Fit a straight line for the following data :-

X	2	3	5	8	10
Y	5	6	10	18	21

Soln :-

$$\Sigma Y = na + b\Sigma X$$
$$\Sigma XY = a\Sigma X + b\Sigma X^2$$

X	Y	X^2	XY
2	5	4	10
3	6	9	18
5	10	25	50
8	18	64	144
10	21	100	210
28	60	202	432

$$na + b\Sigma X = \Sigma Y$$
$$a\Sigma X + b\Sigma X^2 = \Sigma XY$$

$$5a + 28b = 60 \quad \rightarrow (1)$$
$$28a + 202b = 432 \quad \rightarrow (2)$$

$(1) \times 28 \Rightarrow 140a + 784b = 1680$
$(2) \times 5 \Rightarrow \underset{(-)}{140a} \underset{(-)}{+} 1010b = \underset{(-)}{2160}$
$\cancel{-}226b = \cancel{-}480$

$$b = \frac{480}{226} = 2.1239$$

Mode - 2
Reg - $\boxed{2}$
Lin - $\boxed{1}$
2, 5 M+
3, 6 M+
Shift - $\boxed{2}$
$\leftarrow$ $\boxed{\rightarrow}$ A B r
1 2 3.
[illegible] A = $\boxed{1}$ = Ans
B = $\boxed{2}$ = Ans

To find :-
$\Sigma x, \Sigma y, \Sigma x^2$
Shift $\rightarrow \boxed{1}$
$\Sigma x^2 \rightarrow \boxed{1}$
$\Sigma x^2 =$

$$b = 2.1239$$

Sub b in (1)

$$5a + 28(2.1239) = 60$$
$$5a + 59.4692 = 60$$
$$5a = 60 - 59.4692$$
$$5a = 0.5308$$
$$a = \frac{0.5308}{5}$$
$$a = 0.1062$$

$$y = 0.1062 + 2.1239x.$$

Curve fitting:-

Let x_i, y_i $(i = 1, 2 \ldots n)$ be a given set of values, where x_i is independent variable and y_i is dependent variable. The problem of curve fitting is to find out analytic expression of the form $y = f(x)$.

It is useful in the study of correlation and regression. The relationship between two variables by simple algebraic expressions.

Example:- polynomial, exponential, log functions.

It may be used to estimate the values of one variable corresponds to another variable.

Fitting a straight line:-

The equation of straight line is

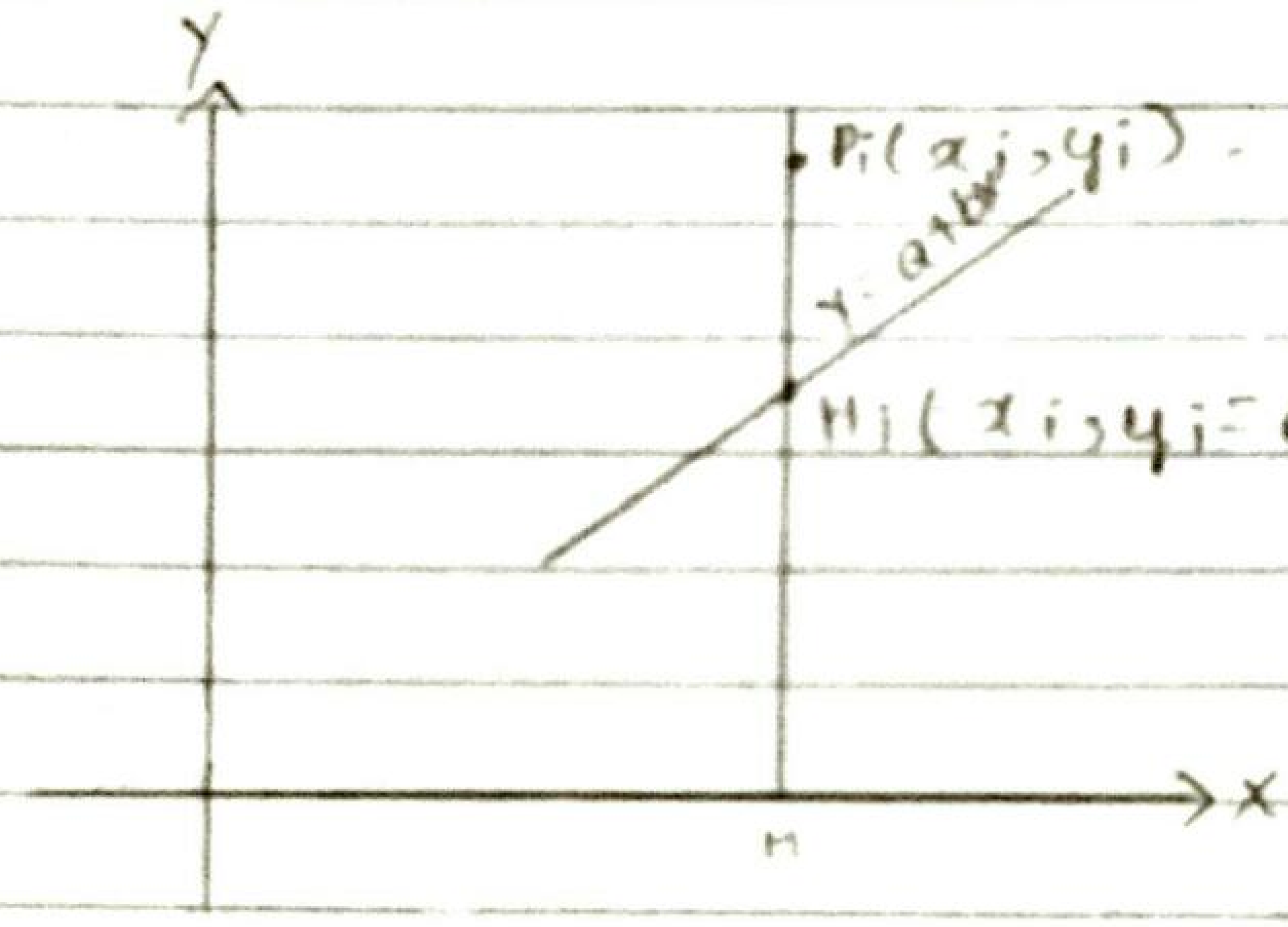
$$y = a + bx$$

where a and b are constants.

To fit a straight line for the given set of n points

$$(x_i, y_i)\ i = 1, 2 \ldots n.$$

The problem is to be determined the constants a and b show that the line is $y = a + bx$ the line of best fit.



Let $P_i(x_i, y_i)$ be any point in the above scattered diagram $P_i H_i$ is perpendicular to x-axis ~~to~~ meeting the line $y = a + bx$ in H_i.

Since H_i is on the line $y = a + bx$ the corresponding H_i are $(x_i, a + bx_i)$

$$P_iH_i = P_iM - H_iM$$

$$= y_i - (a + bx_i)$$

$$P_iH_i = y_i - a - bx_i$$

Scanned by TapScanner

It is called the error of estimate or the residual for y.

According to the principles of least squares we have to determine a and b.

$$\therefore E = \sum_{i=1}^{n} [PiHi]^2$$

$$[PiHi = y_i - a - bx_i]$$

$$= \sum_{i=1}^{n} (y_i - a - bx_i)^2$$

$\therefore$ E is minimum and equating to 0. The partial derivatives of E with respect to a and b. We get the normal equations

$$\frac{\partial E}{\partial a} = 0$$

$$E = \sum_{i=1}^{n} (y_i - a - bx_i)^2$$

$$\frac{\partial E}{\partial a} = 2\sum_{i=1}^{n} (y_i - a - bx_i)(-1)$$

$$-2\sum_{i=1}^{n} (y_i - a - bx_i) = 0$$

$$-2 \neq 0$$

$$\sum_{i=1}^{n} (y_i - a - bx_i) = 0$$

$$\sum_{i=1}^{n} y_i - na - b\sum_{i=1}^{n} x_i = 0$$

$$\sum_{i=1}^{n} y_i = na + b\sum_{i=1}^{n} x_i \rightarrow (1)$$

$$\frac{\partial E}{\partial b} = 0$$

$$E = \sum_{i=1}^{n} (y_i - a - bx_i)^2$$

$$\frac{\partial E}{\partial b} = 2\sum_{i=1}^{n} (y_i - a - bx_i)(-x_i)$$

$$\Rightarrow -2\sum_{i=1}^{n} x_i (y_i - a - bx_i) = 0$$

$$-2 \neq 0$$

$$\sum_{i=1}^{n} x_i y_i - a\sum_{i=1}^{n} x_i - b\sum_{i=1}^{n} x_i^2 = 0$$

$$\sum_{i=1}^{n} x_i y_i = a\sum_{i=1}^{n} x_i + b\sum_{i=1}^{n} x_i^2 \rightarrow (2)$$

Equation (1) and (2) are known as the normal equation for estimating a and b. By solving these two equations, we get the values a and b.

The equation of the form, $Y = a + bx$ the line of the best fit to the given set of points $(x_i, y_i)(i = 1, 2, \dots n)$.

Fitting a second degree parabola :-
(or polynomial)

Let $Y = a + bx + cx^2$ be the second degree parabola for the given set of n points x_i, y_i $(i = 1, 2, \dots n)$. Using the principles of least square, we have to determine the constant a, b and c. Show that $E = \sum_{i=1}^{n} (y_i - a - bx_i - cx_i^2)^2$ is minimum and equating to 0. The partial derivatives of E with respect to a, b and c we getting the normal equations.

$$\frac{\partial E}{\partial a} = 0$$

$$= 2 \sum_{i=1}^{n} (y_i - a - bx_i - cx_i^2)(-1)$$

$$-2 \sum_{i=1}^{n} (y_i - a - bx_i - cx_i^2) = 0$$

$$-2 \neq 0$$

$$\sum_{i=1}^{n} y_i - na - b \sum_{i=1}^{n} x_i - c \sum_{i=1}^{n} x_i^2 = 0$$

$$\sum_{i=1}^{n} y_i = na + b \sum_{i=1}^{n} x_i + c \sum_{i=1}^{n} x_i^2 \rightarrow (1)$$

$$\frac{\partial E}{\partial b} = 0$$

$$= 2 \sum_{i=1}^{n} (y_i - a - bx_i - cx_i^2)(-x_i)$$

$$-2x_i \sum_{i=1}^{n} (y_i - a - bx_i - cx_i^2) = 0$$

$$-2 \neq 0$$

$$\sum_{i=1}^{n} x_i y_i - a \sum_{i=1}^{n} x_i - b \sum_{i=1}^{n} x_i^2 - c \sum_{i=1}^{n} x_i^3 = 0$$

$$\sum_{i=1}^{n} x_i y_i = a \sum_{i=1}^{n} x_i + b \sum_{i=1}^{n} x_i^2 + c \sum_{i=1}^{n} x_i^3 \rightarrow (2)$$

Scanned by TapScanner

$$\frac{\partial E}{\partial c} = 0$$

$$= 2\sum_{i=1}^{n}(y_i - a - bx_i - cx_i^2)(-x_i^2)$$

$$-2x_i^2\sum_{i=1}^{n}(y_i - a - bx_i - cx_i^2) = 0$$

$$-2 \neq 0$$

$$\sum_{i=1}^{n} x_i^2 y_i - a\sum_{i=1}^{n} x_i^2 - b\sum_{i=1}^{n} x_i^3 - c\sum_{i=1}^{n} x_i^4 = 0$$

$$\sum_{i=1}^{n} x_i^2 y_i = a\sum_{i=1}^{n} x_i^2 + b\sum_{i=1}^{n} x_i^3 + c\sum_{i=1}^{n} x_i^4$$

$\longrightarrow (3).$

Equation (1), (2) and (3) are known as the normal equations for estimating a, b and c. For given set of n points $i = 1, 2 \ldots n$. The equations (1), (2) and (3) can be solved for a, b and c.

The equation $y = a + bx + cx^2$ is the parabola of best fit for the given set of points.

Power curve ($y = ax^b$)

The curve of the form is,

$$y = ax^b$$

Taking log on both side,

$$\log y = \log a + b \log x.$$

Let $U = \log y$, $A = \log a$, $V = \log x$.

$U = A + bV$

This is a linear equation in U and V.

The normal equations for estimating A and b.

$$\Sigma U = nA + b\Sigma V$$
$$\Sigma UV = A\Sigma V + b\Sigma V^2$$

The equation can be solved for estimating A and b, where,

a - Antilog of A.

With the values of a and b, we get the lines $y = ax^b$ is the curve of the best fit for the set of n points.

Fitting of exponential curve

(i) $y = ab^x$ (ii) $y = ae^{bx}$

(i) $y = ab^x$

Taking log on both sides we get,

$$\log y = \log a + x \log b$$
$$U = A + Bx$$

where

$$\log y = U$$
$$\log a = A$$
$$\log b = B$$

$Y = ab^x$ This is a linear equation in x and U. The normal equation for estimating A and B is

$$\Sigma U = nA + B\Sigma x$$
$$\Sigma xU = A\Sigma x + B\Sigma x^2$$

Solving these equations, we get the values of A & B,
where

$a = \text{Antilog}(A)$
$b = \text{Antilog}(B)$

with this values of a and b,
$Y = ab^x$ is the curve of best fit for the given set of n points.

(ii) $y = ae^{bx}$

Taking log on both sides we get,

$$\log y = \log a + bx \log e$$
$$\log y = \log a + x(b \log e)$$
$$U = A + Bx$$

where

$U = \log y$ $\qquad$ $\log e = 0.4343$
$A = \log a$
$B = b \log e$

This is a linear equation in x & V. The normal equation for estimating A & B are

$$\Sigma U = nA + B\Sigma x$$
$$\Sigma xU = A\Sigma x + B\Sigma x^2$$

Then finding A & B consequently

Scanned by TapScanner

$a = \text{Antilog } A$

$b = B/\log e$ with these values of a & b, $y = ae^{bx}$ is the curve of the best fit to the given set of "n" points.

Fit a straight line to the following data by the method of least squares.

x	1	2	3	4	5
y	1.8	5.1	9	14	19.

Soln :-

Equation of the straight line

$$y = a + bx$$

$$na + b\Sigma x = \Sigma y$$

$\times x$, $$a\Sigma x + b\Sigma x^2 = \Sigma xy$$

x	y	x^2	xy
1	1.8	1	1.8
2	5.1	4	10.2
3	9	9	27
4	14	16	56
5	19	25	95
15	48.9	55	190

$$5a + 15b = 48.9 \rightarrow (1)$$

$$15a + 55b = 190 \rightarrow (2)$$

Scanned by TapScanner

$(1) \times 3 \Rightarrow 15a + 45b = 146.7$

$$\begin{array}{r} \overset{(-)}{15a} \overset{(-)}{+} 55b = \overset{(-)}{+}190 \\ \hline -10b = -43.3 \end{array}$$

$$b = \frac{43.3}{10} = 4.33$$

Sub $b = 4.33$ in (1)

$$5a + 15(4.33) = 48.9$$

$$5a + 64.95 = 48.9$$

$$5a = 48.9 - 64.95$$

$$5a = -16.05$$

$$a = \frac{-16.05}{5}$$

$$a = -3.21$$

$$y = -3.21 + 4.33x$$

Fitting a second degree parabola (Polynomial) for the following data:-

x	1	2	3	4	5	6	7
y	2.3	5.2	9.7	16.5	29.4	35.5	54.4

Soln:-

Mode - 2 times

Reg - 2

Quad - 3

$$y = a + bx + cx^2$$

$$\Sigma y = na + b\Sigma x + c\Sigma x^2$$

$$\Sigma xy = a\Sigma x + b\Sigma x^2 + c\Sigma x^3$$

$$\Sigma x^2y = a\Sigma x^2 + b\Sigma x^3 + c\Sigma x^4.$$

x	y	x^2	x^3	x^4	xy	x^2y
1	2.3	1	1	1	2.3	2.3
2	5.2	4	8	16	10.4	20.8
3	9.7	9	27	81	29.1	87.3
4	16.5	16	64	256	66	264
5	29.4	25	125	625	147	735
6	35.5	36	216	1296	213	1278
7	54.4	49	343	2401	380.8	2665.6
28	153	140	784	4676	8486	5053

$$7a + 28b + 140c = 157 \rightarrow (1)$$
$$28a + 140b + 784c = 848.6 \rightarrow (2)$$
$$140a + 784b + 4676c = 5053 \rightarrow (3)$$

(1) x 2

$(1) \times 4 \Rightarrow 28a + 112b + 560c = 612$

$(2) \Rightarrow \underline{28a + 140b + 784c = 848.6}$

$-28b - 224c = -236.6$

$28b + 224c = 236.6 \rightarrow (4)$

(2) & (3)

$(2) \times 5 \Rightarrow 140a + 700b + 3920c = 4243$

$(3) \Rightarrow \underline{140a + 784b + 4676c = 5053}$

$-84b - 756c = -810$

$84b + 756c = 810 \rightarrow (5)$

$(4) \times 3 \Rightarrow 84b + 672c = 709.8$

$(5) \Rightarrow \underline{84b + 756c = 810}$

$-84c = -100.2$

$c = \dfrac{100.2}{84}$

$c = 1.1929$

Sub $c = 1.1929$ in (4)

$28b + 224(1.1929) = 236.6$

$28b + 267.2096 = 236.6$

$28b = 236.6 - 267.2096$

$28b = -30.6096$

$b = -\dfrac{30.6096}{28}$

Scanned by TapScanner

$b = -1.0932$

Sub b in (1)

$7a + 28(-1.0932) + 140(1.1929) = 153$

$7a - 30.6096 + 167.006 = 153$

$7a + 136.3964 = 153$

$7a = 16.6036$

$a = \frac{16.6036}{7}$

$a = 2.3719$

The equation is

$Y = 2.3719 - 1.0932x + 1.1929x^2$.

= 97.7608.

Fit a power curve to the following data.

X	1	2	3	4	5	6	7
Y	150	263	310	385	427	504	612

Procedure :-

Let the equation of the power curve is $Y = a x^b$.

Taking log on both sides,

$$\log y = \log a + b \log x.$$

where,

$$U = \log y$$
$$A = \log a$$
$$V = \log x$$
$$U = A + bV$$

This is the linear equation in U & V.

The normal equations are,

$$\Sigma U = nA + b\Sigma V$$

$$\Sigma UV = A\Sigma V + b\Sigma V^2$$

where,

$$a = \text{Antilog } A$$

solving these equations & substituting the values of a & b in $y = ax^b$, we get the required equation of power curve.

x	y	V=log x	U=log y	UV	V^2
1	150	0	2.1761	0	0
2	263	0.3010	2.4200	0.7284	0.0906
3	310	0.4771	2.4914	1.1886	0.2276
4	385	0.6021	2.5855	1.5567	0.3625
5	427	0.6990	2.6304	1.8386	0.4886
6	504	0.7782	2.7024	2.1030	0.6056
7	612	0.8451	2.7868	2.3551	0.7142
28	2651	3.7025	17.7926	9.7704	2.4891

$$\Sigma U = nA + b\Sigma V$$

$$\Sigma UV = A\Sigma V + b\Sigma V^2$$

$$7A + 3.7025b = 17.7926 \rightarrow (1)$$

$$3.7025A + 2.4891b = 9.7704 \rightarrow (2)$$

(1)&(2)

$(1) \times 3.7025 \Rightarrow 25.9175A + 13.7085b = 65.8771$

$(2) \times 7 \Rightarrow 25.9175A + 17.4237b = 68.3928$

$$-3.7152b = -2.5157$$

$$b = \frac{2.5157}{3.7152}$$

$$b = 0.6771$$

Sub $b = 0.6771$ in (1)

$(1) \Rightarrow 7A + 3.7025b = 17.7926$

$7A + 3.7025(0.6771) = 17.7926$

$7A + 2.5070 = 17.7926$

$7A = 17.7926 - 2.5070$

$A = \frac{15.2856}{7}$

$A = 2.1837$

$a = \text{Antilog } A$

$= \text{Antilog}(2.1837)$

$a = 152.6511.$

The equation of the power curve is

$y = 152.6511\, x^{0.6771}.$

Fit an exponential curve of the form $y = ab^x$.

x	5	6	7	8	9	10
y	120	150	175	230	260	210.

Procedure :-

Let the equation of an exponential curve is $y = ab^x$.

Taking log on both sides,

$\log y = \log a + x \log b$

where,

$U = \log y \; ; A = \log a ; B = \log b$

$U = A + Bx$

Normal equation,

$\Sigma U = nA + B\Sigma x$

$\Sigma xU = A\Sigma x + B\Sigma x^2.$

x	y	U = log y	Ux	x^2
5	120	2.0792	10.396	25
6	150	2.1761	13.0566	36
7	175	2.2430	15.701	49
8	230	2.3617	18.8936	64
9	260	2.4150	21.735	81
10	210	2.3222	23.222	100
45		13.5972	103.0042	355

$$\Sigma U = nA + B\Sigma x.$$
$$\Sigma xU = A\Sigma x + B\Sigma x^2.$$

$$6A + 45B = 13.5972 \rightarrow (1)$$
$$45A + 355B = 103.0042 \rightarrow (2)$$

(1) & (2)

$(1) \times 45 \Rightarrow 270A + 2025B = 611.874$

$(2) \times 6 \Rightarrow 270A + 2130B = 618.0252$

$$-105B = -6.1512$$
$$B = \frac{6.1512}{105}$$
$$B = 0.0586.$$

Sub $B = 0.0586$ in (1)

$(1) \Rightarrow 6A + 45B = 13.5972$

$$6A + 45(0.0586) = 13.5972$$
$$6A + 2.637 = 13.5972$$
$$6A = 13.5972 - 2.637$$
$$A = \frac{10.9602}{6}$$

$A = 1.8267$

$a = \text{Antilog } A$

$a = \text{Antilog } (1.8267)$

$a = 67.0965$

$b = \text{Antilog } B$

$b = \text{Antilog } (0.0586)$

$= 1.1445$

$Y = (67.0965)(1.1445)^x$.