

## ⇒ Regression

The word regression is used in a quite different sense. It investigates the dependence of one variable, conventionally called the dependent variable, on one or more other variables called independent variables.

The dependent variable is assumed to be a random variable whereas the independent variables are assumed to have fixed values i.e. they are chosen non-randomly.

- \* The relation between the expected value of the dependent variable and the independent variable is called regression relation.
- \* When the dependence of a variable on two or more than two independent variables is studied, it is called multiple regression.
- \* When the dependence is represented by a straight line equation, the regression is said to be linear, otherwise it is said to be curvilinear.
- \* The dependent variable is also called regressand, the predictand, the response or the explained variable.
- \* The independent or the non-random variable is also referred to as regressor, the predictor, the explanatory variable.

## ⇒ Simple Linear Regression Model

$$Y_i = \alpha + \beta X_i + \epsilon_i$$

The linear relationship b/w the dependent variable  $Y_i$  and the value  $X_i$  of the regressor  $X$  is

Where  $X_i$  are fixed or predetermined value.  
 $Y_i$  are obs randomly drawn from a population  
 $\alpha$  &  $\beta$  are population parameter.  
 $\epsilon_i$  are error component.

$\alpha$  is the intercept and the slope  $\beta$  is called regression coefficient. which may be +ve or -ve

Regression line of  $Y$  on  $X$ .

$$Y = a + bX$$

$$b = \frac{n \sum XY - (\sum X)(\sum Y)}{n \sum X^2 - (\sum X)^2}$$

$$a = \bar{Y} - b\bar{X}$$

Regression line of  $X$  on  $Y$

$$X = a + bY$$

$$b = \frac{n \sum XY - (\sum X)(\sum Y)}{n \sum Y^2 - (\sum Y)^2}$$

$$a = \bar{X} - b\bar{Y}$$