

What do we mean by correlation?

It shows the degree of linear association between two variables
Nifty $\uparrow$ Sensex $\uparrow$
Nifty $\downarrow$ Sensex $\downarrow$

Why do we study?

"Stock Market involve wide application of Correlation"

CH-17

CORRELATION & REGRESSION

[Avg. w. 5 marks]

Correlation

Introduction

Correlation Coefficient (r)

x	y	x	y	x	y
$\uparrow$	$\uparrow$	$\uparrow$	$\downarrow$	$\uparrow$	$\times$
$\downarrow$	$\downarrow$	$\downarrow$	$\uparrow$	$\uparrow$	$\times$
positive correlation		Negative Correlation		No Correlation	
$0 < r \leq 1$		$-1 \leq r < 0$		$r = 0$	

$$-1 \leq r \leq 1$$

$r = -1$	Perfectly negative Correlation
$-1 < r < 0$	Negative Correlation
$r = 0$	No Correlation
$0 < r < 1$	positive Correlation
$r = +1$	Perfectly positive Correlation

Straight Line Inproportional always

TYPE I: $ax + by + c = 0$
 $r_{xy} = -1$ or $+1$

$$2x + y = 5$$

$$+ 2x = -y + 5 \quad \text{Ans } -1$$

If sign of two variables are same then ans. is $+1$ or opposite then ans. is -1

$$-2x + 3y = 10$$

$$-2x = -3y + 10 \quad \text{Ans } +1$$

What do we mean by correlation?

It shows the degree of linear association between two variables
 Nifty $\uparrow$ Sensex $\uparrow$
 Nifty $\downarrow$ Sensex $\downarrow$

Why do we study?

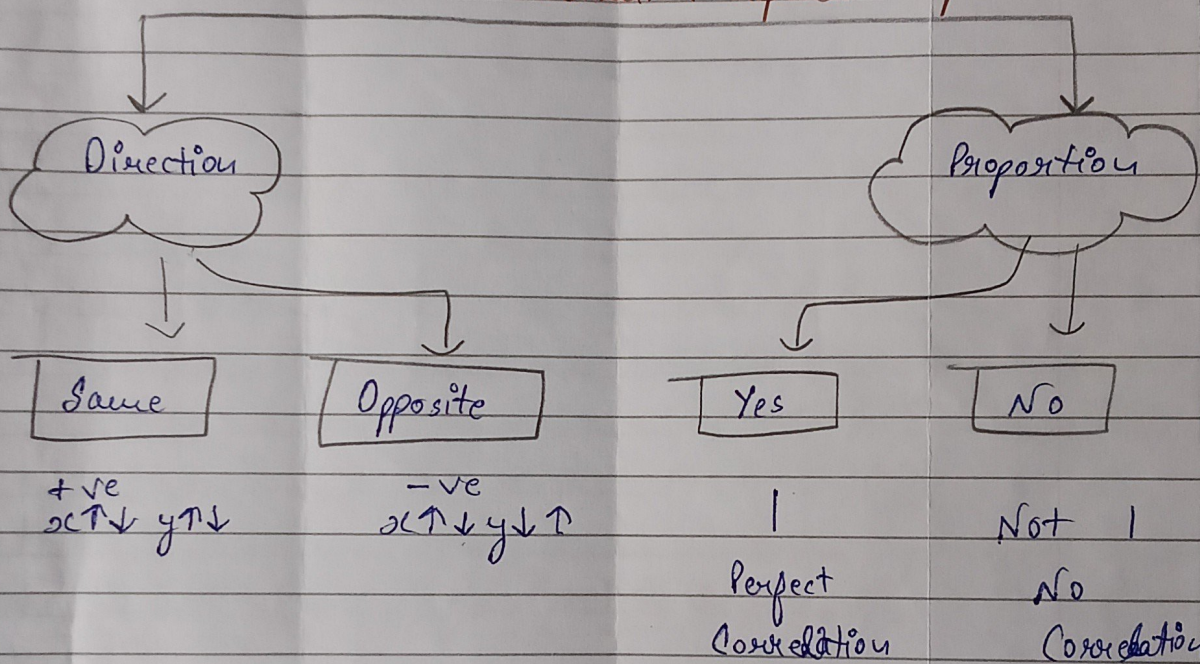
"Stock Market involve wide application of Correlation"

CH-17

CORRELATION & REGRESSION

[Avg. w. 5 marks]

Correlation depends upon



x
x
x
Correlation
 $r=0$

$r = -1$ Perfectly negative Correlation
 $-1 < r < 0$ Negative Correlation
 $r = 0$ No Correlation
 $0 < r < 1$ positive Correlation
 $r = +1$ Perfectly positive Correlation

Straight Line Inproportional always

TYPE I: $ax + by + c = 0$
 $rx + y = -1$ or $+1$

$$2x + y = 5$$

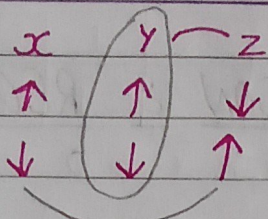
$$+ 2x = -y + 5 \quad \text{Ans } -1$$

$$-2x + 3y = 10$$

$$-2x = -3y + 10 \quad \text{Ans } +1$$

If sign of two variables are same then ans. is $+1$ or opposite then ans. is -1

TYPE 2:



$$r_{xy} = 0.59$$

$$2y + 5z = 10$$

$$r_{yz} = -0.59$$

Ans

If the signs are same then ans is positive & if opposite then ans is negative

$$r_{ab} = -0.26$$

$$3a - 2c = 10$$

$$r_{bc} = -0.26$$

Ans

Bivariate Data → Analysis of two datas at one time
or
table.

P - Column

↓ ↓ ↓ ↓

Maths →	0-10	10-20	20-30	30-40
Stats ↓				
→ 0-10	10	12	15	16
→ 10-20	8	10	9	10
→ 20-30	8	11	6	3

$$\bullet \text{ No. of cells} = p \times q$$

$$= 4 \times 3 = 12$$

$$\bullet \text{ No. of Marginal Distributions} = 2$$

$$\bullet \text{ No. of Conditional Distributions} = p + q$$

$$= 4 + 3 = 7$$

M.D (Maths)

$$0-10 \rightarrow 26$$

$$10-20 \rightarrow 33$$

$$20-30 \rightarrow 30$$

$$30-40 \rightarrow 29$$

$$118$$

M.D 2 (Stats)

$$0-10 \rightarrow 53$$

$$10-20 \rightarrow 37$$

$$20-30 \rightarrow 28$$

$$118$$

Maths 10-20

Stats

0-10

12

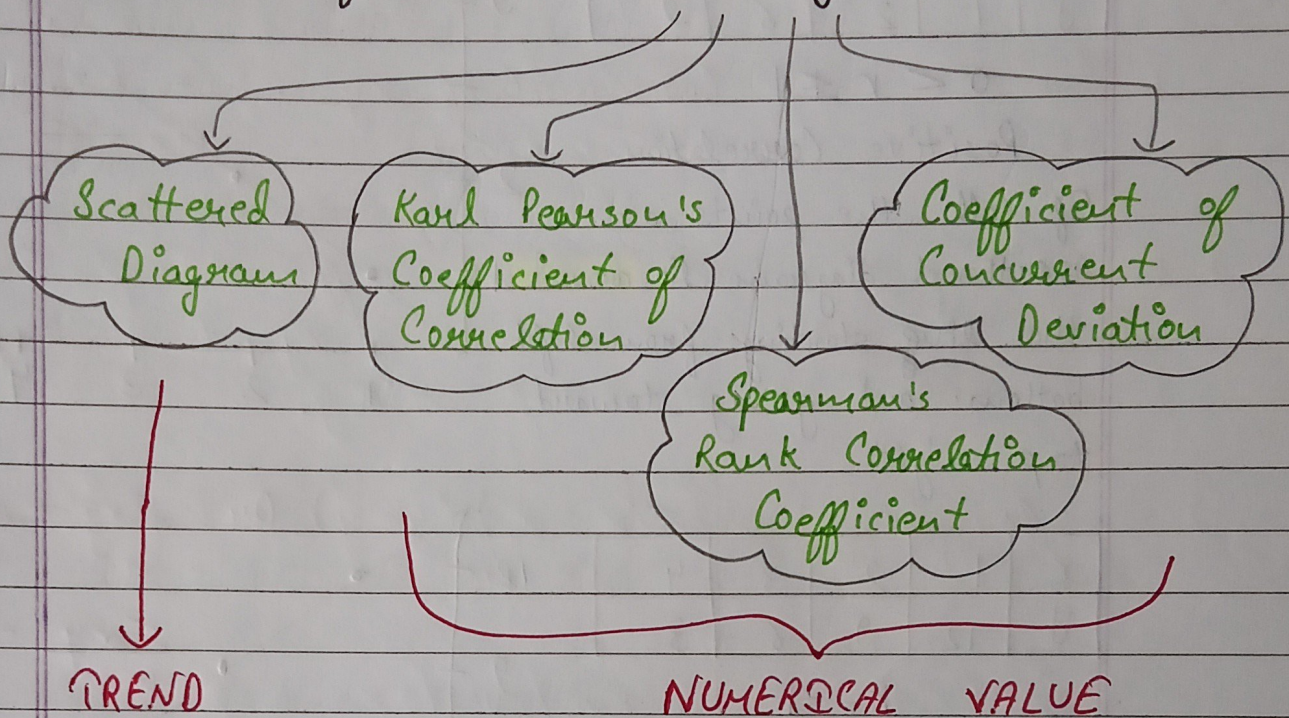
10-20

10

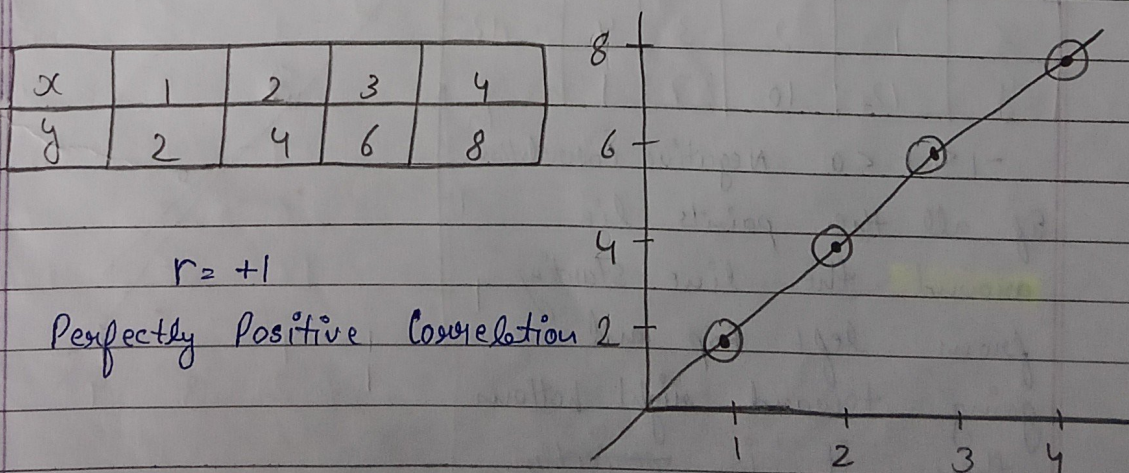
20-30

11

Methods of Correlation Analysis



1. Scattered Diagram - If all the points on scattered diagram lie on the line starting from left bottom and going toward top right then



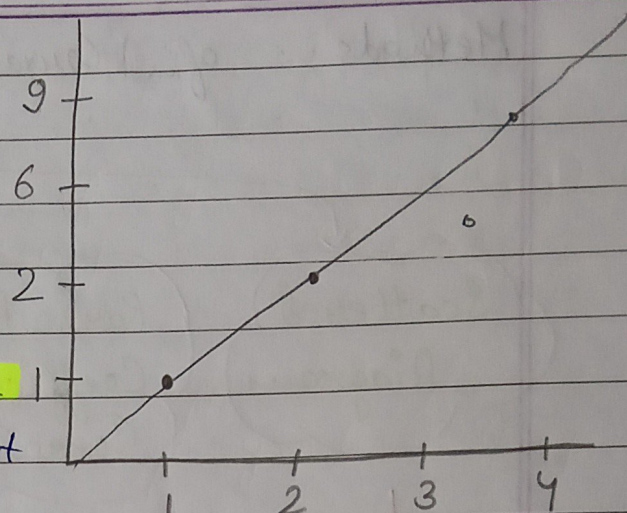
This method only tells us the nature of correlation & not the magnitude

x	1	2	3	4
y	1	2	6	9

$$0 < r \leq 1$$

Positive Correlation

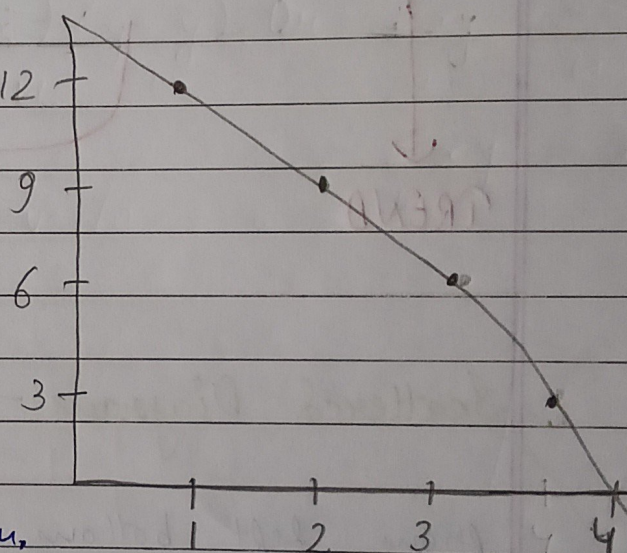
If all the points on scattered diagram lie around the line starting from left bottom and going toward top right



x	1	2	3	4
y	12	9	6	3

$$r = -1$$

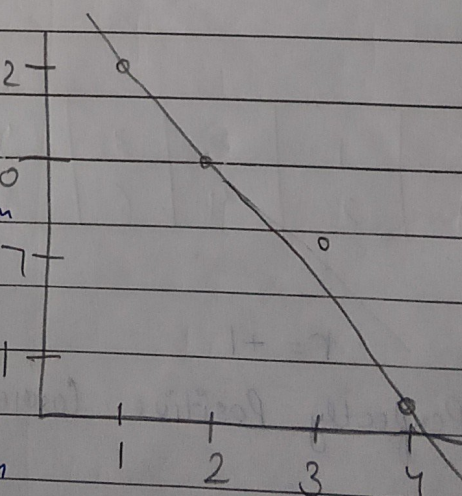
If all the points lie on the line starting from left top and going toward right bottom then it is perfectly negative correlation.



x	1	2	3	4
y	12	10	7	1

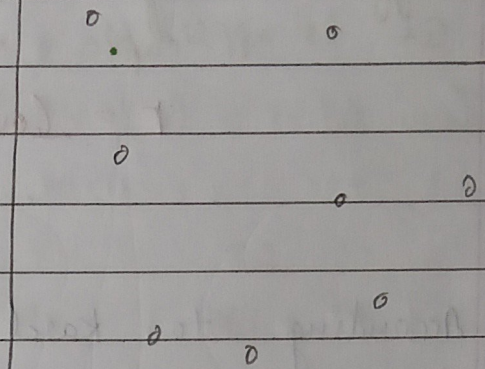
$$-1 \leq r < 0 \text{ Negative Correlation}$$

If all the points lie around the line starting from left top and going toward right bottom then it is perfectly Negative Correlation.



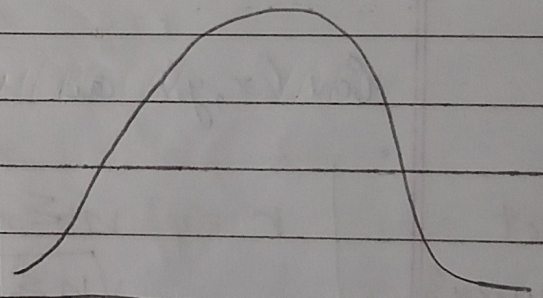
x	1	2	3	4
y	10	6	19	1

$r = 0$ [Spurious Corr.]
 No Correlation



x	1	2	3	4	5
y	2	4	10	4	2

Curvi Linear Correlation
 $r = 0$



KP

2. Coefficient of Correlation (Product Moment Corr. Coeff.)
It is applicable only for Linear variable

$$r = \frac{\text{Covariance of } xy}{\sigma_x \times \sigma_y} = \frac{\text{Cov}(x, y)}{\sigma_x \times \sigma_y}$$

According to Karl Pearson, coefficient of Correlation is the ratio of covariance of xy to the product of standard Deviations of x & y .

$$\text{Cov.}(x, y) = \frac{\sum(x - \bar{x})(y - \bar{y})}{n} \text{ Or } = \frac{\sum xy}{n} - \frac{\sum x}{n} \cdot \frac{\sum y}{n}$$

Sathi me variation

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

Cov(x, y) can be +ve/-ve/0

Direct
Formula
of KP

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

E.g. 1

x	y	xy	$\bar{x} = 6.25$
2	10	20	$\bar{y} = 13.5$
5	12	60	$n = 4$
8	14	112	
10	18	180	
25	54	372	

$$\begin{aligned} \text{Cov.}(xy) &= \frac{\sum xy}{n} - \bar{x} \cdot \bar{y} \\ &= \frac{372}{4} - 6.25 \times 13.5 \\ &= 8.625 \end{aligned}$$

E.g. 2

x	y	xy	$\bar{x} = 3.5$
1	5	5	$\bar{y} = 8.75$
3	8	24	$n = 4$
4	10	40	
6	12	72	
14	35	141	

$$\begin{aligned} \text{Cov.}(xy) &= \frac{\sum xy}{n} - \bar{x} \cdot \bar{y} \\ &= \frac{141}{4} - 3.5 \times 8.75 \\ &= 4.625 \end{aligned}$$

Ex. 3-

x	y	x^2	y^2	xy	$\bar{x} = 3.5$ $\bar{y} = 8.75$ $n = 4$
1	5	1	25	5	
3	8	9	64	24	
4	10	16	100	40	
6	12	36	144	72	
14	35	62	333	141	

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

$$= \frac{4 \times 141 - 14 \times 35}{\sqrt{4 \times 62 - 196} \cdot \sqrt{4 \times 333 - 1225}}$$

$$= \frac{74}{\sqrt{52} \times 107}$$

$$\rightarrow D.O.C. = 52 \times 107 = \sqrt{\div} = \times 74$$

$$= 0.9920$$

Kuch angrezi Baatein

~~Ex~~ $\sum (x - \bar{x})(y - \bar{y})$ = Sum of product of deviations

$\sum (x - \bar{x})^2$ = Sum of squares of deviation

$\sum x^2$ = Sum of squares of x

$(\sum x)^2$ = square of sum of x

Properties of Correlation Coefficient (K.P.)

1. Correlation Coefficient is a unit free quantity i.e. It has no unit.
2. $-1 \leq r \leq 1 \Rightarrow r^2 \leq 1$ $[\sigma_x \sigma_y \geq \text{Cov.}(x, y)]$
3. K.P. Coefficient is independent of both change in origin as well as change in scale but the sign needs to be considered.

$$\begin{array}{ccc} r_{xy} = 0.58 & & \\ + 2x & + y + 5 & r = 0.58 \\ \text{Same sign} \rightarrow & \text{No change} & \\ \text{Opp. sign} \rightarrow & \text{Change the sign} & \end{array}$$

- o It is the most accurate & acceptable method of correlation

4. Coefficient of determination = r^2 $\frac{\text{Explained Var.}}{\text{Total Var.}} \times 100$ (Explained variance)

5. Coefficient of Non-determination = $1 - r^2$ (Unexplained variance)
 $\frac{\text{Unexplained Var.}}{\text{Total Var.}} \times 100$

6. Nature of correlation entirely depends upon nature of covariance.

$$\begin{array}{ll} \text{If } \text{Cov.} = -ve & , r = -ve \\ \text{If } \text{Cov.} = +ve & , r = +ve \end{array}$$

3. Spearman's Rank Coefficient of Correlation

It is represented by r_0

It is important method for ordinal data / qualitative data

①
High to Low
or
Low to High } Same Rank

$$r_0 = 1 - \frac{6 \sum d^2}{n(n^2-1) \text{ or } (n^3-n)}$$

①① If the Rank are exactly opposite then $r_0 = -1$

$\sum d^2$ = Sum of Squares of difference in ranks of x & y .

①①① If the ranks are exactly same then r_0 will always be 1

it will always 0

e.g.

x	y	R_x	R_y	$d = R_x - R_y$	d^2
15	5	3	5	-2	4
18	13	2	1	+1	1
10	10	4	2	+2	4
9	6	5	4	+1	1
20	9	1	3	-2	4
				$\sum d = 0$	$\sum d^2 = 14$

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

$$= 1 - \frac{6 \times 14}{5 \times 24} \rightarrow \text{D.O.C.} = 6 \times 14 \div 5 \div 24 + 1 =$$

$$= 0.3$$

Formula for Common Ranks

$$r_0 = 1 - \frac{6 \left[\sum d^2 + \frac{\sum (m^3 - m)}{12} \right]}{n(n^2-1) \text{ or } n^3-n}$$

m - the times no. repeated or common ranks
Jitno kai beech tie hai

eg.	x	y	R _x	R _y	d	d ²
	40	20	1.5	2	-0.5	0.25
	38	20	4.5	2	2.5	6.25
	40	18	1.5	4	-2.5	6.25
	32	20	6	2	4	16
	38	16	4.5	5	-0.5	0.25
	39	5	3	6	-3	9
					0	38

$$r = 1 - \frac{6 \left[\sum d^2 + \frac{\sum (m^3 - m)}{12} \right]}{n(n^2 - 1)}$$

$$r = 1 - \frac{6 \left[38 + \frac{(2^3 - 2) + (2^3 - 2) + (3^3 - 3)}{12} \right]}{6 \times 35}$$

$$r = -0.1714$$

$$\text{D.O.C.} \rightarrow 6 + 6 + 24 \div 12 + 38 \times 6 \div 6 \div 35 + / - + 1$$

Example - The sum of sq. of difference in ranks of the marks obtained by the 15 students in two different subjects was found to be 80, find the Spearman's Rank Correlation Coefficient.

Solve - $n = 15$
 $\sum d^2 = 80$

$$6 \times 80 \div 15 \div 224 + / - + 1 = 0.657$$

$$r = 0.657$$

Example 2 - If the sum of squares of difference in Ranks of certain number of states was found to be 63 and the Rank Correlation was -0.125 .

(a) 9 (b) ~~7~~ (c) 8 (d) 6

Solve - $\sum d^2 = 63$

$$r = -0.125$$

$$-0.125 = 1 - \frac{6 \times 63}{n(n^2-1)}$$

[Do direct by knock out method]

$$-0.125$$

Example 3 - The value of Rank Correlation b/w age & spectacles numbers of 12 persons was found to be 0.55. For one of the candidates the difference in the rank was wrongly taken as 9 instead of 6. If the mistake is rectified find the correct correlation coefficient.

Solve -

$$n = 12$$

$$r = 0.55$$

$$0.55 = 1 - \frac{6 \sum d^2}{12 \times 143}$$

$$\frac{6 \sum d^2}{12 \times 143} = 1 - 0.55$$

$$\text{D.O.C} \rightarrow 1 - 0.55 \times 12 \times 143 \div 6 \quad \sum d^2 = 128.7$$

$$128.7 - 9^2 + 6^2 = 83.7$$

$$6 \times 83.7 \div 12 \div 143 + 1 = 0.707$$

4. Coefficient of Concurrent Deviations

$$r = +/- \sqrt{\frac{2c - m}{m}}$$

applicable if $\sqrt{\quad}$ value is -ve

$c \rightarrow$ No. of concurrences [count + sign in last column]
 $m = n - 1$, no. of observations compared

Ex: Calculate the coefficient of concurrent Deviations

x	y	C_x	C_y	$C = C_x \times C_y$
58	100	-	-	+
64	105	+	+	+
49	98	-	-	+
36	96	+	-	+
39	107	+	+	+
40	109	+	+	+
44	100	+	-	-
59	97	+	-	-

$$C = 5$$

$$m = 8 - 1$$

$$= 7$$

$$r = + \sqrt{\frac{2 \times 5 - 7}{7}}$$

$$r = \sqrt{\frac{3}{7}}$$

$$r = 0.654$$

e.g.

If the coefficient of concurrent deviation was found to be $\frac{1}{\sqrt{5}}$ and the number of concurrent was 6, then find the number of observations

$$\rightarrow r = \frac{1}{\sqrt{5}}$$

$$c = 6$$

$$n = ?$$

$$\frac{1}{\sqrt{5}} = \sqrt{\frac{2 \times 6 - m}{m}}$$

$$\frac{1}{5} = \frac{12 - m}{m}$$

$$m = 60 - 5m$$

$$6m = 60$$

$$m = 10$$

$$n = m + 1$$

$$n = 10 + 1$$

$$n = 11$$

Ans

Regression

We use slope point form to make regression equations
Regression Analysis (Regression Equations)

$x \rightarrow$ Given, $y \rightarrow$ find
 $y \rightarrow$ Given, $x \rightarrow$ find

Regression
Regression

of y on x
of x on y

x	y
Advertisement	Sales
10	130
12	140
15	125
18	135
22	150

? ——— 200
25 ——— ?

What is diff. b/w corr. & reg.
 $\rightarrow$ Correlation only tells us that if x increases, whether y will increase or decrease.

However, regression is a step forward.
it tells the exact value of y , if exact value of x is given.

Regression Line of y on x

$$y = a + b_{yx}x$$

$b_{yx} \Rightarrow$ Regression Coefficient of y on x

$$b_{yx} = \frac{\text{Cov.}(x, y)}{\text{Var.}(x)} = \frac{\text{Cov.}(x, y)}{\sigma_x^2}$$

$$b_{yx} = \frac{r \sigma_y}{\sigma_x}$$

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

To find a in this equation
Replace $(x, y) \rightarrow (\bar{x}, \bar{y})$

Q.

x	y	$\Sigma x = 75$
10	5	$\Sigma y = 51$
12	8	$\Sigma xy = 830$
15	10	$\Sigma x^2 = 1,193$
18	13	$n = 5$
20	15	$\bar{x} = 15$
↓		$\bar{y} = 10.2$
25	?	

$$y = a + b_{yx} x$$

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

$$b_{yx} = \frac{5 \times 830 - 75 \times 51}{5 \times 1,193 - (75)^2} = \frac{325}{340} = 0.9558$$

$$y = a + 0.9558 x$$

$$10.2 = a + 0.9558 \times 15$$

$$a = -4.137$$

$$D.O.C \rightarrow 0.9558 \times 15 + 10.2$$

$$y = -4.137 + 0.9558 x$$

$$y = -4.137 + 0.9558 \times 25$$

$$y = 19.75$$

Regression Line of x on y

$$x = a + b_{xy} y$$

$b_{xy} \rightarrow$ Regression coefficient of x on y

$$b_{xy} = \frac{\text{Cov.}(x, y)}{\text{Cov. } y} = \frac{\text{Cov.}(x, y)}{\sigma_y^2}$$

$$b_{xy} = \frac{r \sigma_x}{\sigma_y}$$

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

To find a in this equation
Replace $(x, y) \rightarrow (\bar{x}, \bar{y})$

Q.

x	y	$\sum x = 55$
5	10	$\sum y = 95$
8	15	$\sum xy = 1,172$
12	20	$\sum y^2 = 1,993$
13	22	$\bar{x} = 11$
17	28	$\bar{y} = 19$
	$\downarrow$	$n = 5$
	35	

$$\begin{aligned} b_{xy} &= \frac{n \sum xy - \sum x \sum y}{n \sum y^2 - (\sum y)^2} \\ &= \frac{5 \times 1,172 - 55 \times 95}{5 \times 1,993 - (95)^2} = \frac{635}{940} = 0.6755 \end{aligned}$$

$$x = a + b_{xy} y$$

$$11 = a + 0.6755 \times 19$$

$$a = -1.8345$$

$$x = a + b_{xy} y$$

$$x = -1.8345 + 0.6755 \times 35$$

$$x = 21.808$$

Q.

x	y	$y = ?$	$x = 30$
15	6		
18	7	$\Sigma x = 78$	$\bar{x} = 19.5$
20	10	$\Sigma y = 32$	$\bar{y} = 8$
25	9	$\Sigma xy = 641$	$n = 4$
		$\Sigma x^2 = 1,574$	

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

$$= \frac{4 \times 641 - 78 \times 32}{4 \times 1,574 - (78)^2} = \frac{68}{212} = 0.3207$$

$$y = a + b_{yx} x$$

$$a = 1.74635$$

$$y = 1.74635 + 0.3207 \times 30$$

$$y = 11.36735$$

Q.

Room

	Maths (x)	Stats (y)
Avg	35	41
SD	1.5	2.4

$$r = 0.78 \quad M = 40 \quad S = ?$$

$$Y = a + b_{yx} X$$

$$b_{yx} = \frac{r \sigma_y}{\sigma_x}$$

$$= \frac{0.78 \times 2.4}{1.5} = 1.248$$

$$a = -2.68$$

$$Y = a + b_{yx} X$$

$$Y = -2.68 + 1.248 \times 40$$

$$Y = 47.24$$

Properties of Regression Coefficient

1. b_{yx} , b_{xy} and r all will always have same sign.

$$2. b_{yx} = \frac{r \sigma_y}{\sigma_x}, \quad b_{xy} = \frac{r \sigma_x}{\sigma_y}$$

$$b_{yx} \times b_{xy} = \frac{r \cancel{\sigma_y}}{\cancel{\sigma_x}} \times \frac{\cancel{\sigma_x}}{\cancel{\sigma_y}}$$

$$b_{yx} \times b_{xy} = r^2$$
$$\sqrt{b_{yx} \times b_{xy}} = r$$

Correlation Coefficient is the GM of the two regression coefficient.

Example $b_{yx} = 0.98$, $b_{xy} = 0.75$

$$r^2 = b_{yx} \times b_{xy}$$
$$r = \sqrt{0.98 \times 0.75}$$
$$\boxed{r = 0.85}$$

Example $b_{yx} = -1.2$, $b_{xy} = -0.65$

$$r^2 = b_{yx} \times b_{xy}$$
$$r = \sqrt{-0.65 \times -1.2}$$
$$\boxed{r = -0.88}$$

Example $\frac{b_{xy} + b_{yx}}{2} \geq r$

$$\frac{b_{xy} + b_{yx}}{2} \geq \sqrt{b_{xy} \times b_{yx}}$$

$$AM \geq GM$$

Definitely True

3. $r^2 \leq 1 \Rightarrow b_{xy} \times b_{yx} \leq 1$

[This condition is used to identify which line is x on y and which line is y on x]

4. The solution of the two regression lines is A.M. of x and y i.e. $(\bar{x}, \bar{y})$

Example If the two regression lines are these find arithmetic mean of x and y . Find $\bar{x}, \bar{y}$.

$$2x - y = 5$$

$$3x - 2y = 7$$

(A) 1, 2

(B) 2, 3

~~(C) 3, 1~~

(D) 4, 2

[Do by knock out or solve the equation]

5. Regression Coefficients are independent of Δ in origin but they do depend on Δ in scale, Δ in sign.

u & x
 $u = a + bx$
 $\downarrow$ $\searrow$ scale
 origin

v & y
 $v = c + dy$
 $\downarrow$ $\searrow$ scale
 origin

$b_{yx} \rightarrow$ Given

$b_{vu} \rightarrow ?$

$$b_{vu} = \frac{d}{b} \times b_{yx}$$

$b_{xy} \rightarrow$ Given

$b_{uv} \rightarrow ?$

$$b_{uv} = \frac{b}{d} \times b_{xy}$$

Q.1 If $b_{yx} = 1.27$

$$2u + 3x = 10, \quad 3v + y = 7$$

find $b_{vu} = ?$

$$2u = -3x + 10$$

$$3v = -y + 7$$

$$u = \frac{-3x + 10}{2}$$

$$v = \frac{-y + 7}{3}$$

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

$$d = \frac{-1}{3}$$

$$b_{vu} = \frac{d}{b} \times b_{yx}$$

$$b_{vu} = \frac{-1/3}{-3/2} \times 1.27$$

$$b_{vu} = \frac{-1 \times 2}{3 \times 3} \times 1.27$$

$$b_{vu} = 0.2822$$

Q.2 If $b_{xy} = 1.5$

$$3u + x = 10, \quad -3v + 4y = 6$$

find $b_{uv} = ?$

$$u = \frac{-1x + 10}{3}$$

$$v = \frac{4y - 6}{3}$$

$$b = \frac{-1}{3}$$

$$d = \frac{4}{3}$$

$$b_{uv} = \frac{b}{d} \times b_{xy}$$

$$b_{uv} = \frac{-1/3}{4/3} \times 1.5 = \frac{-1 \times 3}{3 \times 4} \times 1.5$$

$$b_{uv} = -0.375$$

6. r^2 is called as coefficient of determination (accounted)

$$r^2 = \frac{\text{Explained Variance}}{\text{Total Variance}}$$

$1 - r^2$ is called as coefficient of non determination (unaccounted)

$$1 - r^2 = \frac{\text{Unexplained Variance}}{\text{Total Variance}}$$

Agriculture production & Rainfall $\rightarrow +0.8$

$$\square r = 0.8 \quad r^2 = (0.8)^2 = 0.64 = 64\%$$

$$\square r = 0.55 \quad \text{COD} = 30.25\%$$

$$\text{COND} = 69.75\%$$

Q. If $r = 0.62$ then how much percentage of variation is accounted for?

$$\text{COD} = (0.62)^2 \times 100 = 38.44\%$$

$$\text{COND} = (1 - r^2) \times 100 =$$

$$= (1 - 0.62^2) \times 100$$

$$= 61.56\%$$

7. If $r = -1$ or $+1$, then the two regression lines are coincident (ek he upar ek)

If $r = 0$, then the two regression lines are perpendicular to each other.

8. The slope of the line y on x is ' b_{yx} '
The slope of the line x on y is ' $\frac{1}{b_{xy}}$ '

example

$5x + 3y = 10$ then y on x is

$$3y = -5x + 10$$

$$y = \frac{-5x}{3} + \frac{10}{3}$$

$$Y = a + b_{yx} x$$

$$b_{yx} = \frac{-5}{3}$$

example

$x - 3y = 12$ then x on y is

$$x = 3y + 12$$

$$b_{xy} = 3$$

$$X = a + b_{xy} y$$

$$\text{Slope} = \frac{1}{3}$$

9. Formula for Probable error (PE)

$$PE = 0.6745 \frac{(1-r^2)}{\sqrt{n}}$$

[Range of Population $\Rightarrow r \pm PE$]
 r

Q.1

$r = 0.7$, $n = 49$ then $PE \rightarrow 0.049$

$$D.O.C. \rightarrow 0.7 \pm +1 - +1 \times 0.6745 \div 7 =$$

10. The method to find two regression equation is called as 'method of least square'.

example - find which regression eq. is x on y & y on x ?

$$5x - 3y = 12$$

$$2x - 3y = 6$$

$$b_{xy} - r x = a + b_{xy} y$$

$$b_{yx} - r y = a + b_{yx} x$$

1st x on y

$$5x = 3y + 12$$

$$x = \frac{3y}{5} + \frac{12}{5}$$

$$b_{xy} = \frac{3}{5}$$

2nd y on x

$$3y = 2x - 6$$

$$y = \frac{2x}{3} - \frac{6}{3}$$

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

$$b_{xy} \times b_{yx} \leq 1$$

$$\frac{3}{5} \times \frac{2}{3} \leq 1$$

$$0.4 \leq 1$$

So, the assumption is right the first equation is x on y and second equation is y on x

example $4x - 7y = 12$

$$x - 5y = 8$$

1st y on x

2nd x on y

$$Y = a + b_{yx} x$$

$$X = a + b_{xy} y$$

$$y = \frac{4x - 12}{7}$$

$$x = 5y + 8$$

$$b_{yx} = \frac{4}{7}$$

$$b_{xy} = 5$$

$$b_{xy} \times b_{yx} \leq 1$$

$$5 \times \frac{4}{7} \leq 1$$

$$2.8 \leq 1$$

Not Satisfied

Assumption 1 is not right so we do the opposite of it

$$b_{xy} = \frac{7}{4}$$

$$b_{yx} = \frac{1}{5}$$

$$b_{xy} \times b_{yx} \leq 1$$

$$\frac{7}{4} \times \frac{1}{5} \leq 1$$

$$0.35 \leq 1$$

Assumption 2 is right that equation 1 is x on y and equation 2 is y on x .