Covariance

Eg:

(X) Age (year)	Weight (kg) (Y)
20	75
18	63
15	45
14	40
25	78

$X \uparrow \quad y \uparrow$

$X \downarrow \quad y \downarrow$

$$\left\{ \begin{array}{cc} X\uparrow & y\downarrow \\ X\downarrow & y\uparrow \end{array} \right\}$$

Covariance

① Quantify the relationship between X & y

Numerical value

Population

$$Cov(X, y) = \sum_{i=1}^{N} \frac{(x-\bar{x})(y-\bar{y})}{N}$$

Sample

Degree of freedom

{Bessels correction}

$$Cov(X, y) = \sum_{i=1}^{n} \frac{(x-\bar{x})(y-\bar{y})}{n-1}$$

X	Y
Economic Growth (%)	NIFTY 50 Index (%)
2.1	8
2.5	12
4.0	14
3.6	10

Covariance

$$Cov(x,y) = \frac{\sum_{i=1}^{n}(x-\bar{x})(y-\bar{y})}{n-1}$$

+ve $\begin{cases} x\uparrow & y\uparrow \\ x\downarrow & y\downarrow \end{cases}$

-ve $\begin{cases} x\uparrow & y\downarrow \\ x\downarrow & y\uparrow \end{cases}$

X	Y	$\bar{X}$	$\bar{y}$	$X-\bar{X}$	$y-\bar{y}$
2.1	8	3.1	11	-1	-3 ←
2.5	12	3.1	11	-0.6	1
4.0	14	3.1	11	0.9	3
3.6	10	3.1	11	0.5	-1

$$Cov(x,y) = \frac{(-1)(-3) + (-0.6)\times(1) + (0.9)\times(3) + (0.5)(-1)}{4-1}$$

$$= \frac{3 - 0.6 + 2.7 + 0.5}{3}$$

$Cov(x,y) = \frac{4.6}{3} = 1.533$ → +ve value

$\boxed{+1000 \quad X \quad -1000}$

Satisfying 2 conditions

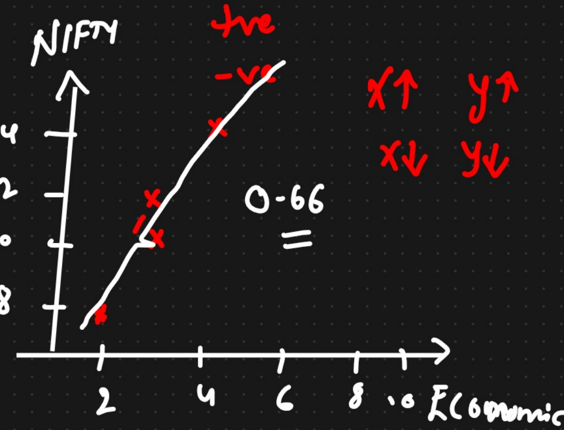
① Economic Growth ↑ NIFTY 50 Index Growth ↑

② Economic Growth ↓ NIFTY 50 Index ↓
Growth

$\{-1 \text{ to } 1\}$

② Pearson Correlation Coefficient

+100
-50



$x\uparrow \quad y\uparrow$

$x\downarrow \quad y\downarrow$

Economic Growth (%)	NIFTY 50 Index Growth (%)
2.1	8
2.5	12
4.0	14
3.6	10

$$\rho_{(x,y)} = \frac{Cov(x,y)}{\sigma_x * \sigma_y} = \frac{1.533}{(0.8981)(2.58)}$$

$$= \frac{1.533}{2.317}$$

$$= 0.66 \text{ +ve}$$

$$\sigma_x = \sum_{i=1}^{n} \sqrt{\frac{(x-\bar{x})^2}{n-1}}$$

$$= \frac{(-1)^2 + (-0.6)^2 + (0.9)^2 + (0.5)^2}{3}$$

$$= \frac{1 + 0.36 + 0.81 + 0.25}{3}$$

Answer

$$= \sqrt{0.806} = 0.8981$$

$[-1 \text{ to } 1]$

$$\sigma_y = \sum_{i=1}^{n} \sqrt{\frac{(y-\bar{y})^2}{n-1}}$$

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

$$= \frac{9 + 1 + 9 + 1}{3} = \frac{20}{3} = 6.66$$

$$\sqrt{6.66} = \underline{\underline{2.58}}$$

<u>Disadvantage</u>

1. It is not able to capture the non linear properties

$\Downarrow$

Spearman Rank Correlation

3. <u>Spearman Rank Correlation</u>

$$r_s = \frac{Cov(R_x, R_y)}{\sigma R_x * \sigma R_y}$$

X	Y	Rx	Ry

Economic Growth (%)	NIFTY 50 %
2.1	8
2.5	12
4.0	14
3.6	10
X	X

3	4
2	2
1	1
4	3