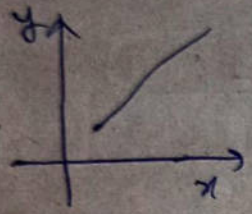
Unit 6-

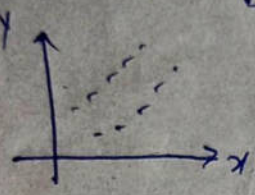
Correlation & Regression

① Correlation Coefficient - (r)

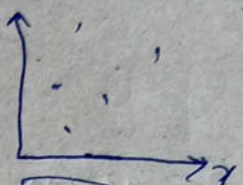
$$-1 \leq r \leq 1$$



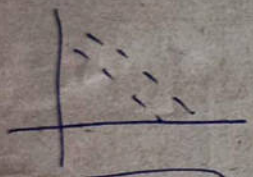
$r = 1$



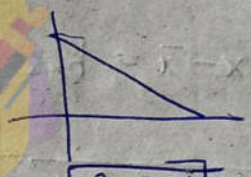
$0 \leq r \leq 1$



$r = 0$



$-1 \leq r < 0$



$r = -1$

② formula to calculate correlation coefficient

cov = covariance

$$r = \frac{cov(x,y)}{\sigma_x \sigma_y}$$

$$E(x) = \frac{\sum x}{n}$$

$$cov(x,y) = E(xy) - E(x)\,E(y)$$

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

$$\sigma_x = \sqrt{Ex^2 - (E(x))^2} = \sqrt{\frac{\sum x^2}{n} - \left(\frac{\sum x}{n}\right)^2}$$

(i) if $Cov(x,y) > 0$ — trend in 1st & 3rd quad

(ii) if $Cov(x,y) < 0$ — 2nd & 4th

(iii) if $Cov(x,y) = 0$ — evenly distributed.

Regression –

Line of regression –

① x on y :-

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

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

② y on x :-

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

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

if eqn are given then

① $(\bar{x}, \bar{y})$ = intersection of both eqn

② $r = \sqrt{b_{xy} \cdot b_{yx}}$