\left(\frac{P_1}{P_0} \times 100\right) \times \left(\frac{P_1}{P_0} \times 100\right)_B \times \ldots \times \left(\frac{P_1}{P_0} \times 100\right)_n\right]^{1/n}$$

* When not mentioned in Q), use simple aggregative

ii) <u>Weighted method</u>

a) <u>Weighted aggregative PIN:-</u>

- In this method, different important is given to different products by assigning weights to the product
- Usually quantity consumed/ sold are used as weight
- Various formulas under this method are:-

1) <u>Las peyre's PIN:- (Base year quantity as weight)</u>

$$Po_1 = \frac{\Sigma p_1 q_0}{\Sigma p_0 q_0} \times 100$$

2) <u>Paas che's PIN:- (Current yr. Qty as weight)</u>

$$Po_1 = \frac{\Sigma p_1 q_1}{\Sigma p_0 q_1} \times 100$$

3) <u>Dorbish – Bowley PIN:-</u>

$$Po_1 = \frac{L+P}{2}$$

4) <u>Fischer's PIN (most ideal, most used)</u>

$$Po_1 = \sqrt{L \times P}$$

5) <u>Marshall- Edgeworth PIN:-</u>

$$Po_1 = \frac{\Sigma p_1 (q_{0+q_1}) \times 100}{\Sigma p_o (q_0 + q_1)} \times 100$$

6) <u>Kelly's PIN:-</u>

$Po_1 = \frac{\sum p_1 q}{\sum p_o q} \times 100$, where $q = \frac{q_1+q_0}{2}$

7) <u>Walsh's PIN:-</u>

$Po_1 = \frac{\sum p_1 q}{\sum p_o q} \times 100$, where $q = \sqrt{q_1 q_0}$

(b) <u>Weighted relatives method</u>

i) <u>Using A.M</u>,

$$Po_1 = \frac{\sum\left(\frac{P_1}{P_0} \times 100\right) \times W}{\sum W}$$

ii) <u>Using G.M</u>,

$$Po_1 = \left[\left(\frac{P_1}{P_0} \times 100\right)_A^{WA} \times \left(\frac{P_1}{P_0} \times 100\right)_B^{WB} \times \ldots\ldots \times \left(\frac{P_1}{P_0} \times 100\right)_n^{Wn}\right]^{\frac{1}{\sum w}}$$

∗ <u>Quantity Index Numbers (Q.I.N)</u>

• In P.I.N, prices of different products over different periods were compared taking quantity as weight. In Q.I.N, quantities of different products over different periods will be compared taking price as weights.

• Therefore all formula for Q.I.N can be derived from P.I.N by interchanging price and quantity.

i) Laspeyre's QIN= $Po_1 = \frac{\sum q_1 p_0}{\sum q_0 p_0} \times 100$

ii) Paasche's QIN= $Po_1 = \frac{\sum q_1 p_1}{\sum q_0 p_1} \times 100$

.....and so on.....

∗ <u>Value Index Number (V.I.N)</u>

$Vo_1 = \frac{\sum p_1 q_1}{\sum p_0 q_0} \times 100$

∗ Cost of living index (CLI)/ general index /Inflation index /consumer price index (CPI)

i) CLI= Laspeyre's PIN= $\frac{\sum p_1 q_1}{\sum p_0 q_0}$

ii) CLI= Group index no.

$= \frac{\sum IW}{\sum W}$,

Where, I= index no. of given groups

W= weights

∗ <u>Dearness allowance:-</u>

• While Comparing salary/wages of a person over 2 years, it is assumed that salary /wages must increase in the same proportion as increase in the C.L.I. (Inflation index).

• Dearness allowance = expected salary – actual current salary

Expected salary= Base yr's salary × $\frac{\text{current CLI}}{\text{Base CLI}}$

∗ <u>Purchasing Power:-</u>

Purchasing power of money is equivalent to the quantity of products that we can buy with the available money.

Ex.

Let's assume the C.P.I. of year 2020 (with 2010 as base year) is 125.

Purchasing power of rupee 1 (of 2020) =

$\frac{1}{125}$= 0.008 units

For buying 0.008 units in 2010, amount required= 0.008× 100= Rs.0.8

∴ purchasing power of Rs.0.8 (of 2010)= 0.008 units

∴ purchasing power of Rs.1 (2020)

Purchasing power of Rs.0.8 (2010)

∴ purchasing power of Rs. 1 (of 2020 as compared to 2010)= Rs. 0.8.

∴ Purchasing power = $\frac{1}{CLI} \times 100$

Real Wages:-

- The amount of wages received in cash (actual) is the nominal wages
- However, if effect of inflation is removed from nominal wages, we get real wages.

Real Wages = $\frac{nominal\ wages}{current\ CLI} \times 100$

Ex.

	CLI	Salary
2010	100	10000
2020	125	12000

Real wages (2020) = $\frac{12000}{125} \times 100$

= Rs.9, 600

Meaning:-

Considering inflation, receiving salary of Rs.12, 000 in 2020 is equivalent to receiving salary of Rs.9600 in 2010.

∴ Loss in real wages = Rs.10, 000 – Rs.9600

= Rs.400.

∗ Chain index Numbers:-

- In link relatives we compare prices of current year with prices of year before c.y.
- If in the questions prices of the products are missing, however we here to calculate the P.I.N. of each year compared to a fixed base year using link relatives –

$$\text{Chain index} = \frac{\text{Link relative of current year} \times \text{chain index of previous year}}{100}$$

Ex.	2010	2011	2012	2013	2014
LR		110	115	120	110
PIN	100	110	126.5	151.8	166.98

- Chain index nos. are simply P.I.N., however in the absence of prices, we use link relatives, to find chain index nos.

∗ <u>Shifting & splicing of index numbers</u>

a) <u>Shifting:-</u>

For a given list of index numbers with a fixed base, if the base year is shifted to a fixed base, if the base year is shifted to a new year, then the new index numbers of all the years can be calculated as follows:-

New index no. = $\frac{old\ index\ no.\ \times 100}{fold\ index\ no. of\ new\ base\ year.}$

b) <u>Splicing of 2 index no. series:-</u>

When there are 2 diff. index no. series with 2 diff. base years, the process of combining the 2 series with a common base year is splicing of the 2 index no. series (page 19.12→ module) (example of splicing table)

∗ <u>Deflating time series using index numbers</u>

When the effect of inflation has to be removed from a certain value then we calculate its real value by deflating its current value.

i) Deflated value = $\frac{current\ value}{P.I. of\ current\ yr.} \times 100$

ii) Real wages = $\frac{nominal\ wages}{P.I.N} \times 100$

iii) Real GNP = $\frac{GNP\ at\ current\ price}{PIN} \times 100$

Test of Adequacy

There are 4 tests which can be used to determine which of the formulae for calculating P.I.N. are more adequate.

i) Unit test:-
   It states that the formula should be independent of the unit in which price are quoted.
   All satisfy this except simple aggregative index

ii) Time reversal test:-
   Deals with interchange of base year this test states that the product of the 2 index numbers calculated by taking ratio of current year prices on base year prices and then by taking ratio of base year prices on current year prices should be 1 (unity).
   (i.e. $Po_1 \times P_{10} = 1$

The following formulae satisfy this test:-

i) Simple aggregative
ii) Simple average of price relatives using G.M
iii) Weighted average of price relative using GM
iv) Fischer's PIN
v) Marshal- Edgeworth PIN
vi) Walsh PIN
vii) Kelly's PIN

Note: - Laspeyre & Passche do not satisfy this test.

iii) Factor Reversal test:-
   It states that the P.I.N and the Q.I.N should multiply into V.I.N.
   (i.e. $P_{01} \times Q_{10} = V_{01}$)
   Only Fischer's Index no. satisfies this test

↓

Ideal index number

↓

As it satisfies test ii) & iii)

iv) Circular Test:-

This test is an extension of time reversal test as per this test:-

$P_{01} \times P_{12} \times P_{20} = 1 \quad OR \quad P_{01} \times P_{12} = P_{02}$

Deals with shift ability of base year

This test is satisfied by

- Simple average of price relatives by GM.
- Weighted aggregative with fixed weights
- Simple aggregative method

Note: - Laspeyre, Paasche, Fischer do not satisfy this test.

(1) Log of PIN using GM of PR is the A.M. of log price relatives.

(2) Log P_{01}= Σ log PR

(3) Log P_{01}= ΣW log PR

General Formulae

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

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

3) $a^2 - b^2 = (a+b)(a-b)$

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

5) $(a-b)^2 = (a+b)^2 - 4ab$

6) $a^2 + b^2 = (a+b)^2 - 2ab$

7) $a^2 + b^2 = (a-b)^2 + 2ab$

8) $(a+b)^3 = a^3 + b^3 + 3a^2b + 3ab^2$

$= a^3 + b^3 + 3ab(a+b)$

9) $(a-b)^3 = a^3 - b^3 - 3a^2b + 3ab^2$

$= a^3 - b^3 - 3ab(a-b)$

10) $a^3 + b^3 = (a+b)(a^2 - ab + b^2)$

11) $a^3 - b^3 = (a-b)(a^2 + ab + b^2)$

12) *If* $a + b + c = 0$, *then* $a^3 + b^3 + c^3 = 3abc$

* Sum of first n natural nos = $\frac{n(n+1)}{2}$
(1+2+3+....n)

* Sum of squares of first n natural nos = $\frac{n(n+1)(2n+1)}{6}$
$(1^2 + 2^2 + 3^2 + \cdots n^2)$

* Sum of cubes of first n natural nos = $\left(\frac{n(n+1)}{2}\right)^2$