QUANTITATIVE APTITUDE

CONCEPTS, FORMULAS AND TRICKS MARATHON

All Chapters



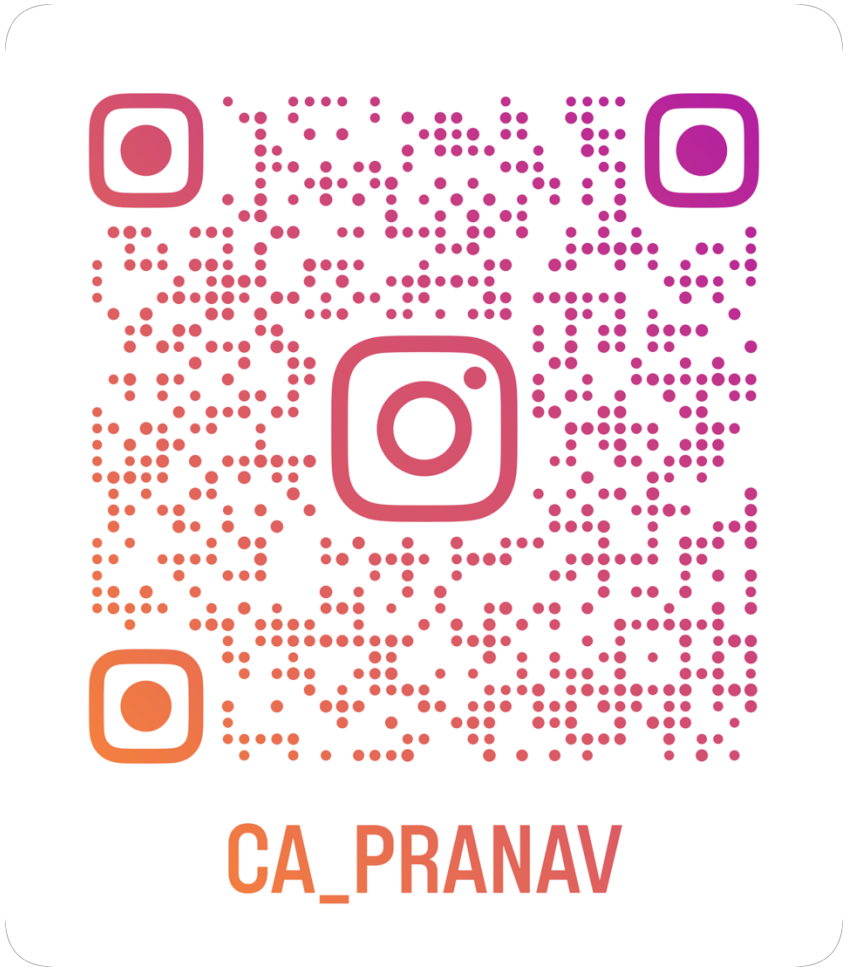
CA Pranav Popat



Join my Telegram Channel for NOTES



QA REVISION LIBRARY



Follow Insta for connecting socially

CA FOUNDATION QA

Chp1 Ratio Proportion Indices Logarithm

Concept, Formula and Tricks Marathon



Concepts

Ratio Basics	
Definition 60kg / 70kg = 6:7	• A ratio is a comparison of the sizes of **two or more** quantities of the **same kind** by division.
Terms	• The quantities **a** and **b** are called the **terms** of the ratio, **a** is called the first term or **antecedent** and **b** is called the **second term** or **consequent**.
Simplest Form of Ratio 240/360 ÷10 = 24/36 ÷6 = 4/6 = 2/3; 2.5/3.5 × 2/2 = 5/7	• Both terms of a ratio can be multiplied or divided by the same (non–zero) number. • Usually, a ratio is expressed in lowest terms (or simplest form).
Order of Terms 2:3 3:2 (not same)	• The order of the terms in a ratio is important.
Quantities of same kind	• Ratio exists only between quantities of the same kind. • Example: There is no ratio between the weight of one child and age of another
Quantities of same unit	• Quantities to be compared (by division) must be in the same units. • Example: 150 gm and 2 kg cant be directly compared, convert 2kg into 2000 gm (2 kg → 2000 g)

Equivalent like fractions	• *To compare two ratios, convert them into equivalent like fractions i.e. ratios with same denominator*
Increase or Decrease of quantity by ratios	• *If a quantity increases or decreases in the ratio **a : b** then new quantity is **b/a** times of original quantity* • *The fraction by which the original quantity is multiplied (i.e. b/a) to get a new quantity is called the **factor multiplying ratio**.*

$\frac{2}{3} + \frac{5}{6}$

$\frac{4}{6} + \frac{5}{6}$

$= \frac{9}{6}$

$\frac{2}{3} : \frac{5}{6}$

$\frac{4}{6} : \frac{5}{6}$

4 : 5

weight 7 : 9 increase

orig weight 35 kg

new weight $= 35 \times \frac{9}{7} = 45$ kg

$\frac{7}{9} = \frac{35}{x}$

$7x = 315$

$x = 45$ kg

Properties of Ratio

Inverse Ratio $\frac{7}{9} \longrightarrow \frac{9}{7}$	• *One ratio is the inverse of another if their **product is 1.*** • *Thus b : a is the inverse of **a : b** and vice-versa.*
Compounding	• *The ratio **compounded** of the two ratios a : b and c : d is ac : bd.* • *Compounding two or more ratios means **multiplying them***
Duplicate Ratio, Triplicate Ratio	• *A ratio compounded of itself is called its duplicate ratio.* • $a^2 : b^2$ *is the **duplicate ratio** of a:b* • $a^3 : b^3$ *is the **triplicate ratio** of a:b*
Sub-Duplicate Ratio, Sub-Triplicate Ratio	• $\sqrt{a} : \sqrt{b}$ *is the **sub-duplicate ratio** of a:b* • $\sqrt[3]{a} : \sqrt[3]{b}$ *is the **sub-triplicate ratio** of a:b*
Commensurable (Rational no.)	• *If the ratio of two similar quantities can be expressed as a ratio of two integers, the quantities are said to be commensurable.* • *Otherwise, they are said to be incommensurable* • *Example of Incommensurable -* $\sqrt{3} : \sqrt{2}$
Continued Ratio 3:2:3	• ***Continued Ratio:** is the relation or comparison between the magnitudes of **three or more** quantities of same kind.* • *The continued ratio of three similar quantities a, b, c can be written as **a:b:c***

$\frac{\sqrt{3}}{\sqrt{2}}$ irrational no. Incommensurable

$\frac{A}{B} = \frac{4}{5} \times \frac{3}{3} = \frac{12}{15}$ $\frac{B}{C} = \frac{3}{4} \times \frac{5}{5} = \frac{15}{20}$

12:15:20

Proportion Basics

Definition $\frac{a}{b} = \frac{c}{d}$ $\frac{4}{5} = \frac{8}{10}$	• *An **equality** of **two ratios** is called a proportion.* • *Four quantities a, b, c, d are said to be in proportion if* $a : b = c : d$ *or* $a : b :: c : d$
Terms $a : b = c : d$ a, d extreme b, c means	• *The quantities a, b, c, d are called **terms** of the proportion; a, b, c and d are called its first, second, third and fourth terms respectively.* • ***Terms** or proportion can also be called as **Proportional*** • ***First and fourth** terms are called **extremes** (or extreme terms).* • ***Second** and **third** terms are called **means** (or middle terms).*
Cross Product Rule $\frac{a}{b} = \frac{c}{d} \Rightarrow ad = bc$	• *If a : b = c : d are in proportion then* $ad = bc$ • *Product of Extremes = Product of Means*
Continued Proportion $\frac{a}{b} = \frac{b}{c}$	• *Three quantities a, b, c of the same kind (in same units) are said to be in continuous proportion if* $\mathbf{a : b = b : c}$ $\frac{a}{b} = \frac{b}{c} \Rightarrow b^2 = ac$ *a = first proportional, c = third proportional and b is mean proportional (because b is GM of a and c)*

$b^2 = ac \Rightarrow b = \sqrt{ac}$

Properties of Proportion

Invertendo	If $a : b = c : d$, then $b : a = d : c$
Alternendo	If $a : b = c : d$, then $a : c = b : d$
Componendo	If $a : b = c : d$, then $a+b : b = c+d : d$
Dividendo	If $a : b = c : d$, then $a-b : b = c-d : d$
Componendo and Dividendo	If $a : b = c : d$, then $\frac{a+b}{a-b} = \frac{c+d}{c-d}$ or $\frac{a-b}{a+b} = \frac{c-d}{c+d}$
Addendo	If $a:b = c:d = e:f = \ldots = k$, then $\frac{a+c+e+\ldots}{b+d+f+\ldots} = k$
Subtrahendo	If $a:b = c:d = e:f = \ldots = k$, then $\frac{a-c-e-\ldots}{b-d-f-\ldots} = k$

Indices Basics

Base	Number which is raised to some power is called as Base
Power	Number of times Base is multiplied by itself
Index	Entire Number including Base and Power is Index
Indices	Plural of Index
Example	$3^4 = 81$ Here Base is 3, Power is 4 and Index is 3^4 and 81 is the result
Base	Number which is raised to some power is called as Base
Standard Results	• Any base raised to the power zero is defined to be 1 $a^0 = 1$ • Roots can also be expressed in the form of power $\sqrt{9} = 3$ or $9^{\frac{1}{2}} = 3$ $\sqrt[r]{a} = a^{\frac{1}{r}}$
Power Shifting Punch	• If $6^3 = x \Rightarrow 6 = x^{\frac{1}{3}}$ • If $5^{\frac{3}{2}} = y \Rightarrow 5 = y^{\frac{2}{3}}$

POPAT

Law of Indices

Law 1	*If two or more terms with same base are multiplied, we can make them one term having the same base and power as sum of all powers.* $a^m \times a^n = a^{m+n}$
Law 2	*If two or more terms with same base are in division, we can make them one term having the same base and power as difference of power.* $\frac{a^m}{a^n} = a^{m-n}$
Law 3	*If a term having power is raised to another power, we can do product of powers to simplify the expression* $(7^2)^3 = 7^6$ $\left(a^m\right)^n = a^{m \times n}$
Law 4	*If a product of two or more terms is raised to power, we can split the two terms with same individual power to each one of them.* $\left(a \times b\right)^n = a^n \times b^n$

$(7 \times 2)^3 = 7^3 \times 2^3$

$\left(\frac{7}{2}\right)^3 = \frac{7^3}{2^3}$

POPAT

Logarithm Basics

Meaning of Log	The logarithm of a number to a given base is the **index or the power** to which the **base must be raised** to **produce** the **number**, i.e. to make it equal to the given number
Mathematical Explanation of Log	If $a^x = n$ then $\log_a n = x$ Example: If $3^4 = 81$ then $\log_3 81 = 4$
Conditions under Logarithm Function	• Log can be calculated only for Positive Number • Base should be positive and not equal to 1 $n > 0, a > 0, a \neq 1$
Standard Results	• Log of a number with same base as number is equal to 1 $\log_a a = 1$ • Log of 1 (one) for any base is equal to zero $\log_a 1 = 0$

$\log_3 9 = 2$

$\log_2 1$

$8^0 = 1$

PRANAV POPAT

Law of Logarithm

Law 1	• Logarithm of the product of two numbers is equal to the sum of the logarithms of the numbers to the same base $\log_a mn = \log_a m + \log_a n$
Law 2	• The logarithm of the quotient of two numbers is equal to the difference of their logarithms to the same base $\log_a \frac{m}{n} = \log_a m - \log_a n$
Law 3	• Logarithm of the number raised to the power is equal to the index of the power multiplied by the logarithm of the number to the same base $\log_a m^n = n \log_a m$
Change of Base Theorem	• If the logarithm of a number to any base is given, then the logarithm of the same number to any other base can be determined from the following relation $\log_b m = \frac{\log m}{\log b} = \frac{\log_a m}{\log_a b}$ old base = b, new base = a
Special Relation	$\log_b a \times \log_a b = 1$

CA FOUNDATION QA

Chp2 Equations

Concept, Formula and Tricks Marathon

CA PRANAV POPAT

Concepts

$2x - 3 = 9$

$x = 6$

Equation Basics

Definition	• *It is a mathematical statement of **equality***
Solution of Equation or Root of Equation	• *The value of variable (say x) that satisfies a given equation*
Degree of an equation	• *The highest power of variable in a given equation*

$2x + 10 = 12$

$2x^2 - 10x = 10$

Simple Equation

Description	• *Equation of one degree and having one unknown variable is simple.* • *A simple equation has only one root.* • *It can be solved directly (No Method Needed)*
Format of Equation	$ax + b = 0$ *where, a is coefficient of x, b is constant,* $a \neq 0$

Concepts

Simultaneous Linear Equations (2 var)	
Format of Equation (two unknowns)	$2x+3y+10=0$ $\quad a_1x+b_1y+c_1=0$ $5x-4y+18=0$ $\quad a_2x+b_2y+c_2=0$ *where, a is coefficient of x, b is coefficient of y, c is constant, $a \neq 0$*
Methods of Solution	*To solve linear equation in two variables, we need two such equations* • ***Elimination Method:*** *In this method two given linear equations are reduced to a linear equation in one unknown by eliminating one of the unknowns and then solving for the other unknown.* • ***Substitution Method:*** *equation is written in the form of one variable in LHS and that value is substituted in other equation.* • ***Cross Multiplication Method:*** *Formula based method* $\frac{x}{b_1c_2-b_2c_1}=\frac{y}{c_1a_2-c_2a_1}=\frac{1}{a_1b_2-a_2b_1}$
Format of Equation (three unknowns)	$a_1x+b_1y+c_1z+d_1=0$ $a_2x+b_2y+c_2z+d_2=0$

Methods of Solution	*Elimination method – first eliminate one variable from the three equations. Then further eliminate second variable.*

Concepts

Quadratic Equation	
Description	• Equation having **degree = 2** is called as Quadratic Equation • QE will have two roots/ solutions usually denoted by α, β
Equation Format	$ax^2 + bx + c = 0$ where, a is coefficient of x^2, b is coefficient of x, c is constant, $a \neq 0$
Trial and Error Method	$ax^2 + bx + c = 0$ • In this method value of **b** is split into two parts • Split is done in such a way that product of those two values is equal to **ac** • There will be four terms, we take common and form two factors • Solving factors we will get roots of the equation
Direct Formula	$\frac{-b \pm \sqrt{b^2 - 4ac}}{2a}$
Sum of Roots	$\alpha + \beta = -\frac{b}{a}$
Product of Roots	$\alpha\beta = \frac{c}{a}$
Construct a Quadratic Equation	$x^2 - (\alpha + \beta)x + \alpha\beta = 0$

Discriminant to find nature of roots of QE

Discriminant b^2-4ac of QE is the mathematical expression which is used to understand nature of roots of QE, it is expressed as below:

Condition	Nature of Roots
$b^2-4ac=0$	Real and Equal
$b^2-4ac<0$	Imaginary
$b^2-4ac>0$	Real and Unequal
$b^2-4ac>0$ and a perfect square	Real, Unequal and Rational
$b^2-4ac>0$ and not a perfect square	Real, Unequal, and Irrational

Difference of Roots

$$\alpha-\beta=\frac{\sqrt{D}}{a}$$ where, D is discriminant of QE.

Conjugate Pairs

- If one of the root of the equation is $m+\sqrt{n}$
- The other one is surely $m-\sqrt{n}$
- This pair of irrational roots are called as conjugate pairs

Important Algebraic Identities

$$(a+b)^2=a^2+b^2+2ab$$
$$(a-b)^2=a^2+b^2-2ab$$
$$(a+b)(a-b)=a^2-b^2$$
$$a^2+b^2=(a+b)^2-2ab$$
$$(a+b)^2-(a-b)^2=4ab$$
$$(a+b)^3=a^3+b^3+3ab(a+b)$$
$$a^3+b^3=(a+b)^3-3ab(a+b)$$

Concepts

Cubic Equation	
Description	• Equation having **degree = 3** is called as Cubic Equation • Cubic Equation will have three roots
Format of Equation	$ax^3 + bx^2 + cx + d = 0$ where, a is coefficient of x^3, b is coefficient of x^2, c is coefficient of x, d is constant, $a \neq 0$
Sum and Product of roots	If α, β, γ are the roots of cubic equation then Sum of roots, $\alpha + \beta + \gamma = \frac{-b}{a}$ Product of roots, $\alpha\beta\gamma = \frac{-d}{a}$

CA FOUNDATION QA

Chp3 Linear Inequalities

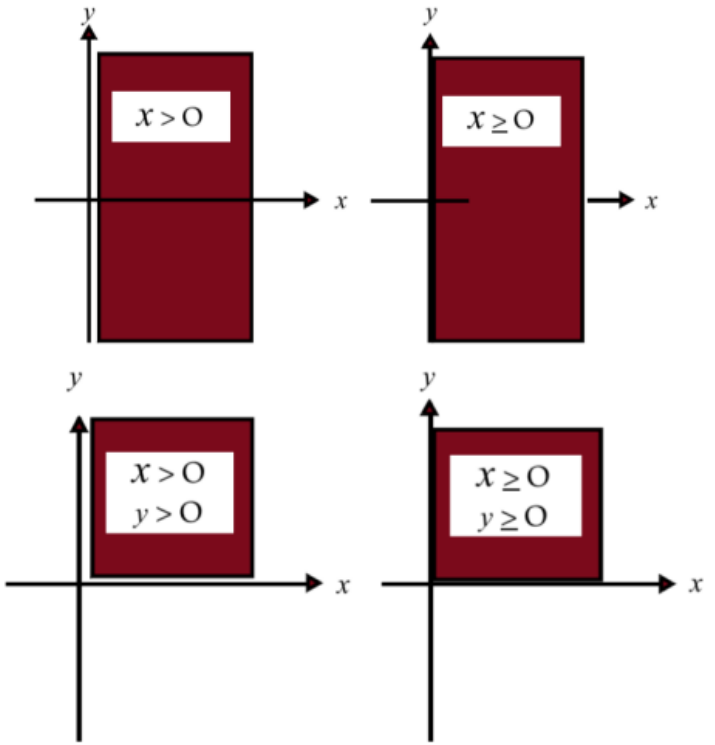
Concept, Formula and Tricks Marathon



Concepts

Meaning of Inequalities	• *Inequalities are statements where two quantities are unequal but a relationship exists between them.* • *These type of inequalities occur in business whenever there is a limit on supply, demand, sales etc.* • *Example* – *if a producer requires a certain type of raw material for his factory and there is an upper limit in the availability of that raw material, then any decision which he takes about production should involve this constraint also.*
Solution and Solution Space	• *The value of variable that satisfies the equation is called as solution of the equation* • *The values of the variables that satisfy an inequality are called the solution space, and is abbreviated as S.S.*
Linear Inequality	• *Any linear function (degree 1) that involves an inequality sign is a linear inequality.* • *It may be of one variable, or, of more than one variable.*

Drawing and Shading of Solution Space

Linear Inequalities in one variable $x > 3$ $(3, \infty)$ $x \geq 3$ $[3, \infty)$	*Example:* $x > 0$  $x \leq 0$ 
Basics of Shading in a plane (two variables)	

$2x+3y \leq 10$

How to solve Linear Inequalities in two variables

Step 1	We replace the inequality sign by an equal sign (just to find points/ coordinates) $2x+3y=10$
Step 2	Find coordinates of the line by first taking $x = 0$ to get y and then taking $y = 0$ to get x. If line passes through origin take points other than 0. $x=0, y=10/3$ | $y=0, x=5$ $(0, 10/3)$ $(5, 0)$
Step 3	Using points obtained in step 2 – join the points by a straight line. This line will be used as basis for shading.
Step 4	Do the shading above or below / right or left of line using below rules:

Scenario	Shading side from line
more than sign towards x	Right
Less than sign towards x	Left
more than sign towards y	Up
less than sign towards y	Down

Note: before applying above rule please ensure that coefficients are positive by shifting them from LHS to RHS or vice-versa.

Concept, Formula and Tricks Marathon QA
CA PRANAV POPAT

CA FOUNDATION QA

Chp4 Math for Finance

Concept, Formula and Tricks Marathon

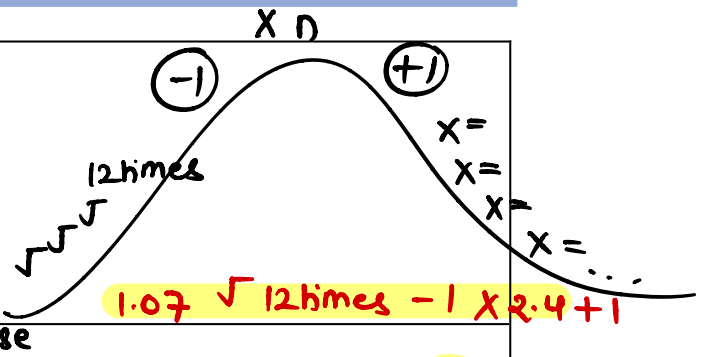


Calculator Tricks

Calculator Tricks & Basics

Power (Integer)	2000 (1·07)⁴ → 1·07 X = = = → 2 3 4 → 2621·59 Base × = = = … (square, cube)
n^{th} power (Non-Integer)	Base √ √ √ …12times.. −1 × n +1 × = × = × = …12times
n^{th} root	Base √ √ √ …12times.. −1 ÷ n +1 × = × = × = …12times
Reciprocal of any number	x = 2000 / (1·06)³ → 1·06 X = = ÷ = X2000 → 1679·238 ÷ =
Use of Memory Function	

M +	*It stores value as it is without changing sign*
M −	*It stores value after changing sign*
MRC	*It adds up and give total of all the values stored in memory*



x	2	4	8
f	5	3	5

$\frac{\Sigma fx}{\Sigma f} = \frac{62}{13} = 4.76$

2000 (1·07)^2.4

2000 = 1000 (1+i)⁵

2 = (1+i)⁵

$SD = \sqrt{\frac{660}{8} - (7)^2}$

660 ÷ 8 M+ 7X = M− MRC √ 5.78

Concept, Formula and Tricks Marathon QA
CA PRANAV POPAT

Concepts

$(1+i) = 2^{1/5}$ $\sqrt[5]{2}$ 1.1487

Basics

Reasons to pay/ receive Interest	Opportunity Cost	▪ To lend money to others, we sacrifice the return on investing that money somewhere else
	Inflation	▪ Time Factor: Due to inflation a given amount of money buys fewer goods in the future than it will now
	Liquidity Preference	▪ After lending, money is not available for immediate use
	Risk Factor	▪ There is always a risk that borrower will not repay at all.

Basic Terms	Interest	Interest is the **price paid** by a borrower for the **use of a lender's** money.
	Principal	Principal is initial value of **lending** (or **borrowing**).
	Rate of Interest % p.a.	The rate at which the interest is charged for a defined length of time for use of principal generally on a **yearly basis** is known to be the rate of interest.
	Accumulated Balance (A)	Accumulated amount is the **final value** of an investment. It is the **sum total** of principal and interest earned.

Concepts

Partied involved in Interest related Transactions like Loan or Investment	*Lender*	*Person/ Financial Institution like Bank, NBFC who lends money to borrower. Interest is income for lender.*
	Borrower	*Person who borrow money from lender. Interest is expense for borrower.*
	Investor	*Person who invest money to earn return. Interest is cash inflow for investor. Example: Depositors who opened Fixed Deposit Account with Bank are Investor.*
	Investee	*Person/ Financial Institution like Bank, NBFC where Investor can deposit their money for investment purpose to earn interest.*

CA PRANAV POPAT

Simple Interest

Concept	▪ Simple interest is the interest computed on the principal **for the entire period** of borrowing. ▪ It is calculated on the **principal amount only** and not on interest previously earned. ▪ Value of Interest **remains constant** for each year ▪ Difference between Amounts of any two consecutive year is equal to one year Interest.
Formula of Simple Interest	$SI = \frac{P.r.t}{100}$ 4yr 3m = 51m/12 where, P = principal value, r = rate of interest per annum, t = time in years (if time is given in months convert it to year by dividing 12)
Formula of Amount as per Simple Interest	$A = P + SI$ $A = P + \frac{P.r.t}{100} = P(1 + \frac{rt}{100})$

Concept, Formula and Tricks Marathon QA

Example. $P = 480{,}000$, $r = 10\%$ p.a., $t = 4$ years. Prepare a passbook for 4 years using SI.

Sol

Passbook

Date	Particulars	Dr.	Cr.	Balance
1/4/25	deposit cash	-	480,000	480.000
31/3/26	Interest credited (480,000 × 10%)	-	48,000	528,000 (A_1)
31/3/27	Interest credited (480.000 × 10%)	-	48.000	576,000 (A_2)
31/3/28	"	-	48,000	624,000 (A_3)
31/3/29	"	-	48,000	672000 (A_4)

$A_2 - A_1 = 48000$ $A_3 - A_2 = 48000$ $A_4 - A_3 = 48000$

CA PRANAV POPAT

Compound Interest

Basics	▪ *Compound Interest is interest computed on **latest accumulated balance**.* ▪ *It is calculated on the **principal** and **interest previously earned.*** ▪ *Value of Interest keeps on **increasing** in every next period.* ▪ *Value of CI depends on **conversion period** also.*

CA PRANAV POPAT

Master Example You invested ₹ 200,000 in an account paying interest at 10% p.a. for 3 years. Prepare a passbook both by SI and CI approach

Sol

Simple Int.

Date	Particulars	Cr.	Balance
1/4/25	deposit	200,000	200,000
31/3/26	Interest (200,000 × 10%)	20,000	220,000
31/3/27	Interest (200,000 × 10%)	20,000	240,000
31/3/28	Interest (200,000 × 10%)	20,000	260,000

Compound Int.

Date	Particulars	Cr.	Balance
1/4/25	deposit	200,000	200,000
31/3/26	Interest (200,000 × 10%)	20,000	220,000
31/3/27	Interest (220,000 × 10%)	22,000	242,000
31/3/28	Interest (242000 × 10%)	24,200	266,200

Conversion Period	***Conversion Period***: *Period for which interest is computed*

Conversion Period	***Compounding Frequency***	***No of Conversion Period per year (nocppy)***
1 month	*Compounded Monthly*	*12*
3 months	*Compounded Quarterly*	*4*
6 months	*Compounded Semi Annually*	*2*
12 months	*Compounded Annually*	*1*

CA PRANAV POPAT

Concept, Formula and Tricks Marathon QA

Masku Example. You invest 200,000 in an account that pay interest @ 10% p.a. compound interest. Prepare pass book in these scenarios

(i) compounded annually (ii) compounded half yearly

Solution

CI annually

Date	Particulars	Cr.	Balance
1/4/25	deposit	200,000	200,000
31/3/26	Interest (200,000 × 10%)	20,000	220,000
31/3/27	Interest (220,000 × 10%)	22,000	242,000
31/3/28	Interest (242000 × 10%)	24200	266,200

Time = 3 yr,

n = 3

rate = 10%

i = 10%

CI half yearly

Date	Particulars	Cr.	Balance
1/4/25	deposit	200,000	200,000
30/9/25	Interest (200,000 × 10%/2)	10,000	210,000
31/3/26	Interest (210,000 × 10%/2)	10500	220500
30/9/26	Interest (220500 × 10%/2)	11025	231525
31/3/27	Interest (231525 × 10%/2)	11576.25	243101.25
30/9/27	Interest (243101.25 × 10%/2)	12155.0625	255256.3125
31/3/28	Interest 255256.3125 × 10%/2	12762.815625	268019.128125

n = 6

i = 5%

CA PRANAV POPAT

Formula for Amount as per Compound Interest	$A = P(1+i)^n$ *where,* *P = Initial Principal, i = adjusted interest rate n = no. of periods*
Adjusted Interest Rate (i)	*Adjusted Interest Rate is the rate of interest adjusted based on compounding frequency. Example, if rate is 10% and interest is computed half-yearly then adjusted rate is* $\frac{10\%}{2} = 5\%$ $i = \frac{r\%}{nocppy}$
Number of Periods (n)	*Number of periods means number of conversion periods during the entire tenure of loan/ investment based on compounding frequency. Example, if time is 3 years and interest is compounded quarterly then number of periods is* $3 years \times 4 = 12$ $n = t \times noccpy$

PRANAV POPAT

Formula for Compound Interest	$CI = A - P$ $CI = P(1+i)^n - P$ $CI = P[(1+i)^n - 1]$ *where,* *P = initial principal, i = adjusted interest rate, n = no. of periods*
Trick for Amount as per Compound Interest	$\boxed{P}\boxed{+}\boxed{i}\boxed{\%}\boxed{+}\boxed{i}\boxed{\%}\boxed{+}$...*n times* *Suitable when value of n is small or rates are changing*
Formula of Amount of CI when rates are changing	$A = P(1+i_1)(1+i_2)(1+i_3)...$ *where* i_1, i_2, i_3 *are rates for first, second, third period etc.*

$P(1.1)(1.12)(1.08)$

CA PRANAV POPAT

Effective Rate of Interest	*Equivalent **annual rate** of interest compounded annually if interest is compounded **more than once a year.** Effective rate is not dependent on Principal.* $$E = [(1+i)^n - 1]$$
CI Concept in WDV Depreciation	$$A = P(1-i)^n$$ *where, P = Historical Cost of Asset, A = Scrap Value/ Residual value of asset, n = no. of periods, i = Depreciation %*
Difference between CI and SI for 2 years	Pi^2 or Pr^2
Special Note on CI and SI	*When question is silent about type of interest – CI or SI, we follow below approach:* • *If similar pattern is seen elsewhere then take CI or SI accordingly and check options* • *If options are not matching then recalculate by other way round* • *If questions is of unseen pattern then first try with SI, if options do not match then try CI*

Types of Cashflows

Single Cashflow/ Lumpsum	When deposits or payments are made in total at once (not in parts) then type of payment is called single or lumpsum. Example: Fixed Deposit, Loan with one time repayment etc.
Annuity	Annuity can be defined as installment having below features: • Constant/ same value • Regular payments/ receipts • Specified period of time
Types of Annuities	Annuity Regular — First payment/receipt at the end of the period Annuity Due — First payment/receipt at the beginning of the period If question is silent, we take annuity as regular.

Future Value

Future Value – Single Cashflow	▪ Future value is the **cash value** of an **investment** at some time in the future. ▪ It is **tomorrow's value** of today's money **compounded** at the rate of interest.

Formula for FV of Single Cashflow	$$FV = CF(1+i)^n$$ *where, CF = single cashflow for which FV is to be calculated, i = adjusted interest rate, n = no. of periods*
FV of Annuity Regular	▪ *To calculate **final maturity value** of an investment like RD where sum is invested in the annuity pattern starting at the end of each period.* ▪ *To calculate the final value of Sinking Fund or Savings amount to achieve the **target maturity value.***
Formula for Future Value - Annuity Regular	$$FVAR = A_i \times FVAF(n,i)$$ $$FVAR = A_i \times \left\{\frac{[(1+i)^n - 1]}{i}\right\}$$ *where, **FVAR** = Future Value of Annuity Regular, **A_i** = Annuity Value (Installment), **FVAF** = Future Value Annuity Factor, **i** = adjusted interest rate, **n** = no. of periods*

$i = 8\%$ $n = 6$

$\frac{(1.08)^6 - 1}{0.08} = 7.3359$

FV of Annuity Due	▪ *To calculate final maturity value of an investment like RD where sum is invested in the annuity pattern* ***at the beginning of each period*** ▪ *To calculate final maturity value of an investment like RD where sum is invested in the annuity pattern at the beginning of each period*
Formula for Future Value - Annuity Due	$$FVAD = A_i \times FVAF(n,i) \times (1+i)$$ $$FVAD = A_i \times \left\{ \frac{[(1+i)^n - 1]}{i} \right\} \times (1+i)$$ ***where, FVAD****= Future Value of Annuity Due,* A_i *= Annuity Value (Installment),* ***FVAF*** *= Future Value Annuity Factor,* ***i*** *= adjusted interest rate,* ***n*** *= no. of periods*
Sinking Fund	▪ *It is the* ***fund credited*** *for a specified purpose by way of* ***sequence of periodic payments*** *over a* ***time-period*** *at a specified interest rate.* ▪ *Interest is compounded at the* ***end of every period****.* ▪ *Size of the sinking fund deposit is same as Future Value of Annuity*

Sinking Fund	*A **sinking fund** is a fund created by setting aside a fixed amount of money periodically (every year, half-year, etc.) so that a **remembered target amount is accumulated in the future**.* *It is commonly used for: Replacement of assets, Repayment of loans/debentures, Future business expansion, Any planned large expense* $Sinking\ Fund\ Value = Sinking\ Fund\ Installment \times \left\{\frac{[(1+i)^n - 1]}{i}\right\}$ $= Sinking\ Fund\ Installment \times FVAF(n,i)$ $then,\ Sinking\ Fund\ Installment = Sinking\ Fund\ Value \times \left\{\frac{i}{[(1+i)^n - 1]}\right\}$ $\left\{\frac{i}{[(1+i)^n - 1]}\right\}$ *is called Sinking Fund Factor. And it is reciprocal of Future Value Annuity Factor.*

PRANAV POPAT

Present Value

Present Value of Single Cashflow	▪ *Present value is **today's value** of tomorrow's money **discounted** at the interest rate*
Formula for PV of Single Cashflow	$$PV = CF \times \frac{1}{(1+i)^n}$$ CF ÷ (1+i) = = = n times *where, CF = Single Cashflow for which PV is to be calculated, i = adjusted interest rate, n = no. of periods*
Single Cashflow Compounding and Discounting Factors	*Compounding (Adding the interest)* : $(1+i)^n$ *Discounting (Removing the interest)* : $\frac{1}{(1+i)^n}$
Present Value – Annuity Regular	***Use:** To calculate **loan amount** when periodic installments value are given and vice versa.* ***Application:** Leasing, Capital Expenditure etc.*

Formula for PV of Annuity Regular	$$PVAR = A_i \times PVAF(n,i)$$ $$PVAR = A_i \times \left[\frac{1}{i} \times \left\{1 - \frac{1}{(1+i)^n}\right\}\right]$$ ***where, PVAR*** = *Present Value of Annuity Regular,* A_i = *Annuity Value (Installment),* ***PVAF*** = *Present Value Annuity Factor,* ***i*** = *adjusted interest rate,* ***n*** = *no. of periods*
Calculator Trick for PVAF	$1+i$ ÷ = = ...*n times* GT $1+i$ ÷ = $M+$ = $M+$ = $M+$ ...*n times* MRC
Formula for Present Value of Annuity Due	$$PVAD = \left[A_i \times PVAF\left\{(n-1),i\right\}\right] + A_i$$

POPAT

Present Value when payments are neither single nor annuity	*If payments are neither single nor annuity but given as some random non equal installments then Present Value to be calculated separately for each installment and annuity based formula cannot be used. Example: NPV Questions*
How to decide if we need to use FV or PV?	

Situation	*Usage*	*Key Words*
If money is moving into future	*Use Future Value*	*Accumulated Value, Maturity Value, Future Worth, Value at the end, amt. of Annuity, Sinking Fund etc.*
If money is being brought to today	*Use Present Value*	*Present worth, today's value, single sum, discounted value, loan amount, Bond Price, today's value of lease etc.*

CA PRANAV POPAT

Applications of TVOM & Other Concepts

Leasing	▪ ***Lessor:*** *Owner of Asset, who gives asset on rent. Lease Rentals are income for Lessor* ▪ ***Lessee:*** *User of the asset who has taken asset on rent. Lease Rentals are expense for Lessee* ▪ ***Use of TVOM****: Present Value of Annuity (**Lease Rentals**) are compared with asset cash down price to decide if leasing is preferable or not.*
Capital Expenditure Decisions	▪ *Present value of **future benefits** due to **new asset** are compared with purchase value of asset, to decide whether asset to purchase or not.*

PRANAV POPAT

Valuation of Bond

- *Present value of **interest income** and **maturity value** is compared with the issue price of bond*
- *Terms*

Bond	*It is a debt security. Type of loan taken by company from public. Like debentures*
Face Value/ Par Value	*Value written on the document of bond. This value is used to calculate Interest Amount*
Issue Price	*Actual payment made to purchase the bond*
Maturity Value	*Amount to be received on redemption or maturity of bond*
Bond Interest Rate	*This rate is used to find amount of periodic interest earned from the bond.* *Bond Interest is = face value × bond interest rate*
Discounted Rate	*It is used to calculate Present value. It is given in question in the name of cost of capital or expected rate of return.*

PV of Perpetuity	*Perpetuity: An annuity that continues till infinite period of time is called as Perpetuity.* $$PVP = \frac{A_i}{i}$$ ***where, PVP** = Present Value of Perpetuity, **A_i** = Annuity Value (Installment), **i** = adjusted interest rate*
PV Growing Perpetuity	*A stream of cashflows that grows at constant rate forever is known as growing perpetuity.* $$PVGP = \frac{A_i}{i - g}$$ ***PVGP** = Present Value of Growing Perpetuity; **A_i** = Annuity Value (Installment); **i** = adjusted interest rate; **g** = growth rate*

PRANAV POPAT

Net Present Value	***Formula***	*NPV = Present Value of Cash Inflows – Present Value of Cash Outflows*
	Decision Base	*If* ***NPV ≥ 0****, accept the proposal,* *If* ***NPV < 0****, reject the proposal*
Real Rate of Return	*Real Rate of Return = Nominal Rate of Return – Rate of Inflation*	

CA PRANAV POPAT

CAGR	• *Compounded Annual Growth rate is used to show annual growth as per CI.* • *In the CAGR question, value of beginning period is to be taken as principal and value of end period is to be taken as amount. Taking number of periods from question, we will calculate missing i as per amount as per CI formula* $A = P(1+i)^n$ • *There is a direct formula also* $CAGR(t_0, t_n) = \left(\frac{value\ for\ t_n}{value\ for\ t_0} \right)^{\frac{1}{t_n - t_0}} - 1$ *where,* t_0 *is beginning period and* t_n *is end period*

PRANAV POPAT

CA FOUNDATION QA

Chp5 Permutations and Combinations

Concept, Formula and Tricks Marathon

CA PRANAV POPAT

Concepts

Rules of Counting	
Addition Rule/ OR Rule	*It there are two alternative jobs which can be done in 'm' ways and in 'n' ways respectively then either of two jobs can be done in* $(m+n)$ *ways*
Multiplication Rule / AND Rule	*If certain thing may be done in 'm' different ways and when it has been done, a second thing can be done in 'n ' different ways then total number of ways of doing both things simultaneously is* $(m \times n)$ *ways*

Concept of Factorial	
Meaning	*Factorial of an integer "**n**" is product of all integers from 1 to n both inclusive.*
Formula	$n! = n(n-1)(n-2)...3 \times 2 \times 1$ or $n! = 1 \times 2 \times 3...(n-2) \times (n-1) \times n$
Symbol	$n!$ or $\lfloor n$ 5! or ⌊5
Special Trick	$n! = n(n-1)!$ $n! = n(n-1)(n-2)!$ $\frac{6!}{4!} = \frac{6 \times 5 \times 4!}{4!} = 30$
Factorial of 0	$0! = 1$ *(Proof will be discussed using Permutations)*

Concept of Permutations

Meaning	• *The ways of* ***arranging*** *or* ***selecting + arranging*** • ***smaller or equal*** *number of persons or objects* • *from a* ***group*** *of persons or* ***collection*** *of objects* • *with due regard being paid to the* ***order of arrangement*** *or selection are called* ***PERMUTATIONS***
Theorem	*Number of Permutations when* ***r objects*** *are chosen out of* ***n different objects*** ${}^{n}P_{r} = \frac{n!}{(n-r)!}$ ***Condition:*** $n \geq r$
Shortcut of Theorem	*To find* ${}^{n}P_{r}$ *do reverse multiplication of n for r times.* *No. of Factors in* ${}^{n}P_{r}$ *is r*
Special Formula	$(n+1)!-n! = n.n!$
Particular case of theorem when n = r	*Number of Permutations when n objects are chosen out of n different objects (or Number of permutations when all the objects are chosen)* ${}^{n}P_{n} = \frac{n!}{(n-n)!} = \frac{n!}{0!} = n!$

n = 4, r = 4

n = 4, r = 2

${}^{7}P_{3} = \frac{7!}{(7-3)!} = \frac{7!}{4!} = \frac{7\times6\times5\times4!}{4!} = 210$

$7\times6\times5 = 210$

${}^{11}P_{4} = 11\times10\times9\times8$

$6! - 5! = 5\times5!$

${}^{6}P_{6} = 6!$

Use of Permutations in Practical Problems

Number Formations	• Any number is formed by arranging the given digits • Hence, for the purpose of calculating number of possible numbers formed, we use permutations. • When question is silent, we assume that digits will not be repeated in forming number.
Conditional Cases	• If there is a zero, it cannot come to first place of the number • If there is a restriction that the numbers formed should be larger than a particular value then we will use counting rules to find solution 507 075 • Further restrictions should be dealt based on question
Summation of possible numbers	1. Find the number of numbers that can be formed 2. Find repetition value of each digit (4) $\text{Repetition of each digit} = \frac{\text{Value of Step 1}}{\text{number of different digits}}$ 3. Find Sum of digits 2, 4, 1, 7 2+4+1+7 = 14 4. Sum of digits × Repetition 14X4 = 56 5. Multiply value of step 4 by 1111, 111, etc. in case of four digit numbers and three digit numbers respectively

56 X 1111

56 X 111

Use of Permutations in Practical Problems

<table>
<tr>
<td><u>Word Formations</u>
ACT, CAT, TAC, TCA</td>
<td colspan="2">

For the purpose of this chapter, we will consider any arrangement of letters as a word,
Even though it may or may not have any meaning or pronunciation

</td>
</tr>
<tr>
<td>ATC, CTA</td>
<td colspan="2">

As word is arrangement of letters, we use permutations to know the number of possible words with the given letters.

</td>
</tr>
<tr>
<td rowspan="2">FLR (AIUE)
4! x 4! = 576

Conditional Cases</td>
<td>When some of the letters/ things are together</td>
<td>

Group of things which are together should be counted as one thing only
Things within group can change their place within themselves, their arrangements also needs to considered
If based on information in questions, things in the group cannot change their places, ignore their arrangement

</td>
</tr>
<tr>
<td>When some of the letters/things are never together</td>
<td>Use this approach
Ways of Never Together
Total ways – Ways of always together</td>
</tr>
</table>

Permutations when things are alike	• *The number of ways of permutations in which n things may be arranged among themselves, taking them* ***all at a time****,* • *when* n_1 *of the things are exactly alike of one kind ,* n_2 *of the things are exactly alike of another kind,* n_3 *of the things are exactly alike of the third kind, and the rest all are different is given by* $$p = \frac{n!}{n_1! n_2! n_3!}$$

COOL

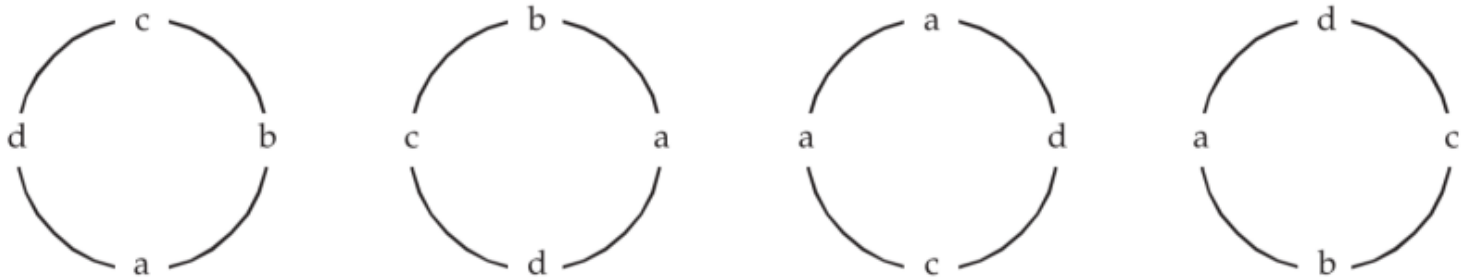
$$\frac{4!}{2!} = 12 \text{ ways}$$

CA PRANAV POPAT

Concept, Formula and Tricks Marathon QA

Circular Permutations

Circular Permutations (Type 1)	• *if we arrange the objects along a* ***closed curve*** *viz., a circle, the permutations are known as circular permutations.* • ***Theorem:*** *The number of circular permutations of n different things* ***chosen all*** *at a time is* $(n-1)!$ ***Note****: this theorem applies only when we choose all of n things*



Here, four ordinary permutations = one circular

Circular Permutations (Type II)	• *number of ways of arranging n persons along a closed curve* ***so that no person has the same two neighbours is*** $\frac{1}{2}(n-1)!$ • *Same formula will apply if ask is to find number of* ***different forms of necklaces/ garlands***

CA PRANAV POPAT

Permutations with Restrictions

Theorem 1	• *Number of permutations of n distinct objects taken r at a time when* ***a particular object is not taken*** *in any arrangement is* $${}^{n-1}P_r$$
Theorem 2	• *Number of permutations of r objects out of n distinct objects when a* ***particular object is always included*** *in any arrangement is* $$r.{}^{n-1}P_{r-1}$$
Relationship between above two theorems	• *Ways when one particular thing always included + Ways when one particular thing is always excluded = Total Ways without restrictions* $${}^{n-1}P_r + r.{}^{n-1}P_{r-1} = {}^{n}P_r$$

PRANAV POPAT

Concept of Combinations

Meaning	• *The number of ways in which* ***smaller or equal*** *number of things are* ***selected*** *from a collection of things* • *where the* ***order*** *of selection or arrangement is* ***not important****, are called combinations.*
Theorem	• *Number of Combinations when* ***r objects*** *are chosen out of* ***n*** ***different*** *objects* $^{n}C_{r} = \frac{n!}{(n-r)!\,r!}$ • ***Conditions*** $n \geq r$, *n is a positive integer*
Linkage of PNC	$^{7}C_{2} = \frac{^{7}P_{2}}{2!} = \frac{7 \times 6}{2} = 21$ $\qquad$ $^{n}C_{r} = \frac{^{n}P_{r}}{r!}$
Standard Results	$^{n}C_{0} = 1$ & $^{n}C_{n} = 1$

$^{10}C_{4} = \frac{^{10}P_{4}}{4!} = \frac{10 \times 9 \times 8 \times 7}{24} = 210$

PRANAV POPAT

Conditional Combinations

Process to solve	• *If the requirement is conditional* • *Split number of objects n into parts* • *Find combinations for each part separately* • *Then add/ multiply as per the situation (using rules of counting)*

Complimentary Combinations

Logic	• *choosing of r things to be selected out of n is same as choosing n-r things for non-selection (rejection)* • *Example: If you want to select 4 things out of 5 it is better to directly choose the one not be selected.*
Formula $^{12}C_8 = {}^{12}C_4$	$^{10}C_7 = {}^{10}C_3$ $\quad {}^nC_r = {}^nC_{n-r}$

Combinations of selecting one or more objects

Combinations of selecting one or more objects out of n objects	• *Manual:* $^nC_1 + {}^nC_2 + {}^nC_3 + \ldots + {}^nC_{n-1} + {}^nC_n$ • *Formula:* $2^n - 1$
Combinations of n different things (zero or more) out of n things at a time	• *Manual:* $^nC_0 + {}^nC_1 + {}^nC_2 + {}^nC_3 + \ldots + {}^nC_{n-1} + {}^nC_n$ • *Formula:* 2^n

Use of Combinations in Geometry

No. of Straight Lines with the given n points	${}^{n}C_{2}$ 2 is used as we need to select two points to make a line
No. of Triangles with the given n points	${}^{n}C_{3}$ 3 is used as we need to select three points to make a triangle
No. of Straight Lines with the given n points where m points are collinear	${}^{n}C_{2} - {}^{m}C_{2} + 1$
No. of triangles with the given n points where m points are collinear	${}^{n}C_{3} - {}^{m}C_{3}$
No. of Parallelogram with the given one set of m parallel lines and another set of n parallel lines	${}^{n}C_{2} \times {}^{m}C_{2}$ Selecting 2 lines from each set of parallel lines
No. of Diagonals with n sides	${}^{n}C_{2} - n$

Special Formula of Combinations

Logic	*number of combinations of (n+1) things taken r at a time =* *combinations of (n+1) things taken r at time in which one specified thing is* ***always included*** *and* + *the number of combinations of (n+1) things taken r at time from which the specified thing is* ***always excluded****.*
Formula	$^{n+1}C_r = {}^{n}C_r + {}^{n}C_{r-1}$

CA FOUNDATION QA

Chp6 Sequence and Series

Concept, Formula and Tricks Marathon



Concepts

Basics	
Sequence	• An ordered collection of numbers $a_1, a_2, a_3, \ldots, a_n$ is a sequence if according to some definite rule or law, there is a definite value of a_n • Types of Sequence: – Finite Sequence: If the number of elements are finite. Ex. 3, 7, 11, 15 – Infinite Sequence: If the number of elements are infinite Ex. 3, 7, 11, 15,…
Series	An expression of the form $a_1 + a_2 + a_3 + \ldots + a_n$ which is the sum of the elements of the sequence is called a series.

PRANAV POPAT

Concept, Formula and Tricks Marathon QA

Arithmetic Progression

Meaning	• AP is a sequence in which each next term is obtained by adding a constant 'd' to the preceding term. • This constant 'd' is called as common difference.
Format	$a, a+d, a+2d, a+3d \ldots a+(n-1)d$
Common Difference	• In AP difference between a term and the preceding term is constant $d = t_2 - t_1 = t_3 - t_2 = \ldots = t_n - t_{n-1}$
General Term of an AP	$t_n = a+(n-1)d$ where, a = first term, d = common difference, n = term position
Calculator Trick for General Term of an AP	$a \pm d = = = \ldots =$ Firste $=$ press will give you second term and so on Note: + or – sign before d depends on sign of d in any question

If three terms a, b, c are in AP	• $b-a=c-b$	$a-b=b-c$	$b=\frac{a+c}{2}$

CA PRANAV POPAT

Sum of first n terms of an AP	$S_n = \frac{n}{2}(a + t_n)$ *or* $S_n = \frac{n}{2}\{2a + (n-1)d\}$ *where, a = first term, d = common difference, n = term position* t_n *is the last term up to which sum is to be calculated* *Note: Both formulas give same results – usage depends on data given in the question [Use first formula if last term is given]*
Calculator trick for Sum of first n terms of an AP	$\boxed{a}\boxed{\pm}\boxed{d}\boxed{=}\boxed{=}\boxed{=}\ldots\boxed{=}\boxed{GT}\boxed{+}\boxed{a}$ *First* $\boxed{=}$ *press will give you second term and so on*
Sum of first n natural numbers / counting nos.	$S = \frac{n(n+1)}{2}$
Sum of first n odd numbers	$S = n^2$
Sum of squares of first n natural numbers	$S = \frac{n(n+1)(2n+1)}{6}$
Sum of cubes of first n natural numbers	$S = \left\{\frac{n(n+1)}{2}\right\}^2$

Standard result:

if p^{th} term is q and q^{th} term is p then

r^{th} term $= p+q-r$

CA PRANAV POPAT

Geometric Progression

Meaning	• *If in a sequence of terms each term is* ***constant multiple*** *of the proceeding term, then the sequence is called a* ***Geometric Progression*** *(G.P).* • *The constant multiplier is called the* ***common ratio****.*
Format	*Let say a = first term and r = common ratio, then GP can be written as* $a, ar, ar^2, \ldots, ar^n$
Common Ratio	*In a GP ratio between a term and the preceding term is constant*
	$\frac{t_2}{t_1} = \frac{t_3}{t_2} = \frac{t_n}{t_{n-1}}$
General Term of a GP	$t_n = ar^{n-1}$ • *where, a = first term, r = common ratio, n = position number of term*
Calculator Trick for General Term of a GP	$r \times a = = = = \ldots =$ *First* $=$ *press will give you second term and so on*

Sum of n terms of a GP	$S_n = \frac{a(1-r^n)}{1-r}$ *or* $S_n = \frac{a(r^n-1)}{r-1}$ *where, a = first term, r = common ratio, n = position number of term*
Calculator Trick for Sum of n terms of a GP	$\boxed{r}\boxed{\times}\boxed{a}\boxed{=}\boxed{=}\boxed{=}\ldots\boxed{=}\boxed{GT}\boxed{+}\boxed{a}$ • *First* $\boxed{=}$ *press will give you second term and so on*
Sum of Infinite Geometric Series	$S_\infty = \frac{a}{1-r}$ • ***Condition:*** *Can be used only if* $-1 < r < 1$
If three terms x, y, z are in GP	• $\frac{y}{x} = \frac{z}{y}$ | $\frac{x}{y} = \frac{y}{z}$ | $y^2 = xz$

CA PRANAV POPAT

CA FOUNDATION QA

Chp7 Set Relation & Functions

Concept, Formula and Tricks Marathon



Concepts

Sets Concepts	
Meaning	• ***Object:*** *In our mathematical language, everything in this universe, whether living or non-living, is called an object.* • ***Sets:*** *Well defined And Distinct Collection of Objects* • ***Elements:*** *Each object of Set*
How to denote	• ***Sets:*** *Capital Letters* • ***Elements:*** *Small Letter*
Descriptive Form	• *when set is written in the form of Paragraph and elements are not listed* *Examples:* • *A = the set of vowels in the English alphabet* • *B = the set of even numbers between 2 and 10 both inclusive* • *P = the set of first six prime numbers*
Roster Form	• *when elements of sets are listed and closed with braces (curly brackets)*
Set Builder/ Algebraic Form (only for numbers)	• *Sets can also be presented using algebraic statements which can be understood by examples below.* • *The method of writing the set is called as Property Method*

Use of Brackets in Sets	***Ordinary ( )***	*(a, b) means this set contains all the elements between a and b exclusive of a and b*	***Open Interval Set***
	Square [ ]	*[a, b] means this set contains all the elements from a and b inclusive of a and b*	***Closed Interval Set***
	Braces { }	*{a, b} means this set contains only two elements a and b*	***Ordinary Set***
Finite and Infinite Sets	• ***Finite Set:*** *if set contains finite number of elements* • ***Infinite Set****: if set contains infinite number of elements where it is impossible to list all. (to show an infinite set in roster form, we use three dots after few elements)*		

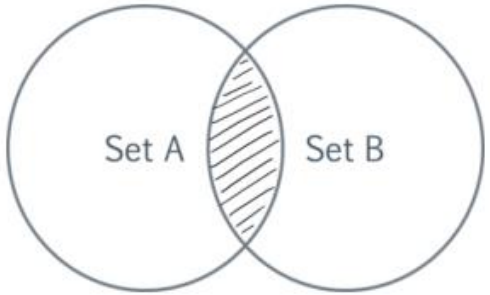
Belongs to	• *an element* ***'a'*** *which is part of Set A can be shown as* $a \in A$ • *If an element b is* ***not part of*** *Set A, then b do not belongs to A can be shown as* $b \notin A$
Subset	• *if every element of Set A is also an element of Set B we say that A is a subset of B* $A \subset B$ • *we can say that B is a super set of A shown as* $B \supset A$

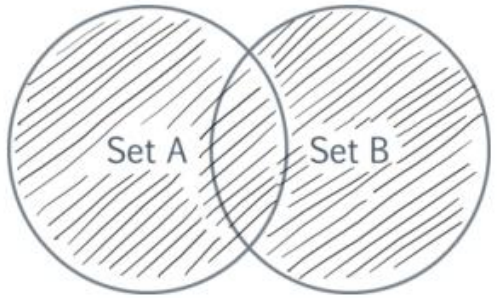
Types of Subsets	• ***Proper Subset:*** *When A is a subset of B and both sets* ***are not equal****, then A is a* ***proper*** *subset of B.* • ***Improper Subset:*** *When A is a subset of B and also B is a subset of A, then both are improper subsets of each other, and this is possible only when they are equal.* $A \subseteq B$
Number of Subsets	• *No. of possible subsets of any set* = 2^n • *No. of proper subsets of any set* = $2^n - 1$

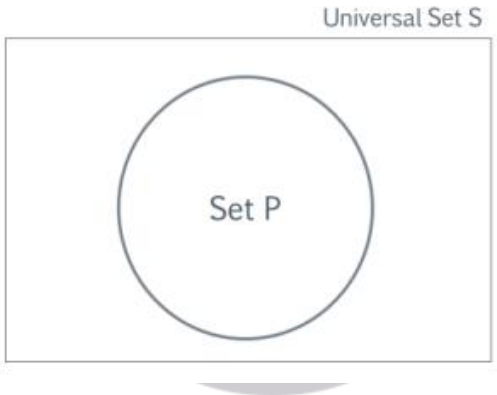
Equal Sets	• *Two sets are equal if they have* ***exactly same*** *elements* • ***Order*** *of elements has* ***no relevance***
Null Set / Empty Set / Void Set	• *A set which contains* ***no elements*** • *It is a proper subset of all sets* • *It has no proper subsets* • *Denoted by* $\{\}$ *or* φ
Cardinal Number	• *The number of distinct elements contained in a finite set A is called its cardinal number.* • *It is shown as* ***n(A)*** ***Example****: If A = {2, 5, 7, 9} then n(A) = 4*

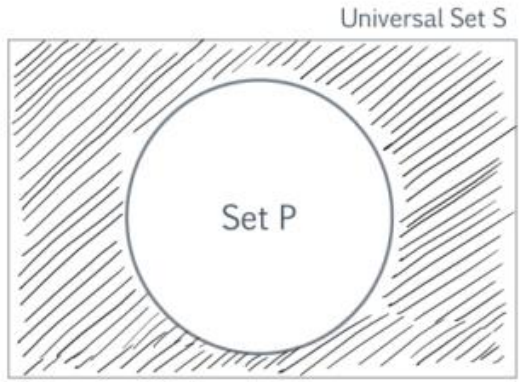
Intersection Sets	• *A new set that contains all the* ***common elements*** *between set A and set B is called as intersection set of set A and B.* • *It is denoted by* $A \cap B$
Union Set	• *A set that contains all the elements of Set A and Set B without repeating the common elements between them is called* ***Union Set*** *of A and B* • *It is denoted by* $A \cup B$
Universal Set	• *The set which contains all the elements under consideration in a particular problem is called the universal set denoted by S.*
Complimentary Set	***Complimentary Set of P:*** • *It is a set that contains all the elements of universe other than P* • *It is denoted by* P' *or* P^c
De-Morgan's Law	$(P \cap Q)' = P' \cup Q'$ ***or*** $(P \cup Q)' = P' \cap Q'$
Set A-B	*It is a set that contains all the elements of A which are not common with B. A-B Set can also be called as* ***Only A***
Set B-A	*It is a set that contains all the elements of B which are not common with A. B-A Set can also be called as* ***Only B***

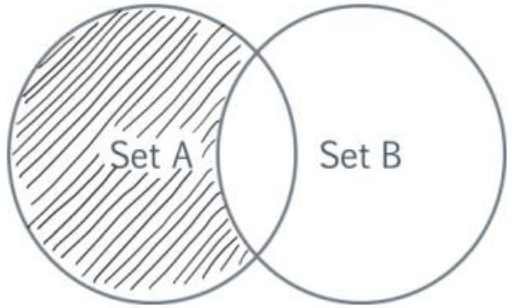
Equivalent Sets	• *Two sets having same cardinal number are called as Equivalent Sets* • *If n(A) = n(B) then A and B are equivalent sets*
Power Set	• *The collection of all possible subsets of a given set A is called the power set of A.* • *It is denoted by* ***P(A)***

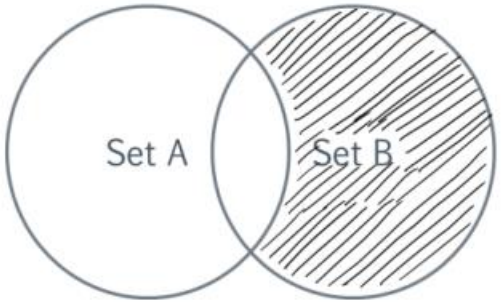












Venn Diagrams

Rules in Venn Diagram	• ***Ordinary Set*** *: Circle* • ***Intersection Set*** *: Shade Common Portion of two Circles* • ***Union Set*** *: Shade Full Portion of Both Circles* • ***Universal Set*** *: Rectangle* • ***Complimentary Set of P*** *: Shade Rectangle other than P* • ***A – B*** *: Shade portion of A other than common portion with B* • ***B – A*** *: Shade portion of B other than common portion with A*
2 Sets Formula	$n(A \cup B) = n(A) + n(B) - n(A \cap B)$
3 Sets Formula	$n(A \cup B \cup C) = n(A) + n(B) + n(C) - n(A \cap B)$ $-n(B \cap C) - n(C \cap A) + n(A \cap B \cap C)$
Word Problem Tip 1	*Total Quantity given in the question may be taken as* ***Universal*** *or* ***Union*** *–will depend on data available in the question and situation.*
Word Problem Tip 2	*If we are using venn diagram to solve word problem, we should put counts from inside to outside in the diagram.*

Relations

Ordered Pair	*Two elements a and b, listed in a* ***specific order****, form an ordered pair, denoted by* ***(a, b)***
Cartesian Product of Sets:	• *If A and B are two non-empty sets,* • *then the set of all ordered pairs (a, b) such that a belongs to A and b belongs to B , is called the Cartesian product of A and B, denoted by* $A \times B$
How to Denote Product Set	$A \times B = \{(a,b) : a \in A, b \in B\}$
Why Product Set	$n(A \times B) = n(A) \times n(B)$

Relation Set	• *Relation set from* ***A to B*** *is any* ***subset of product set*** $A \times B$ • *containing only those elements which* ***satisfy a given relation*** *between both the elements of ordered pair* • *Format:* $R : A \to B = \{(a,b) : a \text{ is related to } b, a \in A, b \in B\}$

PRANAV POPAT

<table>
<tr><td>Types of Relations</td><td>
<table>
<tr><td>Reflexive</td><td>If relation sets contains ordered pair in the form of (a,a), (b,b) and so on</td></tr>
<tr><td>Symmetric</td><td>If relation set contains an ordered pair (a,b) it must also contain (b,a)</td></tr>
<tr><td>Transitive</td><td>If relation set contains an ordered pair (a,b) and (b,c) it must also contain (a,c)</td></tr>
<tr><td>Equivalence</td><td>If a relation is Reflexive, Symmetric and Transitive then it is called as Equivalence</td></tr>
</table>
</td></tr>
<tr><td>Number of Relations between two sets</td><td>• 2^n where n = no. of elements in the product set</td></tr>
<tr><td>Terms related to Relation</td><td>
<table>
<tr><td>Domain of Relation</td><td>Set of all first coordinates</td></tr>
<tr><td>Range of Relation</td><td>Set of all second coordinates</td></tr>
<tr><td>Identity Relation</td><td>$\{(a,a) : a \in A\}$</td></tr>
<tr><td>Inverse Relation</td><td>If R be a relation on A, then the relation R^{-1} on A, defined by $R^{-1} = \{(b,a) : (a,b) \in R\}$ is called inverse relation on A</td></tr>
<tr><td>Domain of R^{-1}</td><td>Range of R</td></tr>
<tr><td>Range of R^{-1}</td><td>Domain of R</td></tr>
</table>
</td></tr>
</table>

Functions

<table>
<tr><td>Function (Practical Meaning)</td><td>it is a processor of numbers which takes input number, process it and gives output</td></tr>
<tr><td>Function Set</td><td>Function Set:
It is a relation set with the condition that
No distinct ordered pairs of set have same first element</td></tr>
<tr><td>Denote a Function</td><td>if f is a function defined from Set A to Set B it is denoted as $f : A \rightarrow B$</td></tr>
<tr><td>Check Function using Mapping</td><td><table><tr><th>Mapping</th><th>Considered as Function</th></tr><tr><td>One to One</td><td>Function</td></tr><tr><td>One to Many</td><td>Not a Function</td></tr><tr><td>Many to One</td><td>Function</td></tr></table></td></tr>
</table>

CA PRANAV POPAT

Terms Used in Function	*If a function is defined from Set A to Set B i.e., $f : A \rightarrow B$*

Domain	*Set A = First Set = Set from where first elements (inputs) of ordered pair are taken*
Codomain of Function	*Set B = Second Set = Set from where second elements (outputs) are taken*
Range of Function	*Set of those elements of Codomain which are part of Function Set. It is a subset of Codomain. It may or may not be equal to Codomain.*
Preimage	*Input or First element in an ordered pair of Function Set*
Image	*Output or Second element in an ordered pair of Function Set*

CA PRANAV POPAT

Types of Function – Based on Mapping	***One-One Function (Injective)***	• *Let $f : A \to B$, if different elements in A have different images in B, then f is said to be a one-one* • *Also called as injective function or mapping.*
	Many-One Function	• *Let $f : A \to B$, if two or more elements in A have common image in B, then f is said to be many-one*
Types of Function – Based on Range	***Onto Function (Surjective)***	• *Let $f : A \to B$, if every element in B has at least one pre-image in A, then f is said to be an onto function.* • *Also called as Surjective Function* • *In an onto function, Range = Codomain*
	Into Function	• *Let $f : A \to B$, If even one element in B codomain is not having any pre-image in A, then f is said to be an into function.* • *In other words, if even one is single in Codomain Set* • *In an onto function, $Range \subset Codomain$*

Example 1

	Example
(1)	$A = \{2, 5, 7, 11\}$, $B = \{8, 20, 28, 44, 48\}$, $f(x) = 4x$. Evaluate the type of Function.

$\{(2,8), (5,20), (7,28), (11,44)\}$

above f is **one-one** as all elements of domain have different images

since one element of codomain (set B) i.e. 48 has no preimage f is said to be **into.**

Here range is a subset of codomain.

	Example
(2)	$A = \{6, 8, 14, 10\}$, $B = \{3, 7, 5, 4\}$, $f(x) = x/2$. Evaluate the type of Function.

$\{(6,3), (8,4), (14,7), (10,5)\}$

all elements of domain have different images so f is **one-one**

all elements of codomain have atleast one pre image i.e. all are paired so f is **onto**

CA PRANAV POPAT

Example

(3) $A = \{-2, -1, 0, 1, 2\}$, $B = \{0, 1, 2, 4\}$, $f(x) = x^2$. Evaluate the type of Function.

$\{(-2,4)\ (-1,1)\ (0,0)\ (1,1)\ (2,4)\}$

as -2 and 2 have common image 4, also -1 and +1 have common image 1 so f is many to one

Here one element of codomain i.e. 2 which has no preimage hence f is into.

Example

(4) $A = \{-2, -1, 0, 1, 2\}$, $B = \{0, 1, 4\}$, $f(x) = x^2$. Evaluate the type of Function.

$\{(-2,4)\ (-1,1)\ (0,0)\ (1,1)\ (2,4)\}$

many to one

all elements of codomain are having atleast one preimage so f is onto.

also here codomain = range

	one one	many one
onto	example 2	example 4
into	example 1	example 3

Bijection Function	• *A one-one and onto function is said to be bijective.* • *It is also called as one-to-one correspondence.* ***Injective + Surjective = Bijective***
Identity Function	• *Identical = Same* • *If in a function set values of preimage and image are same for all ordered pairs.*
Constant Function	• *If in a function the value of image remains constant for any value of preimage*
Equal Functions	• *Two functions f and g are said to be equal, written as* $f = g$ • *if they have the same domain and they satisfy the condition* $f(x) = g(x)$, *for all x.*
Even and Odd Functions	• *A function is even if* $f(-x) = f(x)$ • *A function is odd if* $f(-x) = -f(x)$

Composition of Functions	• $fog = fog(x) = f[g(x)]$ • $gof = gof(x) = g[f(x)]$
Inverse Functions	• *In a function a set of preimages when used as input gives us images, now to obtain such a function which can be used in reverse way i.e., using image values as input and gives pre-images as output.* • *Inverse is possible only if the function is bijective.*
Steps to obtain inverse of a function	1. *Write your function in the form of y:* $y = f(x)$ 2. *From above expression, find the value of x:* $x = \square$ 3. *Interchange value of x and y:* $y = \square$, *now the RHS is inverse function of* $f(x)$

CA PRANAV POPAT

Concept, Formula and Tricks Marathon QA

Limits and Continuity

Concept of Limits	
Application	• *In real life – very important in physics and foundation of all the technologies* • *For CAF Exams – direct MCQ possible and also used in calculus chapter*
Concept of Limits	• *In mathematics, a limit is the value that a function approaches as the input approaches some value.* • *Limits are essential to calculus and mathematical analysis, and are used to define continuity, derivatives, and integrals* • *If the function is continuous and determinant, these limits are simply equal to the value of the function* • *But if there is a discontinuity, we may need to explore other ways to do it.*
How to check if limit exists?	• *Limit is said to exist at any point when both left hand and right hand limit exists and they are equal.* $\lim_{x \to a-} f(x) = \lim_{x \to a+} f(x) = \lim_{x \to a} f(x)$

CA PRANAV POPAT

General limit rules	*Let* $\lim_{x\to a} f(x) = l$ *and* $\lim_{x\to a} g(x) = m$, *then* *i.* $\lim_{x\to a}\{f(x)+g(x)\} = \lim_{x\to a} f(x) + \lim_{x\to a} g(x) = l+m$ *ii.* $\lim_{x\to a}\{f(x)-g(x)\} = \lim_{x\to a} f(x) - \lim_{x\to a} g(x) = l-m$ *iii.* $\lim_{x\to a}\{f(x).g(x)\} = \lim_{x\to a} f(x).\lim_{x\to a} g(x) = l\times m$ *iv.* $\lim_{x\to a}\{f(x)/g(x)\} = \lim_{x\to a} f(x) / \lim_{x\to a} g(x) = l/m$ *v.* $\lim_{x\to a}\{k.f(x)\} = k.\lim_{x\to a} f(x) = k\times l$
Some Important Limits	*i.* $\lim_{x\to 0}\frac{1}{x}$ *do not exist as LHL and RHL are not same* *ii.* $\lim_{x\to 0}\frac{e^x-1}{x} = 1$ *iii.* $\lim_{x\to 0}\frac{a^x-1}{x} = \log_e a \text{ or } \ln a$ *iv.* $\lim_{x\to 0}\frac{\log_a(1+x)}{x} = \log_a e$ *v.* $\lim_{x\to 0}\frac{\log_e(1+x)}{x} = \log_e e = 1$ *vi.* $\lim_{x\to\infty}\left(1+\frac{1}{x}\right)^x = e$ *vii.* $\lim_{x\to 0}(1+x)^{\frac{1}{x}} = e$

Concept of Continuity

Application	• A function is continuous if it is not undefined at any point • A function is continuous at $x = a$, if below conditions are met: i. $f(x)$ is defined at $x = a$ ii. $\lim_{x \to a^-} f(x) = \lim_{x \to a^+} f(x)$ iii. $\lim_{x \to a} f(x) = f(a)$
Properties	• The sum, difference and product of two continuous functions is a continuous function. • The quotient of two continuous functions is continuous function provided the denominator is not equal to zero.

CA PRANAV POPAT

CA FOUNDATION QA

Chp8 Differentiation & Integration

Concept, Formula and Tricks Marathon



Concepts

Differentiation - Basics

Meaning	• *the process of finding the derivative of a continuous function* • *It is defined as the limiting value of the ratio of the change (increment) in the function corresponding to a small change (increment) in the independent variable (argument) as the later tends to zero*
Many Names but same meaning	*Derivative of Function,* $\frac{dy}{dx}$, $f'(x)$, y', y_1, *Gradient of Curve, Slope of Tangent*
Use in Economics	• *To increase profits of company, management may decide to increase production but they are concerned about how the increase will affect of cost of production. Economists use Marginal Cost Function to understand this variation* • *Differentiation is also used to obtain Marginal Revenue, Marginal Profit, Marginal Propensity to Consume, Marginal Propensity to Save etc.*

POPAT

Standard Results / Basic Formulas

Function	Derivative of Function
$f(x)$	$f'(x)$
x^n	nx^{n-1}
e^x	e^x
a^x	$a^x \log_e a$
Constant	0
e^{ax}	ae^{ax}
$\log x$	$\frac{1}{x}$

CA PRANAV POPAT

Laws of Differentiation

Function	Derivative of Function
$if\ h(x)=c.f(x)$	$h'(x)=c.f'(x)$
$if\ h(x)=f(x)\pm g(x)$	$h'(x)=f'(x)\pm g'(x)$
$if\ h(x)=f(x)\times g(x)$	$h'(x)=f(x).g'(x)+g(x).f'(x)$
$if\ h(x)=\frac{f(x)}{g(x)}$	$h'(x)=\frac{\left[g(x).f'(x)-f(x).g'(x)\right]}{\left[g(x)\right]^2}$

Chain Rule

Basics	• *It is used for finding differentiation in case of composite functions* • *Here we will differentiate from outer to inner function* • $if\ h(x)=f[g(x)],\ take\ g(x)=t,\ then\ h'(x)=f'(t)\times g'(x)$

PRANAV POPAT

Implicit Functions

Meaning	• A Function in the form of $f(x,y)=0$ is known as implicit function • In this type of function y cannot be directly defined as a function of x
Process to differentiate	• We don't need to convert implicit function to explicit function • Here we do differentiation of both sides with respect to x and try to find out $\frac{dy}{dx}$

Parametric Equations

Meaning	• When both variables x and y are expressed in terms of a third variable (parameter), the involved equations are called as parametric equations
Process to differentiate	• We can write $\frac{dy}{dx}=\frac{dy}{dt}\times\frac{dt}{dx}$ • So we will separate differentiate both in terms of t and then put it in the above shown form.

Logarithmic Differentiation

Use	• When equation is in the form where variable is in power of another variable, we need to take log on both sides then only it can be differentiated

Higher Order/ Degree Differentiation

First Order Differentiation	$\frac{dy}{dx} = f'(x) = y' = y_1$
Second Order Differentiation	$\frac{d}{dx}\left(\frac{dy}{dx}\right) = \frac{d^2y}{dx^2} = y'' = y_2$
Third Order Differentiation	$\frac{d}{dx}\left(\frac{d^2y}{dx^2}\right) = \frac{d^3y}{dx^3} = y''' = y_3$

CA PRANAV POPAT

Application of Differentiation in Cost and Revenue Functions (Economics)

Marginal Cost Function	Marginal Cost Function is derivative of Cost Function $M(x) = C'(x)$
Min / Max Point of Marginal Cost Function	To obtain min or max point we differentiate Marginal Cost Function and obtain $M'(x)$ and equate it to 0 to get value of x This value of x will be then substituted in $M''(x)$ if it result to negative value then Marginal Cost is maximum at x and if it result to positive value then Marginal Cost is minimum at x
Min / Max Point of Average Cost Function	The process will be similar to what we used in Marginal Cost Function.
Marginal Revenue Function	Marginal Revenue Function is derivative of Total Revenue Function
Break Even Point	It is the level of output x where Cost and Revenue Functions are equal $C(x) = R(x)$
Marginal Propensity to Consume (MPC)	If $C = F(Y)$ is a consumption function where Y is Total Income then marginal propensity is equal to $\frac{dC}{dY}$
Marginal Propensity to Save (MPS)	If S is saving i.e. difference between Income and Consumption then MPS is equal to $\frac{dS}{dY}$

Integration – Basics

Meaning	• Integration is the reverse process of differentiation • If $f'(x)$ is the differentiation of $f(x)$ then $f(x)$ is the integration of $f'(x)$ • Integration of any function $f(x)$ wrt to dx is denoted as $\int f(x)dx$ • It was invented to find area bounded by curves
Basic Formulas	• $\int x^n dx = \frac{x^{n+1}}{n+1} + c$ • $\int 1.dx = \int dx = x + c$ • $\int e^x dx = e^x + c$ • $\int e^{ax} dx = \frac{e^{ax}}{a} + c$ • $\int \frac{dx}{x} = \log x + c$ • $\int a^x dx = \frac{a^x}{\log_e a} + c$ Note: In the answer for all integral sums, we add +c (constant of integration) since the differentiation of constant is always zero

Integration – Rules, Methods and Formulas

Elementary Rules	• $\int c\, f(x)\,dx = c\int f(x)\,dx$ where c is constant • $\int \{f(x)\,dx \pm g(x)\,dx\} = \int f(x)\,dx \pm \int g(x)\,dx$
Method of Substitution	• When a function cannot be integrated directly, we transform the function into another in such a way that it can be readily integrated • Here usually we take inner function as t and differentiate it with respect to x x then find value of $dx\,dx$ in the form of $dt\,dt$

CA PRANAV POPAT

Standard Formulas

- $\int \frac{dx}{x^2-a^2} = \frac{1}{2a}\log\frac{x-a}{x+a}+c$
- $\int \frac{dx}{a^2-x^2} = \frac{1}{2a}\log\frac{a+x}{a-x}+c$
- $\int \frac{dx}{\sqrt{x^2+a^2}} = \log\left(x+\sqrt{x^2+a^2}\right)+c$
- $\int \frac{dx}{\sqrt{x^2-a^2}} = \log\left(x+\sqrt{x^2-a^2}\right)+c$
- $\int e^x\{f(x)+f'(x)\}dx = e^x f(x)+c$
- $\int \sqrt{x^2+a^2}\,dx = \frac{x}{2}\sqrt{x^2+a^2}+\frac{a^2}{2}\log\left(x+\sqrt{x^2+a^2}\right)+c$
- $\int \sqrt{x^2-a^2}\,dx = \frac{x}{2}\sqrt{x^2-a^2}-\frac{a^2}{2}\log\left(x+\sqrt{x^2-a^2}\right)+c$
- $\int \frac{f'(x)}{f(x)}dx = \log f(x)+c$

Integration by Parts

- $\int uv\,dx = u\int v\,dx - \int\left[\frac{du}{dx}\int v\,dx\right]dx$, *where u and v are two functions*
- *To decide which function to be taken as u, ILATE rule is used*

I	L	A	T	E
Inverse Trigonometric	Logarithmic	Algebraic	Trigonometric	Exponential
$\sin^{-1}x, \cos^{-1}x$	$\log x$	$x^2, \sqrt{x}$	$\sin x, \cos x$	$e^x, 2^x$

Method of Partial Fractions	• ***Type I:*** $\int \frac{(3x+2)dx}{(x-2)(x-3)}$ • ***Type II:*** $\int \frac{(3x+2)dx}{(x-2)^2(x-3)}$ • ***Type III:*** $\int \frac{3x^2-2x+5}{(x-1)^2(x^2+5)}dx$

CA PRANAV POPAT

Definite Integration

Basics	• Integration we dealt with above was indefinite integral integration as we take integration of changes for all possible values of x • When we deal with integration as x changes from a to b the value of integral changes from $f(a)$ to $f(b)$ • This is shown mathematically as $\int_a^b f(x)dx = f(b) - f(a)$ • b is called upper limit and a is lower limit of integration. In definite integration constant c should not be added
Properties of Definite Integration	• $\int_a^b f(x)dx = \int_a^b f(t)dt$ • $\int_a^b f(x)dx = -\int_b^a f(x)dx$ • $\int_a^b f(x)dx = \int_a^c f(x)dx + \int_c^b f(x)dx, a < c < b$ • $\int_0^a f(x)dx = \int_0^a f(a-x)dx$ • $\int_{-a}^a f(x)dx = \begin{cases} 2\int_0^a f(x)dx, if\ f(-x) = f(x) \\ 0, if\ f(-x) = -f(x) \end{cases}$

CA FOUNDATION QA

Chp9 Number Series, Coding, Decoding, Odd Man Out

Concept, Formula and Tricks Marathon



Concepts

Number Series	• *In this type of questions, we need to identify the missing term of the series real according to a **specific pattern** of the series rule to form its code.* • *The students are required to detect the missing number of the series and answer the questions accordingly.*
Alphabet Series	• *Alphabet series consists of letters of the alphabet placed in a specific pattern.*
Coding and Decoding	• *Coding: converting original message into code* • *Decoding: converting code into original message* • *Types: Number Coding, Letters Coding etc.*
Odd Man Out	• *In these questions, four or five elements are given, out of which one does not belong to the group, we need to identify it.*
Classification of MCQs	*Quickly Identifiable* — *Pattern which can be identified instantly during exams* *Not Quickly Identifiable (Advance)* — *Pattern which required lot of thinking and even after that there is no surety that it will click on spot*

CA FOUNDATION QA

Chp10 Direction Test

Concept, Formula and Tricks Marathon



Concepts

Direction Test

Question Requirements

- Here questions consist of a sort of direction puzzle.
- A successive follow-up of direction is formulated and
- the students are required to ascertain the final direction or
- direction with respect to starting points or
- other related problems.

Types of Question

Type	Requirement	Remark
Type I	Finding final direction	Distance can be ignored
Type II	Finding the direction with respect to original point	Distances are important
All other varieties	Multiple	No Remark

CA FOUNDATION QA

Chp11 Seating Arrangements

Concept, Formula and Tricks Marathon



Concepts

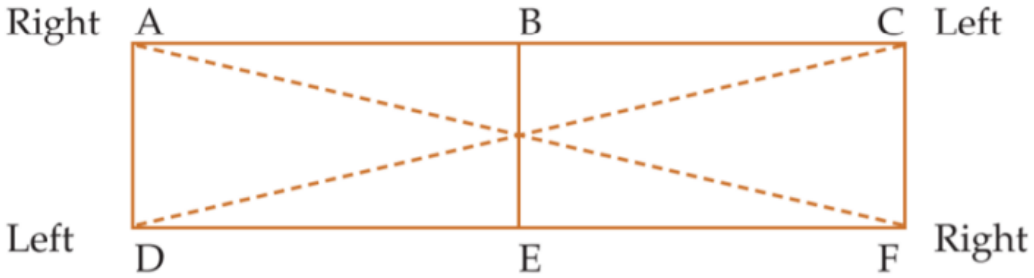
Seating Arrangements - Basics

<table>
<tr><td>Meaning</td><td>The process of making a group of people to sit as per a prefixed manner is called Seating Arrangement</td></tr>
<tr><td>Question Requirement</td><td>In these questions, some conditions are given on the basis of which students are required to arrange objects, either in a row or in a circular order.</td></tr>
<tr><td>Types</td><td>Linear – One Row, Linear – Two Rows, Circular</td></tr>
<tr><td>Pro-Tip of forming diagram</td><td>

We need to make the diagram taking persons Name/ Alphabet Code
First make multiple mini-diagrams using multiple small information.
Try to connect these and make a master diagram.
Be Careful about number of objects and their sense in the context
In some cases if information is not very clear, we may need to make multiple diagrams and then decide the correct one eventually

</td></tr>
<tr><td>Linear Arrangement (Single Row)</td><td>
The arrangement is done only on one ‘axis’ and hence, the position of persons or objects is important in terms of order/ position.
Useful Interpretations
<table>
<tr><td>A is immediate left of P</td><td>Then it is sure no one is between A and P and A is just left of P</td></tr>
<tr><td>A is left of P
•</td><td>In this case A may be on exact left of P or may be at left after some places</td></tr>
</table>
</td></tr>
</table>

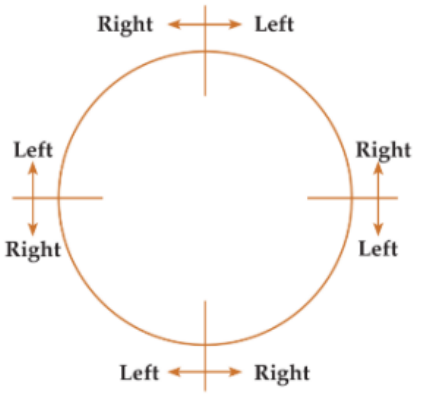
Two Rows Linear Arrangement



- *A is sitting opposite to D*
- *B is sitting opposite to E*
- *C is sitting opposite to F*
- *D and C are sitting at diagonally opposite positions*
- *A and F are sitting at diagonally opposite positions*

Circular Arrangements

When person facing centre	*Clockwise = Left* *Anticlockwise = Right*
When person facing opposite to Centre	*Clockwise = Right* *Anticlockwise = Left*



POPAT

CA FOUNDATION QA

Chp12 Blood Relations

Concept, Formula and Tricks Marathon



Concepts

Important Relations	***Grandfather's son*** — *Uncle or Father* ***Grandfather's only son*** — *Father* ***Brother's Daughter*** — *Niece* ***Brother's Son*** — *Nephew* ***Uncle's Son/ Daughter*** — *Cousin* ***Brother's wife*** — *Sister-in-law* ***Sister's Husband*** — *Brother-in-law* ***Son's wife*** — *Daughter-in-law*

CA PRANAV POPAT

Pro Tips to create chart in Blood Relations	• *Same **Generation** in Same **Row*** • *Distinct marking for Gender (as soon as you know)* | *Male* | + | *Female* | – | • *If there is a **couple put brackets*** • *Make a empty box if name not given* • *If the information is not in proper linkage **hold that statement**, go to next one and then come back again* • *Convert your question requirement into a **fill in the blanks format*** • *If you are getting two answers for a question – just chill, check the options* • *For **pointing based question**, use empty boxes and also create arrows of relation if needed.*

PRANAV POPAT

CA FOUNDATION QA

Chp13 Statistical Description of Data

Concept, Formula and Tricks Marathon



Concepts

<table>
<tr><td>Definition of Statistics</td><td>

Plural Sense: Any data – quantitative or qualitative used for statistical analysis.
Singular Sense: Scientific method of collecting, analyzing, and presenting data to draw statistical inferences. It is also called as Science of Averages or Science of Counting

</td></tr>
<tr><td>Origin of Word</td><td>
<table>
<tr><th>Language</th><th>Actual Word</th><th>Memorize by</th></tr>
<tr><td>Latin</td><td>Status</td><td>Latus</td></tr>
<tr><td>Italian</td><td>Statista</td><td>Pasta</td></tr>
<tr><td>German</td><td>Statistic</td><td>Breadstick</td></tr>
<tr><td>French</td><td>Statistique</td><td>Barbeque</td></tr>
</table>
</td></tr>
<tr><td>Publication</td><td>
<table>
<tr><td>Koutilya's Arthashastra</td><td>Record of Birth and Deaths
Chandragupta's reign
4th Century B.C</td></tr>
<tr><td>Abu Fezal's Ain-i-Akbari</td><td>Record on Agriculture
Akbar Reign
16th Century A.D.</td></tr>
<tr><td>First Census</td><td>Egypt 300 BC to 2000 BC
By Pharaoh</td></tr>
</table>
</td></tr>
<tr><td>Application of Statistics</td><td>

Economics: Demand Analysis, Future Projection etc.
Business Management: Decision making using quantitative techniques not intuition

</td></tr>
</table>

	• *Industry and Commerce: Profit maximization using business data – sales, purchase, market etc. by consulting experts*
Limitation of Statistics	• *It deals with aggregate data and not individual data* • *Quantitative data can only be used, however for qualitative – it needs to be converted into quantitative* • *Projections are based on conditions/ assumptions and any change in that will change the projection. Example: Future projections of sales* • *Sampling based conclusions are used, improper sampling leads to improper results. Random Sampling is must.*
Data	• *Quantitative Information shown as number* • *Primary: first time collected by agency/ investigator* • *Secondary: collected data used by different person/ agency*
Variable	• *Measurable Data – Value can vary*

Discrete Variable	• *When a variable assumes a finite or count ably infinite isolated values.* • *Example: no. of petals in a flower, no. of road accident in locality*
Continuous Variable	• *When a variable assumes any value from the given interval (can also be in decimals, fractions).* • *Example: height, weight, sale, money*

<table>
<tr><td>Special point about Money</td><td>

- *In general, Money is a discrete variable*
- *However, from commercial point of view, money is a continuous variable.*

</td></tr>
<tr><td>Attribute</td><td>

- *Qualitative Characteristics. Example: gender of a baby, the nationality of a person, the colour of a flower etc.*

</td></tr>
<tr><td>Collection of Primary Data – Interview Method</td><td>
<table>
<tr><th>Method</th><th>Details</th></tr>
<tr><td>Personal Interview</td><td>

- *Where data is collected directly from respondents.*
- *Highly Accurate – Low Coverage*
- *Example: Natural Calamity, Door to Door Survey*

</td></tr>
<tr><td>Indirect Interview</td><td>

- *When reaching respondent is difficult, data is collected by contacting associated persons.*
- *Highly Accurate – Low Coverage*
- *Example: Rail accident*

</td></tr>
<tr><td>Telephone Interview</td><td>

- *Data is collected over phone*
- *Quick and non-expensive method*
- *Low Accuracy – High Coverage*

</td></tr>
</table>
</td></tr>
<tr><td>Collection of Primary Data – Mailed Questionnaire Method</td><td>

- *In this method well drafted and soundly sequenced questionnaire,*
- *covering all the important aspects of the data requirement is sent to respondent for filling.*
- *Here coverage is wide but amount of non-responses will be maximum*

</td></tr>
</table>

Collection of Primary Data – Observation Method	• *In this method data is collected by direct observation or using instrument.* • *For example: data on height and weight for a group of students.* • *Although more accurate but it is time consuming, low coverage and laborious method.*
Collection of Primary Data – Questionnaire Filled and sent by Enumerators	• *Mix of Interview and Mailed Questionnaire* • *Enumerator means a Person who directly interacts with respondent and fills the questionnaire.* • *It is generally used in case of Surveys and Census.*
Sources of Secondary Data	

International Sources	*World Health Organization (WHO), International Monetary Fund (IMF), International Labor Organization (ILO), World Bank*
Government Sources	*In India – Central Statistics Office (CSO), Indian Agricultural Statistics by the Ministry of Food and Agri, National Sample Survey Office- NSSO, Regulators – RBI, SEBI, RERA, IRDA*
Private or Quasi-govt. sources	*Indian Statistical Institute (ISI), Indian Council of Agriculture, NCERT*

Concept, Formula and Tricks Marathon QA

<table>
<tr><td>Scrutiny of Data</td><td>

checking accuracy and consistency of data
There is no rule for it, one must apply his intelligence, patience and experience while scrutinizing the given information.
Internal Consistency: When two or more series of related data are given, we should check consistency among them.

</td></tr>
<tr><td>Presentation of Data – Classification / Organization of Data</td><td>
Classification or Organisation: putting data in a neat, precise, and condensed form, making it comparable, suitable for analysis, more understandable.
<table>
<tr><td>Chronological/ Temporal/ Time Series Data</td><td>Data arranged based on Time
Example: Revenues YoY i.e year on year</td></tr>
<tr><td>Geographical or Spatial Series Data</td><td>Arrangement based on regions
Example: Country wise Revenue of a global company</td></tr>
<tr><td>Qualitative or Ordinal Data</td><td>Based on some attribute
Nationality Wise Medal Winners in Olympics</td></tr>
<tr><td>Quantitative or Cardinal Data</td><td>Based on some variable
Example: Frequency Distribution of a Data</td></tr>
</table>
</td></tr>
<tr><td>Mode of Presentation of Data – Textual</td><td>

This method comprises presenting data with the help of a paragraph or several paragraphs.
This is not a suitable mode of presentation as it is dull, monotonous and non-comparable.

</td></tr>
</table>

CA PRANAV POPAT

<table>
<tr><td>Mode of Presentation of Data – Tabular Form</td><td>

When data is shown in the form of Table.
Useful in easy comparison
Complicated data can be presented
Table is must to create a diagram
No analysis possible without table
Components of Table

</td></tr>
<tr><td>Components of Table</td><td>
<table>
<tr><th>Description</th><th>Name of Component of Table</th></tr>
<tr><td>Entire Upper Part</td><td>Box Head</td></tr>
<tr><td>Upper Part describing columns and sub-columns</td><td>Caption</td></tr>
<tr><td>Left part of the table describing rows</td><td>Stub</td></tr>
<tr><td>Main Data of Table</td><td>Body</td></tr>
<tr><td>Source of Data at the bottom of Table</td><td>Footnote</td></tr>
</table>
</td></tr>
<tr><td>Mode of Presentation of Data – Diagrams</td><td>

Can be used by educated and uneducated section of society
Hidden trend can be traced
If priority is accuracy, then tabulation is better

</td></tr>
</table>

Line Diagram	• *Time Series is generally in x axis* • *For wide fluctuation – log chart or ratio chart is used* • *Two or more series of same unit – Multiple Line Chart* • *Two or more series of different unit – Multiple Axis Chart*
Bar Diagram	• *Bar means rectangle of same width and of varying length drawn horizontally or vertically* • *For comparable series – multiple or grouped bar diagrams can be used* • *For data divided into multiple components – subdivided or component bar diagrams* • *For relative comparison to whole, percentage bar diagrams or divided bar diagrams* • *Vertical Bar Diagram: Useful for Data varying over Time and Quantitative Data* • *Horizontal Bar Diagram: Useful for Data varying over Space and Qualitative Data*
Pie Chart	• *Used for circular presentation of relative data (% of whole)* • *Summation of values of all components/segments are equated to 360 Degree (total angle of circle)* • ***Segment angle =*** $\frac{(segment\ value \times 360°)}{(total\ value)}$

Statistical Description of Data – Frequency Distribution

<table>
<tr><td>Frequency and Distribution</td><td>

Frequency means number of times a particular observation is repeated.
Frequency Distribution is table which contains observation or class intervals in one column and corresponding frequency in the other.
Definition: A frequency distribution may be defined as a

tabular representation of statistical data, usually in an ascending order,
relating to a measurable characteristic
according to individual value or a group of values of the characteristic under study.

</td></tr>
<tr><td>Types of Frequency Distribution</td><td>
<table>
<tr><td>Ungrouped/ Simple Frequency Distribution</td><td>

When there are limited number of distinct observations, frequency can be assigned to each one of them.
This distribution is simple

</td></tr>
<tr><td>Grouped Frequency Distribution</td><td>

When there are large no. of observations, grouping is done among them (generally in ascending order).
Each group is called as class interval and frequency is assigned to group and not individual values,
this is called Grouped Frequency Distribution

</td></tr>
</table>
</td></tr>
</table>

Concept, Formula and Tricks Marathon QA

Class Limit

- *For a class interval CL is the minimum and maximum value the class interval may contain*
- *Minimum Value – Lower Class Limit*
- *Maximum Value – Upper Class Limit*

Class Interval	Frequency	LCL	UCL
10-19	10	10	19
20-29	5	20	29
30-39	8	30	39

Classification of Grouped of Frequency Distribution

Mutually Exclusive / Overlapping Classification

Class	LCL	UCL
10-20	10	20
20-30	20	30
30-40	30	40

- *Here UCL an interval and LCL of next interval are same*
- *This is usually applicable for continuous variable.*
- *An observation which is equivalent to common class limit is excluded from the class interval where it is UCL and taken in the class where it is LCL.*

Mutually Inclusive / Non-Overlapping Classification

Class	LCL	UCL
10-19	10	20
20-19	20	30
30-39	30	40

- *There is no common class limits between two intervals.*
- *This is usually applicable to discrete variable.*
- *All observation including UCL and LCL will be taken in the same class interval as there is no confusion.*

CA PRANAV POPAT

Class Boundary

In case of Exclusive / Overlapping Classification

Class Boundary = Class Limit

Class	LCL	UCL	LCB	UCB
10-20	10	20	10	20
20-30	20	30	20	30
30-40	30	40	30	40

In case of Inclusive / Overlapping Classification

Lower Class Boundary

LCB = LCL – 0.5

UCB = UCL + 0.5

Class	LCL	UCL	LCB	UCB
10-19	10	19	9.5	19.5
20-29	20	29	19.5	29.5
30-39	30	39	29.5	39.5

Mid-Point / Class Mark / Mid Value of Class Interval

$$\frac{LCL+UCL}{2} \qquad \frac{LCB+UCB}{2}$$

- *Useful in calculation of AM, GM, HM, SD in case of grouped frequency distribution*

Class Length/ Width or Size

UCB – LCB only

Cumulative Frequency

- *Less than type: It shows no. of observations less than UCB*
- *More than type: It shows no. of observations more than UCB*

Class Interval	Freq.	UCB	Less than type CF	More than type CF	Total of both CF
44-48	3	48.5	3	33	36
49-53	4	53.5	7	29	36
54-58	5	58.5	12	24	36
59-63	7	63.5	19	17	36
64-68	9	68.5	28	8	36
69-73	8	73.5	36	0	36
Total	36				

Frequency Density

$$\frac{\text{Class Frequency}}{\text{Class Length of class}}$$

Relative Frequency

$$\frac{\text{Class frequency}}{\text{Total Frequency}}$$

It can take values between 0 and 1 exclusive

Percentage Frequency

$$\frac{\text{Class frequency}}{\text{Total Frequency}} \times 100$$

Frequency Dist. Diagram – Histogram

- *It is a convenient way to represent FD*
- *Comparison between frequency of two different classes possible*
- *It is useful to calculate mode also*

Frequency Polygon	• *Usually preferable for ungrouped frequency distribution* • *Can be used for grouped also but only if class lengths are even*
Ogives/ Cumulative Frequency	• *This graph can be made by both type of Cumulative Frequency and called as Less than Ogive or More than Ogive* • *It can be used for calculating quartiles, median*
Frequency Curve	• *It is a limiting form of Area Diagram (Histogram) or Frequency Polygon* • *It is obtained by drawing smooth and free hand curve though the mid points* • *Most used curve is Bell Shaped*

CA PRANAV POPAT

Theory Concepts not in ICAI Study Material

Descriptive Statistics vs Inferential Statistics	***Descriptive Statistics***	• *When a statistic is used only to describe scores in a sample (and not used to make inferences about populations), that is a descriptive use.* • *For example, a teacher may compute the mean test score for her class. She simply uses this number to think about how well her students did.*
	Inferential Statistics	• *Inferential use of statistics occurs when a data analyst uses information from a sample (such as a mean or correlation) to make inferences or guesses about values of the corresponding mean or correlation in a* ***population*** • *Statistical inference only works well when sample sizes are reasonably large* • *For example, in a political poll, a polling organization obtains the percent of people who state an intention to vote for Candidate X for a sample of votes, perhaps selected by random digit telephone dialing from the population of all registered voters.*

Simple Classification vs Manifold Classification	***Simple Classification***	• *When based on only one attribute, the given data is classified into two classes, which is known as Simple Classification.* • *For example, when the population is divided into literate and illiterate, it is a simple classification.*
	Manifold Classification	• *When based on more than one attribute, the given data is classified into different classes, and then sub-divided into more sub-classes, which is known as Manifold Classification.* • *For example, when the population is divided into literate and illiterate, then sub-divided into male and female, and further sub-divided into married and unmarried, it is a manifold classification.*
Ideographs/ Pictograms	• *A symbol that represents an idea or a thing, rather than the sounds of a word, is called an ideograph.* • *A smiley face emoji is an ideograph that represents happiness.* • *Many street signs are ideographs, meant to convey a specific meaning without using any words.*	

Concept, Formula and Tricks Marathon QA

Qualitative Data Types	***Nominal Data***	• *Nominal data are used to **label** variables without any quantitative value.* • *Common examples include hair color, nationalities, names of people, and so on.* • *Nominal means Name means Labels*
	Ordinal Data	• *Ordinal word comes from order* • *Ordinal scales are often used for measures of satisfaction, happiness, and so on.* • *Example: high-low-medium, strong-weak, etc.*
Research Data	***Research data can be placed into two broad categories:***	
	Quantitative data	• *Quantitative data are used when a researcher is trying to quantify a problem, or address the "what" or "how many" aspects of a research question.* • *It is data that can either be counted or compared on a numeric scale. For example, it could be the number of first year students at Macalester, or the ratings on a scale of 1-4 of the quality of food served at Cafe Mac. This data are usually gathered using instruments, such as a questionnaire which includes a ratings scale or a thermometer to collect weather data. Statistical analysis software, such as SPSS, is often used to analyze quantitative data.*

CA PRANAV POPAT

Concept, Formula and Tricks Marathon QA

	Qualitative data	• *Qualitative data describes qualities or characteristics.* • *It is collected using questionnaires, interviews, or observation, and frequently appears in narrative form.* • *For example, it could be notes taken during a focus group on the quality of the food at Cafe Mac, or responses from an open-ended questionnaire.* • *Qualitative data may be difficult to precisely measure and analyze. The data may be in the form of descriptive words that can be examined for patterns or meaning, sometimes through the use of coding.*
Sturges' Rule	*In the early 20th century, German statistician Herbert Sturges formulated a method (now called Sturges' Rule) of choosing the optimum number of bins in a histogram that minimize the potential for these pitfalls. His formula is simple:* $k = 1 + 3.322 \log n$ *where:* *k = the number of bins* *n = the number of observations in the data set.*	

CA PRANAV POPAT

Sampling

Disclaimer

- This topic is uses the concept of probability, theoretical distribution & combinations also
- Getting detailed understanding is not required
- If you get overview and able to solve Theory MCQs – that is also fine

CA PRANAV POPAT

Need of Sampling	• *There are situations when we would like to know about a vast, infinite universe or population.* • *But some important factors like time, cost, efficiency, vastness of the population make it almost impossible to go for a complete enumeration of all the units constituting the population.* • *In that case, we take choose to select a representative part of the population and infer/ conclude about the unknown universe on the basis of our knowledge from the known sample.*

CA PRANAV POPAT

Term - Population	*Population/ Universe*	• *It may be defined as the aggregate of all the units under consideration.* • *The number of units belonging to a population is known as population size.* • • *A population may be finite or infinite.* • *Population may also be regarded as existent or hypothetical.*
	Existent Population	• *A population consisting of real objects is known as an existent population.* • *Example: Population of lamps produced by a manufacturer*
	Hypothetical Population	• *A population that exists just hypothetically is known as a hypothetical or an imaginary population.* • *Example: the population of heads when a coin is tossed infinitely*

Concept, Formula and Tricks Marathon QA

Term – Sample	• *A sample may be defined as a part of a population so selected with a view to representing the population in all its characteristics.* • *Selection of a proper representative sample is pretty important because statistical inferences about the population are drawn only on the basis of the sample observations.* • *The number of units belonging to a sample is known as sample size.* • *The units forming the sample are known as "Sampling Units"* • *A detailed and complete list of all the sampling units is known as a "Sampling Frame"*
Term – Parameter	• *A parameter may be defined as a characteristic of a population based on all the units of the population.* • *Statistical inferences are drawn about population parameters based on the sample observations drawn from that population.* • *Example: Population Mean, Population Variance etc.*
Term – Statistic/ Estimator	• *A statistic may be defined as a statistical measure of sample observation and as such it is a function of sample observations.* • *If the sample observations are denoted by* $x_1, x_2, x_3, \ldots, x_n$ *then a statistic T may be expressed as* $T = f(x_1, x_2, x_3, \ldots, x_n)$ • *A statistic is used to estimate a particular population parameter.* • *The estimates of population can be population mean, variance or SD etc.*

CA PRANAV POPAT

Sampling Fluctuations	• *Value of the sample statistics (say mean, variance, proportion etc.) may vary from sample to sample as the sampling units of one sample may be different from that of another sample. The variation in the values of a statistic is termed as "Sampling Fluctuations".*

Sampling Distribution, Expectation, Standard Error	• *If it is possible to obtain the values of a statistic (T) from all the possible samples of a fixed sample size along with the corresponding probabilities, then we can arrange the values of the statistic, which is to be treated as a random variable, in the form of a probability distribution.* • *Such a probability distribution is known as the sampling distribution of the statistic.* • *The mean of the statistic, as obtained from its sampling distribution, is known as "Expectation" and the standard deviation of the statistic T is known as the "Standard Error (SE)" of T.*

POPAT

Standard Error Formula	• *SE can be regarded as a measure of precision achieved by sampling.* • *SE is inversely proportional to the square root of sample size* • *Formula: When Statistic is Mean* *In case of Simple Random Sampling With Replacement (SRS WR)* $SE(\bar{x}) = \frac{\sigma}{\sqrt{n}}$ *In case of Simple Random Sampling Without Replacement (SRS WOR)* $SE(\bar{x}) = \frac{\sigma}{\sqrt{n}} \cdot \sqrt{\frac{N-n}{N-1}}$ • *Formula: When Statistic is Proportion* *In case of Simple Random Sampling With Replacement (SRS WR)* $SE(p) = \sqrt{\frac{Pq}{n}}$ *In case of Simple Random Sampling Without Replacement (SRS WOR)* $SE(p) = \sqrt{\frac{Pq}{n}} \cdot \sqrt{\frac{N-n}{N-1}}$ *The factor* $\sqrt{\frac{N-n}{N-1}}$ *is known as finite population correction (fpc)*

Concept, Formula and Tricks Marathon QA

Example:

A population comprises the following units: a, b, c, d, e. Draw all possible samples of size three without replacement.

Example:

A population comprises 3 member 1, 5, 3. Draw all possible samples of size two

i. *with replacement*
ii. *without replacement*

Find the sampling distribution of sample mean in both cases.

Formula for number of possible samples	• *If population size is n and sample size is r then number of possible samples in both cases are below* ○ *In case of without replacement:* ${}^{n}C_{r}$ ○ *In case of with replacement:* n^{r}

CA PRANAV POPAT

Sample Survey

Meaning	• Sample Survey is the study of the unknown population on the basis of a proper representative sample drawn from it. How can a part of the universe reveal the characteristics of the unknown universe? The answer to this question lies in the basic principles of sample survey.	
Basic Principles of Sample Survey	Law of Statistical regularity	• if a sample of fairly large size is drawn from the population under discussion at random, then on an average the sample would possess the characteristics of that population.
	Principle of Inertia	• The results derived from a sample, according to the principle of inertia of large numbers, are likely to be more reliable, accurate and precise as the sample size increases, provided other factors are kept constant. • This is a direct consequence of the first principle.
	Principle of Optimization	• The principle of optimization ensures that an optimum level of efficiency at a minimum cost or the maximum efficiency at a given level of cost can be achieved with the selection of an appropriate sampling design.
	Principle of Validity	• The principle of validity states that a sampling design is valid only if it is possible to obtain valid estimates and valid tests about population parameters. • Only a probability sampling ensures this validity.

Concept, Formula and Tricks Marathon QA

Complete Enumeration	*When complete information is collected for all the units belonging to a population, it is defined as complete enumeration or census.*
Preference to Sample Survey over complete enumeration	*In most cases, we prefer sample survey to complete enumeration due to the following factors:*

Speed	*It is quick as only a part of vast population in enumerated*
Cost	*Cost of Collection for each unit is higher in sample survey as trained personnel are employed but in terms of total cost sample survey is less expensive due to low volume*
Reliability	*Date under sample survey is more reliable because of trainer enumerators, better supervision, and application of modern technique*
Accuracy	• *There is no sampling error in complete enumeration but there are chances of non-sampling errors like errors in recording, biases, wrong interpretation etc.* • *In sample survey, sampling error exists but can be reduced to great extent by increasing sample size and following sampling design strictly. Also non sampling errors can be reduced to a desirable degree.*

CA PRANAV POPAT

	Necessity	• When it comes to destructive sampling where the items get exhausted like testing the length of life of electrical bulbs or sampling from a hypothetical population like coin tossing, there is no alternative to sample survey.
	Note: If the population size is not large, there is no need for sampling.	

Errors in Sample Survey

Meaning	• Errors or biases in a survey may be defined as the deviation between the value of population parameter as obtained from a sample and its observed value.

CA PRANAV POPAT

Concept, Formula and Tricks Marathon QA

<table>
<tr>
<td>Sampling Errors</td>
<td>

Since only a part of the population is investigated in a sampling, every sampling design is subjected to this type of errors.
The factors contributing to sampling errors

<table>
<tr><td>Errors arising out due to defective sampling design</td><td>If a non- probabilistic sampling design is followed, the bias or prejudice of the sampler affects the sampling technique</td></tr>
<tr><td>Errors arising out due to substitution</td><td>A very common practice among the enumerators is to replace a sampling unit by a suitable unit in accordance with their convenience when difficulty arises in getting information from the originally selected unit. Since the sampling design is not strictly adhered to, this results in some type of bias.</td></tr>
<tr><td>Errors owing to wrong choice of statistic</td><td>One must be careful in selecting the proper statistic while estimating a population characteristic.</td></tr>
<tr><td>Variability in the population</td><td>Errors may occur due to variability among population units beyond a degree. This could be reduced by following somewhat complicated sampling design like stratified sampling, Multistage sampling etc.</td></tr>
</table>
</td>
</tr>
</table>

CA PRANAV POPAT

Concept, Formula and Tricks Marathon QA

Non-Sampling Errors	• *This type of errors happen both in sampling and complete enumeration.* • *Factors responsible: lapse of memory, preference for certain digits, ignorance, psychological factors like vanity, non- responses on the part of the interviewees wrong measurements of the sampling units, communication gap between the interviewers and the interviewees, incomplete coverage etc.*

Types of Sampling

Probability Sampling	• *In this type of sampling there is always a fixed, pre assigned probability for each member of the population to be a part of the sample taken from that population.* • *Some important probability sampling are Simple random sampling, Stratified sampling, Multi Stage sampling, Multi Phase Sampling, Cluster Sampling and so on.*
Non-Probability Sampling	• *No probability is attached to the member of the population and as such it is based entirely on the judgement of the sampler.* • *Non-probability sampling is also known as Purposive or Judgement Sampling.*
Mixed Sampling	• *Mixed sampling is based partly on some probabilistic law and partly on some pre decided rule. Systematic sampling belongs to this category.*

CA PRANAV POPAT

Commonly used Sampling Process

Simple Random Sampling (SRS)	• *When the units are selected independent of each other in such a way that each unit belonging to the population has an equal chance of being a part of the sample, the sampling is known as Simple random sampling or just random sampling.* • *It may be with or without replacement* • *Simple random sampling is a very simple and effective method of drawing samples provided (i) the population is not very large (ii) the sample size is not very small and (iii) the population under consideration is not heterogeneous* • *Simple random sampling is completely free from Sampler's biases* • *All the tests of significance are based on the concept of simple random sampling.*

PRANAV POPAT

Stratified Sampling	• *If the population is large and heterogeneous, then we consider a somewhat, complicated sampling design known as stratified sampling.* • *It comprises dividing the population into a number of strata or sub-populations in such a way that there should be very little variations among the units comprising a stratum and maximum variation should occur among the different strata.* • *The purpose of stratified sampling are (i) to make representation of all the sub populations (ii) to provide an estimate of parameter not only for all the strata but also and overall estimate (iii) reduction of variability and thereby an increase in precision.* • *We consider "Proportional allocation" or "Bowley's allocation" where the sample sizes for different strata are taken as proportional to the population sizes.*
	• *When the strata-variances differ significantly among themselves, we take recourse to "Neyman's allocation" where sample size vary jointly with population size and population standard deviation.* • *Stratified sampling is not advisable if (i) the population is not large (ii) some prior information is not available and (iii) there is not much heterogeneity among the units of population.*

Concept, Formula and Tricks Marathon QA

Multi Stage Sampling	• *In this type of complicated sampling, the population is supposed to compose of first stage sampling units, each of which in its turn is supposed to compose of second stage sampling units, each of which again in its turn is supposed to compose of third stage sampling units and so on till we reach the ultimate sampling unit.* • *The coverage in case of multistage sampling is quite large. It also saves computational labour and is cost-effective.* • *It adds flexibility into the sampling process which is lacking in other sampling schemes.* • *Compared to stratified sampling, multistage sampling is likely to be less accurate.*

CA PRANAV POPAT

CA PRANAV POPAT

Concept, Formula and Tricks Marathon QA

Systematic Sampling	• *It refers to a sampling scheme where the units constituting the sample are selected at regular interval after selecting the very first unit at random i.e., with equal probability.* • *Systematic sampling is partly probability sampling in the sense that the first unit of the systematic sample is selected probabilistically and partly non- probability sampling in the sense that the remaining units of the sample are selected according to a fixed rule which is non-probabilistic in nature.* • *Systematic sampling is a very convenient method of sampling when a complete and updated sampling frame is available.* • *It is less time consuming, less expensive and simple as compared to the other methods of sampling.* • *Systematic sampling has a severe drawback because if there is an unknown and undetected periodicity in the sampling frame and the sampling interval is a multiple of that period, then we are going to get a most biased sample*

CA PRANAV POPAT

CA FOUNDATION QA

Chp14 Measures of Central Tendency & Dispersion

Concept, Formula and Tricks Marathon



Concepts

Central Tendency - Basics	
Meaning	• *Central Tendency is the **tendency** of a given set of observations to **cluster** around a single **central or middle value.*** • *The **single value** that **represents** the given set of observations is described as a measure of central tendency.*
Different Measures of Central Tendency	• *Arithmetic Mean (AM)* • *Median (Me)* • *Mode (Mo)* • *Geometric Mean (GM)* • *Harmonic Mean (HM)*
Types of Formula based Questions	• *Discrete Observations* • *Simple Frequency Distribution* • *Grouped Frequency Distribution*

POPAT

Concept, Formula and Tricks Marathon QA

Arithmetic Mean

Discrete Observations	$\bar{x} = \frac{x_1 + x_2 + x_3 + \ldots + x_n}{n}$ $\quad$ $\bar{x} = \frac{\sum x}{n}$
Frequency Distribution	$\bar{x} = \frac{\sum fx}{N}$

In case of simple frequency distribution	x = individual values
In case of grouped frequency distribution	x = mid-point of class intervals
N = total number of observations	$N = \sum f$

Assumed Mean / Step-Deviation Method	AM using assumed mean / step deviation method $\bar{x} = A + \frac{\sum fd}{N} \times C$,where $d = \frac{x - A}{C}$, A is assumed mean, C is class length
Property 1	If all the observations are constant, AM is also constant
Property 2	the algebraic sum of deviations of a set of observations from their AM is zero
Property 3	AM is affected both due to change of origin and scale If $y = a + bx$ then $\bar{y} = a + b\bar{x}$
Property 4	Combined AM $\quad \bar{x}_c = \frac{n_1\bar{x}_1 + n_2\bar{x}_2}{n_1 + n_2}$

CA PRANAV POPAT

General Review	• *AM is best measure of central tendency* • *AM is based on all observations* • *AM is affected by sampling fluctuations* • *AM is amenable to mathematical property* • *AM cannot be used in case of open end classification*

CA PRANAV POPAT

Median

Discrete Observations

- If n = odd, then middle term
- If n = even, average of two middle terms

Simple Frequency Distribution

- First make column of less than cumulative frequency
- Apply same formula as discrete

Grouped Frequency Distribution

Median in case of grouped frequency distribution

Step 1	Prepare a less than type cumulative frequency distribution
Step 2	Calculate $\frac{N}{2}$ and check between which class boundaries it falls and call it as Median Class
Step 3	l_1: LCB of Median Class; N_u: Cum Freq. of Median Class; N_l: Cum. Freq. of Pre-Median Class; C: Class length of Median Class
Step 4	Apply Formula $Me = l_1 + \left(\frac{\frac{N}{2} - N_l}{N_u - N_l}\right) \times C$ or $Me = l_1 + \left(\frac{\frac{N}{2} - N_l}{f}\right) \times C$

Property 1	*For a set of observations, the sum of absolute deviations is minimum, when the deviations are taken from the median.* $\sum \lvert x - Me \rvert$ *is minimum*
Property 2	*Median is also affected by both change of origin and scale.*
General Review	• *Median is also called as positional average* • *Median is not based on all observations* • *Median is not affected by sampling fluctuations* • *Median is best measure of central tendency in case of open-end classification*

CA PRANAV POPAT

Partition Values

Meaning

- *These may be defined as **values dividing** a given **set of observations** into number of **equal parts***
- *When we want to divide the given set of observations into two equal parts, we consider median, similarly there are quartiles, deciles, percentiles*

Name of PV	No. of equal parts	No. of PVs	Symbol
Median	2	1	Me
Quartile	4	3	Q_1, Q_2, Q_3
Decile	10	9	$D_1, D_2, \ldots, D_9$
Percentile	100	99	$P_1, P_2, \ldots, P_{99}$

PRANAV POPAT

Formula – Discrete Observations

- *Rank Calculation* $(n+1)p^{th}$ *term*
- *Value of p depends on partition value*

#	Median	Quartile	Decile	Percentile
First	1/2	1/4	1/10	1/100
Second		2/4	2/10	2/100
…				
Last		3/4	9/10	99/100

Quartiles Grouped FD

Quartiles in case of Grouped Frequency Distribution: Steps are like median with few modifications.

1^{st} Quartile	3^{rd} Quartile
Find Q_1 class using $\frac{N}{4}$	Find Q_3 class using $\frac{3N}{4}$
$Q_1 = l_1 + \left(\frac{\frac{N}{4} - N_l}{N_u - N_l}\right) \times C$	$Q_3 = l_1 + \left(\frac{\frac{3N}{4} - N_l}{N_u - N_l}\right) \times C$

Mode

<table>
<tr><td>Meaning</td><td>Mode is the value that occurs the maximum number of times</td></tr>
<tr><td>Special Thing about Mode</td><td>

If two or more observations are having maximum frequency then there are multiple modes [multimodal distribution]
If there are exactly two modes then distribution is called as Bimodal Distribution
If all observations are having same frequency then distribution has no mode
We can say that Mode is not rigidly defined

</td></tr>
<tr><td>Grouped Frequency Distribution</td><td>

Find Modal Class: Class with highest frequency and obtain below values
<table>
<tr><td>f_{-1}</td><td>f_0</td><td>f_1</td></tr>
<tr><td>frequency of pre modal class</td><td>frequency of the modal class</td><td>frequency of the post modal class</td></tr>
</table>

Apply Formula
$$Mo = l_1 + \left(\frac{f_0 - f_{-1}}{2f_0 - f_{-1} - f_1} \right) \times C$$

</td></tr>
</table>

Property 1	*If all the observations are constant, mode is also constant*
Property 2	*Mode is also affected both due to change of origin and scale*
General Review	• *Mode is not based on all observations* • *Mode is not rigidly defined/ uniquely defined* • *Mode is not amenable to Mathematical Property*

Relationship between Mean, Median and Mode

In case of Symmetric Distribution	*Mean = Median = Mode*
In case of Moderately Skewed Distribution (Empirical relationship)	*Mean – Mode = 3 (Mean – Median)* *OR Mode = 3 Median – 2 Mean*

CA PRANAV POPAT

Geometric Mean

Definition	For a given set of n **positive observations**, the geometric mean is defined as the n^{th} root of the product of the observations
Formula – Discrete	$G = \left(x_1 \times x_2 \times \ldots \times x_n\right)^{1/n}$
Formula – Frequency Distribution	$G = \left(x_1^{f_1} \times x_2^{f_2} \times \ldots \times x_n^{f_n}\right)^{1/N}$
Property 1	Logarithm of G for a set of observations is the AM of the logarithm of the observations $\log G = \frac{1}{n}\sum \log x$
Property 2	If all the observations are constant, GM is also constant
Property 3	If $z = xy$, then GM of z = GM of $x \times$ GM of y
Property 4	If $z = x/y$, then GM of $z = \frac{GM \text{ of } x}{GM \text{ of } y}$

POPAT

Harmonic Mean

Definition	For a given set of **non-zero** observations, harmonic mean is defined as the **reciprocal of the AM of the reciprocals of the observation**
Formula – Discrete	$H = \frac{n}{\Sigma(\frac{1}{x})}$
Formula – Frequency Distribution	$H = \frac{N}{\Sigma(\frac{f}{x})}$
Property 1	If all observations are constant HM is also constant
Property 2	$\text{Combined } HM = \frac{n_1 + n_2}{\frac{n_1}{H_1} + \frac{n_2}{H_2}}$

PRANAV POPAT

Use of GM and HM

Both	Both are used for calculating average rates
GM	Appropriate for rates having percentages
HM	Appropriate for rates other than percentages

Relationship between AM, GM, and HM

Relation

Scenario	Relation
When all the observations are same/ identical	$AM = GM = HM$
When observations are distinct	$AM > GM > HM$
In question is silent	$AM \geq GM \geq HM$

Special Relation

If there are only two observations:

$$AM \times HM = (GM)^2$$

Weighted Average

When to use	*If the observations are not of equal importance and we need to treat observations according to their hierarchical importance, then we use Weighted Average*	
Formulas	*Weighted AM*	$\frac{\sum wx}{\sum w}$
	Weighted GM	$\left(x_1^{w_1} \times x_2^{w_2} \times x_3^{w_3} \times \ldots \times x_n^{w_n}\right)^{\frac{1}{\sum w}}$
	Weighted HM	$\frac{\sum w}{\sum\left(\frac{w}{x}\right)}$

CA PRANAV POPAT

Measures of Dispersion

Meaning of Measure of Dispersion	• *Dispersion for a given set of observations may be defined as* • *the **amount of deviation** of the observations,* • *usually, from an **appropriate** measure of **central tendency***	
Types of Measure of Dispersion	*Absolute Measures of Dispersion*	• *These are with units* • *These are not useful for comparison of two variables with different units.* • *Example: Range, Mean Deviation, Standard Deviation, Quartile Deviation*
	Relative Measures of Dispersion	• *These are unit free measures* • *These are useful for comparison of two variables with different units.* • *Example: Coefficient of Range, Coefficient of Mean Deviation, Coefficient of variation, Coefficient of Quartile Deviation*

Range

Discrete – Formula	$L - S$ *where,* *L: Largest Observation, S: Smallest Observation*
Grouped Frequency Distribution – Formula	$L - S$ *where,* *Largest Observation =* ***UCB*** *of last class interval,* *Smallest Observation =* ***LCB*** *of first-class interval*
Coefficient of Range	$\frac{L-S}{L+S} \times 100$
Property 1	• ***Not affected*** *by change of* ***origin*** • *Affected by* ***change of scale (only value)*** • ***No impact of sign*** *of change of scale* • *Note:* ***Measure of Dispersion can never be negative***
General Review	• *Not Based on All Observations* • *Easy to Compute*

Mean Deviation

Meaning	• Mean deviation is defined as the • **arithmetic mean** of the • **absolute deviations** of the observations • from an **appropriate measure** of central tendency
Formula – Discrete	$MD_A = \frac{1}{n}\Sigma \lvert x - A \rvert$ where, A = Appropriate Central Tendency Measure
Formula – Frequency Distribution	$MD_A = \frac{1}{N}\Sigma f \lvert x - A \rvert$
Coefficient of Mean Deviation	Coefficient of Mean Deviation: $\frac{\text{Mean Deviation about } A}{A} \times 100$
Property 1	Mean Deviation takes its **minimum value** when deviations are taken from **Median**
Property 2	Change of Origin – **No Affect**, Change of Scale – **Affect of value not sign**

General Review	• *Based on* ***all observations*** • *Improvement over Range* • ***Difficult to compute*** • ***Not amenable to Mathematical Property*** *because of usage of* ***Modulus***

CA PRANAV POPAT

Concept, Formula and Tricks Marathon QA

Standard Deviation

Meaning	• *Improvement over Mean Deviation* • *It is defined as the **root mean square deviation** when the deviations are taken from the AM of the observations*
Formula – Discrete	$\sigma_x = SD_x = \sqrt{\frac{\sum(x-\bar{x})^2}{n}}$ $\sigma_x = SD_x = \sqrt{\frac{\sum x^2}{n} - (\bar{x})^2}$
Formula – Frequency Distribution	$\sigma_x = SD_x = \sqrt{\frac{\sum f(x-\bar{x})^2}{N}}$ $\sigma_x = SD_x = \sqrt{\frac{\sum fx^2}{N} - (\bar{x})^2}$
Coefficient of Variation	$\frac{SD_x}{\bar{x}} \times 100$

CA PRANAV POPAT

SD for any two numbers	$SD = \frac{Range}{2}$
SD for first n natural numbers	$s = \sqrt{\frac{n^2 - 1}{12}}$
Property 1	*If all the observations are constant, SD is* ***ZERO***
Property 2	*No effect of change of origin but affected by change of scale in the magnitude (ignore sign)*
Property 3	$SD_c = \sqrt{\frac{n_1 s_1^{\,2} + n_2 s_2^{\,2} + n_1 d_1^{\,2} + n_2 d_2^{\,2}}{n_1 + n_2}}$ $d_1 = \bar{x}_c - \bar{x}_1, d_2 = \bar{x}_c - \bar{x}_2$

CA PRANAV POPAT

Quartile Deviation

Formula	$QD_x = \frac{Q_3 - Q_1}{2}$
Calculation	Quartiles are calculated same as we studied in Central Tendency
Coefficient of Quartile Deviation	$\frac{Q_3 - Q_1}{Q_3 + Q_1} \times 100$
General Review	• It is the **best measure** of dispersion for **open-end** classification • It is also **less affected** due to sampling fluctuations • Like other measures of Dispersion, **QD** is also not affected by change of origin but affected by scale ignoring sign
Relationship between SD, MD and QD	$4SD = 5MD = 6QD$ or $SD : MD : QD = 15 : 12 : 10$

POPAT

CA FOUNDATION QA

Chp15 Probability

Concept, Formula and Tricks Marathon



Concepts

Probability - Basics

History	• *first application of probability was made by a group of mathematicians in Europe* • *about three hundred years back* • *to enhance their chances of winning* • *in different games of* ***gambling***
Types	• ***Subjective Probability*** – *personal judgement* – *experience* – *influenced by the* ***personal belief, attitude, and bias*** – *of the person applying it* • ***Objective Probability*** – *Based on* ***Rules and Maths*** – *This is what we are going to discuss*
Random Experiment	***Experiment*** — *A performance that produces certain results* ***Random Experiment*** — *An experiment is defined to be random if the results of the experiment depend* ***on chance*** *only.* ***Examples*** — *Tossing a coin, throwing a dice, drawing cards from a pack*

Events	*The results or outcomes of a* ***random experiment*** *are known as* ***Events***

Event at its simplest form	*Simple Event or Elementary Event*
Event that can be sub-divided into further events	*Composite Event or Compound Event*

Classical Probability

General	• *Also called Prior Definition of Probability, this formula is Event (Result) Based.* • *It is given by Bernoulli and Laplace.*
Formula of Probability	$P(A) = \frac{\text{no. of favorable outcomes}}{\text{total number of outcomes}}$
Formula to obtain total outcomes	*Total number of possible outcomes:* p^q *where,* p *= no. of events in one trial,* q *= no. of trials*
Conditions	• *Value of Probability:* $0 \le P(A) \le 1$ • *If* $P(A) = 1$, *A is a sure event* • *If* $P(A) = 0$, *A is an impossible event*

<table>
<tr><td>Complimentary Probability</td><td>

- *Probability of* ***non-occurrence*** *of an event A is denoted by* $P(A')$ ***or*** $P(\bar{A})$ *is called as complimentary event of A.* $P(\bar{A}) = 1 - P(A)$

</td></tr>
<tr><td>Odds in Favour of an Event A</td><td>$\dfrac{\text{number of favorable events to } A}{\text{number of unfavorable events to } A}$</td></tr>
<tr><td>Odds against an Event A</td><td>$\dfrac{\text{number of unfavorable events to } A}{\text{number of favorable events to } A}$</td></tr>
<tr><td>Playing Cards</td><td>

Suites	*Spades*	*Hearts*	*Diamond*	*Clubs*
Symbol	♠	♡	♢	♣
Count	13	13	13	13

Ranks	*Ace*	*King*	*Queen*	*Jack*	*Numbers*
Symbol	*A*	*K*	*Q*	*J*	*10, 9, 8, 7, 6, 5, 4, 3, 2*
Count	4	4	4	4	4

</td></tr>
<tr><td>Limitations</td><td>

- *Applicable only when events are finite and are equally likely*
- *Limited application of this definition like in tossing coin, throwing dice, cards etc.*

</td></tr>
<tr><td>Usage of Combinations in Probability</td><td>

If only one object is to be selected	*No need to use combination*
If more than one object is to be selected	*Use combinations to calculate both favorable outcome and total outcome*

</td></tr>
</table>

Set Based Probability

Terms

Term in Probability	Term in Sets	Explanation
Sample Space	Universal Set	Set all the possible events for a random experiment
Sample Points	Elements	Each event of a Random experiment is termed as Sample Point
Event Set A	Ordinary Set A which is a subset of Universal Set	Event Set A which is under consideration for probability calculations is defined as a non-empty subset of Set S (Sample Space) containing all favorable sample points for A

Formula of Probability Event A

$$P(A) = \frac{\text{number of sample points is Set } A}{\text{number of sample points is Sample Space } S} = \frac{n(A)}{n(S)}$$

Why Sets?

- To use concept of **union, intersection** in probability
- To use Sets formula for **complicated Probability Questions**

2-Sets Formula	$n(A \cup B) = n(A) + n(B) - n(A \cap B)$
3-Sets Formula	$n(A \cup B \cup C) = n(A) + n(B) + n(C) - n(A \cap B) - n(B \cap C) - n(C \cap A) + n(A \cap B \cap C)$
Types of Sets of Events	*Mutually Exclusive or Incompatible Events*: *Events are exclusive if they **cannot occur simultaneously** implies* $P(A \cap B) = 0$
	Exhaustive Events: *Events are exhaustive if **one of them must necessarily occur** implies* $P(A \cup B) = 1$
	Equally likely or Equi-Probable Events: *Events are equally likely if all of them **have same probability** (also called mutually symmetric events) implies* $P(A) = P(B)$

CA PRANAV POPAT

Addition Theorem of Probability

Theorem	Formula	Applicable for
1	$P(A \cup B) = P(A) + P(B)$	For two mutually exclusive events
2	$P(A_1 \cup A_2 \cup A_3 \cup \ldots) = P(A_1) + P(A_2) + P(A_3) + \ldots$	For any number of mutually exclusive events
3	$P(A \cup B) = P(A) + P(B) - P(A \cap B)$	For any two events
4	$P(A \cup B \cup C) = P(A) + P(B) + P(C) - P(A \cap B) - P(B \cap C) - P(C \cap A) + P(A \cap B \cap C)$	For any three events

Expected Frequency

Expected Frequency of Event $A = P(A) \times N$ (total number of outcomes)

Probability of A – B or B – A

Probability of **Only A**	$P(A - B)$	$P(A \cap B')$	$P(A) - P(A \cap B)$
Probability of **Only B**	$P(B - A)$	$P(A' \cap B)$	$P(B) - P(A \cap B)$

Conditional Probability

Dependent Events	If occurrence of one event is **influenced** by occurrence of another event, then two events are dependent. Conditional Probability is applicable only for dependent events
Independent Events	Two events are said to be independent if occurrence of one event do not influence the occurrence of other. Conditional Probability is **not applicable**
Formula for Conditional Probability	Conditional Probability event B given that event A has already been occurred: $P(B/A)=\frac{P(B\cap A)}{P(A)}$ where $P(A)\neq 0$
	Conditional Probability event B given that event A has already been occurred: $P(A/B)=\frac{P(A\cap B)}{P(B)}$ where $P(B)\neq 0$

CA PRANAV POPAT

Compound Theorem for Dependent Events	$P(A \cap B) = P(A/B) \times P(B)$ $P(A \cap B) = P(B/A) \times P(A)$ *Joint Probability = Unconditional Probability of one event × Conditional Probability of another event*
Compound Theorem for Independent Events	*In case of independent events, Conditional and Unconditional are same* $P(A \cap B) = P(A) \times P(B)$ *If A and B are independent events, then their complementary events are also independent and same theorem can be applied* $P(A' \cap B) = P(A') \times P(B)$ $P(A \cap B') = P(A) \times P(B')$ $P(A' \cap B') = P(A') \times P(B')$
De-Morgan's Law Application	$P(A \cup B \cup C) = 1 - P(A \cup B \cup C)'$ $P(A \cup B \cup C) = 1 - P(A' \cap B' \cap C')$ $P(A \cup B \cup C) = 1 - P(A') \times P(B') \times P(C')$

Random Variable | Probability Distribution

Random Variable	*It is a variable defined on Sample Space of a random experiment that can take any value (Real Number)*
Probability Distribution	*It is defined as the statement/ table that shows* • *no. of different value taken by Random Variable and* • *their corresponding probabilities* *Sum of all probabilities of distribution will always be equal to 1*
Expected Value	*It is also mean of probability distribution.* $\mu = E(X) = \sum PX$
Variance	$\sigma^2 = E(X^2) - \left[E(X)\right]^2$ $\sigma^2 = \sum PX^2 - \left[\sum PX\right]^2$

PRANAV POPAT

Statistical Definition of Probability

Concept	• *It is based on the concept of* ***Relative Frequency*** • *Concept:* – *Relative Frequency on infinite no. of trials is equal to probability of an event* – $Relative\ frequency = \frac{class\ frequency}{total\ frequency}$
Formula	$P(A) = \lim_{n \to \infty} \frac{f_A}{n}$

CA PRANAV POPAT

CA FOUNDATION QA

Chp16 Theoretical Distribution

Concept, Formula and Tricks Marathon



Concepts

Modern Probability

<table>
<tr><td>Approaches of Probability</td><td>

- *Classical Probability*
- *Set Based Probability*
- *Modern Probability (Axiomatic)*

</td></tr>
<tr><td>Modern Probability</td><td>

- *Unlike other, here probability is not calculated by a simple ratio*
- *Probability is calculated by using a function defined on random variable*

</td></tr>
<tr><td>Probability Function</td><td>

Probability = $f(X)$ where X is a random variable

Types of Probability Function

<table>
<tr><th>Discrete</th><th>Continuous</th></tr>
<tr><td>Probability Mass Function</td><td>Probability Density Function</td></tr>
</table>

</td></tr>
<tr><td>Parameter</td><td>characteristics that are used to define a given population</td></tr>
</table>

PRANAV POPAT

Binomial Distribution

When to use	• When random variable is discrete • Probability is not very low (like 1/2, 2/3, 1/4 around)
Binomial Variate	$X \sim B(n,p)$ Number of Parameters: 2 Binomial Distribution: **bi-parametric discrete** probability distribution

Probability Mass Function	$f(x) = P(X = x) = {}^{n}C_{x} p^{x} q^{n-x}$ X can take whole number value from **0 to n**
Mean	$\mu = np$
Variance	$\sigma^2 = npq$ • Mean is always more than variance • Maximum Value of Variance $= \frac{n}{4}$

<table>
<tr><td>Mode</td><td>
<table>
<tr><td colspan="2">$Find\ (n+1)p$</td></tr>
<tr><td>If result is integer, then there are two modes.
$\mu_0 = (n+1)p$
$\mu_0 = (n+1)p-1$</td><td>If result is non-integer, then there is only one mode

Largest integer contained in $(n+1)p$</td></tr>
</table>
</td></tr>
<tr><td>Additive Property</td><td>$If\ X \sim B(n_1,p)\ and\ Y \sim B(n_2,p)\ then\ X+Y \sim B(n_1+n_2,p)$</td></tr>
</table>

CA PRANAV POPAT

Poisson Distribution

When to use	• When random variable is discrete • Probability is very low (like 1%, 2% around)
Poisson Parameter	• Poisson distribution has only one parameter • Denoted by m • It can either be given in question or can be calculated using: $m = np$
Poisson Variate	$X \sim P(m)$
Probability Mass Function	$f(x) = P(X = x) = \frac{e^{-m}.m^{x}}{x!}$ X can take whole number value from **0 to n**
Mean	$\mu = m$
Variance	$\sigma^2 = m$

PRANAV POPAT

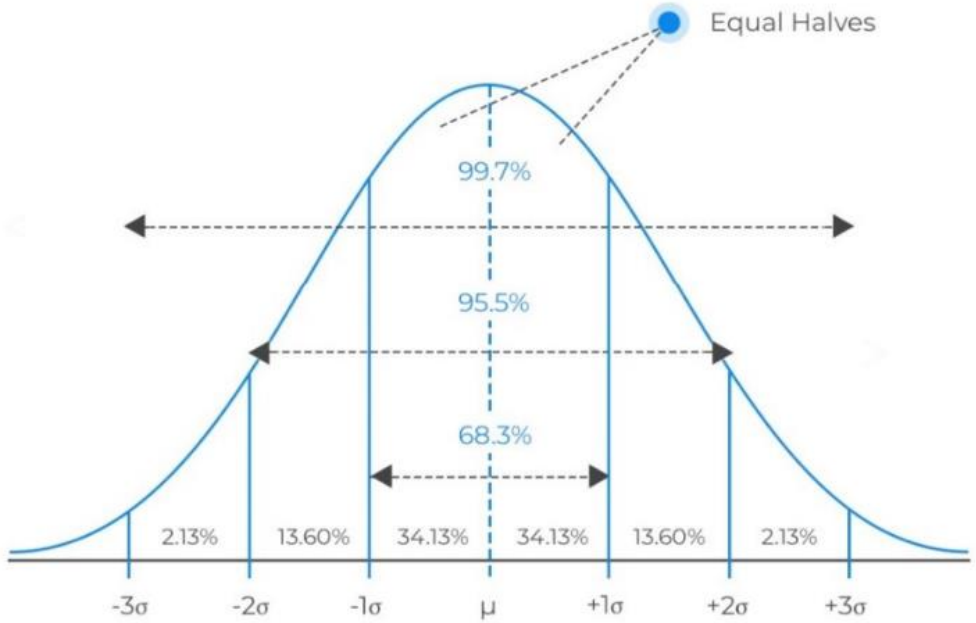
<table>
<tr><td>Mode</td><td>
<table>
<tr><td colspan="2">Find m</td></tr>
<tr><td>If result is integer, then there are two modes. $\mu_0 = m$ $\mu_0 = m-1$</td><td>If result is non-integer, then there is only one mode Largest integer contained in m</td></tr>
</table>
</td></tr>
<tr><td>Additive Property</td><td>If $X \sim P(m_1)$ and $Y \sim P(m_2)$ then $X+Y \sim P(m_1+m_2)$</td></tr>
</table>

CA PRANAV POPAT

Normal Distribution

When to use	• When random variable is continuous • It is bi-parametric continuous probability distribution
Normal Variable	$X \sim N(\mu, \sigma^2)$
Special Point about Mean	Mean = Median = Mode = μ (symmetric distribution)
Variance	Here σ^2 is a parameter and will be given in question
Mean Deviation	0.8σ
Quartiles	$Q_1 = \mu - 0.675\sigma$ $Q_3 = \mu + 0.675\sigma$
Quartile Deviation	$QD = 0.675\sigma$
Ratio between QD:MD:SD	10:12:15
Probability Density Function	$\frac{1}{\sigma\sqrt{2\pi}} e^{-(\frac{x-\mu}{\sigma})^2 \times \frac{1}{2}}$ This is used only for the purpose of obtaining mean and SD comparing the function given in question with this function.

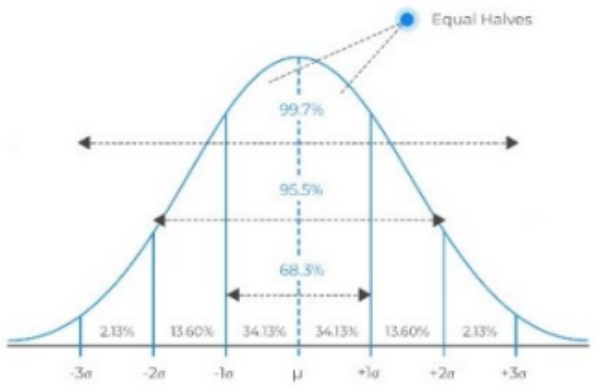
Normal Curve	*Curve formed by normal distribution. It is also called as Probability Curve*
Shape of Normal Curve	*Bell Shaped*
Points of Inflexion	*Points of normal curve at which curve changes its curvature i.e. convex to concave and vice-versa. Formula is* $\mu-\sigma$ *and* $\mu+\sigma$
Area under Normal Curve	*The total area of the normal curve or for that any probability curve is taken to be unity i.e. one.*
Tails of Normal Curve	*the two tails of the normal curve extend indefinitely on both sides of the curve and both the left and right tails never touch the horizontal axis.*
Symmetric Distribution	*The line drawn through* $x=\mu$ *has divided the normal curve into two parts which are equal in all respect.*



Popular Probability Intervals in Normal Curve

From	To	Area/ Probability
μ	$\mu+\sigma$	34.135%
$\mu+\sigma$	$\mu+2\sigma$	13.59%
$\mu+2\sigma$	$\mu+3\sigma$	2.14%
$\mu+3\sigma$	∞	0.135%

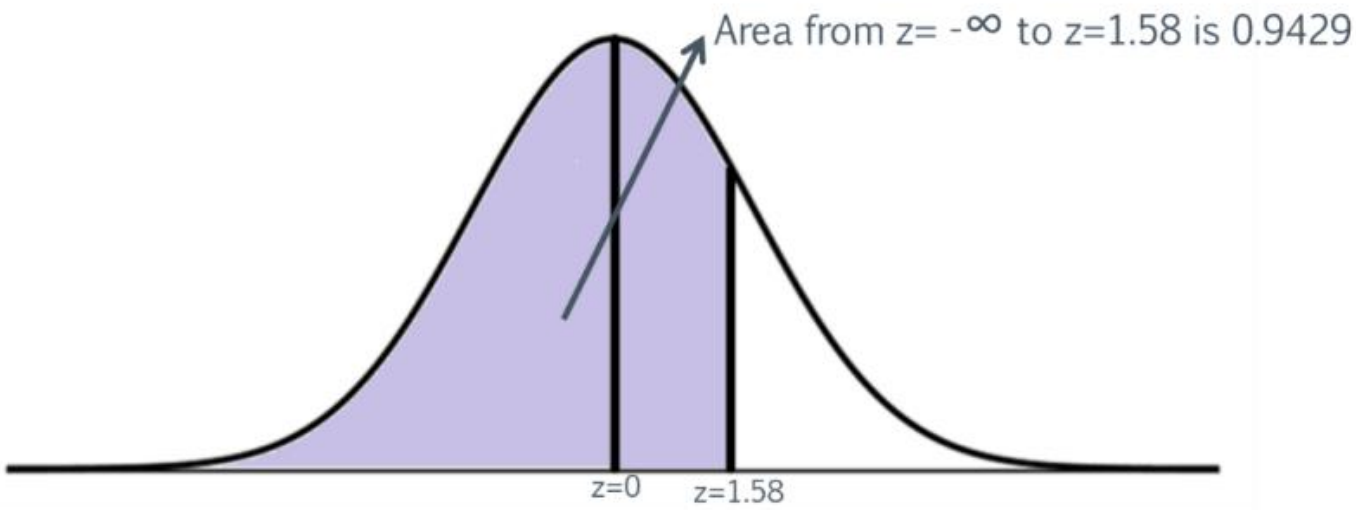
From	To	Area/ Probability
$\mu-\sigma$	$\mu+\sigma$	68.27%
$\mu-2\sigma$	$\mu+2\sigma$	95.45%
$\mu-3\sigma$	$\mu+3\sigma$	99.73%



Standard Normal Distribution

When to use	• When probability calculation is outside the popular intervals • Normal Distribution is converted to Standard Normal Distribution so that Z Table can be used for probability calculation
Conditions	Normal Distribution where $\mu=0$ and $\sigma=1$
Convert X to Z	$Z=\frac{X-\mu}{\sigma}$

μ	$Mean = Median = Mode = 0$
Variance	$\sigma^2 = 1$ *and* $\sigma = 1$
Points of Inflexion	*-1 and +1*
Mean Deviation	*0.8*
Quartile Deviation	*0.675*
Use of Z Table	*This table gives us the probability of values from z = 0 to any value of z. (Area from center)*
Cumulative Distribution Function	$\varphi(k) = P(X \leq k)$ $\varphi(k)$ *gives the area from* $-\infty$ *to point k in a standard distribution*



Z Table

Example:

To find the area under the curve between the mean and a point 2.24 standard deviations to the right of the mean, look up the value opposite 2.2 and under 0.04 in the table; 0.4875 of the area under the curve lies between the mean and a z Value of 2.24

x	0.00	0.01	0.02	0.03	0.04	0.05	0.06	0.07	0.08	0.09
0.0	0.0000	0.0040	0.0080	0.0120	0.0160	0.0199	0.0239	0.0279	0.0319	0.0359
0.1	0.0398	0.0438	0.0478	0.0517	0.0557	0.0596	0.0636	0.0675	0.0714	0.0753
0.2	0.0793	0.0832	0.0871	0.0910	0.0948	0.0987	0.1026	0.1064	0.1103	0.1141
0.3	0.1179	0.1217	0.1255	0.1293	0.1331	0.1368	0.1406	0.1443	0.1480	0.1517
0.4	0.1554	0.1591	0.1628	0.1664	0.1700	0.1736	0.1772	0.1808	0.1844	0.1879
0.5	0.1915	0.1950	0.1985	0.2019	0.2054	0.2088	0.2123	0.2157	0.2190	0.2224
0.6	0.2257	0.2291	0.2324	0.2357	0.2389	0.2422	0.2454	0.2486	0.2517	0.2549
0.7	0.2580	0.2611	0.2642	0.2673	0.2704	0.2734	0.2764	0.2794	0.2823	0.2852
0.8	0.2881	0.2910	0.2939	0.2967	0.2995	0.3023	0.3051	0.3078	0.3106	0.3133
0.9	0.3159	0.3186	0.3212	0.3238	0.3264	0.3289	0.3315	0.3340	0.3365	0.3389
1.0	0.3413	0.3438	0.3461	0.3485	0.3508	0.3531	0.3554	0.3577	0.3599	0.3621
1.1	0.3643	0.3665	0.3686	0.3708	0.3729	0.3749	0.3770	0.3790	0.3810	0.3830
1.2	0.3849	0.3869	0.3888	0.3907	0.3925	0.3944	0.3962	0.3980	0.3997	0.4015
1.3	0.4032	0.4049	0.4066	0.4082	0.4099	0.4115	0.4131	0.4147	0.4162	0.4177
1.4	0.4192	0.4207	0.4222	0.4236	0.4251	0.4265	0.4279	0.4292	0.4306	0.4319
1.5	0.4332	0.4345	0.4357	0.4370	0.4382	0.4394	0.4406	0.4418	0.4429	0.4441
1.6	0.4452	0.4463	0.4474	0.4484	0.4495	0.4505	0.4515	0.4525	0.4535	0.4545
1.7	0.4554	0.4564	0.4573	0.4582	0.4591	0.4599	0.4608	0.4616	0.4625	0.4633
1.8	0.4641	0.4649	0.4656	0.4664	0.4671	0.4678	0.4686	0.4693	0.4699	0.4706
1.9	0.4713	0.4719	0.4726	0.4732	0.4738	0.4744	0.4750	0.4756	0.4761	0.4767
2.0	0.4772	0.4778	0.4783	0.4788	0.4793	0.4798	0.4803	0.4808	0.4812	0.4817
2.1	0.4821	0.4826	0.4830	0.4834	0.4838	0.4842	0.4846	0.4850	0.4854	0.4857
2.2	0.4861	0.4864	0.4868	0.4871	0.4875	0.4878	0.4881	0.4884	0.4887	0.4890
2.3	0.4893	0.4896	0.4898	0.4901	0.4904	0.4906	0.4909	0.4911	0.4913	0.4916
2.4	0.4918	0.4920	0.4922	0.4925	0.4927	0.4929	0.4931	0.4932	0.4934	0.4936
2.5	0.4938	0.4940	0.4941	0.4943	0.4945	0.4946	0.4948	0.4949	0.4951	0.4952
2.6	0.4953	0.4955	04956	0.4957	0.4959	0.4960	0.4961	0.4962	0.4963	0.4964
2.7	0.4965	0.4966	0.4967	0.4968	0.4969	0.4970	0.4971	0.4972	0.4973	0.4974
2.8	0.4974	0.4975	0.4976	0.4977	0.4977	0.4978	0.4979	0.4979	0.4980	0.4981
2.9	0.4981	0.4982	0.4982	0.4983	0.4984	0.4984	0.4985	0.4985	0.4986	0.4986
3.0	0.4987	0.4987	0.4987	0.4988	0.4988	0.4989	0.4989	0.4989	0.4990	0.4990

Concept, Formula and Tricks Marathon QA

CA PRANAV POPAT

CA FOUNDATION QA

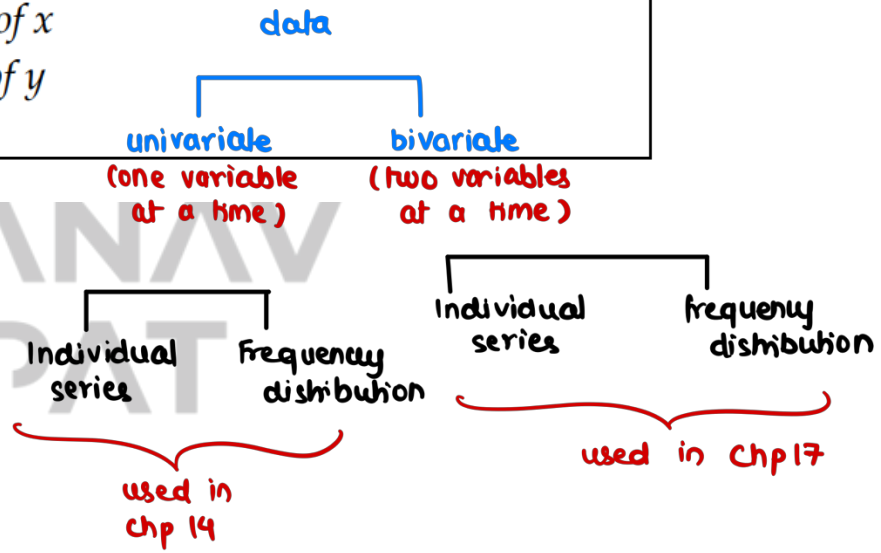
Chp17 Correlation Regression

Concept, Formula and Tricks Marathon



Concepts

Bivariate Data	
Definition	• *When data are collected on two variables **simultaneously**, they are known as **bivariate data*** • *and the corresponding frequency distribution, derived from it, is known as **Bivariate Frequency Distribution***
Marginal Distribution	• *It is the frequency distribution of **one variable** (x or y) across the other variable's **full range of values*** • ***Number of Marginal Distribution = 2***
Conditional Distribution	• *It is the frequency distribution of **one variable** (x or y) across a **particular sub-population** of the other variable.* • *No. of Conditional Distributions = m + n* *m = no. of class interval of x* *n = no. of class interval of y* • *Number of Cells = m × n*



PRANAV POPAT

Example:

Prepare a Bivariate Frequency table for the following data relating to the marks in Statistics (x) and Mathematics (y):

(15, 13),	(1, 3),	(2, 6),	(8, 3),	(15, 10),	(3, 9),	(13, 19),
(10, 11),	(6, 4),	(18, 14),	(10, 19),	(12, 8),	(11, 14),	(13, 16),
(17, 15),	(18, 18),	(11, 7),	(10, 14),	(14, 16),	(16, 15),	(7, 11),
(5, 1),	(11, 15),	(9, 4),	(10, 15),	(13, 12)	(14, 17),	(10, 11),
(6, 9),	(13, 17),	(16, 15),	(6, 4),	(4, 8),	(8, 11),	(9, 12),
(14, 11),	(16, 15),	(9, 10),	(4, 6),	(5, 7),	(3, 11),	(4, 16),
(5, 8),	(6, 9),	(7, 12),	(15, 6),	(18, 11),	(18, 19),	(17, 16)
(10, 14)						

Take mutually exclusive classification for both the variables, the first class interval being 0-4 for both.

Concept, Formula and Tricks Marathon QA

		MARKS IN MATHS					
Y / X		0-4	4-8	8-12	12-16	16-20	Total
MARKS IN STATS	0–4	I (1)	I (1)	II (2)			4
	4–8	I (1)	IIII (4)	~~IIII~~ (5)	I (1)	I (1)	12
	8–12	I (1)	II (2)	IIII (4)	~~IIII~~ I (6)	I (1)	14
	12–16		I (1)	III (3)	II (2)	~~IIII~~ (5)	11
	16–20			I (1)	~~IIII~~ (5)	III (3)	9
	Total	3	8	15	14	10	50

Marginal Distribution of Marks in Statistics

Marks	No. of Students
0-4	4
4-8	12
8-12	14
12-16	11
16-20	9
Total	50

Conditional Distribution of Marks in Statistics for Students having Mathematics Marks between 8 to 12

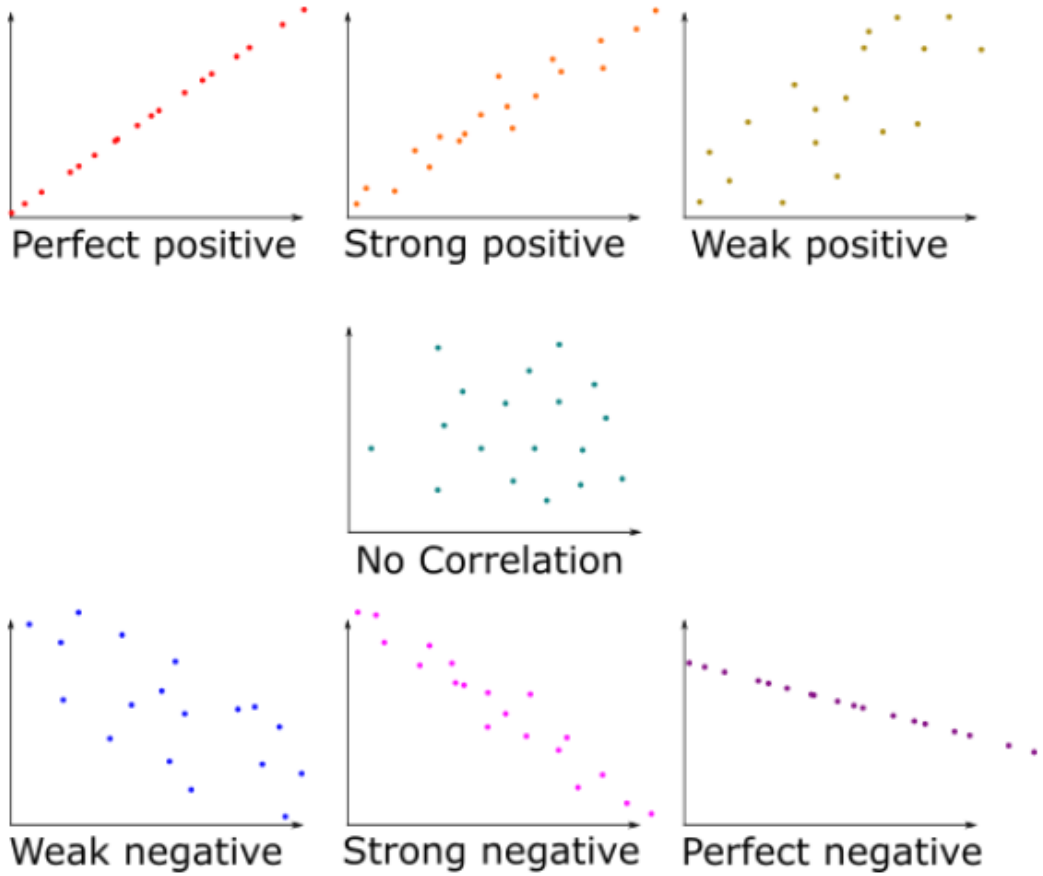
Marks	No. of Students
0-4	2
4-8	5
8-12	4
12-16	3
16-20	1
Total	15

CA PRANAV POPAT

Correlation

Meaning	• *In bivariate data, if **change in one** variable causes **change in another** variable either **directly or inversely**, then the two variables are known to be associated or correlated.* • ***Positive Correlation:** If two variables move in the **same direction*** • ***Negative Correlation:** if the two variables move in the **opposite directions***
Measurement of Correlation	• ***What to Measure*** – *Nature* – *Relative Strength (High, Low)* – *Value* • ***Methods*** – *Scatter diagram* – *Karl Pearson's Product moment correlation coefficient* – *Spearman's rank correlation co-efficient* – *Co-efficient of concurrent deviations*

Scatter Diagram

Theory about Scatter Diagram	• *It helps us to find* ***Nature*** *and* ***Relative Strength*** *of Correlation* • *It is useful for* ***Non-Linear*** *Correlation also* • *It* ***cannot*** *be used to determine* ***value*** • *Diagrams are* ***time taking***
Scatter Diagram and their Interpretations	

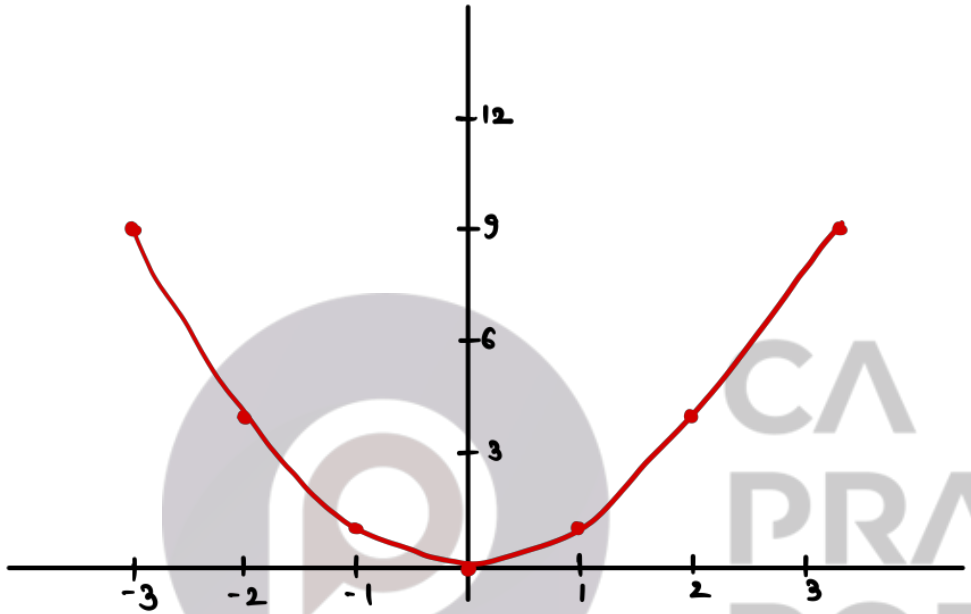
Example. If $y = x^2$ and value of x are given below as

x	-3	-2	-1	0	1	2	3

Draw a scatter diagram between x and y.

Sol

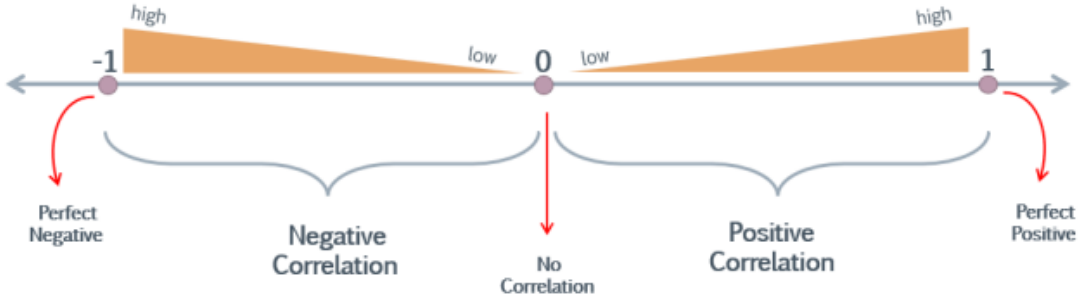
x	-3	-2	-1	0	1	2	3
y	9	4	1	0	1	4	9



In this example, x and y are not in linear correlation. They are correlated but type of correlation is non linear.

Concept, Formula and Tricks Marathon QA
CA PRANAV POPAT
CA PRANAV POPAT

Karl Pearson's Correlation Coefficient

<table>
<tr><td>Correlation Coefficient</td><td>
› Correlation Coefficient:
– It is a unit less value
– It is used to understand nature as well as accurate strength
– It is denoted by r_{xy}


</td></tr>
<tr><td>Formula</td><td>$$r_{xy} = \frac{Cov(x,y)}{(\sigma_x \times \sigma_y)}$$</td></tr>
<tr><td>Formula of Covariance</td><td>$$Cov(x,y) = \frac{\Sigma(x-\bar{x})(y-\bar{y})}{n} \text{ or } \frac{\Sigma xy}{n} - \bar{x}.\bar{y}$$</td></tr>
<tr><td>Property 1</td><td>The Coefficient of Correlation is a unit-free measure</td></tr>
<tr><td>Property 2</td><td>Value lies from -1 to +1</td></tr>
<tr><td>Property 3</td><td>
<table>
<tr><td>Change of Origin</td><td>No impact</td></tr>
<tr><td>Change of Scale</td><td>No impact of value, but if change of scale of both variables are of different sign then sign of r will also change</td></tr>
</table>
</td></tr>
</table>

Example Find correlation coefficient as per product moment with given data (8,3), (12,7) (6,2) (15,4)

x	8	12	6	15
y	3	7	2	4

Sol

$$r_{xy} = \frac{cov(x,y)}{SD_x \times SD_y}$$

$$cov(x,y) = \frac{\Sigma xy}{n} - \bar{x}\cdot\bar{y}$$

$\Sigma xy = 8\times3\ M+\ 12\times7\ M+\ 6\times2\ M+\ 15\times4\ M+\ MRC$

$= \frac{180}{4} - \frac{41}{4}\times\frac{16}{4}$

$= 180 \div 4\ M+\ 41 \div 4 \times 16 \div 4\ M-\ MRC$

$= 4$

$$SD_x = \sqrt{\frac{\Sigma x^2}{n} - \bar{x}^2}$$

$\Sigma x^2 = 8X=M+\ 12X=M+\ 6X=M+\ 15X=M+\ MRC$

$= 469$

$$= \sqrt{\frac{469}{4} - \left(\frac{41}{4}\right)^2}$$

$469 \div 4\ M+\ 41 \div 4 X = M-\ MRC\ \sqrt{}$

$= 3.49106$

$$SD_y = \sqrt{\frac{\Sigma y^2}{n} - \bar{y}^2}$$

$\Sigma y^2 = 3X=M+\ 2X=M+\ 7X=M+\ 4X=M+\ MRC$

$= 78$

$$= \sqrt{\frac{78}{4} - \left(\frac{16}{4}\right)^2} =$$

$78 \div 4\ M+\ 16 \div 4 X = M-\ MRC\ \sqrt{}$

$= 1.8708$

$$r_{xy} = \frac{4}{3.49106 \times 1.8708} = 0.61245$$

$4 \div 3.49106 \div 1.8708 =$

Concept, Formula and Tricks Marathon QA
CA PRANAV POPAT
CA PRANAV POPAT

If two variables are related by a linear equation then correlation between them will always be perfect.

if $y = a + bx$

then $r = +1$ if $b > 0$

$r = -1$ if $b < 0$

b is called as slope of equation.

" product of SD cannot be less than covariance "

Reason: Bcoz if product of SD is smaller than covariance then r_{xy} will exceed 1 which is not allowed.

CA PRANAV POPAT

Spearman's Rank Correlation Coefficient

Usage	• *find the level of* ***agreement (or disagreement)*** *between two judges so far as assessing a* ***qualitative characteristic (attribute)*** *is concerned* • *Use in case of ranks*
Formula (Regular)	$r_R = 1 - \frac{6\Sigma d^2}{n(n^2-1)}$ *d = difference in ranks*
Formula (In case of Tie)	*Spearman's Rank Correlation Coefficient (in case of tied values)*
	$r_R = 1 - \frac{6\left(\Sigma d^2 + A\right)}{n(n^2-1)}$ *here A is adjustment value* $A = \frac{\Sigma\left(t^3 - t\right)}{12}$ *where t = tie length (calculate t value for each of the ties)*

PRANAV POPAT

Concept, Formula and Tricks Marathon QA

Example:

Compute the coefficient of rank correlation between sales and advertisement expressed in thousands of rupees from the following data:

Sales	*90*	*85*	*68*	*75*	*82*	*80*	*95*	*70*
Advertisement	*7*	*6*	*2*	*3*	*4*	*5*	*8*	*1*

When data is not given in ranking we first need to convert it into ranking data based on given value

R_x	2	3	8	6	4	5	1	7
R_y	2	3	7	6	5	4	1	8

$\Sigma d^2 = 4$

$r_R = 1 - \frac{6 \times 4}{8(8^2-1)} = 1 - \frac{24}{504} = 0.9524$

CA PRANAV POPAT

Example:

Compute the coefficient of rank correlation between Eco. marks and stats. Marks as given below:

Eco Marks	*80*	*56*	*50*	*48*	*50*	*62*	*60*
Stats Marks	*90*	*75*	*75*	*65*	*65*	*50*	*65*

Eco Marks (x)	*80*	*56*	*50*	*48*	*50*	*62*	*60*
Stats Marks (y)	*90*	*75*	*75*	*65*	*65*	*50*	*65*
R_x	1	4	5 (5.5*)	7	6 (5.5*)	2	3 (* $\frac{5+6}{2}$)
R_y	1	2 (2.5#)	3 (2.5#)	4 (5@)	5 (5@)	7	6 (5@)

$\frac{2+3}{2}$

@ $\frac{4+5+6}{3} = 5$

When there is a tie among values we take average ranks

CA PRANAV POPAT

Concept, Formula and Tricks Marathon QA

R_x	1	4	5.5	7	5.5	2	3
R_y	1	2.5	2.5	5	5	7	5

$$\Sigma d^2 = 44.5$$

In case of tie, rank correlation formula is modified

$$r_R = 1 - \frac{6\ (\Sigma d^2 + A)}{n\ (n^2 - 1)}$$

where, A = adjustment value of Ties

Tie description	Tie length (t)	$\frac{t^3 - t}{12}$
x (50, 50)	2	$\frac{2^3 - 2}{12} = 0.5$
y (75, 75)	2	$\frac{2^3 - 2}{12} = 0.5$
y (65, 65, 65)	3	$\frac{3^3 - 3}{12} = 2.0$
		$A = \Sigma\left(\frac{t^3 - t}{12}\right) = 3.0$

$$r_R = 1 - \frac{6(\Sigma d^2 + A)}{n(n^2-1)}$$

$$= 1 - \frac{6\ (44.5 + 3)}{7\ (7^2 - 1)}$$

$$= 0.15178$$

CA PRANAV POPAT

Coefficient of Concurrent Deviations

Usage	*A very **quick, simple** and **casual** method of finding correlation when we are not serious about the magnitude of the two variables*
Formula	$r_c = \pm\sqrt{\pm\left(\frac{2c-m}{m}\right)}$ *where c is number of concurrent deviations (same direction)* *m is number of pairs compared (equals to n-1)*

CA PRANAV POPAT

Concept, Formula and Tricks Marathon QA

Example: Find the coefficient of concurrent deviations from the following data.

Year	1990	1991	1992	1993	1994	1995	1996	1997
Price	25	28	30	23	35	38	39	42
Demand	35	34	35	30	29	28	26	23
d_x	NA	+	+	−	+	+	+	+
d_y	NA	−	+	−	−	−	−	−

concurrent

no. of total pairs = 8 (n = 8)

no. of pairs compared ⇒ m = n − 1 = 8 − 1 = 7

no. of concurrent pairs ⇒ c = 2

Formula of correlation coefficient as per concurrent deviation

$$r_c = \pm\sqrt{\pm\left(\frac{2c-m}{m}\right)}$$

if $\left(\frac{2c-m}{m}\right)$ is +ve then both sign will be +ve

if $\left(\frac{2c-m}{m}\right)$ is −ve then both sign will be −ve

CA PRANAV POPAT

Apply formula in example.

$$r_c = \pm\sqrt{\pm\left(\frac{2\times 2-7}{7}\right)}$$

$$= \pm\sqrt{\pm\left(\frac{-3}{7}\right)}$$

as $\frac{2c-m}{m} = \frac{-3}{7}$ i.e. negative so

$$r_c = -\sqrt{-\left(\frac{-3}{7}\right)}$$

$$= -\sqrt{\frac{3}{7}} = -0.6546$$

CA PRANAV POPAT

Regression Basics

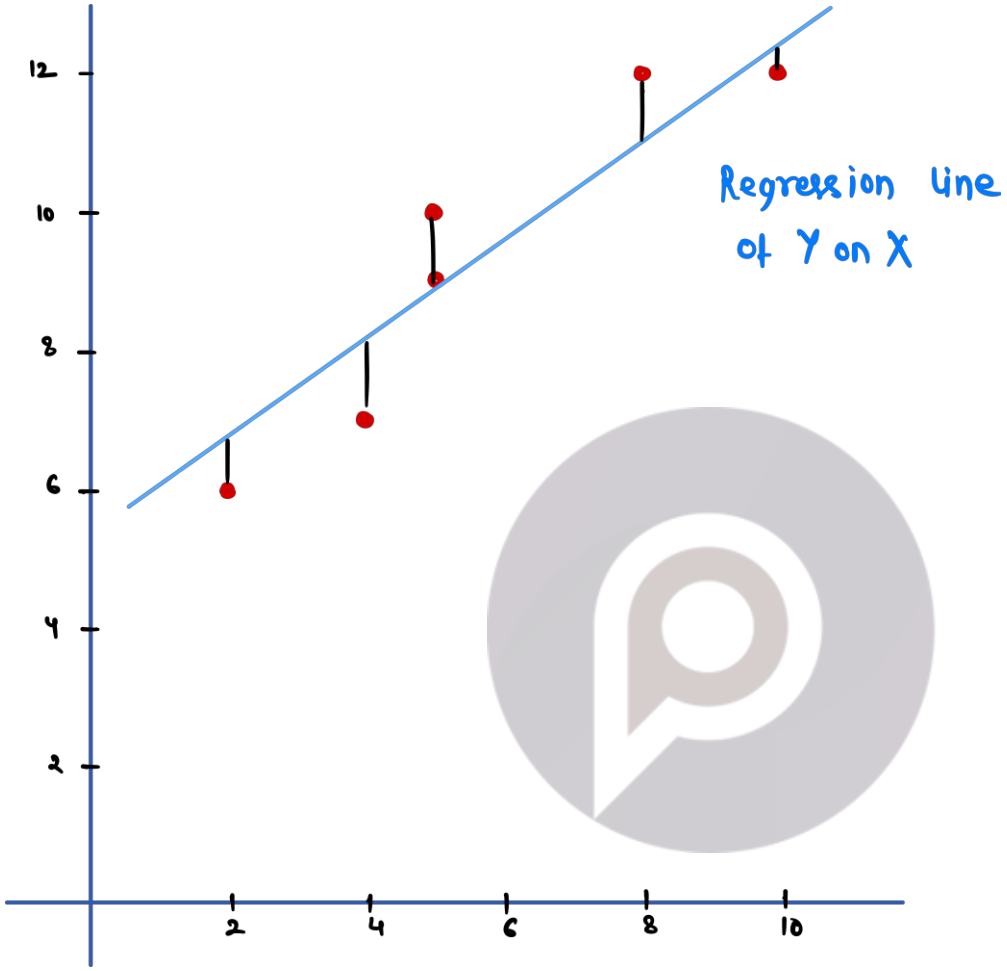
Meaning	*Estimation of one variable for a **given value** of another variable on the basis of an **average mathematical relationship** between the two variables*
Requirements	• *Estimation of Y when X is given* • *Estimation of X when Y is given*
General Points	*Perfect Correlation*: • *When linear relationship exists between two variables, correlation is perfect.* • *Perfect Correlation is represented by a linear equation and this equation can be used for regression purpose directly.* • *Same equation can be used in both ways* *Imperfect Correlation*: • *In case of imperfect correlation there is no definite line and equation* • *We will use method of least square to estimate both regression lines/ equations*

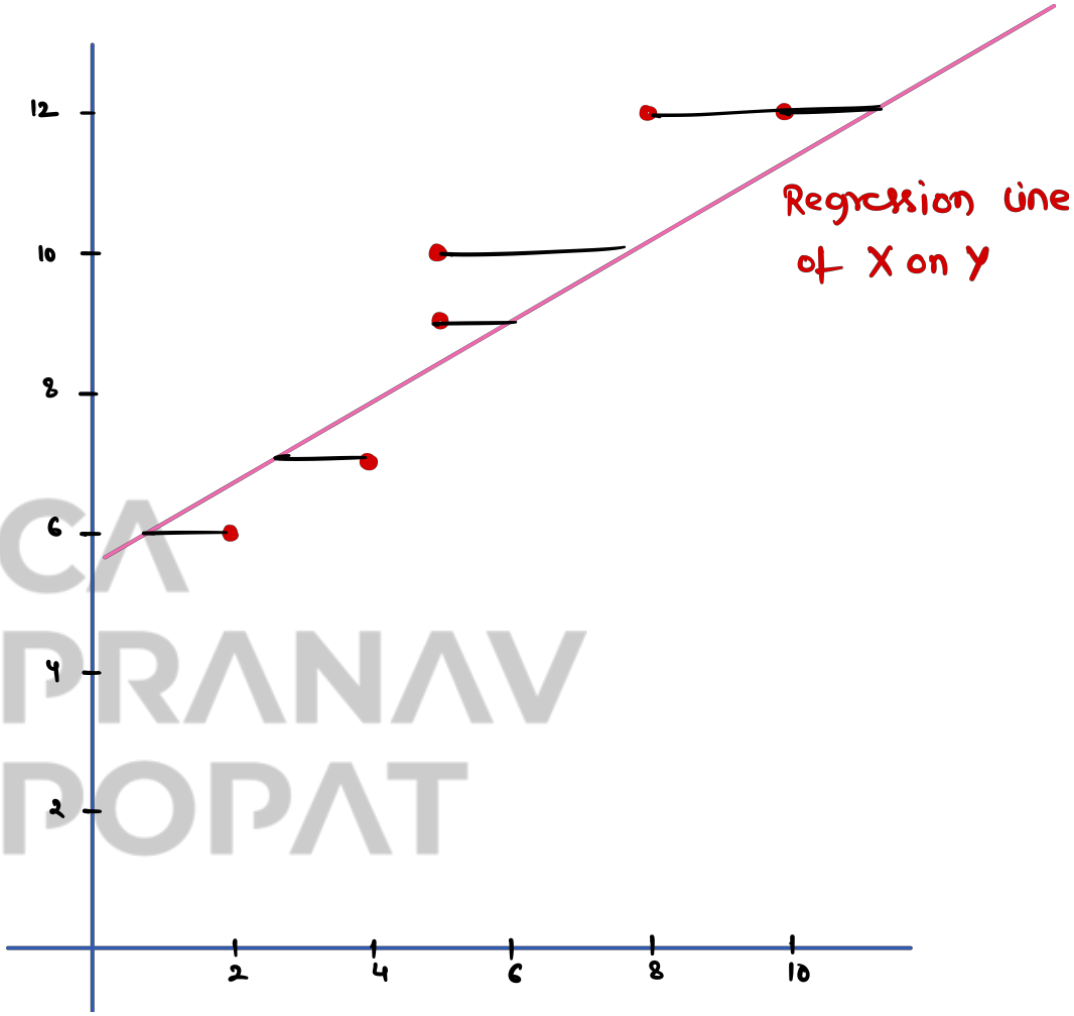
Concept, Formula and Tricks Marathon QA

Find the two regression equations from the following data:

x	2	4	5	5	8	10
y	6	7	9	10	12	12

Hence estimate y when x is 13 and also estimate x when y is 15.





CA PRANAV POPAT

Concept, Formula and Tricks Marathon QA

Part (i) estimate y when $x = 13$

lets form regression eq. of Y on X

$(y - \bar{y}) = b_{yx}(x - \bar{x})$

$\bar{y} = \frac{56}{6} = 9.3333 \quad \bar{x} = \frac{34}{6} = 5.6666$

$$b_{yx} = \frac{cov(x,y)}{var(x)} = \frac{\frac{\Sigma xy}{n} - \bar{x}.\bar{y}}{\frac{\Sigma x^2}{n} - (\bar{x})^2}$$

$$= \frac{\frac{351}{6} - 9.3333 \times 5.6666}{\frac{234}{6} - (5.6666)^2}$$

$$= \frac{5.6119}{6.8896} = 0.8145$$

Reg eq of Y on X,

$y - 9.3333 = 0.8145(x - 5.6666)$

$y = 0.8145x - 4.6154 + 9.3333$

$y = 0.8145x + 4.7179$

if $x = 13$, $y = 0.8145(13) + 4.7179$

$y = 15.3064$ (estimated)

Part (ii) estimate x if $y = 15$

Reg eq of x on y,

$(x - \bar{x}) = b_{xy}(y - \bar{y})$

$\bar{x} = 5.6666 \quad \bar{y} = 9.3333$

$$b_{xy} = \frac{cov(x,y)}{var\ of\ y} = \frac{5.6119}{\frac{\Sigma y^2}{n} - (\bar{y})^2} = \frac{5.6119}{\frac{554}{6} - (9.3333)^2}$$

$= 1.0745$

Reg eq of x on y,

$x - 5.6666 = 1.0745(y - 9.3333)$

Estimated value of x if $y = 15$

$x - 5.6666 = 1.0745(15 - 9.3333)$

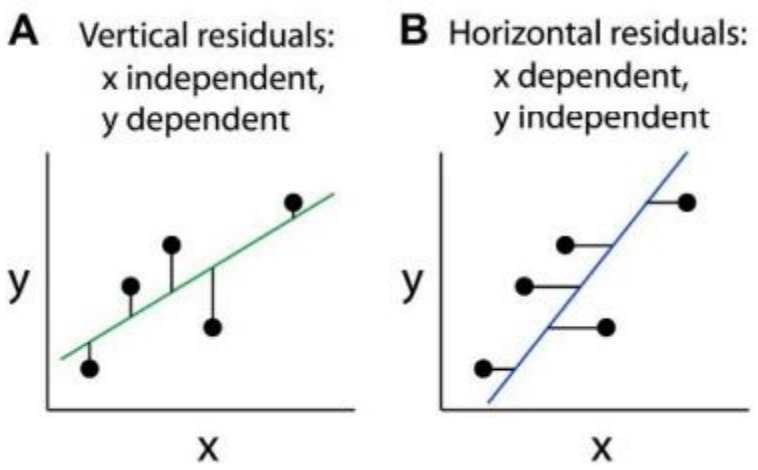
$x = 11.7555$

CA PRANAV POPAT

Formula of Regression Equations/ Lines	*Estimation of Y when X is given*	• *Use Regression line of* ***Y on X*** • *Equation Format:* $Y - \bar{Y} = b_{yx}(X - \bar{X})$ b_{yx} *is regression coefficient of Y on X*
	Estimation of X when Y is given	• *Use Regression line of* ***X on Y*** • *Equation Format:* $X - \bar{X} = b_{xy}(Y - \bar{Y})$ b_{xy} *is regression coefficient of X on Y*
Regression Coefficient	*Regression Coefficient of Y on X*	$b_{yx} = r.\frac{SD_y}{SD_x}$ *and* $b_{yx} = \frac{cov(x,y)}{(SD_x)^2}$
	Regression Coefficient of X on Y	$b_{xy} = r.\frac{SD_x}{SD_y}$ *and* $b_{xy} = \frac{cov(x,y)}{(SD_y)^2}$

PRANAV POPAT

Property 1	*Change of Origin/ Scale for Regression Coefficients: Origin no impact, Scale impact of both magnitude and sign.* $b_{vu} = b_{yx} \times \frac{change\ of\ scale\ of\ y}{change\ of\ scale\ of\ x}$ $b_{uv} = b_{xy} \times \frac{change\ of\ scale\ of\ x}{change\ of\ scale\ of\ y}$
Property 2	*Two regression lines (if not identical) will intersect at the point [means]* $(\bar{x}, \bar{y})$
Property 3	*Correlation Coefficient is the GM of regression coefficients:* $r_{xy} = \pm\sqrt{b_{xy} \times b_{yx}}$ *Note:* r_{xy}, b_{xy}, b_{yx} *all will have same sign*



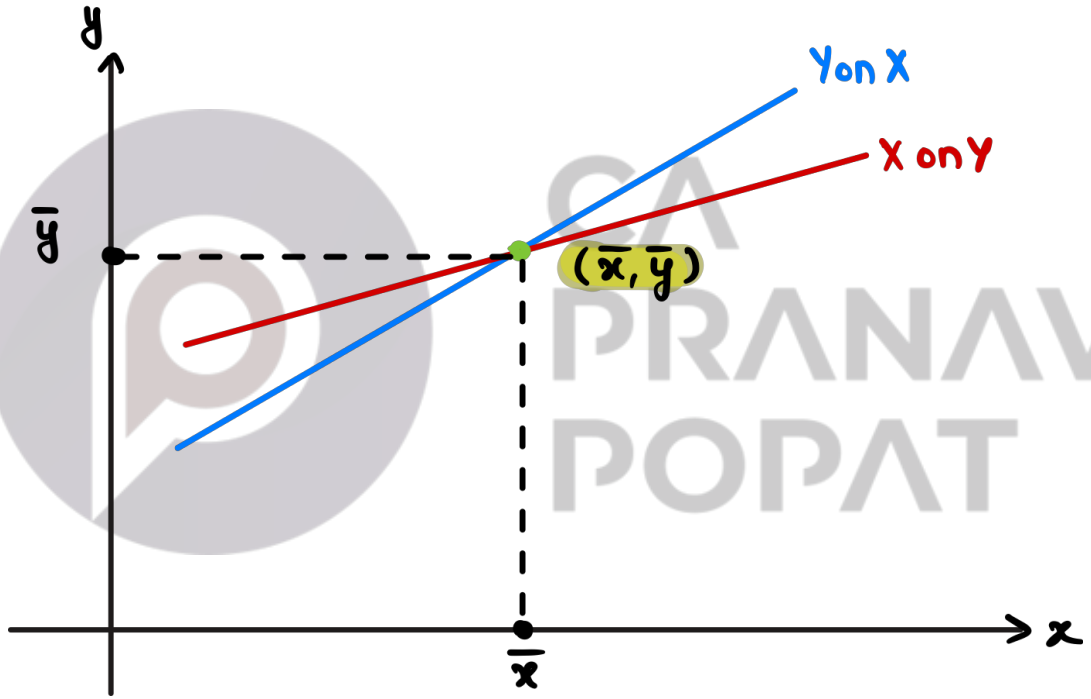
Concept, Formula and Tricks Marathon QA

dependent variable = variable whose value is to be estimated

independent variable = variable whose value is given

Regression equation	dependent variable	independent variable
Y on X	Y	X
X on Y	X	Y

Intersection of two regression lines:



CA PRANAV POPAT

Example:

For the variables x and y, the regression equations are given as $7x - 3y - 18 = 0$ and $4x - y - 11 = 0$

i. *Find the arithmetic means of x and y.*
ii. *Identify the regression equation of y on x.*
iii. *Compute the correlation coefficient between x and y.*
iv. *Given the variance of x is 9, find the SD of y.*

CA PRANAV POPAT

Concept, Formula and Tricks Marathon QA

Part (i) given regression equations are

$7x - 3y - 18 = 0$ and $4x - y - 11 = 0$

as per property, we know intersection point of regression equations are at $\bar{x}, \bar{y}$ so we can solve equations to get $\bar{x}, \bar{y}$

$$7x - 3y - 18 = 0 \quad \text{——} \quad (i)$$

$$4x - y - 11 = 0 \quad \text{——} \quad (ii)$$

Multiply eq. (ii) by 3,

$$12x - 3y - 33 = 0$$

$$-(7x - 3y - 18 = 0)$$

$$5x - 15 = 0$$

$$x = \frac{15}{5} = 3$$

put value of x in eq (ii)

$$4(3) - y - 11 = 0$$

$$y = 1$$

CA PRANAV POPAT

CA PRANAV POPAT

Example 3 (part ii)

Steps to identify regression equations:

1. Assume any one eq. as X on Y and other as Y on X
2. Based on assumption, write X on Y in $x = a + by$ format and write Y on X in $y = a + bx$
3. obtain b_{xy} and b_{yx} from above
4. Calculate r. If r is invalid then assumption is wrong. If r is valid i.e. between -1 and $+1$ then assumption is correct.

$7x - 3y - 18 = 0$ → X on Y

$4x - y - 11 = 0$ → Y on X

(1) X on Y | Y on X

(2) $7x = 3y + 18$ | $4x - 11 = y$

$x = \frac{3}{7}y + \frac{18}{7}$ | $y = 4x - 11$

(3) $b_{xy} = \frac{3}{7}$ | $b_{yx} = 4$

(4) $r = \sqrt{b_{xy} \times b_{yx}} = 1.309 \longrightarrow$ invalid

⇒ assumption is wrong

CA PRANAV POPAT

Concept, Formula and Tricks Marathon QA

Example 3 (Part iii)

$7x - 3y - 18 = 0$

Y on X

$y = a + bx$

$3y = 7x - 18$

$y = -\frac{18}{3} + \frac{7}{3}x$

$b_{yx} = \frac{7}{3}$

$4x - y - 11 = 0$

X on Y

$x = a + by$

$4x = y + 11$

$x = \frac{11}{4} + \frac{y}{4}$

$b_{xy} = \frac{1}{4}$

$r = \sqrt{\frac{7}{3} \times \frac{1}{4}} = 0.7637$

Example 3 (Part iv)

Take $b_{yx} = r \times \frac{SD_y}{SD_x}$

$\frac{7}{3} = 0.7637 \times \frac{SD_y}{\sqrt{9}}$

$SD_y = 9.165$

OR

Take $b_{xy} = r \times \frac{SD_x}{SD_y}$

$\frac{1}{4} = 0.7637 \times \frac{\sqrt{9}}{SD_y}$

$SD_y = 9.1644$

CA PRANAV POPAT

Coefficient of Determination *Accounted Variance/ Explained Variance*	$(r_{xy})^2$
Coefficient of Non-Determination *Unaccounted Variance/ Unexplained Variance*	$1-(r_{xy})^2$
Interpretation of value of $(r_{xy})^2$	***Interpretation of value of*** $\mathbf{r^2}$ • *It explains the percentage of variation in dependent variable due to variation in independent variable* • *Example: if x is demand and y is price and* $r_{xy} = 0.8$ *then* $r^2 = 0.64$ ***Interpretation:*** • *It means 64% of variation in* ***demand (x)*** *is due to variation in* ***price (y)*** *and remaining 36% variation is due to other factors.* • *Accounted Variance/ Explained Variance = 64%* • *Unaccounted Variance/ Unexplained Variance = 36%*

CA FOUNDATION QA

Chp18 Index Number

Concept, Formula and Tricks Marathon



Concepts

Theory of Index Numbers	
Meaning	• *Index numbers are convenient devices for **measuring relative changes (generally in %)** of differences from **time to time** or from **place to place*** • *Series of numerical figures which show relative position* • *Index Numbers show percentage changes rather than absolute amounts of change*
Data Selection	• *It **depends on the purpose** for which the index is used.* • *Index numbers are often constructed from the **sample**. **Random sampling,** and if need be, a **stratified random sampling** can be used to ensure that sample is representative.* • *Data should be **comparable** by ensuring consistency in selection method.*
Base Period	• *It is a **point of reference** in comparing various data.* • *Standard point of comparison.* • *The period should be **normal**.* • *It should be **relatively recent*** • *Choice of suitable base period is a temporary solution*
Use of Averages	• *The **geometric mean is better** in averaging relatives,* • *But for most of the index's **arithmetic mean is used because of its simplicity***

Price/ Quantity/ Value Relative	*For Individual Commodity,* $\frac{\textbf{Current Period Price/ Quantity/ Value}}{\textbf{Base Period Price/ Quantity/ Value}}$
Link Relative	$\frac{P_1}{P_0}, \frac{P_2}{P_1}, \frac{P_3}{P_2}, \ldots, \frac{P_n}{P_{n-1}}$ *Same can be created for quantities also*
Chain relatives/ Chain Indices	*When the above relatives are in respect to a fixed base period these are also called the chain relatives* $\frac{P_1}{P_0}, \frac{P_2}{P_0}, \frac{P_3}{P_0}, \ldots, \frac{P_n}{P_0}$
Formula for Chain Index (when direct data is not available)	$\frac{\textbf{Link relative of current year} \times \textbf{Chain Index of previous year}}{\textbf{100}}$ *The chain index is an unnecessary complication unless of course where data for the whole period are not available or where commodity basket or the weights have to be changed.*

PRANAV POPAT

Concept, Formula and Tricks Marathon QA

Example: Find Link Relatives and Chain Indices using the below data:

Year	1991	1992	1993	1994	1995	1996	1997	1998	1999	2000
Price	50	60	62	65	70	78	82	84	88	90

Yr = 1991 1992 1993 1994 and so on.

link relative: 100 | $\frac{60}{50} \times 100$ | $\frac{62}{60} \times 100$ | $\frac{65}{62} \times 100$

= 120 | = 103.33 | = 104.83%

$$\text{chain relative} = \frac{P_n}{P_0} \times 100$$

Yr = 1991 1992 1993 1994

chain relative: 100 | $\frac{60}{50} \times 100$ | $\frac{62}{50} \times 100$ | $\frac{65}{50} \times 100$

= 120 | = 124 | = 130

CA PRANAV POPAT

Limitations of Index Numbers	• *Chances of errors due to Sampling* • *It gives broad trend not real picture* • *Due to many methods, at times it creates confusion*
Usefulness of Index Numbers	• *Index numbers are very useful in deflating (eg. Nominal wages into real)* • *Framing suitable policies in economics and business* • *They reveal trends and tendencies in making important conclusions* • *They are used in time series analysis to study long-term trend, seasonal variations and cyclical developments*
Shifted Price Index	$\frac{\text{Original Price Index}}{\text{Price Index of the year on which it has to be shifted}} \times 100$

Example:

Year	2012	2013	2014	2015	2016	2017
Price Index (Base 2012)	100	112	110	120	124	136

if base is to be shifted from 2012 to 2015. Find new index values with base as 2015

CA PRANAV POPAT

Concept, Formula and Tricks Marathon QA

Sol

New Index value of 2015 = 100

New Index value of 2016 = $\frac{100}{120} \times 124 = 103.333$

shifted price index of 2016 (new) = $\frac{\text{original price index}}{\text{price index of year where base is shifted}} \times 100$

$= \frac{124}{120} \times 100 = 103.33$

shifted price index of 2017 = $\frac{136}{120} \times 100 = 113.33$

shifted price index of 2014 = $\frac{110}{120} \times 100 = 91.66$

Year	2012	2013	2014	2015	2016	2017
Price Index with base 2012	100	112	110	120	124	136
shifted price index with base 2015	83.33	93.33	91.66	100	103.33	113.33

CA PRANAV POPAT

Concept, Formula and Tricks Marathon QA

Example:

	2012	2013	2014	2015
Price Index (A) base 2012	100	120	124	130

	2015	2016	2017	2018
Price Index (B) base 2015	100	136	142	160

Prepare an Index number series by splicing above Index series taking base as 2015

Sol

new index values for price index A taking 2015 as base

2015 : 100

2014 : $\frac{124}{130} \times 100 = 95.38$

2013 : $\frac{120}{130} \times 100 = 92.31$

2012 : $\frac{100}{130} \times 100 = 76.92$

Spliced index no. series with 2015 as base

Year	2012	2013	2014	2015	2016	2017	2018
Index	76.92	92.31	95.38	100	136	142	160

CA PRANAV POPAT

Methods of Index Numbers

Simple Aggregative Method

Method	• *Price Index is expressed as **total of commodity prices** in a given year as a percentage of total of commodity prices in the base year*
Formula	$\frac{\Sigma P_n}{\Sigma P_0} \times 100$
Merits	• *Easy to compute*
Demerits/ Limitations	• *Commodity with higher price will have greater influence in index value* • *price quotations become the concealed (hidden) weights which have no logical significance* • *If units of prices are changed, index will also change*

PRANAV POPAT

Example: Find Price Index of Year 1999 and 2000 taking 1998 as base using Simple Aggregative Index Method.

Commodities	1998	1999	2000
Cheese (per 100 gm)	12.00	15.00	15.60
Egg (per piece)	3.00	3.60	3.30
Potato (per kg)	5.00	6.00	5.70

Formula of Price Index = $\frac{\Sigma P_n}{\Sigma P_0} \times 100$

Price Index for 1999 taking 1998 as base

$= \frac{15 + 3.6 + 6}{12 + 3 + 5} \times 100 = \frac{24.6}{20} \times 100 = 123$

Price Index for 2000 taking 1998 as base

$= \frac{15.6 + 3.3 + 5.7}{12 + 3 + 5} \times 100 = \frac{24.6}{20} \times 100 = 123$

CA PRANAV POPAT

Example: Find Price Index of Year 1999 and 2000 taking 1998 as base using Simple Aggregative Index Method.

Commodities	1998	1999	2000
Cheese (per 100 gm)	12.00	15.00	15.60
Egg (per dozen)	36.00	43.20	39.60
Potato (per kg)	5.00	6.00	5.70

Price Index of 1999 with base 1998

$$= \frac{15 + 43.2 + 6}{12 + 36 + 5} \times 100 = \frac{64.2}{53} \times 100 = 121.13$$

Price Index of 2000 with base 1998

$$= \frac{15.6 + 39.6 + 5.7}{12 + 36 + 5} \times 100 = 114.905$$

CA PRANAV POPAT

Simple Average of Relatives Method

Method	• *Under this method, we invert the actual price for each variable into percentage of the base period. These percentages are called relatives.* • *The index number is the average of all such relatives.*
Formula	$\frac{\Sigma \frac{P_n}{P_0}}{N}$
Merits	• *One big advantage of price relatives is that they are* ***pure numbers****.* • *Price index number computed from relatives will remain the same regardless of the units by which the prices are quoted*
Demerits	• *Inspite of some improvement, the above method has a flaw that it gives* ***equal importance to each of the relatives*** *(Will not be suitable if the commodities do have equal importance in Index)* • *This defect can be remedied by the introduction of an appropriate weighing system*

Concept, Formula and Tricks Marathon QA

Example: Find Price Index of Year 1999 and 2000 taking 1998 as base using Simple average of Price Relatives Index Method.

Commodities	1998	1999	2000
Cheese (per 100 gm)	12.00	15.00	15.60
Egg (per piece)	3.00	3.60	3.30
Potato (per kg)	5.00	6.00	5.70

Price relative data

	1998	1999	2000
Cheese	100 (12/12 x 100)	125 (15/12 x 100)	130 (15.6/12 x 100)
Egg	100	120 (3.6/3 x 100)	110 (3.3/3 x 100)
Potato	100	120 (6/5 x 100)	114 (5.7/5 x 100)
Index overall $\frac{\Sigma(P_n/P_0)}{n}$	100	121.66 $\left(\frac{125+120+120}{3}\right)$	118 $\left(\frac{130+110+114}{3}\right)$

CA PRANAV POPAT

Concept, Formula and Tricks Marathon QA

Example: Find Price Index of Year 1999 and 2000 taking 1998 as base using Simple average of Price Relatives Index Method.

Commodities	*1998*	*1999*	*2000*
Cheese (per 100 gm)	*12.00*	*15.00*	*15.60*
Egg (per dozen)	*36.00*	*43.20*	*39.60*
Potato (per kg)	*5.00*	*6.00*	*5.70*

Price relatives of Egg

	1998	1999	2000
	100	120	110
		$\left(\frac{43.2}{36} \times 100\right)$	$\left(\frac{39.6}{36} \times 100\right)$
Index	100	121.66	118

CA PRANAV POPAT

Weighted Aggregative Index Method

General Points	• Under this method we weigh the price of each commodity by a **suitable factor** often taken as the **quantity or value weight** sold during the base year or the given year or an average of some years. • There are various alternate formulas **(depends on base used)** • Here indices are shown as %

Method Name	Remark	Formula
Laspeyres' Index	Weight – Base Year Quantity	$\frac{\Sigma P_n Q_0}{\Sigma P_0 Q_0} \times 100$
Passche's Index	Weight – Current Year Quantity	$\frac{\Sigma P_n Q_n}{\Sigma P_0 Q_n} \times 100$
Marshall-Edgeworth Index	Weight – Sum of Base Year Quantity and Current Year Quantity	$\frac{\Sigma P_n (Q_0 + Q_n)}{\Sigma P_0 (Q_0 + Q_n)} \times 100$
Fisher's Index	GM of Laspeyres' Index and Paasce's Index	$\sqrt{\frac{\Sigma P_n Q_0}{\Sigma P_0 Q_0} \times \frac{\Sigma P_n Q_n}{\Sigma P_0 Q_n}} \times 100$
Bowley's Index	AM of Laspeyres' Index and Paasce's Index	$\frac{\frac{\Sigma P_n Q_0}{\Sigma P_0 Q_0} + \frac{\Sigma P_n Q_n}{\Sigma P_0 Q_n}}{2}$

Example:

Find Price Index using all weighted aggregative index methods

Commodities	1998 (Base Year)		2020 (Current Year)	
	Price	Quantity	Price	Quantity
A	4	10	5	9
B	20	3	20	4
C	50	2	44	2.5

Laspeyre's Index (Price) = $\frac{\Sigma P_n q_0}{\Sigma P_0 q_0} \times 100$

$= \frac{5 \times 10 + 20 \times 3 + 44 \times 2}{4 \times 10 + 20 \times 3 + 50 \times 2} \times 100$

$= \frac{198}{200} \times 100 = 99$ $\quad 0.99 \times 100 = 99$

CA PRANAV POPAT

Paasche's Price Index $= \frac{\Sigma P_n q_n}{\Sigma P_0 q_n} \times 100$

$= \frac{5 \times 9 + 20 \times 4 + 44 \times 2.5}{4 \times 9 + 20 \times 4 + 50 \times 2.5} \times 100$

$= \frac{235}{241} \times 100 = 97.51$

Marshal edgeworth's Price Index $= \frac{\Sigma P_n (q_0 + q_n)}{\Sigma P_0 (q_0 + q_n)} \times 100$

$= \frac{5 \times 19 + 20 \times 7 + 44 \times 4.5}{4 \times 19 + 20 \times 7 + 50 \times 4.5} \times 100$

$= \frac{433}{441} \times 100 = 98.18$

Note: originally Marshal edgeworth was based on weightage to be taken as avg of current yr and base yr q1y.

CA PRANAV POPAT

Fisher's Index : GM of Laspeyrey's Index and Paasche's Index

or

$$\sqrt{\frac{\Sigma P_n q_0}{\Sigma P_0 q_0} \times \frac{\Sigma P_n q_n}{\Sigma P_0 q_n}} \times 100$$

GM of 99 and 97.51 i.e. $\sqrt{99 \times 97.51}$

$= 98.25$

or

$$\sqrt{\frac{198}{200} \times \frac{235}{241}} \times 100 = 98.252$$

CA PRANAV POPAT

Concept, Formula and Tricks Marathon QA

Bowley's Index: AM of Laspeyres' and Paasche's Index

$$= \frac{\frac{\Sigma P_n q_0}{\Sigma P_0 q_0} + \frac{\Sigma P_n q_n}{\Sigma P_0 q_n}}{2} \times 100$$

OR

$$= \frac{99 + 97.51}{2} = 98.255$$

OR $$= \frac{\frac{198}{200} + \frac{235}{241}}{2} \times 100 = 98.255$$

CA PRANAV POPAT

Weighted average of Relative Method

General Points	• *To overcome the disadvantage of a simple average of relative method, we can use weighted average of relative method* • *Generally weighted arithmetic mean is used although the weighted geometric mean can also be used.* • *It is same as* ***Laspeyres' Index***
Formula	$$\frac{\Sigma \frac{P_n}{P_0} \times P_0 Q_0}{\Sigma P_0 Q_0} \times 100 = \frac{\Sigma P_n Q_0}{\Sigma P_0 Q_0} \times 100$$

CA PRANAV POPAT

Other Topics having practical problems

Cost of Living Index (also called General Index)	• *CLI is defined as the **weighted AM of index numbers** of few groups of basic necessities.* • *AM of group indices gives the General Index* • *Generally, for calculating CLI; food, clothing, house rent, fuel & lightning and miscellaneous groups are taken into consideration.* • *Examples of CLI: WPI, CPI, etc.*
Formula for Deflated Value	$$\textbf{Deflated Value} = \frac{\textbf{Current Value}}{\textbf{Price Index of the current year}}$$ • *It is used to calculate the value in real terms*

CA PRANAV POPAT

Test of Adequacy

Unit Test	• *This test requires that the formula should be independent of the unit in which or for which prices and quantities are quoted.* • *Except for the simple (unweighted) aggregative index all other formulae satisfy this test.*
Time Reversal Test	• *It is a test to determine whether a given method will work both ways in time, forward and backward.* • $P_{01} \times P_{10} = 1$ • *Laspeyres' method and Paasche's method do not satisfy this test, but Fisher's Ideal Formula does.*
Factor Reversal Test	• *This holds when the product of price index and the quantity index should be equal to the corresponding value index.* • *Symbolically* $P_{01} \times Q_{01} = V_{01}$ • *Fisher's Index Number is ideal as it satisfies Unit, Time Reversal and Factor Reversal Test*

POPAT

Circular Test	• *This property therefore enables us to adjust the index values from period to period without referring each time to the original base.* • *It is an extension of time reversal test* • *The test of this* ***shiftability of base*** *is called the circular test.* • *This test is not met by Laspeyres, or Paasche's or the Fisher's ideal index.* • *The weighted GM of relative,* ***simple geometric mean of price relatives*** *and the* ***weighted aggregative with fixed weights meet this test.*** *(These methods are not in syllabus)*
Symbol of Index	• P_{01} *is the index for time 1 on 0* • P_{10} *is the index for time 0 on 1*

CA PRANAV POPAT

Concept Insights (Newly Added)

Stock Market Index	• *It represents the entire stock market. It shows the changes taking place in the stock market.* • *Movement of index is also an indication of average returns received by the investors.* • *With the help of an index, it is easy for an investor to compare performance as it can be used as a benchmark, for e.g. a simple comparison of the stock and the index can be undertaken to find out the feasibility of holding a particular stock.* • *Examples: India – Sensex of BSE, Nifty of NSE*
Bombay Stock Exchange Limited	• *It is the oldest stock exchange in Asia and was established as "The Native Share & Stock Brokers Association" in 1875.* • *The Securities Contract (Regulation) Act, 1956 gives permanent recognition to Bombay Stock Exchange in 1956.*

PRANAV POPAT

	• *BSE became the first stock exchange in India to obtain such permission from the Government under the Act.* • *BSE Sensex contains basket of 30 constituent stocks. The base year of BSE SENSEX is 1978-79 and the base value is 100 which has grown over the years and quoted at about 592 times of base index as on date.* • *As the oldest Index in the country, it provides the time series data over a fairly long period of time (from 1979 onward).*
National Stock Exchange	• *NSE was incorporated in 1992. It was recognized as a stock exchange by SEBI in April 1993 and commenced operations in 1994.* • *NIFTY50 is a diversified 50 stocks Index of 13 sectors of the economy.* • *The base period of NIFTY 50 Index is 3 November 1995 and base value is 1000 which has grown over years and quoted at 177 times as on date.*

CA PRANAV POPAT

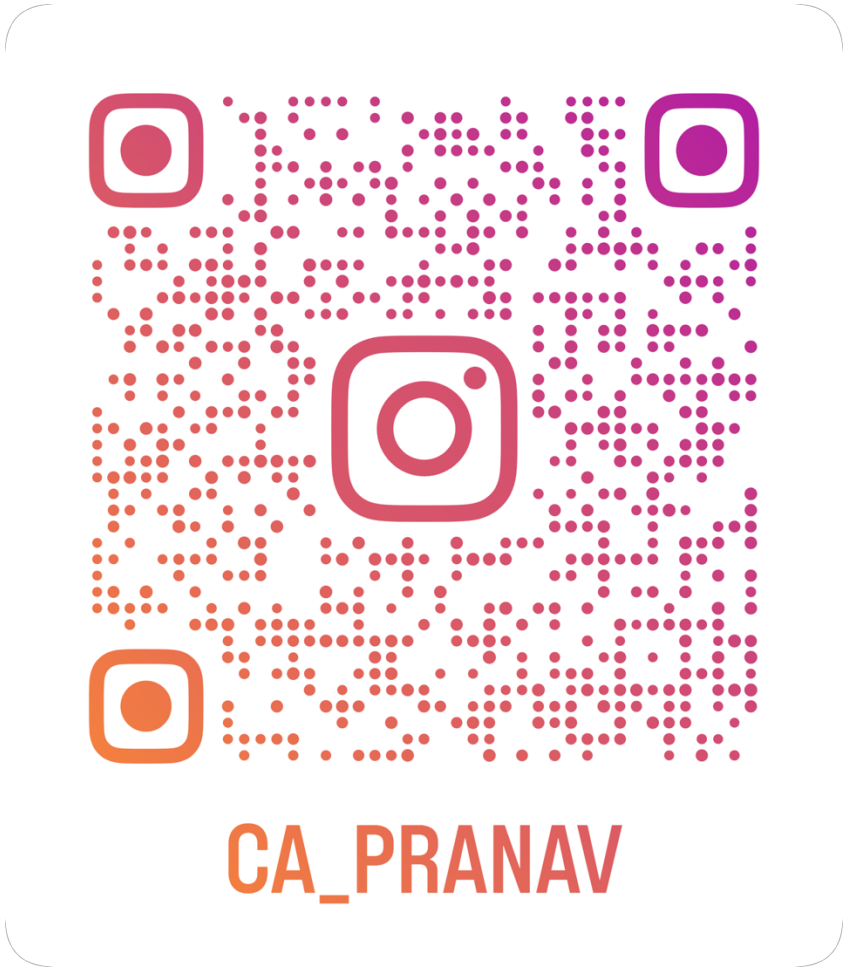
Market Capitalization Weighted Method	*Following steps are involved in calculation of index on a particular date:* • *Calculate market capitalization of each individual company comprising the index.* • *Calculate the total market capitalization by adding the individual market capitalization of all companies in the index.* • *Computing index of next day requires the index value and the total market capitalization of the previous day and is computed as follows:* • $\text{index value today} = \text{index on previous day} \times \frac{\text{total market capitalisation for current day}}{\text{total market capitalisation for previous day}}$ *Note: almost all equity indices worldwide are calculated using the market capitalization weighted method.*
CPI	• *Consumer Price Index/ Cost of living Index or Retail Price Index is the Index which measures the effect of change in prices of basket of goods and services on the purchasing power of specific class of consumer during any current period w.r.t to some base period.*
WPI	• *Whole Sale Price Index - The WPI measures the relative changes in prices of commodities traded in wholesale market.*



Join my Telegram Channel for NOTES



QA REVISION LIBRARY



Follow Insta for connecting socially