Example 2: Fit a least squares quadratic curve to the following data

X	1	2	3	4
Y	1.7	1.8	2.3	3.2

Estimate $Y(2.4)$

Solution: Assume the L.S. quadratic curve (parabola) as

$$Y = a_0 + a_1 X + a_2 X^2$$

The normal equations are

$$\sum Y = N a_0 + a_1 \sum X + a_2 \sum X^2$$

$$\sum XY = a_0 \sum X + a_1 \sum X^2 + a_2 \sum X^3$$

$$\sum X^2 Y = a_0 \sum X^2 + a_1 \sum X^3 + a_2 \sum X^4$$

Here $N = 4$

	X	Y	X^2	XY	X^3	X^4	X^2Y
	1	1.7	1	1.7	1	1	1.7
	2	1.8	4	3.6	8	16	7.2
	3	2.3	9	6.9	27	81	20.7
	4	3.2	16	12.8	64	256	51.2
Total	10	9.0	30	25.0	100	354	80.8

Substituting these sums into normal equations, we have

$$9.0 = 4a_0 + 10a_1 + 30a_2$$

$$25 = 10a_0 + 30a_1 + 100a_2$$

$$80.8 = 30a_0 + 100a_1 + 354a_2$$

Solving $\quad a_0 = 2,\ a_1 = -0.5,\ a_2 = 0.2$

Thus the required L.S. quadratic curve (parabola) is

$$Y(X) = 2 - 0.5X + 0.2X^2$$

Estimate: $\quad Y(2.4) = 2 - 0.5(2.4) + 0.2(2.4)^2 = 1.952$

Inferences based on least square estimates

2. Compute r for the data given below:

X:	1	2	3	4	5	6
Y:	6	4	3	5	4	2

Hint: $N = 6, \sum X = 21, \sum Y = 24,$
$\sum X^2 = 91, \sum Y^2 = 106, \sum XY = 75.$

Ans. $r = -0.68$

$$r = \frac{N\sum XY - \sum X \sum Y}{\sqrt{\left[N\sum X^2 - \left(\sum X\right)^2\right]\left[N\sum Y^2 - \left(\sum Y\right)^2\right]}}$$

Types of Correlation

By ploting a given set of n pairs of random variables (X_i, Y_i), for $i = 1, 2, 3, \ldots, n$, as a scatter diagram, the correlation is said to be

Positve or direct if Y increases as X increases.

Negative or inverse if Y decreases as X increases.

Linear if all the n points lie near a straight line.

Non-linear if the points lie on some non-linear curve.

Examples:

a. Income and expenditure: positively correlated.

b. Age and IQ: negatively correlated.

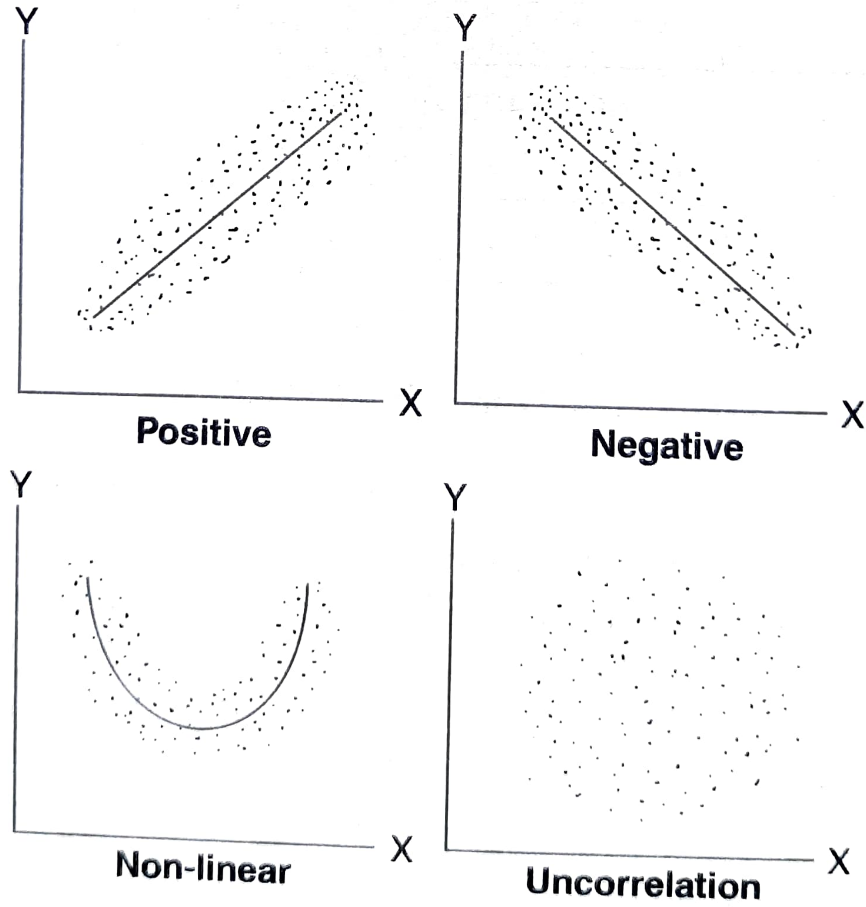

Fig. 30.3

Simple

The correlation between two variables is said to be simple correlation.