__/__/__

* Ch :- 2 A — Measures of Central Tendency

- Average Discover Uniformity in Variability

Average

- **Mathematical** [Based on all obs..]
  - **A.M**
    - best
    - Commonly used
  - **G.M.**
    - Avg. of ~~%~~, %, Groth Rate, Ratio
  - **H.M.**
    - Avg. Speed
- **Positional** [Not affected by Extreme value]
  - **Median**
    - Qualitative Data
    - ogive
  - **Mode**
    - Quickest
    - Histogram
    - Area Diagram

_/_/_

* **A.M.**

$\bar{x} = \frac{\Sigma x_i}{n}$ | $\bar{x} = \frac{\Sigma f_i x_i}{n}$

$x : x_1, x_2, ---x_n$ | $x : x_1, x_2 --- x_n$

| $F : F_1, F_2 --- F_n$

Class	f
$a_1 - b_1$	f_1
$b_1 - c_1$	f_2
'	'
'	'

$\bar{x} = \frac{\Sigma f_i x_i}{n}$ → Class Mark Mid.. value

Ex :-

$x : 10, 13, 14, 18, 20$ | $\bar{x} = 15$

$x - \bar{x} : -5, -2, -1, 3, 5 \rightarrow 0$

$(x-\bar{x})^2 : 25, 4, 1, 9, 25 \rightarrow \underline{64}$ Minimum

$x - 14 : -4, -1, 0, 4, 6$

$(x-14)^2 : 16, 1, 0, 16, 36 \rightarrow 69$

$|x-14| : 4, 1, 0, 4, 6 \rightarrow \underline{15}$ Minimum

$|x-\bar{x}| : 5, 2, 1, 3, 5 \rightarrow 16$

_/_/_

Note :-

1. Sum of deviation from A.M. is zero
2. Sum of squre of deviation from A.M. is Minimum
3. Sum of absolute deviation from Median is Minimum.

* **Combine Mean**

$$\bar{x}_c = \frac{n_1\bar{x}_1 + n_2\bar{x}_2}{n_1+n_2}$$

Weighted Mean

$$\bar{x} = \frac{\Sigma w_i x_i}{\Sigma w_i}$$

Ex:

x	w
1	1
2	2
3	3
4	4
	10

$$\bar{x} = \frac{\Sigma w_i x_i}{\Sigma w_i} = \frac{MRC}{10} = 3$$

Trick $x = \frac{2 \times 4 + 1}{3} = 3$.

_/_/_

Trick :- 1

x	w/f
1	1
2	2
3	3
4	4
|	|
|	|
n	n

$$\bar{x} = \frac{2n+1}{3}$$

Trick :- 2

If nos. are in A.P

$$\bar{x} = \frac{a+l}{2}$$

$x : 1, 2, 3 ---- 10$

$$\bar{x} = \frac{1+10}{2} = 5.5$$

__/__/____

Imp Affected by Change of Origin and Scale.

If $ax + by = c$ then
$a\bar{x} + b\bar{y} = c$

Ex :- $\bar{x} = 30$
$y = 2x - 10$
$\bar{y} = 2\bar{x} - 10$
$\bar{y} = 2 \times 30 - 10$
$= 50$

$\bar{x} = 5$
$2x - 4y = 40$
$2\bar{x} - 4\bar{y} = 40$
$2 \times 5 - 4\bar{y} = 40$
$10 - 4\bar{y} = 40$
$10 - 40 = 4\bar{y} = 0$
$-30 - 4\bar{y}$
$\frac{-30}{4} = \bar{y} = -7.5$

* G.M :-

→ Special Average used only for %, Growth Rate or Ratio.

→ $G.M = \sqrt[n]{x_1 \times x_2 \times \dots x_n}$

$G.M = (\text{Product})^{1/n}$

_/_/_

Note :-

$$n \times A.M = Sum$$
$$G.M^n = Product.$$

- Most difficult to Calculate.

- Cannot be asure if any value is Zero or Negative

- Combine G.M

group of : n_1 n_2
G.M : GM_1 GM_2

$$GM_c = \left(GM_1^{n_1} \times GM_2^{n_2}\right)^{\frac{1}{n_1+n_2}}$$

Ex If $GM(x) = 2$, $GM(y) = 3$
then $GM(x,y) = GM(x)\, GM(y)$
$= 2 \times 3 = 5$

- If $Z = xy$ then $GM_Z = GM(x)\,GM(y)$
- If $Z = \frac{x}{y}$ then $GM_Z = \frac{GM(x)}{GM(y)}$

If nos. are in GP
then $GM = \sqrt{a \times c}$

__/__/____

* H.M.

— Special type of average

— $H.M.(a,b) = \dfrac{2}{\frac{1}{a} + \frac{1}{b}}$

$H.M(a,b,c) = \dfrac{3}{\frac{1}{a} + \frac{1}{b} + \frac{1}{c}}$

Ex:- $H.M(2,3,5) = \dfrac{3}{\frac{1}{2} + \frac{1}{3} + \frac{1}{5}}$

$$\left.\begin{array}{l} 2 \div = M+ \\ 3 \div = M+ \\ 5 \div = M+ \end{array}\right] MRC \div = \times\ 3 = 2.90$$

— **It is used to find**

(I) Avg. Speed when distance is Constant

(II) Avg. Rate when total Interest is Constant

(III) Avg. Price when total Expenditure is Constant

— Cannot Calculate if any value is Zero.

— Combine H.M.

Group of : n_1, n_2
H.M : HM_1, HM_2

$$HM_c = \frac{n_1 + n_2}{\frac{n_1}{HM_1} + \frac{n_2}{HM_2}}$$

* Note

(1) $AM \geq GM \geq H.M$

If all value are same
$AM = GM = HM$

If values are distinct
$AM > GM > HM$

(2) $G.M = \sqrt{A.M \times H.M}$

$G.M^2 = AM \times H.M$

Ex :-

1. 1, 4

AM	GM	HM
2.5	2	1.6

2. 2, 2

AM	GM	HM
2	2	2

_/_/_

* Median

- 2^{nd} best average after A.M
- Qualitative
- not affected by extreme value
- only MCT for open-ends
- Graphically located using ogive
- Affected by change of origin and Scale

If $ax + by = c$ then
$aM_x + bM_y = c$

Ex ① 3, 5, 10, 19, 23
M

② 3, 5, 10, 19, 23, 51
$\frac{10+19}{2} = 14.5$

__/__/____

(3)

x	f
9	3
5	8
2	1
4	2

x	f	cf
2	1	1
4	2	3
5	8	11
9	3	14 = N

$\frac{n+1}{2} = 7.5$

(4)

Class	f	cf
0-5	3	3
5-10	2	5
10-15	10	15
15-20	5	20 = N

$\frac{N}{2} = \frac{20}{2} = 10$

$M = l + \frac{\frac{n}{2} - cf}{f} \times c$

$M = 10 + \frac{10-5}{10} \times 5$

$= 12.5$

l = lower class boundry
C = class width

_/_/_

* Quartiles :

$$Q_i = i\left(\frac{N+1}{4}\right)^{th} \text{ Rank}$$

$$Q_i = l + \frac{i\frac{n}{4} - cf}{f} \times c$$

* Deciles :

$$D_i = i\left(\frac{N+1}{10}\right)^{th} \text{ Rank}$$

$$D_i = l + \frac{i\frac{N}{10} - cf}{f} \times c$$

* Percentiles :

$$P_i = i\left(\frac{N+1}{100}\right)^{th} \text{ Rank}$$

$$P_i = l + \frac{i\frac{N}{100} - cf}{f} \times c$$

— Note : All Partition values (M, Q, D, P) can be located graphically using Ogive.

__/__/__

* Mode

- Highest frequency
- Quickest
- Ill - defined

Ex ① 1, 1, 1, 2, 2, 3

$Z = 1$

② 1, 1, 1, 2, 2, 2, 3, 3

$Z = 1, 2$

- Graphically located using Histogram
- If $ax + by = c$ then

$aZ_x + bZ_y = c$

- $Z = l + \frac{f_m - f_1}{2f_m - f_1 - f_2} \times c$

f_1

$\boxed{f_m}$ $f_m - f_1 = d_1$

f_2 $f_m - f_2 = d_2$

_/_/_

$$Z = l + \frac{d_1}{d_1 + d_2} \times c$$

$$2f_m - f_1 - f_2$$
$$= (f_m - f_1) + (f_m - f_2)$$
$$= d_1 + d_2$$

— Symmetric:

$$\bar{x} = M = Z$$

— Moderately Skewed:

$$Z = 3M - 2\bar{x}$$

Mean − Mode = 3(Mean − Median)

— Note:

(1) Based on all observation: AM, GM, HM

(2) Rigidly defined (unique): Mean, Median

(3) Affected by extreme values: A.M, GM, H.M

(4) Algebric Properties (combine): AM, GM, HM

_/_/_

* Ch :- 2 A

Measures of Dispersion.

→ Dispersion discovers Variability in Uniformity

→ 2^{nd} Order average

→ Not affected by change of origin but affected by change of Scale

→ MCT can take any value but MD is always - NON Negative

_/_/__

Measure of Dispersion (Scatternes)

Absolute (used to Measure.)	Relative (unit free) (used to Compare)
(1) Range = H - L	(1) Coeff. of Range $= \frac{H-L}{H+L} \times 100$
(2) $QD = \frac{Q_3 - Q_1}{2}$	(2) Coeff. of QD $= \frac{(Q_3 - Q_1)}{(Q_3 + Q_1)} \times 100$ $= \frac{Q.D}{Q_2} \times 100$
(3) $MD_A = \frac{\Sigma \lvert x - A \rvert}{n}$ $A = \bar{x}, M, Z$	(3) Coeff. of MD_A $= \frac{MD_A}{A} \times 100$
(4) $S.D = \sqrt{\frac{\Sigma(x-\bar{x})^2}{n}}$ $\boxed{\text{Best}}$	(4) Coeff. of Variance $= \frac{SD}{\bar{x}} \times 100$

_/_/_

Imp Ex 15, 18, 20, 25, 30, 32, 35.

→ Range = H − L = 35 − 15 = 20

Coeff of Range = $\frac{H-L}{H+L} \times 100 = \frac{35-15}{35+15} \times 100$

$= 40.$

→ $Q_1 = \frac{7+1}{4} = 2^{nd}$ rank $= 18$

$Q_3 = 3\left(\frac{7+1}{4}\right) = 6^{th}$ rank $= 32$

$QD = \frac{32-18}{2} = 7$

Coeff. of QD $= \frac{32-18}{32+18} \times 100 = 28$

→ $\bar{x} = 25$

$MD_{\bar{x}} = \frac{\Sigma(x - 25)}{7}$

15 − 25 = −10 M−
18 = −7 M−
20 = −5 M−
30 = 5 M+
32 = 7 M+
35 = 10 M+

MRC ÷ 7
= 6.29 MD

Coeff of MD
$= \frac{6.29}{25} \times 100$
$= 25.14.$

$$SD = \sqrt{\frac{\Sigma(x-\bar{x})^2}{7}}$$

15 - 25 x = M+
18 - 25 x = M+
20 - 25 x = M+
30 - 25 x = M+
32 - 25 x = M+
35 - 25 x = M+

$\sqrt{\text{MRC} \div 7}$
7.05
= SD

7.05 ÷ 25 × 100 = 28.20
Coeff. of variance

* **Range**

- Quickest
- Not based on all observation
- If $ax + by = c$ then

$$|a| R_x = |b| R_y$$

Ex :- $-3x + 4y + 6 = 0$

$$3R_x = 4R_y$$

* **QD**

- Only one which can be used in <u>Open end</u>
- Not based on all observation (Central <u>50</u> %)
- If $ax + by = c$ then

$$|a| QD_x = |b| QD_y$$

_/_/_

* MD

$MD_A = \frac{\Sigma |x-A|}{n}$

$A = \bar{x}, M, Z$

- Minimum for $A = M$
- Based on all observation
- If $ax + by = c$ then

$|a| MD_x = |b| MD_y$

* S.D

- Best, Most Commonly used
- Based on all observation
- If $ax + by = c$ then

$|a| SD_x = |b| SD_y$

- $V(x) = SD^2$

__/__/____

- 'Root mean Square of Deviation from mean'

- $SD = \sqrt{\frac{\Sigma x^2}{n} - \left(\frac{\Sigma x}{n}\right)^2}$

$\sigma = \sqrt{\frac{\Sigma x^2}{n} - (\bar{x})^2}$

$\sigma^2 = \frac{\Sigma x^2}{n} - \bar{x}^2.$

$$\boxed{\Sigma x^2 = (\sigma^2 + \bar{x}^2) \times n}$$

$$\boxed{\Sigma x = \bar{x} \times n}$$

- For two nos. a, b

$SD = \frac{|b-a|}{2} = \frac{Range}{2}$

SD is half of Range

Ex:- 3, 11

$\bar{x} = 7$, $3-7 \times = M+$, $11-7 \times = M+$] - MRC $\div 2 = \sqrt{}$ SD = 4.

Trick $SD = \frac{11-3}{2} = 4.$

$\rightarrow$ SD (3, 3, 11, 11)

$\bar{x} = 7$

$3 - 7 \; x = M+ M+$ } MRC $\div 4 = \sqrt{\;}$

$11 - 7 \; x = M+ M+$ } $= 4$

— If all frequency are same then to find SD consider each values only once.

$\rightarrow$ $x = 1, 2, 3, 4, 5$

$$SD = \sqrt{\frac{5^2 - 1}{12}} = 1.4142$$

— If $x = 1, 2, 3, 4, \text{-----} n$.

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

$$V(x) = \frac{n^2 - 1}{12}$$

* Combine SD

Group of	n_1	n_2
Mean	$\bar{x}_1$	$\bar{x}_2$
SD	S_1	S_2

$$SD_c = \sqrt{\frac{n_1(s_1^2+d_1^2)+n_2(s_2^2+d_2^2)}{n_1+n_2}}$$

$$\bar{x}_c = \frac{n_1\bar{x}_1 + n_2\bar{x}_2}{n_1+n_2}$$

$$d_1 = \bar{x}_c - \bar{x}_1$$

$$d_2 = \bar{x}_c - \bar{x}_2$$

_/_/_

* Ch :- 3A

Correlation Analysis

→ * Correlation :

— Degree of association between two or more Variable

— r = Cofficient of Correlation

$+1 \le r \le 1$

— $r = 0$ No relation

$r = 1$ Perfect Positive Relation

$r = -1$ Perfect Negative Relation

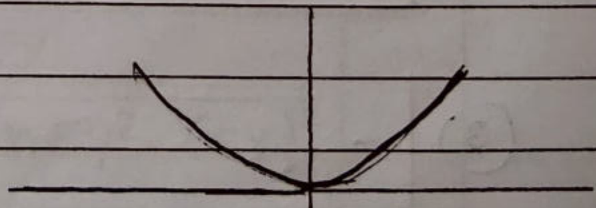
— In Correlation we have only Linear relation.

Ex :-

x	:	-2	-1	1	2
y	:	4	1	1	4

$y = x^2$

No relation



— Not affected by Change of origin (+/-) and Sale (×/÷)

_/_/_

* Method to Calculate Correlation

(1) Scatter Diagram: Cannot Measure value of r

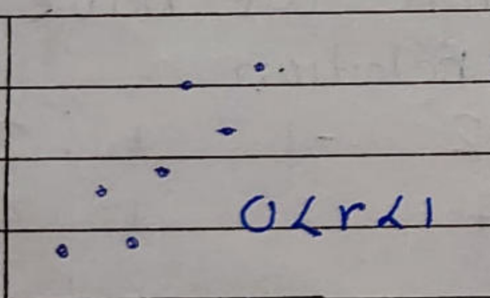
(2) Karl Pearson's Method: Best (Product moment)

(3) Rank Correlation: For Qualitative data

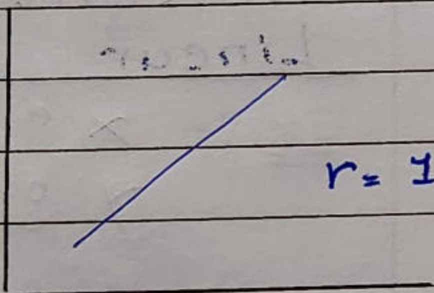
(4) Concurrent Deviation: Quickest Method

(1) Scatter Diagram

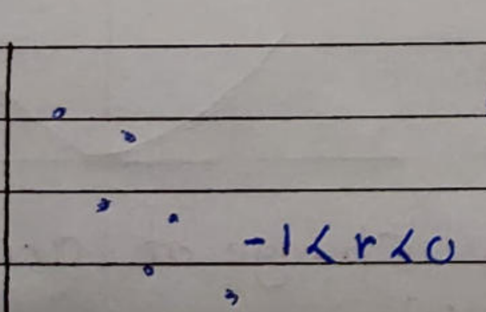
(1)



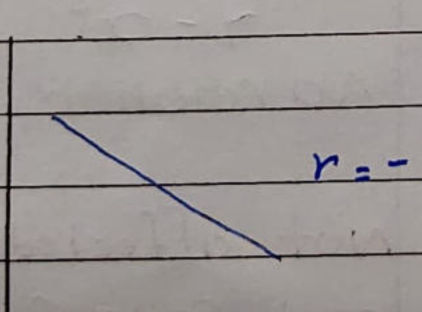
(2)



(3)

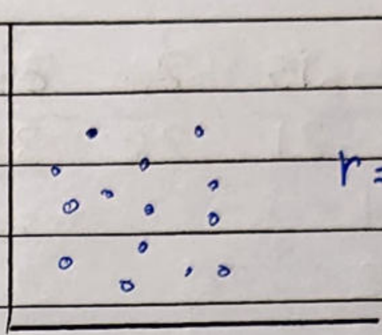


(4)



_/_/_

(5)



$r = 0$

(2) Karl Pearson's Method (Covariance method)

$r = \dfrac{Cov.(x,y)}{S_x . S_y}$	$Cov(x,y) = \dfrac{\Sigma(x-\bar{x})(y-\bar{y})}{n}$
$r = \dfrac{\Sigma(x-\bar{x})(y-\bar{y})}{\sqrt{\Sigma(x-\bar{x})^2}\sqrt{\Sigma(y-\bar{y})^2}}$	$Cov(x,y) = \dfrac{\Sigma xy - n\bar{x}\bar{y}}{n}$
$r = \dfrac{\Sigma xy - n\bar{x}\bar{y}}{\sqrt{\Sigma x^2 - n(\bar{x})^2}\sqrt{\Sigma y^2 - n(\bar{y})^2}}$	$Cov(x,y) = \dfrac{n\Sigma xy - \Sigma x \Sigma y}{n^2}$
$r = \dfrac{n\Sigma xy - \Sigma x \Sigma y}{\sqrt{n\Sigma x^2 - (\Sigma x)^2}\sqrt{n\Sigma y^2 - (\Sigma y)^2}}$	

Ex:-

x :	5	10	15	20	25
y :	7	14	21	28	35

$\Sigma x = 75$

$\Sigma y = 105$

$\Sigma xy = 1925$

$\Sigma x^2 = 1375$

$\Sigma y^2 = 2695$

$$r = \frac{5 \times 1925 - 75 \times 105}{\sqrt{5 \times 1375 - (75)^2}\sqrt{5 \times 2642 - (105)^2}}$$

$$= \frac{1750}{\sqrt{1250}\sqrt{2450}} = 1$$

_/_/_

* Shift of origin

Given - $\Sigma(x-a)^2$, $\Sigma(y-b)^2$

$\Sigma(x-a)(y-b)$ when

$a \neq \bar{x}$

$b \neq \bar{y}$

let $u = x-a$ $\qquad$ $v = y-b$

$$r_{xy} = r_{uv} = \frac{n\Sigma uv - \Sigma u \Sigma v}{\sqrt{n\Sigma u^2 - (\Sigma u)^2}\sqrt{n\Sigma v^2 - (\Sigma v)^2}}$$

$\Sigma u = \Sigma x - na$

$\Sigma v = \Sigma y - nb$

(3) Rank Correlation

- Qualitative data
- Formula for non-repeted Rank

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

where $d = R_x - R_y$

$\boxed{\text{Note :— } \Sigma d = 0}$

_/_/_

Ex:-

X	Y	R_x	R_y	d_i	d^2
10	3	1	3	-2	4
9	5	2	2	0	0
7	9	3	1	2	4

$$R = 1 - \frac{6\Sigma d^2}{n(n^2-1)} = 1 - \frac{6(8)}{3(9-1)}$$

$$= 1 - 2 = -1$$

→ For repeated rank

$$R = 1 - \frac{6(\Sigma d^2 - cf)}{n(n^2-1)}$$

when CF = (Correction Factor)

$$Cf = \Sigma \frac{m^3 - m}{12}$$

If $m = 2$, $Cf = \frac{2^3 - 2}{12} = 0.5$

If $m = 3$, $cf = \frac{3^3 - 3}{12} = 2$

_/_/_

(4) Concurrent Deviation Method

— Quickest Method but cannot take magnitude of Correlation seriously

$$r = \pm\sqrt{\pm\frac{2c-n}{n}}$$

where ; c = no. of Concurrent Deviation

n = Pair of deviation Compared.

Ex: X : 150 + 154 + 160 + 172
Y : 200 − 180 − 170 − 160

$c = 0, \quad n = 3$

$$r = \sqrt{\frac{2(0)-3}{3}} = \sqrt{-1} = -\sqrt{1}$$

$$= -1$$

_/_/__

Imp Note

① n = no. of obs.

② If $c = 0$, $r = -1$

If $c = n$, $r = 1$

If $c = \frac{n}{2}$, $r = 0$

→* Correlation Coefficient is indipendent of Change of Origin and Scale but

If $u = ax \pm c$, $v = by \pm d$

① $r_{uv} = r_{xy}$, If a, b are of Same sign

② $r_{uv} = -r_{xy}$, If a, b are of different Sign

Ex:- $r_{xy} = -0.2$, $u = 3x - 5$

$v = -2y + 100$.

$r_{uv} = 0.2$

_/_/_

$r_{xy} = 0.8$

$3x + 2u + 4 = 0$, $-3y - 8v + 100 = 0$

$2u = -3x - 4$ $\quad$ $v = \frac{-3}{8}y + \frac{100}{8}$

$u = \frac{-3x}{2} - \frac{4}{2}$

$r_{uv} = 0.8$

→ * <u>Imp Note</u>

(1) If $y = a + bx$ where $b > 0$ then $r = +1$

(2) If $y = a + bx$ where $b < 0$ then $r = -1$

__/__/__

→ * Imp Formula

1. Coefficient of determination
$= r^2 \quad / \; 0 \leq r^2 \leq 1$

2. Coefficient of non-determination
$= 1 - r^2$

3. Coefficient of alienation
$= \sqrt{1 - r^2}$

4. Percentage of explained Variance
$= r^2 \times 100$

5. Percentage of unexplained Variance
$= (1 - r^2) \times 100$

6. Standard Error
$= \dfrac{1 - r^2}{\sqrt{n}}$

7. Probable Error
$= \dfrac{1 - r^2}{\sqrt{n}} \times 0.6745$

(8) Probable Error and Standard Error both are use for determining the reliability of Correlation Coefficient

(9) Probable Limit

$$= r - PE,\ r + PE$$

(10) For x, y Correlated Variable.

$$V(x+y) = V(x) + V(y) + 2(\text{cov}(x,y))$$

$$V(x-y) = V(x) + V(y) - 2(\text{cov}(x,y))$$

(11) For x, y Uncorelated Variables

$$V(x \pm y) = V(x) + V(y)$$

__/__/__

→ * Bivariate Data

— Also known as Bivariate frequency distribution on Joint frequency distribution or Two way distribution.

— Ex:

Stat / Math.	0-10	10-20	20-30	Total
0-10	5	3	8	16
10-20	0	1	5	6
20-30.	3	2	8	13
Total	8	6	21	35

— In m rows, n Columns $\underline{n \times m}$ Cells.

— Univariable distribution

① Marginal → 2

② Conditional → m+n

_/_/_

* Ch : 3 B

Regression Analysis

— Regression Analysis

" Estimate or Predict "

— 2 Types of variables

(1) Dependent

(2) Independent

— Correlation is a relative measure
Regression is an absolute measure

Regression Line

Regression Equation of Y on X	Regression Equation of X on Y
→ $y - \bar{y} = b_{yx}(x - \bar{x})$	$x - \bar{x} = b_{xy}(y - \bar{y})$
→ $\hat{y} = a + b_{yx}x$	$\hat{x} = a + b_{xy}y$

→ b_{yx} = Coefficient of Regression of y on x, Slope of Regression of line of y on x	b_{xy} = Coefficient of Regression of x on y Slope of Regression of line of x on y
→ Regression Eq. of y on x is obtain by Scatter diagram by taking sum of squre of vertical distance minimum by the method of least squre	Regression Eq of x on y is obtain by Scatter diagram by taking Sum of squre of ~~te~~ Horizontal distance minimum by the method of least squre

_/_/_

* Regression Coefficient of y on x

$$b_{yx} = \frac{Cov(x,y)}{S_x^2}$$

$$b_{yx} = \frac{\Sigma(x-\bar{x})(y-\bar{y})}{\Sigma(x-\bar{x})^2}$$

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

$$b_{yx} = \frac{n\Sigma xy - \Sigma x \Sigma y}{n\Sigma(x^2) - (\Sigma x)^2}$$

$$b_{yx} = r \cdot \frac{S_y}{S_x}$$

Properties

- Correlation is Symmetric but regression is Not.

- $b_{xy} = r \cdot \frac{S_x}{S_y}$, $b_{yx} = r \cdot \frac{S_y}{S_x}$

$b_{xy} \cdot b_{yx} = r^2$

- $r = \pm \sqrt{b_{xy} \cdot b_{yx}}$

$r_{xy} = r_{yx}$

$b_{xy} \neq b_{yx}$

Ex :- $b_{xy} = 4$, $b_{yx} = 0.14$

$r = \sqrt{4 \times 0.14}$

$r = 0.7483$

- r, b_{xy}, b_{yx} are of Same Sign

- $-1 \leq r \leq 1$ | $-\infty < b_{xy} < \infty$

$0 \leq r^2 \leq 1$ | $-\infty < b_{yx} < \infty$

— If one regression Coefficient is Grater then 1 then other must be less the 1 Since $0 \le r^2 \le 1$

— Note

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

$$x = \bar{x}, \quad y = \bar{y}$$

$(\bar{x}, \bar{y})$ is a Common Solution i.e. Point of intersection of ~~tow~~ two regression line There for, Required regression Equation will be Satisfied by $(\bar{x}, \bar{y})$

— Note :-

In regression analysis, the difference between the observed Value and the estimated Value is known as Error or Residue.

Standard Error of estimate of x on y = $S_x\sqrt{1-r^2}$

Standard Error of estimate of y on x = $S_y\sqrt{1-r^2}$

If $r = \pm 1$ then $SE = 0$

— Regression Coefficient is not affected by change of origin by affected by change of Scale

If $u = ax + b$, $v = cy + d$

then $b_{uv} = b_{xy} \times \frac{a}{c}$

$b_{vu} = b_{yx} \times \frac{c}{a}$

Ex :- $u = 3x + 10$, $v = -4y - 20$

$b_{xy} = 0.8$

$b_{uv} = b_{xy} \times \frac{3}{-4} = 0.8 \times \frac{3}{-4}$

$= -0.6$

_/_/_

(i) Reg. Eq. are

$\underbrace{x + 2y = 5}$ $\underbrace{2x + 3y = 8}$

y on x x on y

$b_{yx} = \frac{-1}{2}$ $b_{xy} = \frac{-3}{2}$

$r^2 = b_{xy} \times b_{yx} = \frac{-1}{2} \times \frac{-3}{2}$

$= \frac{3}{4} = 0.75 \leq 1$

$r^2 = \frac{3}{4}$

$r = -\sqrt{\frac{3}{4}}$

$\underbrace{x + 2y = 5}$ $\underbrace{2x + 3y = 8}$

x on y y on x

$b_{xy} = -2$ $b_{yx} = \frac{-2}{3}$

$r^2 = b_{xy} \times b_{yx} = -2 \times \frac{-2}{3} = \frac{4}{3} > 1$

wrong

Correct : $b_{yx} = \frac{-1}{2}$, $b_{xy} = \frac{-3}{2}$

$r^2 = \frac{3}{4}$, $r = -\sqrt{\frac{3}{4}}$

__/__/__

* Steps to Identify Regression Equation

Step: 1 Consider any one equation of x on y and other of y on x and find b_{xy} and b_{yx}.

Step: 2 (1) If b_{xy}, b_{yx} are of different Sign then data is in-Correct

(2) If of same sign then find $r^2 = b_{xy} \cdot b_{yx}$.

Step: 3 (1) If $r^2 \leq 1$ then assumption is Correct

(2) If $r^2 > 1$ then assumption is in-correct.

– To get Correct value $b_{xy} = \frac{1}{b_{yx}(\text{wrong})}$

$$b_{yx} = \frac{1}{b_{xy}(\text{wrong})}$$

$$r^2 = \frac{1}{r^2(\text{wrong})}$$

_/_/_

Note :-

(1) If $r = 0$ then $cov(x,y) = 0$, $b_{yx} = b_{xy} = 0$

--> Regression equation are
x on y → $x = \bar{x}$ and y on x => $y = \bar{y}$

--> Regression line becomes Perpendicular. (90°, ⊥)

(2) If $r = \pm 1$ then

(0°)

--> Regression Line are identical

$r = 0$	$r = \pm 1$
Angle = 90°	Angle = 0°

(3) In Regression analyses we get a functional relation between two Variable using average.

_/_/_

* ch :- 4

Index Numbers

— Index Number –

- Special type of Avg.
- Expressed as ratio
- Calculated as Percentage
- Unit Free (Pure Number)

* Type of Index Numbers

1. Price Index Number
2. Quantity Index Number (Volume)
3. Value Index Number.
   (Value = Price x Quantity)

__/__/__

* Notation :

0 — Base year

1 — Current year

I_{01} — Index Number for year 1 with base as 0

I_{10} — Index Number for year 0 with base as 1

I_{11} — 100.

* Price Index Number

(1) Price Relative :- PR :-

$$= \frac{P_1}{P_0} \times 100$$

__/__/__

(2) Method of Aggregate :-

(i) Simple Aggregate of P.I.N.

$$I_{01} = \frac{\Sigma P_1}{\Sigma P_0} \times 100$$

- Unweighted
- with Unit rest all are Unit free

(ii) Weighted Aggregate.

$$I_{01} = \frac{\Sigma P_1 w}{\Sigma P_0 w} \times 100$$

- weight is most Important in I.N

_/_/_

- Laspeyre's P.I.N. = $L_{01} = \frac{\Sigma P_1 q_0}{\Sigma P_0 q_0} \times 100$

  $- w = q_0$

- Paasche's P.I.N. = $P_{01} = \frac{\Sigma P_1 q_1}{\Sigma P_0 q_1} \times 100$

  $- w = q_1$

- Fisher's P.I.N. = $F_{01} = \sqrt{L_{01} \times P_{01}}$

  - Best Index Number
  - F_{01} is G.M. of L_{01} and P_{01}

- Bowley's P.I.N. = $B_{01} = \frac{L_{01} + P_{01}}{2}$

  - B_{01} is A.M of L_{01} and P_{01}

- Marshal + Edgeworth P.I.N.

$$= \frac{\Sigma P_1 \left(\frac{q_0 + q_1}{\cancel{2}}\right)}{\Sigma P_0 \left(\frac{q_0 + q_1}{\cancel{2}}\right)} \times 100$$

$$= \frac{\Sigma P_1 q_0 + \Sigma P_1 q_1}{\Sigma P_0 q_0 + \Sigma P_0 q_1} \times 100$$

- Best approximation to Fisher's I.N.

— Note

G.M. is best Average in the Canstruction of I.N. but Practically we use A.M., because G.M. is difficult to ComPute

(3) Relative Method

(i) Simple A.M of P.R

$$I = \frac{\Sigma P.R}{n}$$

(ii) Weighted A.M. of P.R

$$I = \frac{\Sigma P.R \times w}{\Sigma w}$$

(iii) Simple G.M. of P.R

$$I = (\Pi P.R)^{1/n}$$

$$\log I = \frac{1}{n}\left[\Sigma \log \frac{P_1}{P_0} + 2\right]$$

* (Π = Product)

__/__/__

Cost of Living Index Number (CLIN)

- Wholesale P.I.N.,
  Consumer P.I.N.,
  General P.I.N.,

$$CLIN = \frac{\Sigma IW}{\Sigma W} \quad \text{or} \quad \frac{\Sigma P_1 q_0}{\Sigma P_0 q_0} \times 100$$

— Based on Laspeyre's Method

Ex Dearness Allowance (DA)

	Salary	CLIN
2024	50,000	150
2029	55,000	180

150 ⟶ 50,000
180 ⟶ ? = 60,000
− 55,000
5000
(D.A)

__/__/__

* Real Income (Deflation) $= \dfrac{\text{Income}}{\text{CLIN}} \times 100$

Ex:

	Salary	CLIN
2000	40,000	150
1990	25,000	100

Real Income $= \dfrac{40{,}000}{150} \times 100$

$= 26{,}666.67$

* Purchasing Power $= \dfrac{1}{\text{CLIN}} \times 100$

** For Quantity IN, In PIN formula,

$P \leftrightarrow q$

__/__/__

— Base Shift I/V = $\frac{\text{Current Yr I/V}}{\text{New Base yr I/V}} \times 100$

— Fixed and Chain Base Method

- Fixed base Method : Base Year is Fixed

$$FBI/V = \frac{\text{Current yr Price}}{\text{Base yr Price}} \times 100$$

- Chain base Method : Base Year is alredy Previous yr.

CBI/V (Link Related) $= \frac{\text{Current yr Price}}{\text{Previous yr Price}} \times 100$

_/_/_

* FBIN → CBIN

Yr	0	1	2	3
FBIN	80	120	180	200
CBIN	100	150	150	111.11

Note:- To find Chain Base Method take 1^{st} yr IN = 100

* CBIN → FBIN

Yr	0	1	2	3
CBIN	60	150	180	250
FBIN	60	90	162	405

Note:- (1) for 1^{st} yr FBIN = CBIN for that yr.

(2)

$$FBIN = \frac{\text{Previous yr FBIN} \times \text{Current yr CBIN}}{100}$$

* CBIN → CIN

yr	0	1	2	3
CBIN	60	150	180	250
CIN	100	150	270	675

Note:- ① 1^{st} yr CIN = 100

②
$$CIN = \frac{\text{Previous yr CIN} \times \text{Current yr CBIN}}{100}$$

→ **Splicing**

Ex:-

	IN Base 2001	IN Base 2003
2001	100	66.67
2002	120	
2003	150	100
2004	210.	140
2005		180.

→ * **Value Index Number**

$$V = \frac{\Sigma P_1 q_1}{\Sigma P_0 q_0} \times 100$$

Test	Simple Aggregate	L	P	F	B	M.E	weight GM of Price relative	Simple GM of Price relative
Unit Test	✗	✓	✓	✓	✓	✓	✓	✓
Time - reversal test	✓	✗	✗	✓	✗	✓	✓	✓
Factor - reversal test	✗	✗	✗	✓	✗	✗	✗	✗
Circular test	✓	✗	✗	✗	✗	✗	✗	✓

__/__/__

* ch :- 5 A

Probability

- $P(A) = \frac{m}{n}$

* Coin

- 2 Coin (Trick $\frac{11^2}{2^2}$)

$U = \{HH, HT, TH, TT\}$

No. of Head/ Tail	0	1	2
P	$\frac{1}{4}$	$\frac{2}{4}$	$\frac{1}{4}$

- 3 Coin (Trick $\frac{11^3}{2^3}$)

No. of Head/ Tail	0	1	2	3
P	$\frac{1}{8}$	$\frac{3}{8}$	$\frac{3}{8}$	$\frac{1}{8}$

— 4 Coin $\left(\text{Trick } \frac{11^4}{2^4}\right)$

No. of Head / Tail	0	1	2	3	4
P	$\frac{1}{16}$	$\frac{4}{16}$	$\frac{6}{16}$	$\frac{4}{16}$	$\frac{1}{16}$

→ * Cards

52 Cards

Black (26) — (26) Red

♣ Club ♠ SPade — Dimond ♦ · Heart ♥

A, 2, 3, 4, 5, 6, 7, 8, 9, 10, J, Q, k

Face Cards = (J, Q, k) = 3 × 4 = 12

Honour Cards = (A, J, Q, k) = 4 × 4 = 16

_/_/_

→* Dice

— 2 Dice

(1,1), (1,2), (1,3), (1,4), (1,5), (1,6)
(2,1), (2,2), (2,3), (2,4), (2,5), (2,6)
(3,1), (3,2), (3,3), (3,4), (3,5), (3,6)
(4,1), (4,2), (4,3), (4,4), (4,5), (4,6)
(5,1), (5,2), (5,3), (5,4), (5,5), (5,6)
(6,1), (6,2), (6,3), (6,4), (6,5), (6,6)

1

Sum of no.	2	3	4	5	6	7	8	9	10	11	12
times.	1	2	3	4	5	6	5	4	3	2	1

2

F = S	F > S	F < S
$\frac{6}{36}$	$\frac{15}{36}$	$\frac{15}{36}$

3 Product is 12 = $\frac{4}{36}$

(6,2) (2,6) (4,3) (3,4)

__/__/__

→ * Leap Year

366 days = 52 weeks + 2 odd days

M, T
T, W
W, Th
Th, F — 7 ways
F, Sat
Sat, Sun
Sun, M

(1) P(53 Sundays) = 2/7

(2) P(53 Sundays ~~and~~ or 53 Saterdays) = 3/7

(3) P(53 Sundays and 53 saterdays) = 1/2

(4) P(53 Sundays or 53 Fridays) = 4/7

(5) P(53 Sundays and 53 Fridays) = 0.

__/__/__

* Rule of addition

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

$P(A \cup B)$: or, Atleast one

$P(A \cap B)$: and, Both

Ex :- From 1 to 200

P(No. is divisible by 6 or 8)

$P(6 \cup 8)$

$= P(6) + P(8) - P(6 \cap 8)$

$= 33 + 25 - 8$

$= \frac{50}{200} = \frac{1}{4}.$

* Formula

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

2. $P(A') = P(A^c) = 1 - P(A)$

3. $P(A - B) = P(A \cap B') = P(A) - P(A \cap B)$

4. $P(B - A) = P(A' \cap B) = P(B) - P(A \cap B)$

5. $P(A' \cap B') = 1 - P(A \cup B)$

6. $P(A' \cup B') = 1 - P(A \cap B)$

7. $P(A' \cup B) = 1 - P(A \cap B')$

8. $P(A \cup B') = 1 - P(A' \cap B)$

__/__/__

— Odd in the favor of A is
$p:q$ then
A ⁄ \ A'

$$P(A) = \frac{p}{p+q}, \quad P(A') = \frac{q}{p+q}$$

Ex:- Odd in the favor of A is
2 : 3 then

$$P(A) = \frac{2}{5}, \quad P(A') = \frac{3}{5}$$

A' , A
— odd Against A is $m:n$ then

$$P(A) = \frac{n}{m+n}, \quad P(A') = \frac{m}{m+n}$$

Ex:- Odd against A is 3 : 5 then

$$P(A) = \frac{5}{8}, \quad P(A') = \frac{3}{8}$$

_/_/_

* Conditional Probability

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

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

* Rule of Compound Probability

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

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

- Ex :- $P(A) = 0.7$, $P(B) = 0.4$, $P(A \cap B) = 0.3$

$$P(A \cup B) = P(A) + P(B) - P(A \cap B)$$
$$= 0.7 + 0.4 - 0.3$$
$$= 0.8$$

(1) $P(A \cap B') = P(A) - P(A \cap B) = 0.7 - 0.3$
$= 0.4$

(2) $P(A' \cup B) = 1 - P(A \cap B')$
$= 1 - 0.4 = 0.6$

_/_/_

(3) $P(A' \cap B) = P(B) - P(A \cap B)$

$= 0.4 - 0.3$

$= 0.1$

(4) $P(A/B) = \frac{P(A \cap B)}{P(B)} = \frac{0.3}{0.4} = \frac{3}{4}$

(5) $P(A'/B') = \frac{P(A' \cap B')}{P(B')}$

$= \frac{1 - P(A \cup B)}{1 - P(B)} = \frac{1 - 0.8}{1 - 0.4}$

$= \frac{0.2}{0.6} = \frac{1}{3}$

(6) $P(A/B') = \frac{P(A \cap B')}{P(B')}$

$= \frac{0.4}{0.6} = \frac{2}{3}$

_/_/_

* A and B are Independent events

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

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

$P(A \cap B) = P(A) \times P(B) \rightarrow$ Rule of Multiplication for Independent

— **Note :-** If A and B are Independent then.

(1) A and B' are Independent

(2) A' and B are Independent

(3) A' and B' are Independent

- Three events A, B, C are mutually Independent if,

$P(A \cap B) = P(A) \cdot P(B)$

$P(A \cap C) = P(A) \cdot P(C)$

$P(B \cap C) = P(B) \cdot P(C)$

$P(A \cap B \cap C) = P(A) \cdot P(B) \cdot P(C)$

= No. of Condition to check for 3 events to be mutually Independent is 4

* No. of Condition to check for n events to be mutually - Independent =

$$2^n - 1 - n$$

- Notation

$A + B = A \cup B$

$AB = A \cap B$

/ /

→ * Theory :-

* Types :-

1) Subjective :- Baised,
Used in managment

2) Objective :- Unbaised

* Types of Events :-

1) Elementary Event.

Event which cannot be divided into smaller event.

2) Compound (composite) Events :-

Made of two or more elementory events

3) Sure Events :-

$P(A) = 1$; $P(A) = U$

__/__/__

4) Impossible Events :-

$P(A) = 0$; $P(A) = \phi$

5) Chance Events :-

$0 < P(A) < 1$

6) Equally Likely Events :-

A and B are equally Likely events if $P(A) = P(B)$

7) Mutually Exclusive Events :-

If an event A and B are mutually Exclusive then,

1) $P(A \cup B) = P(A) + P(B)$

2) $P(A - B) = P(A)$

3) $P(B - A) = P(B)$

4) $P(A \cup B \cup C) = P(A) + P(B) + P(C)$

8) Collectively Exhaustive Events

$P(A \cup B) = 1$

$P(A \cup B \cup C) = 1$

- Imp Notes:-

If A and B are two events with ~~not~~ non-Zero Probability.

Mutually Exclusive and Independent

$P(A \cap B) = 0$, $P(A \cap B) = P(A) \times P(B)$

$0 = P(A) \times P(B)$

$P(A) = 0$

$P(B) = 0$

Contradiction

Two events with Non-Zero Probability cannot be Simultaneously mutually exclusive and independent both.

→ * Definition of Probability

Classical or mathematical	Empirical or Statistical	Axiomatic
$= \dfrac{m}{n}$	$= \lim\limits_{n \to \infty} \dfrac{m}{n}$	No formula But we have Rule of set
		Eg. $P(A \cup B) = P(A) + P(B) - P(A \cap B)$

→ * Limitation of mathematical/ Classical definition.

- All the outcome must be known in advance.
- If $n \to \infty$ then we cannot use this definition.
- Elementory events are equally likely

__/__/__

* Ch :- 5 B

Random Variablle

Ex : 3 Coins $\left(\frac{113}{2^3}\right)$

x = No. of H	$P = f(x)$
0	1/8
1	3/8
2	3/8
3	1/8

x = Random Variable

$f(x)$ = Probability Function
Frequency Function

— / /

Type of Variable

Discrete variable	Continuous variable
Whole Numbers	Real numbers in Interval.
Finite values	Infinite values
Probability function is known as - Probability mass function.	Probability function is known as - Probability density function.
Condition of PMF (1) $f(x) \geq 0$ (2) $\Sigma f(x) = 1$	Condition of PDF (1) $f(x) \geq 0$ (2) $\int f(x)\,dx = 1$

_/_/_

* **Expected value** $= E(x) = \bar{x}$

- It is also known as mean

- **Formula :-**

$$E(x) = x_1 P_1 + x_2 P_2 + \cdots + x_n P_n$$
$$= \Sigma x_i P_i$$

- **Note :-** (1) If all Probability are equal then,

$E(x) =$ Avg. of x

(2) If all Probability are equal and x valus in A.P then

$$E(x) = \frac{a + l}{2}$$

Ex :- **Dice**

x	$f(x)$
1	1/6
2	1/6
3	1/6
4	1/6
5	1/6
6	1/6

_/_/_

$$E(x) = \frac{1}{6} + \frac{2}{6} + \frac{3}{6} + \frac{4}{6} + \frac{5}{6} + \frac{6}{6}$$

$$= \frac{1+2+3+4+5+6}{6} = \frac{21}{6} = 3.5$$

Trick

$$E(x) = \frac{1+6}{2} = 3.5$$

Ex:- Game fees = ₹6

$6 \to$ ₹ 30

$x - P$

$30 - 1/6$

$E(x) = 30 \times 1/6 = 5 - 6 = -1$

— Note:- (1) For Game $\Rightarrow f(x) = a$

(2) For Game fees Probability first find $E(x)$ and at the end subtract game fees to get actual value of $E(x)$.

_/_/_

Ex:- A box Contains 12 lamps of which 5 are defective, A man Select 3 lamps at random then what is the expected No. of defective lamps

X	P
0	${}^7C_3 = 35/220$
1	${}^5C_2 \times {}^7C_3 = 105/220$
2	${}^5C_2 \times {}^7C_1 = 70/220$
3	${}^5C_3 = 10/220$

$E(x) = 1.25$

Trick

$12 \rightarrow 5$
$3 \rightarrow ?$

$= 1.25$

_/_/_

— Properties of $E(x)$

(1) $E(c) = c$

(2) $E(ax \pm by \pm c) = aE(x) \pm bE(y) \pm c$

$E(x) = 3$, $E(y) = 5$

$E(2x + 3y - 10)$

$= 2(E(x)) + 3(E(y)) - 10$

$= 2(3) + 3(5) - 10 = 6 + 15 - 10.$

$= 11$

(3) If x and y are Independent then,

$$E(x.y) = E(x).E(y)$$

(4) $E(x - \bar{x}) = 0$

$= E(x) - E(\bar{x}) = 0$

$= \bar{x} - \bar{x} = 0$

— Note For Dice

(1) For Sum

$E(\text{one die}) = 3.5$

$E(\text{Sum of two dice}) = E(x+y)$
$= E(x) + E(y) = 3.5 + 3.5 = 2 \times 3.5$

$E(\text{Sum of } n \text{ dice}) = n \times 3.5$

(2) E (Product of two dice)
$E(xy) = E(x) . E(y)$
$= 3.5 \times 3.5$
$= 12.25$

_/_/_

* Standard deviation and Variance

$$\sigma^2 = V(x) = \frac{\Sigma (x-\bar{x})^2}{n} = \text{Avg. of } (x-\bar{x})^2$$

$$V(x) = E((x-\bar{x})^2)$$

$$V(x) = E(x^2) - (E(x)^2)$$

$$\sigma = SD = \sqrt{V(x)}$$
$$= \sqrt{E(x^2) - (E(x)^2)}$$

$$V = SD^2.$$

— Properties of SD and V(x)

(1) $V(c) = 0$, $SD(c) = 0$.

(2) $SD(ax \pm by \pm c) = |a| SD(x) + |b| SD(y)$

$SD_x = 3$, $SD_y = 7$

$SD(-2x - 3y + 4) = 2 \times 3 + 3 \times 7 = 27$

(3). $V(ax \pm by \pm c) = a^2 V(x) + b^2 V(y)$

$V(5x - 2y + 100) = 25 V(x) + 9 V(y)$

$V(x) = 4$, $V(y) = 9$ $= 25 \times 4 + 9 \times 9$

$= 181$

___/___/___

Example:-

x	P	x^2
3	0.2	9
9	0.6	81
10	0.2	100

$E(x) = 8$

$$E(2x+1) = 2E(x)+1$$
$$= 2(8)+1$$
$$= 17.$$

$E(x^2) = 70.4$

$$V(x) = E(x^2) - (E(x)^2)$$
$$= 70.4 - 64$$
$$= 6.4$$

$$SD = \sqrt{6.4}$$
$$= 2.53$$

_/_/_

— Uniform Distribution (For Discrete Variable)

$$X = 1, 2, 3, ----, n$$

$$\text{P.M.F} = f(x) = \frac{1}{n}$$

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

$$V(x) = \frac{n^2-1}{12}$$

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

__/__/__

* Ch :- 6

Theoretical Distribution

- Binomial Distribution

→ Bernoulli Trial

(1) Exactly two possible group of Outcome "Success" and "failure"

(2) Probability of Success in independent of Previous Success

→ Binomial is Bernoulli Trial - repeated n times.

∴ Binomial is an extension of Bernoulli

→ n = no. of times experiment repeated

$X = 0, 1, 2, 3, --- n$

_/_/_

→ X is discrete variable

$X \sim B(n, P)$
↓
Follows

→ P.m.f.

$f(x) = {}^nC_x \, p^x q^{n-x}$

x = no. of Success

P = Probability of Success

q = Probability of Failure

$P + q = 1$

Sum of Power of P, q = n

→ Two Parameters n and P

→ Mean = nP

→ Variance = nPq

mean > variance

→ SD = $\sqrt{nPq}$

_/_/_

→ Maximum Value of Variance $= n \times \frac{1}{2} \times \frac{1}{2} = \frac{n}{4}$

$(p = q = \frac{1}{2})$

Imp $\quad q = \frac{V}{m} \qquad n = \frac{m}{p}$

Ex :- mean = 2 $\qquad V(x) = 1.2$

$np = 2 \qquad npq = 1.2$

$2q = 1.2$

$\therefore n = \frac{2}{0.4} = 5 \qquad q = \frac{1.2}{2} = 0.6$

$P = 0.4$

Ex :- mean = 48 $\qquad V(x) = 19.2$

$q = \frac{19.2}{48} = 0.4$

$n = \frac{48}{0.6} = 80.$

$P = 0.6$

__/__/__

→ * Mode = $(n+1)P$

Uni-Modal	Bi-modal
$(n+1)P$ is any decimal number	$(n+1)P$ is an integer
Mode = Integ Part of $(n+1)P$	Mode = $(n+1)P$, $(n+1)P+1$
Ex :- $n=6$, $P=\frac{1}{4}$	Ex :- $n=7$ $P=\frac{1}{4}$

$(n+1)P = 7 \times \frac{1}{4}$

$= 1.75$

Mode = 1

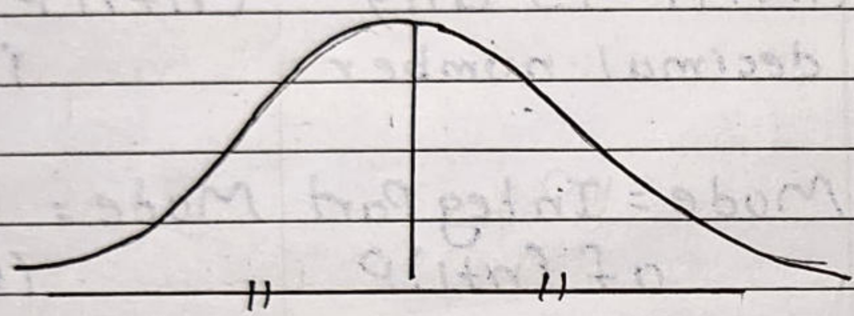
$(n+1)P = 8 \times \frac{1}{4} = 2$

mode = 2, 1

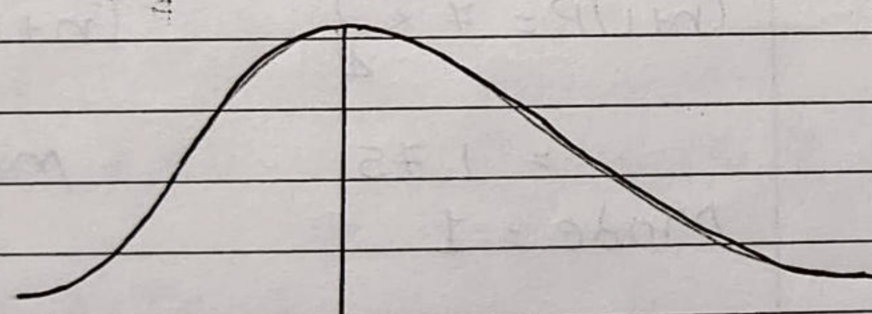
_/_/_

→ Symmetric and Asymmetric

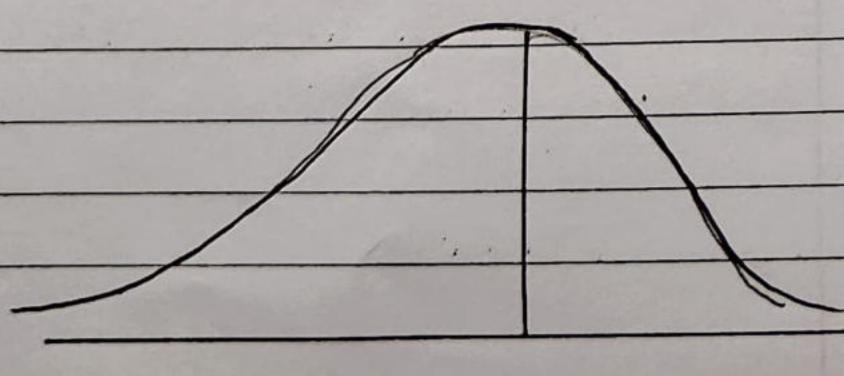
- Symmetric $P = q = 0.5$
(mean = median = mode)



- Positive Skewed $P < q$



- Negative Skewed $P > q$



_/_/_

→ Addition Property

If x, y are Independent variable and
$x \sim B(n_1, P)$, $y \sim B(n_2, P)$
then $x + y \sim B(n_1 + n_2, P)$

Ex :- $x \sim B(10, \frac{1}{2})$

$y \sim B(12, \frac{1}{2})$

$x + y \sim B(22 ; \frac{1}{2})$

— Trick

For Problem of binomial with
$P = q = 1/2$ then

$$f(x) = \frac{{}^{n}C_{x}}{2^n}$$

— Note :-

$x \sim B(n, P)$ is also written as $(p+q)^n$ or

$(q+p)^n$

__/__/__

Bernolli Trial "Success" or "failure"

$\downarrow$ n times

Discrete

Binomial Distribution

$\downarrow$ $n \to \infty$, $P \to 0$, $q \to 1$

Discrete

Poission Distribution

$\downarrow$ Continuous variable $P \to 0.5 \leftarrow q$

Normal Distribution

Note :-

Parameter $\longrightarrow$ Population

Ex Population Mean
μ

Statistics $\xrightarrow[\text{for}]{\text{use}}$ Sample

Ex Sample Mean
$\bar{x}$

__/__/__

— Poissian Distribution

(Rare events / Improbable events distribution)

$n \to \infty, \ P \to 0, \ q \to 1$

$m = \text{mean} = np = \lambda$

$x = 0, 1, 2, 3, ---$

$x \sim P(m)$

→ Discrete Variable

→ P.m.f.

$$f(x) = \frac{e^{-m}\, m^{x}}{x!}$$

$e = 2.7183$

→ Variance = mean = Parameter = m

→ $SD = \sqrt{m}$

→ Mode = m.

Mode = m

Uni-Modal	Bi-modal
m is not an integer	m is an integer
Mode = Integer Part of m	Mode = m, m-1
Ex: m = 3.85 Mode = 3	**Ex:** m = 4 Mode = 4, 3

→ $P \to 0$, Always $P < q$

Always Positive Skewed

→ Addition Property

If x and y are Independent

$x \sim P(m_1)$, $y \sim P(m_2)$ then

$$x + y \sim P(m_1 + m_2)$$

_/_/_

Ex :- $n = 400$

1% defective

$P = 0.01$

$m = n \times P = 400 \times 0.01 = 4$

(Given $e^{-4} = 0.0183$)

(1) $P(\text{none defective}) = P(0)$

$$= \frac{e^{-4}\, 4^0}{0!}$$

$$= e^{-4} = 0.0183$$

(2) $P(1) = \dfrac{e^{-4}\, 4^1}{1!} = 4 \times e^{-4}$

$$= 4 \times 0.0183$$

$$= 0.0732$$

(3) $P(2) = \dfrac{e^{-4}\, 4^2}{2!} = 0.1464$

Note :-

(1) $P(0) = e^{-m}$

(2) $P(1) = m \cdot e^{-m}$

— Ex: $m = 4$, $e^{-4} = 0.0183$

$P(\text{At least } 4)$

$= 1 - P(0, 1, 2, 3)$

$= 1 - \left[e^{-m} + e^{-m} m + e^{-m} \frac{m^2}{2} + e^{-m} \frac{m^3}{6}\right]$

$= 1 - e^{-m} \left[1 + m + \frac{m^2}{2} + \frac{m^3}{6}\right]$

$= 0.5669$

__/__/__

Note :-

Condition to use Poisson Distribution

(1) The Probability of having Success in this time interval is independent of time as well as earlier Success

(2) Probability of having more than one Success is very Low.

(3) Probability of Success in time interval is directly Proportional to time interval length.

Normal Distribution

→ Continuous Variable

→ $x \sim N(\underset{\text{Mean}}{\mu}, \underset{\text{S.D}}{\sigma})$ or $x \sim N(\mu, \underset{\text{Variance}}{\sigma^2})$

→ P.d.f.

$$f(x) = \frac{1}{\sigma\sqrt{2\pi}} e^{-\frac{(x-\mu)^2}{2\sigma^2}}$$

→ $Z = \dfrac{x-\mu}{\sigma}$

$\bar{Z} = \dfrac{\bar{x}-\mu}{\sigma}$ $\qquad \sigma_Z = \dfrac{\sigma_x}{\sigma}$

$= \dfrac{\mu-\mu}{\sigma}$ $\qquad = \dfrac{\sigma}{\sigma}$

$\bar{Z} = 0$ $\qquad \sigma_Z = 1$

$Z \sim N(0,1)$

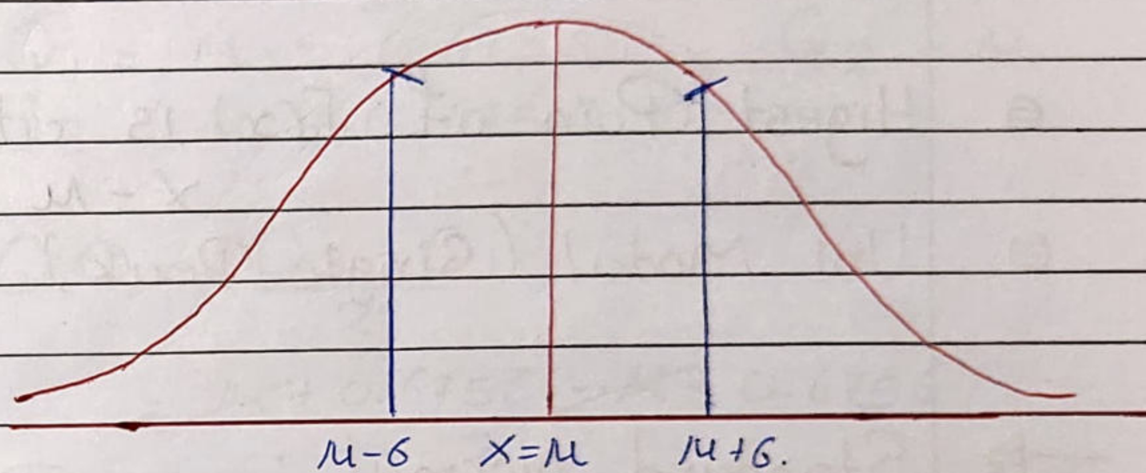
Z is a Standard Normal variable

P.d.f.

$$f(Z) = \frac{1}{\sqrt{2\pi}} e^{-\frac{Z^2}{2}}$$

_/_/_

→ Normal Curve $X \sim N(\mu, \sigma)$



$-\infty < x < \infty$

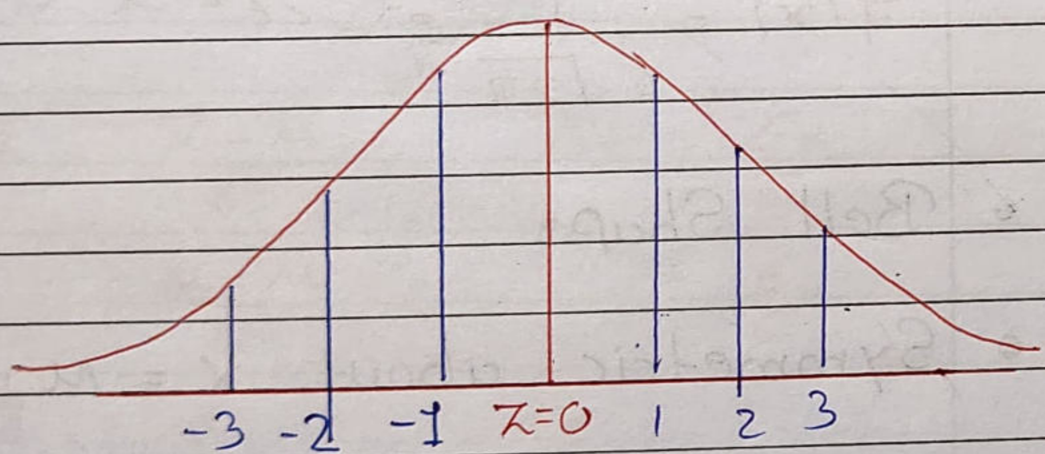
$$f(x) = \frac{1}{\sigma\sqrt{2\pi}} e^{-\frac{(x-\mu)^2}{2\sigma^2}}$$

- Bell Shape
- Symmetric about $x = \mu$
- Total area under the Curve = 1
- Asymptotic ⟹ will never intersect x-axis
- Point of Inflection

$\mu-\sigma$, $\mu+\sigma$.

- Mean = Median = Mode = μ
- Higest Poin of $f(x)$ is at $x = \mu$
- Uni Modal (Single Peak)

→ Standard Normal :



$$f(z) = \frac{1}{\sqrt{2\pi}} e^{-\frac{z^2}{2}}$$

- Symmetric about $Z = 0$
- Point of Inflection
  -1, 1
- Maximum Value at $Z = 0$.

_/_/_

Formula:

$Q_1 = \mu - 0.675\sigma \qquad Q_2 = \mu$

$Q_3 = \mu + 0.675\sigma$

$$Q.D = \frac{Q_3 - Q_1}{2}$$

$$= \frac{\cancel{\mu} + 0.675\sigma \cancel{- \mu} + 0.675\sigma}{2}$$

$$Q.D = 0.675\sigma \approx \frac{2}{3}\sigma$$

$$MD = \frac{4}{5}\sigma$$

$QD : MD : SD = 10 : 12 : 15$

$$15 \times \left(\frac{2}{3}\cancel{\sigma} : \frac{4}{5}\cancel{\sigma} : \cancel{\sigma} \right) = 10 : 12 : 15$$

Ex:- $Q.D = 9$, $MD = ?$

Q.D	:	MD
10	:	12
9	:	

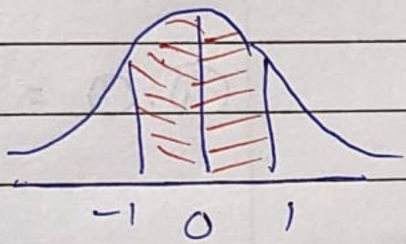
$= 10.8$

_/_/_

Q. In normal distribution, what is the ratio of M.D and S.D?

→ $\frac{MD}{SD} = \frac{12}{15} = 0.8$

→ **I.m.P Probability :**



$P(0 < Z < 1) = 0.3413$

$P(-1 < Z < 1) = 0.6826.$

$P(0 < Z < 2) = 0.4772$

$P(-2 < Z < 2) = 0.9544$

$P(0 < Z < 3) = 0.4987$

$P(-3 < Z < 3) = 0.9974$

Except

0.0026

= 0.26%

_/_/__

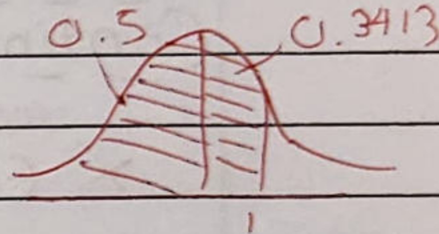
Cumulative Distribution function (C.D.F)

$$F(a) = \phi(a) = P(-\infty < z < a)$$

Ex:- (1) $F(1) = P(Z < 1)$

$= 0.5 + 0.3413$

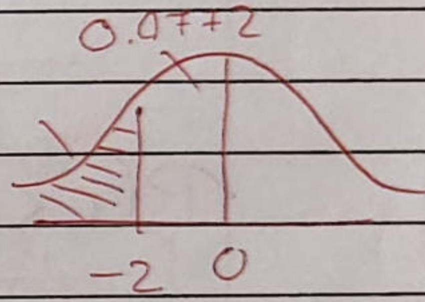
$= 0.8413$



(2) $F(-2) = P(Z < -2)$

$= 0.5 - 0.4772$

$= 0.0228$



Addition Property

If x, y are Independent variable

$$x \sim N(\mu_1, \sigma_1), \; y \sim N(\mu_2, \sigma_2)$$

then $x + y \sim N(\mu_1 + \mu_2, \sqrt{\sigma_1^2 + \sigma_2^2})$

Ex:- $x \sim N(10, 3)$, $y \sim N(12, 4)$

then $x + y \sim N(22, 5)$

→ Normal distribuition is also known as - - Gaussian distribution.

* **Uniform Distribution for Continuous Variable**

$X \in [a, b]$ P.d.f

$$f(x) = \frac{1}{b-a}$$

$$P(c < x < d) = \frac{d-c}{b-a}$$

— Also known as Ractangular distribution.

* Two Methods of Fitting Normal

(1) Ordinate Method

(2) Area Method

→ For Continuous Variable Probability to take any fixed value is Zero. (Since $\frac{1}{\infty} = 0$.