CORRELATION :

Correlation : In a bivariate distribution if the change in one variables affects a change in other variable, the variables are said to correlated.

If the two variables deviate in the same direction i.e., if the increase (or decrease) in one result in a corresponding increase (or decrease) in the other, correlation is said to be positive.

But if they are in the opposite direction i.e., if increase (or decrease) in one variable result in corresponding decrease (or increase) in the other, correlation is said to be Negative.

Examples:

Positive correlation: Income and expenditure

Negative correlation: (i) Price and demand of commodity.

(ii) volume and pressure of a perfect gas.

correlation is said to be perfect if the deviation in one variable is followed by a corresponding and proportional deviation in the others.

Karl Pearson coefficient of correlation:

It is a measure of degree of linear relationship between two variables. Karl pearson developed a formula called "correlation coefficient".

So, correlation coefficient between two random variables X and Y, denoted by $r(X,Y)$ or ρ_{XY} of ρ is a Numerical measure of linear relationship between them and is defined as:—

$$\rho = r(X,Y) = \frac{Cov(X,Y)}{\sigma_X \sigma_Y}$$ → covariance

where,

$$COV(X,Y) = \frac{1}{n} \sum (x_i - \bar{x})(y_i - \bar{y})$$

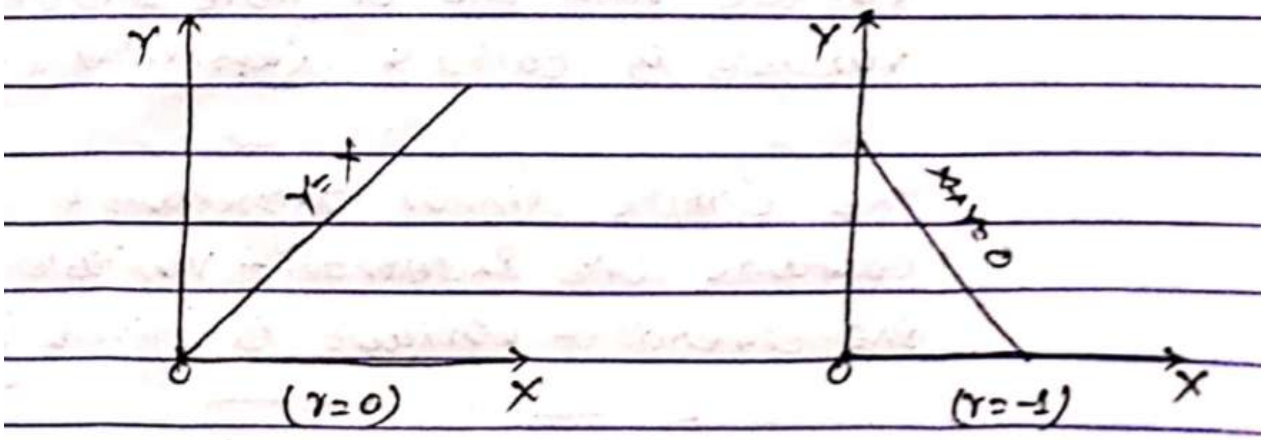
$$\sigma_X = \sqrt{\frac{1}{n} \sum_{i=1}^{n} (x_i - \bar{x})^2}$$

$$\sigma_Y = \sqrt{\frac{1}{n} \sum_{i=1}^{n} (y_i - \bar{y})^2}$$

Limits of correlation coefficient:

Correlation coefficient lies between -1 and $+1$, if $r = +1$, the correlation is perfect and positive and if $r = -1$, correlation is perfect and negative.

$$-1 < r(X,Y) \leq +1$$



If the two variables are independent the correlation coefficient between them is zero, but the converse

is not true.
correlation coefficient is a pure number, i.e. it has no unit.

REGRESSION ANALYSIS:

It is a mathematical measure of the average relationship between two or more variables in terms of the original units of the data.

In Regression analysis, there are two types of variables.
The variable where value is influenced or is to be predicted is called dependent variable and the variable which influences the value or is used for prediction is called Independent variable.

The functional relationship of a dependent variable with one or more Independent variable is called a "Regression equation".

The simple linear regression equation contains one Independent variable and one dependent variable is given by:-

$$Y = B_0 + B_1 X + \epsilon$$

where B_0 (Intercept term) and B_1 (slope) determined using least square method