29-03-2022

Tuesday, March 29, 2022 9:56 AM

Unit-6

* For $Y = a + bX$, the normal equations are –

$$\left.\begin{array}{l} \sum y_i = na + b\sum x_i \\ \sum x_i y_i = a\sum x_i + b\sum x_i^2 \end{array}\right\}$$

* For $X = a + bY$, the normal equations are –

$$\left.\begin{array}{l} \sum x_i = na + b\sum y_i \\ \sum x_i y_i = a\sum y_i + b\sum y_i^2 \end{array}\right\}$$

2nd formula: For $Y = a + bX$,

$$Y - \bar{y} = r\frac{\sigma_y}{\sigma_x}(X - \bar{x})$$

For $X = a + bY$, $X - \bar{x} = r\frac{\sigma_x}{\sigma_y}(Y - \bar{y})$

Q Find the most likely price in Mumbai corresponding to the price of Rs. 70 at Kolkata from the following:

	Kolkata (X)	Mumbai (Y)
Average price	65	67
Standard deviation	2.5	3.5

Correlation co-effecient between the prices of commodities in the two cities is 0.8.

Solution: Given, $\bar{X} = 65$, $\bar{Y} = 67$, $\sigma_x = 2.5$, $\sigma_y = 3.5$, $r_{xy} = 0.8$.

we have –

$$Y - \bar{Y} = r\frac{\sigma_y}{\sigma_x}(X - \bar{X})$$

$$\Rightarrow Y - 67 = 0.8 \times \frac{3.5}{2.5}(X - 65)$$

$$\Rightarrow Y - 67 = \frac{4}{5} \times \frac{7}{5}(X - 65)$$

$$\Rightarrow 25(Y - 67) = 28(X - 65)$$

$$\Rightarrow 25Y = 28X - 1820 + 1675$$

$$\Rightarrow 25(Y-67) = 28(X-65)$$

$$\Rightarrow 25Y = 28X - 1820 + 1675$$

$$\Rightarrow 25Y = 28X - 145$$

$$\Rightarrow Y = \frac{28}{25}X - \frac{29}{5} \quad \checkmark$$

For $X = 70$, $\quad Y = \frac{28}{25} \times 70 - \frac{29}{5} = \frac{28 \times 14}{5} - \frac{29}{5} = \frac{392 - 29}{5}$

$$= \frac{363}{5}$$

$$= 72.6$$

§ Regression co-effecients: 'b', the slope of the regression line of Y on X, is also called the co-effecient of regression of Y on X.

$$b_{YX} = \frac{\mu_{11}}{\sigma_X^2} = r\frac{\sigma_Y}{\sigma_X}$$

Similarly, the regression co-effecient of X on Y,

$$b_{XY} = \frac{\mu_{11}}{\sigma_Y^2} = r\frac{\sigma_X}{\sigma_Y}.$$

Properties :- (a) Correlation co-effecient is the geometric mean of the regression co-effecients.

$$b_{YX} \times b_{XY} = r^2$$

$$\Rightarrow r = \pm\sqrt{b_{XY}\, b_{YX}}$$

(b) If one of the regression co-effecients (b_{YX}) is greater than unity, then the other reg. coeff (b_{XY}) is less than unity.

$$b_{YX} > 1 \Rightarrow \frac{1}{b_{YX}} < 1$$

$$r^2 \le 1$$

$$\Rightarrow b_{yx} \cdot b_{xy} \le 1$$

$$\Rightarrow b_{xy} \le \frac{1}{b_{yx}} < 1 \quad //$$

(c) The modulus value of the arithmetic mean of the regression co-efficients is not less than the modulus value of correlation coff 'r'.

i.e., $\left|\frac{1}{2}(b_{xy}+b_{yx})\right| > |r|$.

(d) Regression co-efficients are independent of change of origin but not of scale.

i.e if $U = \frac{X-a}{h}$, $V = \frac{Y-b}{K}$, $h, K > 0$, then –

$r_{uv} = r_{xy}$ but,

$$\left.\begin{aligned} b_{xy} &= \frac{h}{K} b_{uv} \\ b_{yx} &= \frac{K}{h} b_{vu} \end{aligned}\right\}$$

Eg: $U = \frac{X-64}{2}$ $\quad V = \frac{Y-62}{3}$ $\quad b_{xy} = \frac{2}{3} b_{uv}$, $b_{yx} = \frac{3}{2} b_{vu}$ //

Q) Let X & Y be two random variables, where $b_{xy} = 3$ & $b_{yx} = \frac{1}{3}$. What kind correlation exists betn X & Y ?

(a) Perfect ✓
(b) Non-perfect
(c) Positive & perfect
(d) Negative & perfect

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

$$\Rightarrow r = \pm\sqrt{b_{xy} \times b_{yx}}$$

$$= \pm\sqrt{3 \times \frac{1}{3}}$$

$$= \pm 1$$

§ Angle between two lines of regression :

§ Angle between two lines of regression :

$$Y - \bar{y} = r \frac{\sigma_Y}{\sigma_X}(X - \bar{x}) \qquad\qquad X - \bar{x} = r \frac{\sigma_X}{\sigma_Y}(Y - \bar{y})$$

$$\Rightarrow Y - \bar{y} = \frac{\sigma_Y}{r\sigma_X}(X - \bar{x})$$

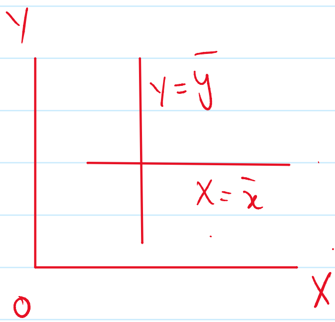
Slopes are $r\frac{\sigma_Y}{\sigma_X}$ & $\frac{1}{r}\frac{\sigma_Y}{\sigma_X}$ resp.

Let θ be the acute angle between the two lines, then -

$$\tan\theta = \left| \frac{r\frac{\sigma_Y}{\sigma_X} - \frac{\sigma_Y}{r\sigma_X}}{1 + r\frac{\sigma_Y}{\sigma_X}\cdot\frac{1}{r}\frac{\sigma_Y}{\sigma_X}} \right| = \left| \frac{\frac{\sigma_Y}{\sigma_X}\left(\frac{r^2-1}{r}\right)}{1 + \frac{\sigma_Y^2}{\sigma_X^2}} \right| = \frac{1-r^2}{|r|}\left(\frac{\sigma_X\sigma_Y}{\sigma_X^2+\sigma_Y^2}\right)$$

$$\Rightarrow \theta = \tan^{-1}\left\{ \frac{1-r^2}{|r|}\left(\frac{\sigma_X\sigma_Y}{\sigma_X^2+\sigma_Y^2}\right) \right\}$$ Ans //

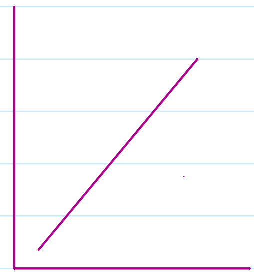
Case (i) $r = 0$, $\theta = \tan^{-1}(\infty) = \frac{\pi}{2}$ $\Rightarrow$ the two lines of regression are perpendicular to each other.



$r = 0$

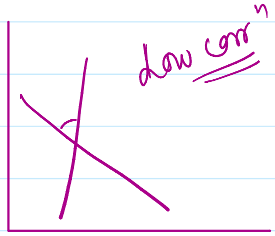
Case (ii) $r = \pm 1$. $\theta = \tan^{-1}(0) = 0°$ or π. The two lines of regression either coincide or they are parallel to each other. But, since both of them pass through $(\bar{x}, \bar{y})$, so they can't be parallel but same.

$(r = -1)$



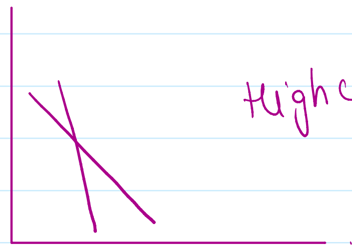
$(r = +1)$

Low degree of corr^n



Low corr^n

$r \in (-1, 0)$



High correlation

$r \in (0, 1)$

High degree of corr^n.