25 November 2020 15:52

Regression Analysis (Linear)

Correlation - Relation between x & y

$y = 2x+3$ $\qquad$ $y = x^2$

$x = y-6$ $\qquad$ $x = e^y$ . . . .

Least Square Method

Regression eq. of y on x (y is dep. x is indp.)

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

b_{yx} = Reg. coeff. of y on x

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

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

[r = correl.
σ_y = S.D. of y
σ_x = S.D of x]

Reg. eq. of x on y [x dep. y indep.]

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

$$b_{xy} = \text{Reg. coeff. of } x \text{ on } y$$

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

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

1. Find Reg. equations.

x	y	x^2	y^2	xy
2	14	4	196	28
5	18	25	324	90
7	19	49	361	133
6	20	36	400	120
$\sum x = 20$	$\sum y = 71$	$\sum x^2 = 114$	$\sum y^2 = 1281$	$\sum xy = 371$

$$\bar{x} = \frac{\sum x}{n} = \frac{20}{4} = 5$$

$$\bar{y} = \frac{\sum y}{n} = \frac{71}{4} = 17.75$$

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

$$= \frac{4 \times 371 - 20 \times 71}{56} = 1.14$$

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

$$= \frac{\quad}{\quad}$$

① Reg. eq. of y on x

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

$$\Rightarrow y - 17.75 = 1.14(x - 5)$$

$$\Rightarrow y = 17.75 + 1.14x - 5.70$$

$$\Rightarrow \boxed{y = 12.05 + 1.14x}$$

Reg. eq. of x on y

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

$\Rightarrow x - 5 = 0.77(y - 17.75)$

$\Rightarrow x = 5 + 0.77y - 13.67$

$\Rightarrow \boxed{x = -8.67 + 0.77y}$

MCQ Range of r is between

1. (i) -1 to 0 (ii) 0 to $+1$
   (iii) -1 to $+1$ ✓ (iv) any value

2. Relation between b_{yx} & b_{xy}
   (i) one is greater than one & the other is less ✓

2. Find the Regression lines and estimate y when $x = 18$ and estimate x when $y = 24$

$[\cancel{x}\, y = 2x - 5]$

x	y	x^2	y^2	xy
9	11			
5	8			

x	y			
4	11			
5	8			
10	14			
13	20			
16	18			
$\Sigma x =$ 53	$\Sigma y =$ 124 71	$\Sigma x^2 =$ 631	$\Sigma y^2 =$ 1105	Σxy 827

y on x

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

$$\frac{\quad}{5} = 10.6$$

$$\bar{y} = \frac{124}{5} = 24.8$$

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

$$= \frac{5 \times 827 - 53 \times 124}{5 \times 631 - (53)^2}$$

$$= \frac{-2437}{346} = -7.04$$

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

$$= \frac{-2437}{\quad}$$

$$\overline{5 \times 1105 - (124)^2}$$

$$= 0.25$$

$\underline{y \text{ on } x}$

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

$$\Rightarrow y - 24.8 = -7.04(x - 10.6)$$

$$\Rightarrow y - 24.8 = -7.04(18 - 10.6)$$

$$\Rightarrow y = 24.8 - 7.04 \times 7.4$$

$$\Rightarrow y = -27.296$$

$\underline{x \text{ on } y}$

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

$$\Rightarrow x - 10.6 = 0.25(y - 24.8)$$

$$\Rightarrow x = 10.6 + 0.25(24 - 24.8)$$

$$= 10.6 + 0.25(-0.8)$$

$$= 10.4$$

3. $\bar{x} = 65$, $\bar{y} = 53$, $\sigma_x = 4.7$, $\sigma_y = 5.2$

$r = 0.78$

Estimate y when $x = 63$

" x " $y = 50$

Sol.

$$b_{yx} = \frac{r\sigma_y}{\sigma_x} = \frac{0.78 \times 5.2}{4.7}$$
$$= 0.86$$

$$b_{xy} = \frac{r\sigma_x}{\sigma_y} = \frac{0.78 \times 4.7}{5.2}$$
$$= 0.71$$

<u>y on x</u>

$$y - \bar{y} = b_{yx}(x - \bar{x})$$
$$\Rightarrow y - 53 = 0.86(x - 65)$$
$$\Rightarrow y = 53 + 0.86(63 - 65)$$
$$\Rightarrow y = 53 - 1.72 \Rightarrow y = 51.28$$

<u>x on y</u>

$$x - \bar{x} = b_{xy}(y - \bar{y})$$
$$\Rightarrow x - 65 = 0.71(y - 53)$$
$$\Rightarrow x - 65 = 0.71(50 - 53)$$
$$\Rightarrow x = 65 + 0.71(-3)$$
$$\Rightarrow x = 65 - 2.13$$
$$\Rightarrow x = 62.87$$

27.11.20 $\bar{x} = 43$, $\bar{y} = 37$

27.11.20 $\bar{x} = 43$, $\bar{y} = 37$

reg. coeff. of y on $x = b_{yx} = 0.59$

" " " x " $y = b_{xy} = 0.72$

Find most likely value of y when $x = 40$ & value of x when $y = 35$

Sol. y on x

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

$$\Rightarrow y - 37 = 0.59(x - 43)$$

$$\Rightarrow y - 37 = 0.59(40 - 43)$$

$$\Rightarrow y = 37 + 0.59(-3)$$

$$\Rightarrow y = 35.23$$

x on y

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

$$\Rightarrow x - 43 = 0.72(y - 37)$$

$$\Rightarrow x - 43 = 0.72(35 - 37)$$

$$\Rightarrow x = 43 + 0.72(-2)$$

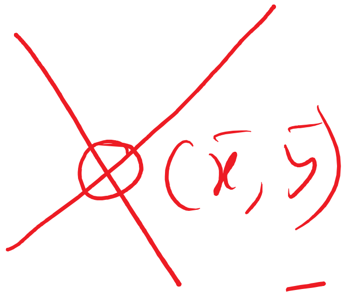
$$\Rightarrow x = 41.56$$

Properties of Regression Equations

x & y — y on x — (1)

x on y — (2)

(1) The intersection pt. of two reg. equations is $(\bar{x}, \bar{y})$



$2x+3y=5$
$x+5y=6$

(2) * $r = \pm\sqrt{b_{yx}\, b_{xy}}$

[If b_{yx} & b_{xy} are both positive, $r = +\sqrt{\ }$; negative, $r = -\sqrt{\ }$]

* Find $\bar{x}, \bar{y}$ from regression equations

$2x+3y=5$ — (1)

$5x+8y=13$ — (2)

(1) $\times 5$ & (2) $\times 2$

$$\cancel{10x}+15y=25$$
$$\underset{-}{\cancel{10x}}\ \underset{-}{+}\ 16y=\underset{-}{26}$$

$$+y=+1$$
$$\Rightarrow y=1$$

$2x + 3\times 1 = 5$

$\Rightarrow 2x + 3 = 5 \Rightarrow 2x = 2$

$\Rightarrow x = 1$

$$\left.\begin{array}{l}\bar{x} = 1 \\ \bar{y} = 1\end{array}\right\}$$

Ex

$2x + 3y - 5 = 0$ —— ①

$x + y - 2 = 0$ —— ②

$$\begin{array}{r} \cancel{2x} + 3y - 5 = 0 \\ -\cancel{2x} \pm 2y \mp 4 = 0 \\ \hline y - 1 = 0 \Rightarrow y = 1 \end{array}$$

$x + 1 - 2 = 0 \Rightarrow x = 1$

$$\left.\begin{array}{l}\bar{x} = 1 \\ \bar{y} = 1\end{array}\right\}$$

Ex

$2x + 3y - 5 = 0$ ( ) Find η

$x + y - 2 = 0$ (reg. eq. of x on y)

Sol.

$2x + 3y - 5 = 0$

$\Rightarrow 3y = 5 - 2x$

$\Rightarrow y = \frac{5}{3} - \frac{2}{3}x$

$3x = 25$

$\Rightarrow x = \frac{25}{3}$

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

$x + y - 2 = 0$ (x on y)

$\Rightarrow x = 2 - y$

$\Rightarrow b_{xy} = -1$

$r = \pm\sqrt{b_{yx}\, b_{xy}}$

$= \pm\sqrt{\left(-\frac{2}{3}\right)(-1)} = -\sqrt{\frac{2}{3}}$

$= -\sqrt{0.67}$

$= -0.82$

Practice sum

Find $(\bar{x}, \bar{y})$ $\quad 2x + y - 3 = 0$

$2x + 3y - 5 = 0$

2.12.20 Regression Equations are

$3x - y - 25 = 0$ & $2x - 3y + 30 = 0$

i) $\bar{x}, \bar{y}$

ii) Find corr. coeff. r

iii) Estimate of y when $x = 17$

& " " x " $y = 25$

Sol.

$3x - y - 25 = 0$ —— 1

$2x - 3y + 30 = 0$ —— 2

1 × 3

$$\begin{aligned} 9x - \cancel{3y} - 75 &= 0 \\ -2x \mp \cancel{3y} \pm 30 &= 0 \\ \hline 7x - 105 &= 0 \end{aligned}$$

$\Rightarrow x = \frac{105}{7} = 15$

$y = 20$

$\left.\begin{aligned} \bar{x} &= 15 \\ \bar{y} &= 20 \end{aligned}\right\}$

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

Let Eq. 1 is reg. of y on x
& Eq. 2 - - - - - - x " y

$3x - y - 25 = 0$

$\Rightarrow -y = 25 - 3x$

$\Rightarrow y = -25 + 3x$

$b_{yx} = +3$

$2x - 3y + 30$

$\Rightarrow 2x = -30 + 3y$

$\Rightarrow x = -\frac{30}{2} + \frac{3}{2}y$

$\Rightarrow b_{xy} = +\frac{3}{2}$

Note: b_{yx} = reg. coeff. of y on x
= coeff. of x from the eq. y on x

b_{xy} = reg. coeff. of x on y
= coeff. of y from the eq. x on y

$$r = \pm\sqrt{b_{yx}\, b_{xy}}$$

$$= \pm\sqrt{(+3)\left(+\frac{3}{2}\right)}$$

$$= +\sqrt{\frac{9}{2}} = +2.12$$

Note $\boxed{-1 \le r \le +1}$

① y ~~on~~ x (x on y)

② x on ~~y~~ (y on x)

$b_{xy} = \frac{1}{3}$ $\quad$ $b_{yx} = \frac{2}{3}$

$$r = \pm\sqrt{\frac{2}{3} \times \frac{1}{3}} = +\sqrt{\frac{2}{9}}$$
$$= +0.47$$

iii) Est. of y when $x = 17$

Note Est. of y is calculated from ~~reg.~~ eq. of y on x

$$2x - 3y + 30 = 0$$
$$\Rightarrow 2 \times 17 - 3y + 30 = 0$$
$$\Rightarrow 34 - 3y + 30 = 0$$
$$\Rightarrow y = 21.33$$

Estimate of x when $y = 25$
(x on y)

$$3x - y - 25 = 0$$
$$\Rightarrow 3x - 25 - 25 = 0 \Rightarrow x = 16.67$$

$$2x + 3y = 5 \quad \text{--- (1)} \qquad r = ?$$
$$x + y = 2 \quad \text{--- (2)}$$

Let (1) is y on x & (2)

is x on y

$$2x + 3y = 5$$

$$\Rightarrow 3y = +5 - 2x$$

$$\Rightarrow y = +\frac{5}{3} - \frac{2}{3}x$$

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

$$x + y = 2$$

$$\Rightarrow x = 2 - y$$

$$\Rightarrow b_{xy} = -1$$

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

$$= \pm\sqrt{\left(-\frac{2}{3}\right)(-1)}$$

$$= -\sqrt{\frac{2}{3}} = -0.82$$

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

Since -0.82 belongs in the range -1 to $+1$, our assumptions are correct

$$r = -0.82$$ ✓

e

$r = -0.82$