CHAPTER 24

Empirical Laws and Curve-fitting

24.1 INTRODUCTION

In many branches of applied mathematics, it is required to express a given data, obtained from observations, in the form of a *Law* connecting the two variables involved. Such a *Law* inferred by some scheme is known as *Empirical Law*. For example, it may be desired to obtain the law connecting the length and the temperature of a metal bar. At various temperatures, the length of the bar is measured. Then, by one of the methods explained below, a law is obtained that represents the relationship existing between temperature and length for the observed values. This relation can then be used to predict the length at an arbitrary temperature.

(2) Scatter diagram. To find a relationship between the set of paired observations x and y (say), we plot their corresponding values on the graph taking one of the variables along the x-axis and other along the y-axis *i.e.* (x_1, y_1), (x_2, y_2), (x_n, y_n). The resulting diagram showing a collection of dots is called a *scatter diagram*. A smooth curve that approximates the above set of points is known as the *approximating curve*.

(3) Curve fitting. Several equations of different types can be obtained to express the given data approximately. But the problem is to find the equation of the curve of '*best fit*' which may be most suitable for predicting the unknown values. The process of finding such an equation of 'best fit' is known as *curve-fitting*.

If there are n pairs of observed values then it is possible to fit the given data to an equation that contains n arbitrary constants for we can solve n simultaneous equations for n unknowns. If it were desired to obtain an equation representing these data but having less than n arbitrary constants, then we can have recourse to any of the four methods : *Graphical method, Method of Least squares, Method of Group averages* and *Method of Moments*. The graphical method fails to give the values of the unknowns uniquely and accurately while the other methods do. *The method of Least squares is, probably, the best to fit a unique curve to a given data.* It is widely used in applications and can be easily implemented on a computer.

24.2 GRAPHICAL METHOD

When the curve representing the given data is a **linear law** $y = mx + c$ *; we proceed as follows :*

(*i*) *Plot the given points on the graph paper* taking a suitable *scale*.

(*ii*) *Draw the straight line of best fit* such that the points are evenly distributed about the line.

(*iii*) *Taking two suitable points* (x_1, y_1) *and* (x_2, y_2) *on the line, calculate m,* the slope of the line and c, its intercept on y-axis.

When the points representing the observed values do not approximate to a straight line, a smooth curve is drawn through them. From the shape of the graph, we try to infer the law of the curve and then reduce it to the form $y = mx + c$.

24.3 LAWS REDUCIBLE TO THE LINEAR LAW

We give below some of the laws in common use, indicating the way these can be reduced to the linear form by suitable substitutions :

(1) *When the law is* $y = mx^n + c$.

Taking $x^n = X$ and $y = Y$ the above law becomes $Y = mX + c$

(2) *When the law is* $y = ax^n$.

Taking logarithms of both sides, it becomes $\log_{10} y = \log_{10} a + n \log_{10} x$

Putting $\log_{10} x = X$ and $\log_{10} y = Y$, it reduces to the form $Y = nX + c$, where $c = \log_{10} a$.

(3) *When the law is* $y = ax^n + b \log x$.

Writing it as $\dfrac{y}{\log x} = a\dfrac{x^n}{\log x} + b$ and taking $x^n/\log x = X$ and $y/\log x = Y$,

the given law becomes, $Y = aX + b$.

(4) *When the law is* $y = ae^{bx}$

Taking logarithms, it becomes $\log_{10} y = (b \log_{10} e)\, x + \log_{10} a$

Putting $x = X$ and $\log_{10} y = Y$, it takes the form $Y = mX + c$ where $m = b \log_{10} e$ and $c = \log_{10} a$.

(5) *When the law is* $xy = ax + by$.

Dividing by x, we have $y = b\dfrac{y}{x} + a$.

Putting $y/x = X$ and $y = Y$, it reduces to the form $Y = bX + a$.

Example 24.1. *R is the resistance to maintain a train at speed V; find a law of the type* $R = a + bV^2$ *connecting R and V, using the following data :*

V (miles/hour) :	10	20	30	40	50
R (lb/ton) :	8	10	15	21	30

Solution. Given law is $R = a + bV^2$ ...(*i*)

Taking $V^2 = x$ and $R = y$, (*i*) becomes

$$y = a + bx \qquad \text{...}(ii)$$

which is a *linear law*.

Table for the values of x and y is as follows :

x	100	400	900	1600	2500
y	8	10	15	21	30

Plot these points. Draw the straight line of best fit through these points (Fig. 24.1)

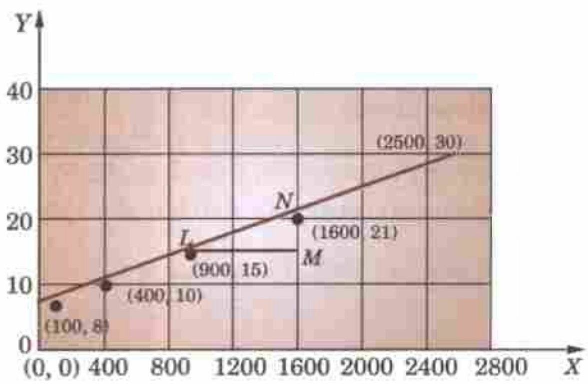

Fig. 24.1

Slope of this line (= b)

$$= \frac{MN}{LM} = \frac{21-15}{1600-900} = \frac{6}{700} = 0.0085 \text{ nearly.}$$

Since L (900, 15) lies on (*ii*),

$\therefore$ $15 = a + 0.0085 \times 900$,

whence $a = 15 - 7.65 = 7.35$ nearly.

Example 24.2. *The following values of x and y are supposed to follow the law* $y = ax^2 + b\log_{10} x$. *Find graphically the most probable values of the constants a and b.*

x	2.85	3.88	4.66	5.69	6.65	7.77	8.67
y	16.7	26.4	35.1	47.5	60.6	77.5	93.4

Solution. Given law is $y = ax^2 + b\log_{10} x$

i.e. $$\frac{y}{\log_{10} x} = a\frac{x^2}{\log_{10} x} + b \quad ...(i)$$

Taking $x^2/\log_{10} x = X$ and $y/\log_{10} x = Y$

(i) becomes $Y = aX + b$ ...(ii)

This is a *linear law*. Table for the values of X and Y is as follows :

$X = x^2/\log_{10} x$	17.93	25.56	32.49	42.87	53.75	67.80	80.83
$Y = y/\log_{10} x$	35.59	44.83	52.50	62.90	73.65	87.04	99.56
Points	P_1	P_2	P_3	P_4	P_5	P_6	P_7

Plot these points and draw the straight line of best fit through these points (Fig. 24.2).

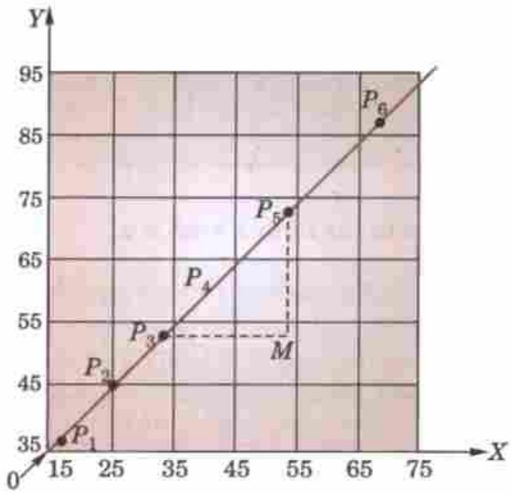

Fig. 24.2

Slope of this line (= a) = $\dfrac{MP_5}{P_3M} = \dfrac{73.65 - 52.50}{53.75 - 32.49} = \dfrac{21.15}{21.26} = 0.99$

Since P_3 lies on (ii), therefore, $52.50 = 0.99 \times 32.49 + b$ whence $b = 20.2$

Hence (i) becomes $y = (0.99)\,x^2 + (20.2)\log_{10} x$.

Example 24.3. *The values of x and y obtained in an experiment are as follows :*

x	2.30	3.10	4.00	4.92	5.91	7.20
y	33.0	39.1	50.3	67.2	85.6	125.0

Th probable law is $y = ae^{bx}$. Test graphically the accuracy of this law and if the law holds good, find the best values of the constants.

Solution. Given law is $y = ae^{bx}$ ...(i)

Taking logarithms to base 10, we have $\log_{10} y = \log_{10} a + (b\log_{10} e)\,x$

Putting $x = X$ and $\log_{10} y = Y$, it becomes $y = (b\log_{10} e)\,X + \log_{10} a$ ...(ii)

Table for the values of X and Y is as under :

$X = x$	2.30	3.10	4.00	4.92	5.91	7.20
$Y = \log_{10} y$	1.52	1.59	1.70	1.83	1.93	2.1
Points	P_1	P_2	P_3	P_4	P_5	P_6

Scale : 1 small division along x–axis = 0.1

10 small divisions along y–axis = 0.1.

Plot these points and draw the line of best fit. As these points are lying almost along a straight line, the given law is nearly accurate (Fig. 24.3).

Now slope of this line ($= b \log_{10} e$)

$$= \frac{MN}{NM} = 0.12$$

whence $$b = \frac{0.12}{\log_{10} e} = 0.12 \times 2.303 = 0.276$$

Since the point L (4, 1.71) lies on (*ii*), therefore, $1.71 = 0.12 \times 4 + \log_{10} a$ whence $a = 17$ nearly.

Hence the curve of best fit is $y = 17\, e^{0.276x}$.

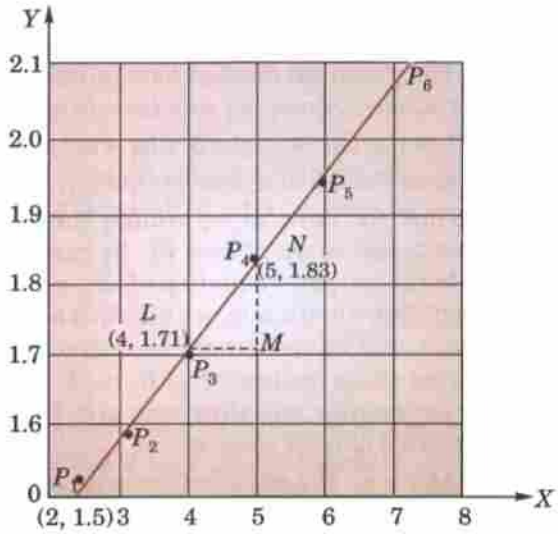

Fig. 24.3

PROBLEMS 24.1

1. If p is the pull required to lift the weight by means of a pulley block, find a linear law of the form $p = a + bw$, connecting p and w, using the following data :

w (lb) :	50	70	100	120
p (lb) :	12	15	21	25

Compute p, when w = 150 lb.

2. The resistance R of a carbon filament lamp was measured at various values of the voltage V and the following observations were made :

Voltage	V...	62	70	78	84	92
Resistance	R...	73	70.7	69.2	67.8	66.3

Assuming a law of the form $R = \dfrac{a}{V} + b$, find by graphical method the best value of a and b.

3. Verify if the values of x and y, related as shown in the following table, obey the law $y = a + b\sqrt{x}$. If so, find graphically the values of a and b.

x :	500	1,000	2,000	4,000	6,000
y :	0.20	0.33	0.38	0.45	0.51

4. The following values of T and l follow the law $T = al^n$. Test if this is so and find the best values of a and n.

$T = 1.0$	1.5	2.0	2.5
$l = 25$	56.2	100	1.56

5. Find the best value of a and b if $y = ax + b \log_{10} x$ is the curve which represents most closely the observed values given below :

x :	2	3	4	5	6
y :	3.74	5.99	7.47	8.92	9.86

6. Fit the curve $y = ae^{bx}$ to the following data :

x :	0	2	4
y :	5.1	10	31.1

(*Coimbatore, 1997*)

7. The following are the results of an experiment on friction of bearings. The speed being constant, corresponding values of the coefficient of friction and the temperature are shown in the table :

t :	120	110	100	90	80	70	60
μ :	0.0051	0.0059	0.0071	0.0085	0.00102	0.00124	0.00148

If μ and t are given by the law $\mu = ae^{bt}$, find the values of a and b by plotting the graph for μ and t.

24.4 PRINCIPLE OF LEAST SQUARES

The graphical method has the obvious drawback of being unable to give a unique curve of fit. *The principle of least squares, however, provides an elegant procedure for fitting a unique curve to a given data.*

Let the curve, $y = a + bx + cx^2 + \ldots + kx^{m-1}$ ...(1)

be fitted to the set of n data points $(x_1, y_1), (x_2, y_2) \ldots (x_n, y_n)$.

Now we have to determine the constants $a, b, c, \ldots k$ such that it represents the curve of best fit. In case $n = m$, on substituting the values (x_i, y_i) in (1), we get n equations from which a unique set of n constants can be found. But when $n > m$, we obtain n equations which are more than the m constants and hence cannot be solved for these constants. So we try to determine those values of $a, b, c, \ldots k$ which satisfy all the equations as nearly as possible and thus may give the best fit. In such cases, we apply the *principle of least squares*.

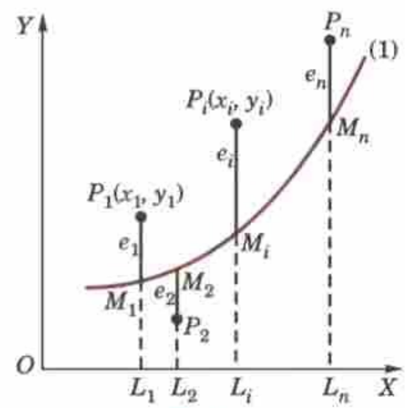

Fig. 24.4

At $x = x_i$, the *observed* (or *experimental*) *value* of the ordinate is $y_i = P_iL_i$ and the corresponding value on the fitting curve (1) is $a + bx_i + cx_i^2 + \ldots + kx_i^m = M_iL_i$ ($= \eta_i$, say) which is the *expected* (or *calculated*) *value* (Fig. 24.4). The difference of the observed and the expected values *i.e.* $y_i - \eta_i (= e_i)$ is called the *error* (or *residual*) at $x = x_i$. Clearly some of the errors $e_1, e_2, \ldots e_n$ will be positive and others negative. Thus to give equal weightage to each error, we square each of these and form their sum *i.e.* $E = e_1^2 + e_2^2 + \ldots e_n^2$.

The curve of best fit is that for which e's are as small as possible i.e., E, the sum of the squares of the errors is a minimum. This is known as the *principle of least squares* and was suggested by *Legendre** in 1806.

Obs. The principle of least squares does not help us to determine the form of the appropriate curve which can fit a given data. It only determines the best possible values of the constants in the equation when the form of the curve is known before hand. The selection of the curve is a matter of experience and practical considerations.

24.5 (1) METHOD OF LEAST SQUARES

For clarity, suppose it is required to fit the curve

$$y = a + bx + cx^2 \qquad \ldots(1)$$

to a given set of observations $(x_1, y_1), (x_2, y_2), \ldots, (x_5, y_5)$. For any x_i, the observed value is y_i and the expected value is $\eta_i = a + bx_i + cx_i^2$ so that the error $e_i = y_i - \eta_i$.

$\therefore$ the sum of the squares of these errors is

$$E = e_1^2 + e_2^2 + \ldots + e_5^2$$
$$= [y_1 - (a + bx_1 + cx_1^2)]^2 + [y_2 - (a + bx_2 + cx_2^2)]^2 + \ldots + [y_5 - (a + bx_5 + cx_5^2)]^2 \qquad \text{[See § 5.12 (3)]}$$

For E to be minimum, we have

$$\frac{\partial E}{\partial a} = 0 = 2[y_1 - (a + bx_1 + cx_1^2)] - 2[y_2 - (a + bx_2 + cx_2^2)] - \ldots - 2[y_5 - (a + bx_5 + cx_5^2)] \qquad \ldots(2)$$

$$\frac{\partial E}{\partial b} = 0 = -2x_1[y_1 - (a + bx_1 + cx_1^2)] - 2x_2[y_2 - (a + bx_2 + cx_2^2)] - \ldots - 2x_5[y_5 - (a + bx_5 + cx_5^2)] \qquad \ldots(3)$$

$$\frac{\partial E}{\partial c} = 0 = -2x_1^2[y_1 - (a + bx_1 + cx_1^2)] - 2x_2^2[y_2 - (a + bx_2 + cx_2^2)] - \ldots - 2x_5^2[y_5 - (a + bx_5 + cx_5^2)] \qquad \ldots(4)$$

Equation (2) simplifies to

$$y_1 + y_2 + \ldots + y_5 = 5a + b(x_1 + x_2 + \ldots + x_5) + c(x_1^2 + x_2^2 + \ldots + x_5^2)$$

i.e.,
$$\Sigma y_i = 5a + b\Sigma x_i + c\,\Sigma x_i^2 \qquad \ldots(5)$$

* See footnote on p. 311.

Equation (3) becomes

$$x_1y_1 + x_2y_2 + \dots + x_5y_5 = a(x_1 + x_2 + \dots + x_5) + b(x_1^2 + x_2^2 + \dots + x_5^2) + c(x_1^3 + x_2^3 + \dots + x_5^3)$$

i.e., $$\Sigma x_i y_i = a\Sigma x_i + b\Sigma x_i^2 + c\Sigma x_i^3 \quad \text{...(6)}$$

Similarly (4) simplifies to $\Sigma x_i^2 y_i = a\Sigma x_i^2 + b\Sigma x_i^3 + c\Sigma x_i^4$ ...(7)

The equations (5), (6) and (7) are known as *Normal equations* and can be solved as simultaneous equations in a, b, c. The values of these constants when substituted in (1) give the desired curve of best fit.

(2) Working procedure

(a) **To fit the straight line $y = a + bx$**

(i) Substitute the observed set of n values in this equation.

(ii) Form normal equations for each constant

i.e., $$\Sigma y = na + b\Sigma x,\ \Sigma xy = a\Sigma x + b\Sigma x^2$$

[The normal equation for the unknown a is obtained by multiplying the equations by the coefficient of a and adding. The normal equation for b is obtained by multiplying the equations by the coefficient of b (*i.e.*, x) and adding.]

(iii) Solve these normal equations as simultaneous equations for a and b.

(iv) Substitute the values of a and b in $y = a + bx$, which is the required line of best fit.

(b) **To fit the parabola : $y = a + bx + cx^2$**

(i) Form the normal equations $\Sigma y = na + b\Sigma x + c\Sigma x^2$

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

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

[The normal equation for c has been obtained by multiplying the equations by the coefficient of c (*i.e.*, x^2) and adding.]

(ii) Solve these as simultaneous equations for a, b, c.

(iii) Substitute the values of a, b, c in $y = a + bx + cx^2$, which is the required parabola of best fit.

(c) In general, the curve $y = a + bx + cx^2 + \dots + kx^{m-1}$ can be fitted to a given data by writing m normal equations.

Example 24.4. *If P is the pull required to lift a load W by means of a pulley block, find a linear law of the form P = mW + c connecting P and W, using the following data :*

P = 12	15	21	25
W = 50	70	100	120

where P and W are taken in kg-wt. Compute P when W = 150 kg. wt. (U.P.T.U., 2007 ; V.T.U., 2002)

Solution. The corresponding normal equations are

$$\left.\begin{aligned}\Sigma P &= 4c + m\Sigma W\\ \Sigma WP &= c\Sigma W + m\Sigma W^2\end{aligned}\right\} \quad \text{...}(i)$$

The values of ΣW etc. are calculated by means of the following table :

W	P	W^2	WP
50	12	2500	600
70	15	4900	1050
100	21	10000	2100
120	25	14400	3000
Total = 340	73	31800	6750

$\therefore$ The equations (i) becomes $73 = 4c + 340m$ and $6750 = 340c + 31800m$

i.e., $$2c + 170m = 365 \quad \text{...}(ii)$$

and $$34c + 3180m = 675 \quad \text{...}(iii)$$

Multiplying (ii) by 17 and substracting from (iii), we get

$$m = 0.1879 \qquad \therefore \text{ from } (ii),\quad c = 2.2785$$

Hence the line of best fit is

$$P = 2.2759 + 0.1879\ W$$

When $W = 150$ kg., $P = 2.2785 + 0.1879 \times 150 = 30.4635$ kg.

Obs. The calculations get simplified when the central values of x is zero. It is therefore, advisable to make the central value zero, if it be not so. This is illustrated by the next example.

Example 24.5. *Fit a second degree parabola to the following data :*

x	0	1	2	3	4
y	1	1.8	1.3	2.5	6.3

(P.T.U., 2006)

Solution. Let $u = x - 2$ and $v = y$ so that the parabola of fit $y = a + bx + cx^2$ becomes

$$v = A + Bu + Cu^2 \qquad ...(i)$$

The normal equations are

$$\Sigma v = 5A + B\Sigma u + C\Sigma u^2 \quad \text{or} \quad 12.9 = 5A + 10C$$
$$\Sigma uv = A\Sigma u + B\Sigma u^2 + C\Sigma u^2 \quad \text{or} \quad 11.3 = 10B$$
$$\Sigma u^2 v = A\Sigma u^2 + B\Sigma u^3 + C\Sigma u^4 \quad \text{or} \quad 33.5 = 10A + 34C$$

Solving these as simultaneous equations, we get

$$A = 1.48, \qquad B = 1.13, \qquad C = 0.55.$$

$\therefore$ (i) becomes, $\qquad v = 1.48 + 1.13u + 0.55u^2$

or $\qquad y = 1.48 + 1.13\ (x - 2) + 0.55\ (x - 2)^2$

Hence $y = 1.42 - 1.07x + 0.55x^2$.

Example 24.6. *Fit a second degree parabola to the following data :*

$x = 1.0$	1.5	2.0	2.5	3.0	3.5	4.0
$y = 1.1$	1.3	1.6	2.0	2.7	3.4	4.1

(V.T.U., 2009 ; Bhopal, 2008)

Solution. We shift the origin to (2.5, 0) and take 0.5 as the new unit. This amounts to changing the variable x to X, by the relation $X = 2x - 5$.

Let the parabola of fit be $y = a + bX + cX^2$. The values of ΣX etc., are calculated as below :

x	X	y	Xy	X^2	X^2y	X^3	X^4
1.0	–3	1.1	–3.3	9	9.9	–27	81
1.5	–2	1.3	–2.6	4	5.2	–8	16
2.0	–1	1.6	–1.6	1	1.6	–1	1
2.5	0	2.0	0.0	0	0.0	0	0
3.0	1	2.7	2.7	1	2.7	1	1
3.5	2	3.4	6.8	4	13.6	8	16
4.0	3	4.1	12.3	9	36.9	27	81
Total	0	16.2	14.3	28	69.9	0	196

The normal equations are

$$7a + 28c = 16.2\ ; \qquad 28b = 14.3\ ; \qquad 28a + 196c = 69.9$$

Solving these as simultaneous equations, we get

$$a = 2.07,\ b = 0.511,\ c = 0.061$$

$\therefore$ $$y = 2.07 + 0.511X + 0.061X^2$$

Replacing X by $2x - 5$ in the above equation, we get

$$y = 2.07 + 0.511\ (2x - 5) + 0.061\ (2x - 5)^2$$

which simplifies to $y = 1.04 - 0.198x + 0.244x^2$. This is the required parabola of best fit.

Example 24.7. *Fit a second degree parabola to the following data :*

x	1989	1990	1991	1992	1993	1994	1995	1996	1997
y	352	356	357	358	360	361	361	360	359

(*U.P.T.U., 2009*)

Solution. Taking $u = x - 1993$ and $v = y - 357$, the equation $y = a + bx + cx^2$ becomes

$$v = A + Bu + Cu^2 \qquad \text{...}(i)$$

x	$u = x - 1993$	y	$v = y - 357$	uv	u^2	u^2v	u^2	u^4
1989	−4	352	−5	20	16	−80	−64	256
1990	−3	360	−1	3	9	−9	−27	81
1991	−2	357	0	0	4	0	−8	16
1992	−1	358	1	−1	1	1	−1	1
1993	0	360	3	0	0	0	0	0
1994	1	361	4	4	1	4	1	1
1995	2	361	4	8	4	16	8	16
1996	3	360	3	9	9	27	27	81
1997	4	359	2	8	16	32	64	256
Total	$\Sigma u = 0$		$\Sigma v = 11$	$\Sigma uv = 51$	$\Sigma u^2 = 60$	$\Sigma u^2 v = -9$	$\Sigma u^3 = 0$	$\Sigma u^4 = 708$

The normal equations are

$$\Sigma v = 9A + B\Sigma u + C\Sigma u^2 \quad \text{or} \quad 11 = 9A + 60C$$

$$\Sigma uv = A\Sigma u + B\Sigma u^2 + C\Sigma u^3 \quad \text{or} \quad 51 = 60B \quad \text{or} \quad B = \frac{17}{20}$$

$$\Sigma u^2 v = A\Sigma u^2 + B\Sigma u^3 + C\Sigma u^4 \quad \text{or} \quad -9 = 60A + 708C$$

On solving these equations, we get $A = \dfrac{694}{231}, B = \dfrac{17}{20}, C = -\dfrac{247}{924}$

$\therefore$ (i) becomes $\qquad v = \dfrac{694}{231} + \dfrac{17}{20}u - \dfrac{247}{924}u^2$

or $\qquad y - 357 = \dfrac{694}{231} + \dfrac{17}{20}(x - 1933) - \dfrac{247}{924}(x - 1933)^2$

or $\qquad y = \dfrac{694}{231} - \dfrac{32861}{20} - \dfrac{247}{924}(1933)^2 + \dfrac{17}{20}x + \dfrac{247 \times 3866}{924}x - \dfrac{247}{924}x^2$

or $\qquad y = 3 - 1643.05 - 998823.36 + 357 + 0.85x + 1033.44x - 0.267x^2$

Hence $\qquad y = -1000106.41 + 1034.29x - 0.267x^2.$

PROBLEMS 24.2

1. By the method of least squares, find the straight line that best fits the following data :

x :	1	2	3	4	5
y :	14	27	40	55	68

(*U.P.T.U., 2008*)

2. Fit a straight line to the following data :

Year x :	1961	1971	1981	1991	2001
Production y :	8	10	12	10	16

(in thousand tons)

and find the expected production in 2006.

3. A simply supported beam carries a concentrated load P (lb) at its mid-point. Corresponding to various values of P, the maximum deflection Y (in) is measured. The data are given below :

P:	100	120	140	160	180	200
Y:	0.45	0.55	0.60	0.70	0.80	0.85

Find a law of the form $Y = a + bP$.

4. The results of measurement of electric resistance R of a copper bar at various temperatures t°C are listed below :

t :	19	25	30	36	40	45	50
R:	76	77	79	80	82	83	85

Find a relation $R = a + bt$ where a and b are constants to be determined by you.

5. Find the best possible curve of the form $y = a + bx$, using method of least squares for the data :

x :	1	3	4	6	8	9	11	14
y:	1	2	4	4	5	7	8	9

(*V.T.U., 2011*)

6. Fit a straight line to the following data

(*a*)

x:	1	2	3	4	5	6	7	8	9
y:	9	8	10	12	11	13	14	16	5

(*Bhopal, 2008*)

(*b*)

x:	6	7	7	8	8	8	9	9	10
y:	5	5	4	5	4	3	4	3	3

(*J.N.T.U., 2008*)

7. Find the parabola of the form $y = a + bx + cx^2$ which fits most closely with the observations :

x:	−3	−2	−1	0	1	2	3
y:	4.63	2.11	0.67	0.09	0.63	2.15	4.58

(*V.T.U., 2006 ; J.N.T.U., 2000 S*)

8. Fit a parabola $y = a + bx + cx^2$ to the following data :

x:	2	4	6	8	10
y:	3.07	12.85	31.47	57.38	91.29

(*V.T.U., 2003 S*)

9. Fit a second degree parabola to the following data :

x:	1	2	3	4	5	6	7	8	9	10
y:	124	129	140	159	228	289	315	302	263	210

(*U.P.T.U., 2009*)

10. The following table gives the results of the measurements of train resistances ; V is the velocity in miles per hour. R is the resistance in pounds per ton :

V:	20	40	60	80	100	120
R:	5.5	9.1	14.9	22.8	33.3	46.0

If R is related to V by the relation $R = a + bV + cV^2$, find a, b, and c. (*U.P.T.U., 2002*)

11. The velocity V of a liquid is known to vary with temperature according to a quadratic law $V = a + bT + cT^2$. Find the best values of a, b and c for the following table :

T:	1	2	3	4	5	6	7
V:	2.31	2.01	3.80	1.66	1.55	1.47	1.41

(*U.P.T.U., MCA, 2010*)

24.6 FITTING OF OTHER CURVES

(1) $y = ax^b$

Taking logarithms, $\log_{10} y = \log_{10} a + b \log_{10} x$

i.e., $\quad Y = A + bX \quad$ where $X = \log_{10} x$, $Y = \log_{10} y$ and $A = \log_{10} a$. $\quad$ (*i*)

$\therefore$ The normal equations for (*i*) are : $\Sigma Y = nA + b\Sigma X$, $\Sigma XY = A\Sigma X + b\Sigma X^2$

from which A and b can be determined. Then a can be calculated from $A = \log_{10} a$.

(2) $y = ae^{bx}$ (*Exponential curve*)

Taking logarithms, $\log_{10} y = \log_{10} a + bx \log_{10} e$

i.e., $\quad Y = A + Bx$ where $Y = \log_{10} y$, $A = \log_{10} a$ and $B = b \log_{10} e$

Here the normal equations are : $\Sigma Y = nA + B\Sigma x$, $\Sigma xY = A\Sigma x + B\Sigma x^2$

from which A, B can be found and consequently a, b can be calculated.

(3) $xy^n = b$ (*or $pv^\gamma = k$*) (*Gas equation*)

Taking logarithms, $\log_{10} x + a \log_{10} y = \log_{10} b \quad$ or $\quad \log_{10} y = \dfrac{1}{a} \log_{10} b - \dfrac{1}{a} \log_{10} x$.

This is of the form $Y = A + BX$

where $X = \log_{10} x$, $Y = \log_{10} y$, $A = \dfrac{1}{a} \log_{10} b$, $B = -\dfrac{1}{a}$.

Here also the problem reduces to finding a straight line of best fit through the given data.

Example 24.8. *Find the least squares fit of the form $y = a_0 + a_1x^2$ to the following data :*

x :	-1	0	1	2
y :	2	5	3	0

(U.P.T.U., 2008)

Solution. Putting $x^2 = X$, we have $y = a_0 + a_1X$

$\therefore$ the normal equations are : $\Sigma y = 4a + a_1\Sigma X$; $\Sigma Y = a_0\Sigma X + a_1\Sigma X^2$.

The values of ΣX, ΣX^2 etc. are calculated below :

x	y	X	X^2	XY
-1	2	1	1	2
0	5	0	0	0
1	3	1	1	3
2	0	4	16	0
	$\Sigma y = 10$	$\Sigma X = 10$	$\Sigma X^2 = 18$	$\Sigma XY = 5$

$\therefore$ the normal equations become $\quad 10 = 400 + 6a_1$; $5 = 600 + 18a_1$

Solving these equations we get, $\quad a_0 = 4.167$, $a_1 = -1.111$.

Hence the curve of best fit is $\quad y = 4.167 - 1.111X$ *i.e.,* $y = 4.167 - 1.111x^2$.

Example 24.9. *An experiment gave the following values :*

v (ft/min) :	*350*	*400*	*500*	*600*
t (min) :	*61*	*26*	*7*	*26*

It is known that v and t are connected by the relation $v = at^b$. Find the best possible values of a and b.

Solution. We have $\quad \log_{10} v = \log_b a + b \log_{10} t$

or $\quad y = A + 6X$, where $X = \log_{10} t$, $y = \log_{10} v$, $A = \log_{10} a$

$\therefore$ the normal equations are

$$\Sigma Y = 4A + b\Sigma X \qquad ...(i)$$

$$\Sigma XY = A\Sigma X + b\Sigma X^2 \qquad ...(ii)$$

Now ΣX etc. are calculated as in the following table :

v	t	$X = \log_{10} t$	$y = \log_{10} v$	XY	X^2
350	61	1.7853	2.5441	4.542	3.187
400	26	1.4150	2.6021	3.682	2.002
500	7	0.8451	2.6990	2.281	0.714
600	2.6	0.4150	2.7782	1.153	0.172
Total ·		4.4604	10.6234	11.658	6.075

$\therefore$ Equations (*i*) and (*ii*) become

$$4A + 4.46b = 10.623 \text{ ; } 4.46A + 6.075b = 11.658$$

Solving these, $A = 2.845$, $b = -0.1697$

$\therefore$ $a = \text{antilog } A = \text{antilog } 2.845 = 699.8$.

Example 24.10. *Predict the mean radiation dose at an altitude of 3000 feet by fitting an exponential curve to the given data :*

Altitude (x) :	*50*	*450*	*780*	*1200*	*4400*	*4800*	*5300*
Dose of radiation (y) :	*28*	*30*	*32*	*36*	*51*	*58*	*69*

(S.V.T.U., 2007 ; J.N.T.U., 2003)

Solution. Let $y = ab^x$ be the exponential curve.

Then $\log_{10} y = \log_{10} a + x \log_{10} b$

or $Y = A + Bx$ where $Y = \log_{10} y, A = \log_{10} a, B = \log_{10} b$

$\therefore$ the normal equations are

$$\Sigma Y = 7A + B\,\Sigma x \quad ...(i)$$

$$\Sigma x\, Y = A\,\Sigma x + B\,\Sigma x^2 \quad ...(ii)$$

Now Σx etc. are calculated as follows :

x	y	$Y = \log_{10} y$	xY	x^2
50	28	1.447158	72.3579	2500
450	30	1.477121	664.7044	202500
780	32	1.505150	1174.0170	608400
1200	36	1.556303	1867.5636	1440000
4400	51	1.707570	7513.3080	19360000
4800	58	1.763428	8464.4544	23040000
5300	69	1.838849	9745.8997	28090000
Σ = 16980		11.295579	29502.305	72743400

$\therefore$ equations (i) and (ii) become

$$11.295579 = 7A + 16980B$$

$$29502.305 = 16980A + 72743400B$$

Solving these equations, we get $A = 1.4521015$, $B = 0.0000666289$

$\therefore$ $\log_{10} y = Y = 1.4521015 + 0.0000666289x$

Hence y (at $x = 3000$) = 44.874 *i.e.* 44.9 approx.

Example 24.11. *The pressure and volume of a gas are related by the equation $pv^\gamma = k$, γ and k being constants. Fit this equation to the following set of observations :*

p (kg/cm²)	:	0.5	1.0	1.5	2.0	2.5	3.0
v (litres)	:	1.62	1.00	0.75	0.62	0.52	0.46

(V.T.U., 2011)

Solution. We have $\log_{10} p + \gamma \log_{10} v = \log_{10} k$

or $$\log_{10} v = \frac{1}{\gamma} \log_{10} k - \frac{1}{\gamma} \log_{10} p \quad \text{or} \quad Y = A + BX$$

where $X = \log_{10} p, Y = \log_{10} v, A = \dfrac{1}{\gamma} \log_{10} k, B = -\dfrac{1}{\gamma}$.

$\therefore$ the normal equations are

$$\Sigma Y = 6A + B\Sigma X \quad ...(i)$$

$$\Sigma XY = A\Sigma X + B\Sigma X^2 \quad ...(ii)$$

Now ΣX etc. are calculated as follows :

p	v	$X = \log_{10} p$	$Y = \log_{10} v$	XY	X^2
.5	1.62	– 0.3010	0.2095	– 0.0630	0.0906
1.0	1.00	0.0000	0.0000	– 0.0000	0.0000
1.5	0.75	0.1761	– 0.1249	– 0.0220	0.0310
2.0	0.62	0.3010	– 0.2076	– 0.0625	0.0906
2.5	0.52	0.3979	– 0.2840	– 0.1130	0.1583
3.0	0.46	0.4771	– 0.3372	– 0.1609	0.2276
Total		1.0511	– 0.7442	– 0.4214	0.5981

$\therefore$ equations (i) and (ii) become

$$6A + 1.0511B = -0.7442$$

$$1.0511A + 0.5981B = -0.4214$$

Solving these, we get $A = 0.0132$, $B = -0.7836$.

$\therefore \quad \gamma = -1/B = 1.1276$ and $k = \text{antilog}(A\gamma) = \text{antilog}(0.0168) = 1.039$.

Hence the equation of best fit is $pv^{1.276} = 1.039$.

PROBLEMS 24.3

1. If V (km/hr) and R (kg/ton) are related by a relation of the type $R = a + bV^2$, find by the method of least squares a and b with the help of the following table :

V :	10	20	30	40	50
R :	8	10	15	21	30

(*Indore, 2008*)

2. Using the method of least squares fit the curve $y = ax + bx^2$ to following observations :

x :	1	2	3	4	5
y :	1.8	5.1	8.9	14.1	19.8

3. Fit the curve $y = ax + b/x$ to the following data :

x :	1	2	3	4	5	6	7	8
y :	5.4	6.3	8.2	10.3	12.6	14.9	17.3	19.5

(*U.P.T.U., 2010*)

4. Estimate y at $x = 2.25$ by fitting the *indifference curve* of the form $xy = Ax + B$ to the following data :

x :	1	2	3	4
y :	3	1.5	6	7.5

(*J.N.T.U., 2003*)

5. Find the least square curve $y = ax + b/x$ for the following data :

x :	1	2	3	4
y :	– 1.5	0.99	3.88	7.66

(*Madras, 2003*)

6. Predict y at $x = 3.75$, by fitting a *power curve* $y = ax^b$ to the given data :

x :	1	2	3	4	5	6
y :	298	4.26	5.21	6.10	6.80	7.50

(*J.N.T.U., 2003*)

7. Fit the curve of the form $y = ae^{bx}$ to the following data :

x :	77	100	185	239	285
y :	2.4	3.4	7.0	11.1	19.6

(*V.T.U., 2011 S ; J.N.T.U., 2006*)

8. Obtain the least squares fit of the form $f(t) = ae^{-3t} + be^{-2t}$ for the data :

x :	0.1	0.2	0.3	0.4
$f(t)$:	0.76	0.58	0.44	0.35

(*U.P.T.U., 2008*)

9. The voltage v across a capacitor at time t seconds is given by the following table :

t :	0	2	4	6	8
v :	150	63	28	12	5.6

Use the method of least squares to fit a curve of the form $v = ae^{kt}$ to this data.

10. Using method of least squares, fit a relation of the form $y = ab^x$ to the following data :

x :	2	3	4	5	6
y :	144	172.8	207.4	248.8	298.5

(*Tiruchirapalli, 2001*)

24.7 METHOD OF GROUP AVERAGES

Let the straight line, $\qquad y = a + bx \qquad$ ...(1)

fit the set of n observations

$(x_1, y_1), (x_2, y_2), \ldots, (x_n, y_n)$ quite closely. (Fig. 24.5)

When $x = x_1$, the *observed* (*or experimental*) *value of* $y = y_1 = L_1P_1$ and from (1),

$$y = a + bx_1 = L_1M_1,$$

which is known as the *expected* (or *calculated*) *value* of y at L_1.

Then e_1 = observed value at L_1 – expected value at L_1

$$= y_1 - (a + bx_1) = M_1P_1,$$

which is called the *error* (or *residual*) at x_1. Similarly the errors for the other observations are

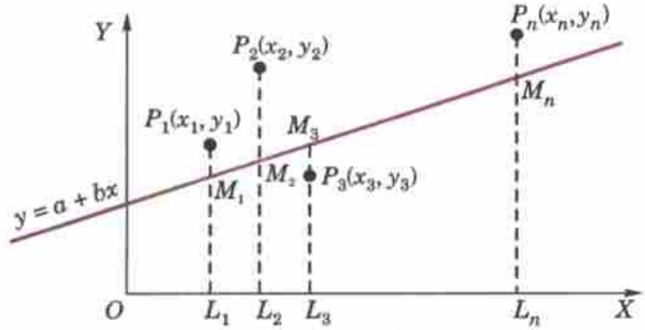

Fig. 24.5

$$e_2 = y_2 - (a + bx_2) = M_2P_2$$

....

$$e_n = y_n - (a + bx_n) = M_nP_n$$

Some of these errors may be positive and others negative.

The method of group averages is based on the assumption that the sum of the residuals is zero. To find the constants a and b is (1), we require two equations. As such we divide the data into two groups : the first containing k observations

$$(x_1, y_1), (x_2, y_2) \ldots (x_k, y_k) ;$$

and the second group having the remaining $n - k$ observations

$$(x_{k+1}, y_{k+1}), (x_{k+2}, y_{k+2}), \ldots, (x_n, y_n).$$

Assuming that the sum of the errors in each group is zero, we get

$$\{y_1 - (a + bx_1)\} + \{y_2 - (a + bx_2)\} + \ldots \{y_k - (a + bx_k)\} = 0$$
$$\{y_{k+1} - (a + bx_{k+1})\} + \{y_{k+2} - (a + bx_{k+2})\} + \ldots + \{y_n - (a + bx_n)\} = 0$$

On simplification, we obtain

$$\frac{y_1 + y_2 + \ldots + y_k}{k} = a + b\,\frac{x_1 + x_2 + \ldots + x_k}{k} \qquad \ldots(2)$$

$$\frac{y_{k+1} + y_{k+2} + \ldots + y_n}{n-k} = a + b\,\frac{x_{k+1} + x_{k+2} + \ldots + x_n}{n-k} \qquad \ldots(3)$$

In (2), $\frac{1}{k}(x_1 + x_2 + \ldots + x_k)$ and $\frac{1}{k}(y_1 + y_2 + \ldots + y_k)$ are simply the average values of x's and y's of the first group. Hence the equations (2) and (3) are obtained from (1) by replacing x and y by their respective averages of the two groups. Solving (2) and (3), we get a and b.

Obs. The main drawback of this method is that a different grouping of the observations will give different values of a and b. In practice, we divide the data in such a way that each group contains almost an equal number of observations.

Example 24.12. *The latent heat of vaporisation of steam r, is given in the following table at different temperatures t :*

t :	*40*	*50*	*60*	*70*	*80*	*90*	*100*	*110*
r :	*1069.1*	*1063.6*	*1058.2*	*1052.7*	*1049.3*	*1041.8*	*1036.3*	*1030.8*

For this range of temperature, a relation of the form $r = a + bt$ is known to fit the data. Find the values of a and b by the method of group averages. (*Madras, 2003*)

Solution. Let us divide the data into two groups each containing four readings. Then we have

t	r	t	r
40	1069.1	80	1049.3
50	1063.6	90	1041.8
60	1058.2	100	1036.3
70	1052.7	110	1030.8
$\Sigma t = 220$	$\Sigma r = 4243.6$	$\Sigma t = 380$	$\Sigma r = 4158.2$

Substituting the averages of t's and r's of the two groups in the given relation, we get

$$\frac{4243.6}{4} = a + b\,\frac{220}{4} \quad i.e.,\ 1060.9 = a + 55b \qquad ...(i)$$

$$\frac{4158.2}{4} = a + b\,\frac{380}{4} \quad i.e.,\ 1039.55 = a + 95b \qquad ...(ii)$$

Solving (*i*) and (*ii*), we obtain

$$a = 1090.26,\ b = -0.534.$$

24.8 FITTING A PARABOLA

We have applied the method of averages to *linear law* involving two constants only. To fit the parabola

$$y = a + bx + cx^2 \qquad ...(1)$$

which contains three constants, to a set of observations, we proceed as follows :

Let (x_1, y_1) be a point on (1) satisfying the given data so that

$$y_1 = a + bx_1 + cx_1^2$$

Then $$y - y_1 = b\,(x - x_1) + c\,(x^2 - x_1^2)$$

or $$\frac{y - y_1}{x - x_1} = b + c\,(x + x_1)$$

Putting $x + x_1 = X$ and $(y - y_1)/(x - x_1) = Y$, it takes the linear form

$$Y = b + cX.$$

Now b and c can be found as before.

Example 24.13. *The corresponding values of x and y are given by the following table :*

x :	87.5	84.0	77.8	63.7	46.7	36.9
y :	292	283	270	235	197	181

Solution. Taking $x = 84$, $y = 283$ as a particular point on $y = a + bx + cx^2$,

we get $$283 = a + b\,(84) + c\,(84)^2 \qquad ...(i)$$

$\therefore$ $$y - 283 = b\,(x - 84) + c\,[x^2 - (84)^2]$$

or $$\frac{y - 283}{x - 84} = b + c\,(x + 84)$$

i.e., $$Y = b + cX \qquad ...(ii)$$

where $X = x + 84$, $Y = (y - 283)/(x - 84)$.

Now we have the following table of values :

x	y	$X = x + 84$	$Y = (y - 283)/(x - 84)$
87.5	292	171.5	2.571
84.0	283	—	—
77.8	270	161.8	2.097
		$\Sigma X = 333.3$	$\Sigma Y = 4.668$
63.7	235	147.7	2.364
46.7	197	130.7	2.306
36.9	181	120.9	2.166
		$\Sigma X = 399.3$	$\Sigma Y = 6.836$

Substituting the averages of X and Y in (*ii*), we get

$$\frac{4.668}{2} = b + c\,\frac{333.3}{2} \quad i.e.,\ 2.33 = b + 166.65\,c \qquad ...(iii)$$

$$\frac{6.836}{3} = b + c\,\frac{399.3}{3} \quad i.e.,\ 2.28 = b + 131.1\,c \qquad ...(iv)$$

$(iv)-(iii)$ gives $\quad c = 0.0014$

and (iii) gives $\quad b = 2.0967 \quad$ *i.e.*, 2.1 nearly

From (i), we get $\quad a = 96.9988 \quad$ *i.e.*, 97 nearly.

Hence the parabola of fit is

$$y = 97 + 2.1x + .0014x^2.$$

Example 24.14. *The train resistance R (lbs/ton) is measured for the following values of its velocity V (km/hr) :*

V :	20	40	60	80	100
R :	5	9	14	25	36

If R is related to V by the formula $R = a + bV^n$, find a, b, and n.

Solution. To find a, we take the following three values of v which are in G.P. :

$$v_1 = 20, \quad v_2 = 40, \quad v_3 = 80$$

Then $\quad R_1 = 5, \quad R_2 = 9, \quad R_3 = 25$

$\therefore \quad (R_1 - a)(R_3 - a) = (R_2 - a)^2$

whence $\quad a = \dfrac{R_1R_3 - R_2^{\,2}}{R_1 + R_3 - 2R_2} = 3.67$

Thus $R - 3.67 = bV^n$ or $\log_{10}(R - 3.67) = \log_{10} b + n \log_{10} V$

i.e., $\quad Y = k + nX \quad$ ...(i)

where $X = \log_{10} V$, $Y = \log_{10}(R - 3.67)$, $k = \log_{10} b$.

Now we have the following table of values :

V	R	$X = \log_{10} V$	$Y = \log_{10}(R - 3.67)$
20	5	1.3010	0.1238
40	9	1.6021	0.7267
60	14	1.7782	1.0141
		$\Sigma X = 4.6813$	$\Sigma Y = 1.8646$
80	25	1.9031	1.3290
100	36	2.0000	1.5096
		$\Sigma X = 3.9031$	$\Sigma X = 2.8396$

Substituting the averages of X's and Y's in (i), we obtain

$$\frac{1.8646}{2} = k + n\,\frac{4.6813}{2} \qquad i.e.,\ 0.6215 = k + 1.5604\,n \qquad ...(ii)$$

$$\frac{2.8386}{2} = k + n\,\frac{3.9031}{2} \qquad i.e.,\ 1.4193 = k + 1.9516\,n \qquad ...(iii)$$

Solving (ii) and (iii), we get $n = 2.04$, $k = -2.56$ approx.

$$b = \text{antilog } k = \text{antilog } (-2.56) = 0.0028.$$

PROBLEMS 24.4

1. Fit a straight line of the form $y = a + bx$ to the following data by the method of group averages :

x :	0	5	10	15	20	25
y :	12	15	17	22	24	30

(*Tiruchirapalli, 2001*)

2. The weights of a calf taken at weekly intervals are given below :

Age :	1	2	3	4	5	6	7	8	9	10
Weight :	52.5	58.7	65.0	70.2	75.4	81.1	87.2	95.5	102.2	108.4

Find a straight line of best fit.

3. Using the method of averages, fit a parabola $y = ax^2 + bx + c$ to the following data :

x :	20	40	60	80	100	120
y :	5.5	9.1	14.9	22.8	33.3	46.0

4. While testing a centrifugal pump, the following data is obtained. It is assumed to fit the equation $y = a + bx + cx^2$, where x is the discharge in litre/sec and y, head in metres of water. Find the values of the constants a, b, c by the method of group averages.

x :	2	2.5	3	3.5	4	4.5	5	5.5	6
y :	18	17.8	17.5	17	15.8	14.8	13.3	11.7	9

5. By the method of averages, fit a curve of the form $y = ae^{bx}$ to the following data :

x :	5	15	20	30	35	40
y :	10	14	25	40	50	62

(Madras, 2002)

24.9 METHOD OF MOMENTS

Let $(x_1, y_1), (x_2, y_2), \dots (x_n, y_n)$ be the set of n observations such that

$$x_2 - x_1 = x_3 - x_2 = \dots = x_n - x_{n-1} = h \text{ (say)}$$

We define the *moments of the observed values* of y as follows :

m_1, the 1st moment $= h\Sigma y$

m_2, the 2nd moment $= h\Sigma xy$

m_3, the 3rd moment $= h\Sigma x^2 y$ and so on.

Let the curve fitting the given data be $y = f(x)$. Then the *moments of the calculated values* of y are

μ_1, the 1st moment $= \int y\,dx$

μ_2, the 2nd moment $= \int xy\,dx$

μ_3, the 3rd moment $= \int x^2 y\,dx$ and so on.

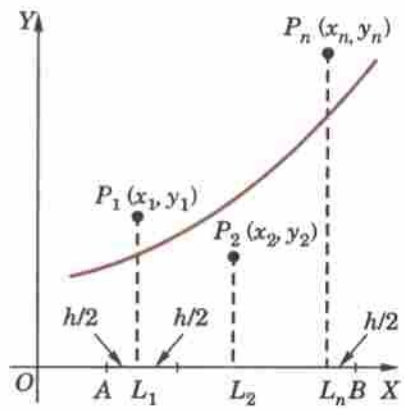

Fig. 24.6

This method is based on the assumption that the moment of the observed values of y are respectively equal to the moments of the calculated values of y i.e., $m_1 = \mu_1, m_2 = \mu_2, m_3 = \mu_3$ *etc. These equations (known as observation equations)* are used to determine the constants in $f(x)$.

m's are calculated from the tabulated values of x and y while μ's are computed as follows :

In Fig. 24.6, y_1 the ordinate of P_1 $(x = x_1)$, can be taken as the value of y at the mid-point of the interval $(x_1 - h/2, x_1 + h/2)$. Similarly, y_n, the ordinate of P_n $(x = x_n)$, can be taken as the value of y at the mid-point of the interval $(x_n - h/2, x_n + h/2)$. If A and B be the points such that

$$OA = x_1 - h/2 \text{ and } OB = x_n + h/2,$$

then
$$\mu_1 = \int y\,dx = \int_{x_1 - h/2}^{x_n + h/2} f(x)\,dx$$

$$\mu_2 = \int_{x_1 - h/2}^{x_n + h/2} xf(x)\,dx$$

and
$$\mu_3 = \int_{x_1 - h/2}^{x_n + h/2} x^2 f(x)\,dx.$$

Example 24.15. *Fit a straight line $y = a + bx$ to the following data by the method of moments :*

x :	1	2	3	4
y :	16	19	23	26

(Madras, 2001 S)

Solution. Since only two constants a and b are to be found, it is sufficient to calculate the first two moments in each case. Here $h = 1$.

$$m_1 = h\Sigma y = 1\,(16 + 19 + 23 + 26) = 84$$
$$m_2 = h\Sigma xy = 1\,(1 \times 16 + 2 \times 19 + 3 \times 22 + 4 \times 26) = 227$$

To compute the moments of calculated values of $y = a + bx$, the limits of integration will be $1 - h/2$ and $4 + h/2$ *i.e.*, 0.5 to 4.5

$$\therefore \quad \mu_1 = 2\int_{0.5}^{4.5} (a + bx)\, dx = \left| ax + b\frac{x^2}{2} \right|_{0.5}^{4.5} = 4a + 10b$$

$$\mu_2 = \int_{0.5}^{4.5} x(a + bx)\, dx = 10a + \frac{91}{3}b.$$

Thus, the observation equations $m_r = \eta_r\,(r = 1, 2)$ are $4a + 10b = 84$; $10a + \frac{91}{3}\,b = 227$

Solving these, $a = 13.02$ and $b = 3.19$.

Hence the required equation is $y = 13.02 + 3.19x$.

Example 24.16. *Given the following data :*

x :	0	1	2	3	4
y :	1	5	10	22	38

find the parabola of best fit by the method of moments.

Solution. Let the parabola of best fit be $y = a + bx + cx^2$ ...(*i*)

Since three constants are to be found, we calculate the first three moments in each case. Here $h = 1$.

$$m_1 = h\Sigma y = 1\,(1 + 5 + 10 + 22 + 38) = 76$$
$$m_2 = h\Sigma xy = 1\,(0 + 5 + 20 + 66 + 152) = 243$$
$$m_3 = h\Sigma x^2y = 1\,(0 + 5 + 40 + 198 + 608) = 851$$

For computing the moments of calculated values of (*i*), the limits of integration will be $0 - h/2$ and $4 + h/2$ *i.e.*, -0.5 and 4.5.

$$\therefore \quad \mu_1 = \int_{-0.5}^{4.5} (a + bx + cx^2)dx = 5a + 10b + 30.4c$$

$$\mu_2 = \int_{-0.5}^{4.5} x\,(a + bx + cx^2)\,dx = 10a + 30.4b + 102.5c$$

$$\mu_3 = \int_{-0.5}^{4.5} x^2\,(a + bx + cx^2)\,dx = 30.4a + 102.5b + 369.1c$$

Thus the observation equations $m_r = \mu_r\,(r = 1, 2, 3)$ are

$$5a + 10b + 30.4c = 76 \;;\; 10a + 30.4 + 102.5c = 243 \;;\; 30.4a + 102.5b + 369.1c = 851$$

Solving these equations, we get $a = 0.4$, $b = 3.15$, $c = 1.4$.

Hence the parabola of best fit is $y = 0.4 + 3.15x + 1.4x^2$.

PROBLEMS 24.5

1. Use the method of moments to fit the straight line $y = a + bx$ to the data :

x :	1	2	3	4
y :	0.17	0.18	0.23	0.32

2. Fit a straight line to the following data, using the method of moments :

x :	1	3	5	7	9
y :	1.5	2.8	4.0	4.7	6.0

(*Madras, 2001*)

3. Fit a parabola of the form $y = a + bx + cx^2$ to the data :

x :	1	2	3	4
y :	1.7	1.8	2.3	3.2

by the method of moments.

4. By using the method of moments, fit a parabola to the following data :

x :	1	2	3	4
y :	0.30	0.64	1.32	5.40

(*Madras, 2000 S*)

24.10 OBJECTIVE TYPE OF QUESTIONS

PROBLEMS 24.6

Fill up the blanks or choose the correct answer in the following problems :

1. The law $y = ax^2 + bx$ converted to linear form is ....
2. The gas equation $pv^r = k$ can be reduced to $y = a + bx$ where $a =$ .... and $b =$ ....
3. The principle of *'least squares'* states that ....
4. $y = ax^b + c$ in linear form is ....
5. To fit the straight line $y = mx + c$ to n observations, the normal equations are
   (*i*) $\Sigma y = n\,\Sigma x + \Sigma cm$, $\Sigma xy = c\,\Sigma x^2 + c\,\Sigma n$.
   (*ii*) $\Sigma y = m\,\Sigma x + nc$, $\Sigma xy = m\,\Sigma x^2 + c\,\Sigma x$.
   (*iii*) $\Sigma y = c\,\Sigma x + m\,\Sigma n$, $\Sigma xy = c\,\Sigma x^2 + m\,\Sigma x$.
6. To fit $y = ab^x$ by least square method, normal equations are ....
7. The observation equations for fitting a straight line by *method of moments* are ....
8. The *method of group averages* is based on the assumption that the sum of the residuals is ....
9. $y = ax^2 + b\log_{10} x$ reduced to linear law takes the form ....
10. Given $\begin{bmatrix} x: & 0 & 1 & 2 \\ y: & 0 & 1.1 & 2.1 \end{bmatrix}$ then the straight line of best fit is ....
11. The *method of moments* is based on the assumption that ....
12. In $y = a + bx$, $\Sigma x = 50$, $\Sigma y = 80$, $\Sigma xy = 1030$, $\Sigma x^2 = 750$ and $n = 10$, then $a =$ ...., $b =$ ....
13. $y = x/(ax + b)$ in linear form is ....
14. If $y = a + bx + cx^2$ and

x:	0	1	2	3	4
y:	1	1.8	1.3	2.5	7.3

   then the first normal equation is :
   (α) $15 = 5a + 10b + 29c$, (β) $15 = 5a + 10b + 31c$
   (γ) $12.9 = 5a + 10b + 30c$ (δ) $34 = 5a + 10b + 27c$.
15. If $y = 2x + 5$ is the best fit for 8 pairs of values (x, y) by the method of least squares and $\Sigma y = 120$, then $\Sigma X =$
   (*a*) 35 (*b*) 40 (*c*) 45 (*d*) 30.