30-03-2022

Wednesday, March 30, 2022 11:01 AM

Unit-6

Q) In a partially destroyed laboratory, record of an analysis of correlation data, the foll^n results only are legible :

Var of $X = 9$. Regression equ's : $8X - 10Y + 66 = 0$

$40X - 18Y = 214$

$8\bar{x} - 10\bar{y} + 66 = 0$

$40\bar{x} - 18\bar{y} = 214$

what are :

(i) the mean values $\bar{X}$ & $\bar{Y}$?

(a) $(\bar{x}, \bar{y}) = (13, 17)$

(b) $(\bar{x}, \bar{y}) = (17, 13)$

(c) $(\bar{x}, \bar{y}) = (13, 13)$

(d) $(\bar{x}, \bar{y}) = (17, 17)$

(ii) r_{xy} ?

(iii) σ_y ?

(a) -0.6

(b) $+0.6$

(c) ± 0.6

(d) $+0.5$

$$Y = \frac{8}{10}X + \frac{66}{10} \quad (b_{yx})$$

$$X = \frac{18}{40}Y + \frac{214}{40} \quad (b_{xy})$$

$10Y = 8X + 66$

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

$$r^2 = \frac{8}{10} \times \frac{18}{40} = \frac{4}{5} \times \frac{9}{20} = \frac{9}{25} \Rightarrow r = \pm\frac{3}{5} = \pm 0.6$$

$\sigma_x = 3$

$\sigma_y = ?$

(a) 2

(b) 3

(c) 4

(d) 6

$$r\frac{\sigma_y}{\sigma_x} = b_{yx} \qquad \Rightarrow 0.6\,\frac{\sigma_y}{3} = \frac{4}{5}$$

$$r\frac{\sigma_x}{\sigma_y} = b_{xy} \qquad \Rightarrow \sigma_y = \frac{12}{5} \times \frac{5}{3}$$

$$\Rightarrow \sigma_y = 4$$

Q) can $Y = 5 + 2.8X$ and $X = 3 - 0.5Y$ be the estimated regression equations of Y on X & X on Y resp ?

YES

NO

$Y = 5 + 2.8X$

$b_{yx} = 2.8$

$X = 3 - 0.5Y$

$b_{xy} = -0.5$

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

$b_{xy} = -0.5$

$b_{yx} b_{xy} < 0$

r

§ Curvilinear regression: $Y = a + bX + cX^2$ ✓ $\quad Y = a + bX + cX^2 + dX^3$

Normal equations are :

$$\boxed{\begin{aligned} \Sigma y_i &= na + b\Sigma x_i + c\Sigma x_i^2 \\ \Sigma x_i y_i &= a\Sigma x_i + b\Sigma x_i^2 + c\Sigma x_i^3 \\ \Sigma x_i^2 y_i &= a\Sigma x_i^2 + b\Sigma x_i^3 + c\Sigma x_i^4 \end{aligned}} \left.\begin{aligned} &+ d\Sigma x_i^3 \\ &+ d\Sigma x_i^4 \\ &+ d\Sigma x_i^5 \end{aligned}\right\}$$

$$\Sigma x_i^3 y_i = a\Sigma x_i^3 + b\Sigma x_i^4 + c\Sigma x_i^5 + d\Sigma x_i^6$$

Q) For 10 randomnly selected observations, the follⁿ data were recorded :

Obs. NO	1	2	3	4	5	6	7	8	9	10
Overtime hrs (X)	1	1	2	2	3	3	4	5	6	7
Additional units (Y)	2	7	7	10	8	12	10	14	11	14

Determine the curve of best fit using the non-linear form :

$$Y = a + bX + cX^2.$$

Solⁿ:

S.NO	X	Y	X^2	X^3	X^4	XY	X^2Y
1	1	2	1	1	1	2	2
2	1	7	1	1	1	7	7
3	2	7	4	8	16	14	28
4	2	10	4	8	16	20	40
5	3	8	9	27	81	24	72
6	3	12	9	27	81	36	108
7	4	10	16	64	256	40	160
8	5	14	25	125	625	70	350
9	6	11	36	216	1296	66	396
10	7	14	49	343	2401	98	686
Σ	34	95	154	820	4774	377	1849

Σ	34	95	154	820	4774	377	1849

The normal equations are —

$\Sigma y_i = na + b\Sigma x_i + c\Sigma x_i^2 \Rightarrow 95 = 10a + 34b + 154c \longrightarrow (1)$

$\Sigma x_i y_i = a\Sigma x_i + b\Sigma x_i^2 + c\Sigma x_i^3 \Rightarrow 377 = 34a + 154b + 820c \longrightarrow (2)$

$\Sigma x_i^2 y_i = a\Sigma x_i^2 + b\Sigma x_i^3 + c\Sigma x_i^4 \Rightarrow 1849 = 154a + 820b + 4774c \longrightarrow (3)$

Solⁿ of (1), (2) & (3) —

$a = 1.80, \quad b = 3.48, \quad c = -0.27.$

$\therefore$ The curve of best fit is $Y = 1.80 + 3.48X - 0.27X^2.$

§ Fitting of a power curve: $Y = aX^b$

Taking log, $\log Y = \log a + b\log X$

$\Rightarrow U = A + bV,$

$$\left.\begin{aligned} U &= \log Y \\ A &= \log a \\ V &= \log X \end{aligned}\right\}$$

Normal equⁿs are —

$$\left.\begin{aligned} \Sigma U &= An + b\Sigma V \\ \Sigma UV &= A\Sigma V + b\Sigma V^2 \end{aligned}\right\}$$

$Y = a + bX$

$$\left\{\begin{aligned} \Sigma y_i &= na + \Sigma x_i \\ \Sigma x_i y_i &= a\Sigma x_i + \Sigma x_i^2 \end{aligned}\right.$$

Ex: $Y = aX^b$ for the follⁿ data:

X :	1	2	3	4	5	6	7	8
Y :	1.0	1.2	1.8	2.5	3.6	4.7	6.6	9.1

$\log Y = \log a + b\log X \Rightarrow U = A + bV.$

$$\left.\begin{aligned} \Sigma U &= An + b\Sigma V \\ \Sigma UV &= A\Sigma V + b\Sigma V^2 \end{aligned}\right\}$$

11/20

X	Y	$U = \log Y$	$V = \log X$	UV	V^2

X	Y	$U = \log Y$	$V = \log X$	UV	V^2

$\log a$

A b