12 - 03 - 2022

Saturday, March 12, 2022 12:59 PM

Unit - 6

§ Correlation:

X & Y

x_i y_i

§ Karl Pearson Correlation co-effecient:

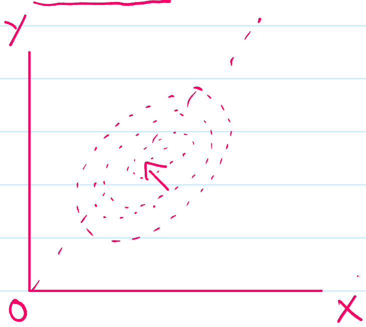
$$r(X,Y) \text{ or } r_{XY} = \frac{Cov(X,Y)}{\sigma_X \sigma_Y},$$

$$Cov(X,Y) = \frac{1}{n}\sum x_i y_i - \bar{x}\bar{y}$$

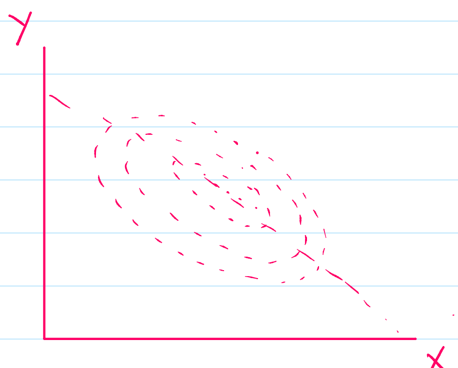
$$\sigma_X^2 = \frac{1}{n}\sum x_i^2 - \bar{x}^2$$

$$\sigma_Y^2 = \frac{1}{n}\sum y_i^2 - \bar{y}^2$$

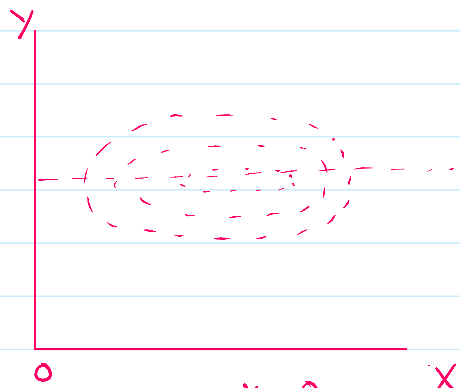
§ Remark :-



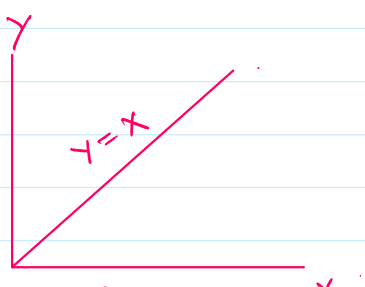
$(r>0)$



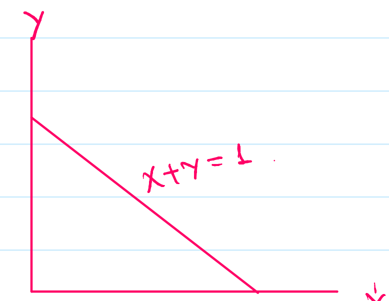
$(r<0)$



$r=0$



$(r=+1)$



$(r=-1)$

Remark 2 :- r_{XY} provides a measure of linear relationship betn X & Y. For non-linear relationship, it is not very suitable.

Remark 3: $Cov(X,Y) = \sigma_{XY}$ $\Rightarrow r_{XY} = \frac{\sigma_{XY}}{\sigma_X \sigma_Y}$.

Remark 4 :- K.P.C.C is also called product-moment correlation co-efficient, since

$$\text{cov}(X,Y) = E\left[\{X - E(X)\}\{Y - E(Y)\}\right] = \mu_{11}.$$

* $-1 \leq r \leq 1$.

Schwartz inequality:

$$r_{XY} = \frac{\text{cov}(X,Y)}{\sigma_X \sigma_Y}$$

$$\text{cov}(X,Y) = \frac{1}{n}\sum(x_i - \bar{x})(y_i - \bar{y})$$

$$\sigma_X^2 = \frac{1}{n}\sum(x_i - \bar{x})^2$$

$$\sigma_Y^2 = \frac{1}{n}\sum(y_i - \bar{y})^2$$

$$= \frac{\frac{1}{n}\sum(x_i - \bar{x})(y_i - \bar{y})}{\left[\frac{1}{n}\sum(x_i - \bar{x})^2 \cdot \frac{1}{n}\sum(y_i - \bar{y})^2\right]^{1/2}}$$

$$\Rightarrow r_{XY}^2 = \frac{\left(\sum_i a_i b_i\right)^2}{\left(\sum_i a_i^2\right)\left(\sum_i b_i\right)^2}, \quad \text{where } a_i = x_i - \bar{x},\ b_i = y_i - \bar{y} \longrightarrow (1)$$

we have Schwartz inequality as follows —

$$\left(\sum_{i=1}^{n} a_i b_i\right)^2 \leq \left(\sum_{i=1}^{n} a_i^2\right)\left(\sum_{i=1}^{n} b_i^2\right),$$

& the sign of equality holds under the condition $\frac{a_1}{b_1} = \frac{a_2}{b_2} = \cdots = \frac{a_n}{b_n}$

$$(1) \Rightarrow r_{XY}^2 \leq \frac{\left(\sum_i a_i^2\right)\left(\sum_i b_i\right)^2}{\left(\sum_i a_i^2\right)\left(\sum_i b_i\right)^2}$$

$$\Rightarrow r_{XY}^2 \leq 1$$

$$\Rightarrow -1 \leq r_{XY} \leq 1$$

$$\Rightarrow r_{XY} \in [-1, 1]$$

Proved//

Hence, the value of r_{XY} cannot exceed unity numerically. It always lies between -1 & 1.

- Corr^n is +ve & perfect if $r_{XY} = +1$
- Corr^n is −ve & perfect if $r_{XY} = -1$.

Q> Let r_{XY} denotes the K.P.C.C of R.Vs X & Y such that $5r_{XY}^2 - 4r_{XY} - 1 = 0$. What type of corr^n exists bet^n X & Y ?

(a) Corr^n is +ve & perfect ✓
(b) Corr^n is −ve & perfect
(c) Corr^n is +ve but not perfect
(d) N.O.T

$5r^2 - 4r - 1 = 0$

$\Rightarrow (5r+1)(r-1) = 0$

$\Rightarrow r = -\frac{1}{5}, 1$

×

Q> r_{XY} $\quad 4r^2 - 4r + 1 = 0$.

(a) corr^n is +ve & perfect
(b) Corr^n is −ve & "
(c) Corr^n is +ve ✓
(d) Corr^n is −ve.

Theorem: Correlation co-effecient is independent of change of origin & scale.

Exp^n: Let X & Y be the two random variables. Also, let us change the origin & scale as follows –

$$U = \frac{X-a}{h}, \quad V = \frac{Y-b}{K}$$

$$r(X,Y) = r(U,V)$$

Cor :- If X, Y are random variables and a, b, c, d are any numbers provided $a, c \neq 0$, then –

$$r(aX+b,\ cY+d) = \frac{ac}{|ac|}\, r(X,Y).$$

Eg: Let $r(X,Y) = 0.5$. Find out $r(2X+3, Y+5)$

$$r(2X+3, Y+5) = \frac{2\cdot 1}{|2\cdot 1|}\, r(X,Y)$$

$$= \frac{2}{2} 0.5$$

$$= \frac{\cdot}{2} \cdots$$

$$= \underline{\underline{0.5}}$$

Pf: $Var(aX+b) = a^2 V(X)$ $\qquad$ $Var(cY+d) = c^2 V(Y)$.

$$Cov(aX+b, cY+d) = ac\, Cov(X,Y)$$

$$\therefore\ r(aX+b, cY+d) = \frac{Cov(aX+b, cY+d)}{\{Var(aX+b)\ Var(cY+d)\}^{1/2}}$$

$$= \frac{ac\, Cov(X,Y)}{|a|\{Var(X)\}^{1/2}\ |c|\{Var(Y)\}^{1/2}}$$

$$= \frac{ac}{|ac|}\ \frac{Cov(X,Y)}{\sigma_X \sigma_Y}$$

$$= \frac{ac}{|ac|} r_{XY} \quad //$$

- If $b = d = 0$, then $r(aX, cY) = \frac{ac}{|ac|} r_{XY}$.

Theorem 2 :- Two independent random variables are uncorrelated.

Pf: If X & Y are independent variables, then $Cov(X,Y) = 0$

$$\Rightarrow r_{XY} = \frac{Cov(X,Y)}{\sigma_X \sigma_Y}$$

$$= 0$$

Converse is not always true. That is, if $r_{XY} = 0$, it may not imply X, Y are independent variables.

$\underline{Y = X^2}$

X	-3	-2	-1	1	2	3	$\Sigma X = 0$
Y	9	4	1	1	4	9	$\Sigma Y = 28$
XY	-27	-8	-1	1	8	27	$\Sigma XY = 0$.

$$\bar{X} = \frac{1}{6}\Sigma X = 0$$

$$\bar{Y} = \frac{1}{6}\Sigma Y = \frac{14}{3}.$$

$$\text{cov}(X,Y) = \frac{1}{n}\Sigma XY - \bar{X}\bar{Y} = \frac{1}{n}\cdot 0 - 0 = 0$$

$$\Rightarrow r_{XY} = \frac{\text{cov}(X,Y)}{\sigma_X \sigma_Y} = 0$$

Q> Calculate the co-efficient of correlation betⁿ X & Y:

X	1	3	4	5	7	8	10
u	−4	−2	−1	0	2	3	5
Y	2	6	8	10	14	16	20
v	−4	−2	−1	0	2	3	5

$$u = \frac{X-5}{1} \qquad r_{XY} = r_{uv}.$$

$$v = \frac{Y-10}{2}$$

u	−4	−2	−1	0	2	3	5	$\Sigma u = 3$
v	−4	−2	−1	0	2	3	5	$\Sigma v = 3$
u^2	16	4	1	0	4	9	25	$\Sigma u^2 = 59$
v^2	16	4	1	0	4	9	25	$\Sigma v^2 = 59$
uv	16	4	1	0	4	9	25	$\Sigma uv = 59$

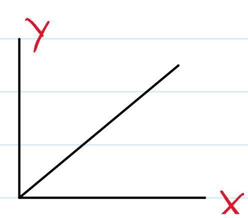
$$\text{cov}(u,v) = \frac{1}{n}\Sigma uv - \bar{u}\bar{v} = \frac{1}{7}\times 59 - \frac{9}{49} \qquad \bar{u} = \frac{3}{7}$$

$$= \frac{413-9}{49} = \frac{404}{49} \qquad \bar{v} = \frac{3}{7}$$

$$\sigma_u^2 = \frac{1}{n}\Sigma u^2 - \bar{u}^2 = \frac{1}{7}\times 59 - \frac{9}{49} = \frac{404}{49} \Rightarrow \sigma_u = \sqrt{\frac{404}{49}}$$

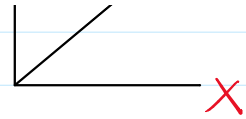
$$\sigma_v^2 = \frac{1}{n}\Sigma v^2 - \bar{v}^2 = \frac{59}{7} - \frac{9}{49} = \frac{404}{49} \Rightarrow \sigma_v = \sqrt{\frac{404}{49}}$$

$$\therefore r_{uv} = \frac{\text{cov}(X,Y)}{\sigma_X \sigma_Y} = \frac{404/49}{\left(\sqrt{\frac{404}{49}}\right)^2} = +1 = r_{XY}$$



⇒ Corrⁿ betⁿ X & Y is perfect & +ve.

$(\sqrt{49}\ )$



$\Rightarrow$ Corr^n bet^n X & Y is perfect & +ve.

~~~~~ x ~~~~~
~~~~~