Sunday, April 9, 2023 6:26 PM

Curve Fitting

Example: Fit a straight line to the following data.

X:	1	2	3	4	6	8
Y:	2·4	3	3·6	4	5	6

Solution. Let the line be $Y = a + bX$

X	Y	X^2	XY
1	2·4	1	2·4
2	3·0	4	6·0
3	3·6	9	10·8
4	4·0	16	16·0
6	5·0	36	30·0
8	6·0	64	48·0
Total 24	24	130	113·2

Using normal equations (9.2*a*), we get

$24 = 6a + 24b$ and $113{\cdot}2 = 24a + 130b$

Solving these equations, we get $a = 1{\cdot}976$ and $b = 0{\cdot}506$.

$Y = a + bX \rightarrow$ Tot. cost to the comp. (in billions)

$\rightarrow$ No. of overtime hrs.

$Y \approx a + bX$

$X = 4.5$ hrs.

$X = 4$ hrs. 15 min.

$Y_i = a + bX_i$ —— ①

$\sum Y_i = \sum_{i=1}^{n} a + \sum bX_i$

$\sum Y_i = na + b\sum X_i$ —— ②

$\sum (X_i Y_i) = \sum (aX_i + bX_i^2)$

$\sum X_i Y_i = a\sum X_i + b\sum X_i^2$ —— ③

Poll Que. The correct option is:

(a) $Y = (1.976)X + 0.506$ (b) $Y = 1.976 + (0.506)X$ (c) $X = 1.976 + (0.506)Y$ (d) *None of these.*

$Y = 1.976 + (0.506)X = 2.482$

[Solve For a Variable Calculator - Symbolab]

Poll Que. The correct option is:

(a) $Y = (1.976)X + 0.506$ (b) $Y = 1.976 + (0.506)X$ (c) $X = 1.976 + (0.506)Y$ (d) *None of these*.

Example: Fit a parabola of second degree to the following data:

X:	0	1	2	3	4
Y:	1	1·8	1·3	2·5	6·3

$Y \approx a + bX + cX^2$

Solution. Let $Y = a + bX + cX^2$ be the second degree parabola.

	X	Y	X^2	X^3	X^4	XY	X^2Y
	0	1·0	0	0	0	0	0
	1	1·8	1	1	1	1·8	1·8
	2	1·3	4	8	16	2·6	5·2
	3	2·5	9	27	81	7·5	22·5
	4	6·3	16	64	256	25·2	100·8
Total	10	12·9	30	100	354	37·1	130·3

$Y_i = a + bX_i + cX_i^2$, $\sum_{i=1}^{n} a$

$\sum Y_i = na + b\sum X_i + c\sum X_i^2$ —— ①

$\sum X_i^2 Y_i = \sum (aX_i^2 + bX_i^3 + cX_i^4)$

$= a\sum X_i^2 + b\sum X_i^3 + c\sum X_i^4$

Using normal equations (9·4), we get

$12{\cdot}9 = 5a + 10b + 30c$; $37{\cdot}1 = 10a + 30b + 100c$; $\times 30$

$100 \times [\,130{\cdot}3 = 30a + 100b + 354c$

Solving these equations, we get $a = 1{\cdot}42$, $b = -\,1{\cdot}07$ and c $0{\cdot}55$. Thus the required equation of the second degree parabola is

$$Y = 1{\cdot}42 - 1{\cdot}07\,X + 0{\cdot}55X^2$$

$C' = aA' + bB'$

$C'' = aA'' + bB''$

Poll Que. The correct option is:

(a) $Y = 1.42 + (1.07)X - (0.55)X^2$ (b) $Y = 1.42 + (1.07)X + (0.55)X^2$

(c) $X = 1.42 - (1.07)Y + (0.55)Y^2$ (d) *None of these.*

Poll Que. The correct option is:

(a) $Y = 1.42 + (1.07)X - (0.55)X^2$ (b) $Y = 1.42 + (1.07)X + (0.55)X^2$

(c) $X = 1.42 - (1.07)Y + (0.55)Y^2$ (d) *None of these*.

Example: Fit an exponential curve of the form $Y = ab^X$ to thefollowing 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

X	Y	$U = \log Y$	XU	X^2
1	1·0	0·0000	0·0000	1
2	1·2	0·0792	0·1584	4
3	1·8	0·2553	0·7659	9
4	2·5	0·3979	1·5916	16
5	3·6	0·5563	2·7815	25
6	4·7	0·6721	4·0326	36
7	6·6	0·8195	5·7365	49
8	9·1	0·9590	7·6720	64
36	30·5	3·7393	22·7385	204

(9·11a) gives the normal equations as

$$3{\cdot}7393 = 8A + 36B$$

and
$$22{\cdot}7385 = 36A + 204B$$

Solving, we get

$B = 0{\cdot}1408$ and $A = -0{\cdot}1662 = \bar{1}{\cdot}8338$

$\therefore$ b = Antilog $B = 1{\cdot}383$ and a = Antilog $A = 0{\cdot}6821$

Hence the equation of the required curve is

$$Y = 0{\cdot}6821(1{\cdot}38)^X$$

$Y = ab^X$

$\log Y = \log(ab^X)$

$= \log a + \log b^X$

$= \log a + X(\log b)$

$\sum U_i = \sum(A + BX_i)$

$Y = A + BX$

$\sum(X_iU_i) = \sum(AX_i + BX_i^2)$

$A = -0.1662$ $A = \log a$

$\boxed{B = 0.1408}$ $a = 10^A$

$a = 10^{-0.1662}$

If $5^3 = 125$ $\log_5 125 = 3$.

$a = e^{-0.1662}$

Angle Between Two Lines of Regression

Poll Que: The angle between two Lines of Regression is given by:

(a) $\tan^{-1}\left[\left|\frac{1-r^2}{r}\right|\left(\frac{\sigma_X\sigma_Y}{\sigma_X^2+\sigma_Y^2}\right)\right]$ (b) $\tan^{-1}\left[\left|\frac{1-r^2}{r}\right|\left(\frac{\sigma_X^2\sigma_Y^2}{\sigma_X^2+\sigma_Y^2}\right)\right]$

(c) $\tan^{-1}\left[\left|\frac{1-r^2}{r}\right|\left(\frac{\sigma_X\sigma_Y}{\sigma_X+\sigma_Y}\right)\right]$ (d) *None of these.*

$Y - \bar{y} = \frac{Cov.(X,Y)}{\sigma_X^2}(X - \bar{x}) = \frac{r\sigma_Y}{\sigma_X}(X - \bar{x}) \rightarrow b_{YX} = m_1$ $b_{XY} = \frac{1}{m_2}$

$r = r(X,Y) = \frac{Cov.(X,Y)}{\sigma_X\sigma_Y}$

$X - \bar{x} = \frac{Cov.(X,Y)}{\sigma_Y^2}(Y - \bar{y})$

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

l_2

$$180^\circ - \theta = 145^\circ \qquad \theta = 35^\circ$$

$l_2 \to m_2$, $l_1 \to m_1$

$$\tan\theta = \left|\frac{m_1 - m_2}{1 + m_1 m_2}\right|$$

$$\tan\theta = \left|\frac{\frac{r\sigma_y}{\sigma_x} - \frac{\sigma_y}{r\sigma_x}}{1 + \frac{r\sigma_y}{\sigma_x}\times\frac{\sigma_y}{r\sigma_x}}\right| = \left|\frac{\frac{r^2\sigma_y - \sigma_y}{r\sigma_x}}{\frac{\sigma_x^2+\sigma_y^2}{\sigma_x^2}}\right|$$

$$= \left|\frac{(r^2-1)\sigma_y}{r\sigma_x} \times \frac{\sigma_x^2}{\sigma_x^2+\sigma_y^2}\right| = \left|\frac{r^2-1}{r}\right| \frac{\sigma_x\sigma_y}{\sigma_x^2+\sigma_y^2}$$

$$\tan\theta = \left|\frac{r^2-1}{r}\right| \frac{\sigma_x\sigma_y}{\sigma_x^2+\sigma_y^2} = \left|\frac{1-r^2}{r}\right| \frac{\sigma_x\sigma_y}{\sigma_x^2+\sigma_y^2}$$

$$|r^2-1| = |1-r^2|$$

Poll Que: If the two variables X and Y are uncorrelated then the lines of regression are:

(a) *parallel* (b) *perpendicular* (c) *coincide* (d) *None of these*

Poll Que: If the two variables X and Y are perfectly correlated $(r = \pm 1)$ then the lines of regression are:

(a) *parallel* (b) *perpendicular* (c) *coincide* (d) *None of these*

$$-1 \le r \le 1 \qquad r = \pm 1$$

$$r = \frac{Cov.(X,Y)}{\sigma_x \sigma_y} = \frac{0}{\sigma_x\sigma_y} = 0$$

$$\tan\theta = 0$$

$$\theta = 0^\circ \text{ or } 180^\circ$$

$$\tan\theta = \left|\frac{1-r^2}{r}\right| \frac{\sigma_x\sigma_y}{\sigma_x^2+\sigma_y^2} = \infty \qquad \theta = \frac{\pi}{2}$$

l_1

l_2

l_1, l_2

180°

l_1

$\longrightarrow l_1, l_2$

180°

$l_2 \longleftrightarrow l_1$

Theorem: ***Two independent variables are uncorrelated.***

$$r = \frac{Cov(X,Y)}{\sigma_X \sigma_Y} = \frac{0}{\sigma_X \sigma_Y} = 0$$

But the converse of the theorem is not true, *i.e.*, two uncorrelated variables may not be independent as the following example illustrates :

If X & Y are indep.

then $r(X,Y) = 0$

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

If $r(X,Y) = 0$ then

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

$$\therefore \qquad r(X, Y) = \frac{\text{Cov}(X, Y)}{\sigma_X \sigma_Y} = 0$$

X & Y are indep. X

$X = Y^2$

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

$r(X,Y)$

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