Chapter- 7:

STATISTICS FOR ECONOMICS

Correlation Analysis: -

1. Meaning of correlation: It studies and measures the intensity of relationship between two or more variables. If the two variables, X and Y change vary in such a way that with a change in value of one variable the values of the other variable also change, then they are said to be correlated.
2. Significance of correlation: correlation has immense utility in statistics.
   i. It helps in determining the degree of relationship between variables.
   ii. We can estimate the value of one variable on the basis of the value of another variable correlation serves the basis of regression.
   iii. Correlation is useful for economists. An economist specifies the relationship between different variables like demand and supply, money supply and price level by way of the correlation.
3. Correlation and causation: It measures co-variation, not causation. It should never be interpreted as implying cause and effect relationship between two variables. The presence of correlation between two variables X and Y simply means that when one variable is found to change in one direction, the value of the other variable is found to change either in same direction or in the opposite direction.

4. Positive and Negative Correlation :-
   Correlation is classified into positive and negative correlation when two variables move in the same direction, i.e. if the value of Y increases or decreases with an increase or decrease in the value of X, they are said to be positively related. On the other hand when two variables move in the opposite direction i.e. if the value of variable 'X' increase or decrease with the decrease or increase in the value of Y variable, they one said to be negatively correlated.

5. Linear and Non- linear correlation:- Correlation may be linear or non-linear. If the amount of change in one variable tends to have a constant relation with the amount of change in the other variable then the correlation is said to be liner. It is represented by a straight line. On the other hand if the amount of change in one variable does not have constant proportional relationship to the amount of change in the other variable, then the correlation is said to be non-linear or curvy-linear.

6. Simple , multiple and partial correlation :-

Correlation may also be simple, multiple and partial correlation. When two variables are studied to determine correlation, it is called simple correlation on the other hand when more than two variables are studied to determine the correlation it is called multiple correlation. When correlation of only two variables is studied keeping other variables constant, it is called partial correlation.

7. Methods of studying correlation :- The correlation between the two variables can be determined by the following three methods:-

a) Scatter diagram
b) Karl Pearson's method of correlation coefficient
c) Spearman's method of Rank correlation.

8. Scatter Diagram: It is a graphic or visual method of studying correlation. To construct a scatter diagram, x. variable is taken on X axis and Y Variable is taken on Y-axis. The cluster of points so plotted is referred to as a scatter diagram. In a scatter diagram, the degree of closeness of scatter points and their overall direction gives us an idea of the nature of the relationship:-

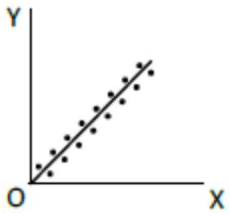
a. If the dots move from left to the right upwards, correlation is said to be positive where as the movements of dots from left to right downward indicates negative correlation.

b. Dots in a straight line indicate perfect correlation.

c. Scattered dots indicate no-correlation. The following diagrams illustrate the idea:-

d. Dots falling close to each other in a straight line indicate high degree of correlation.

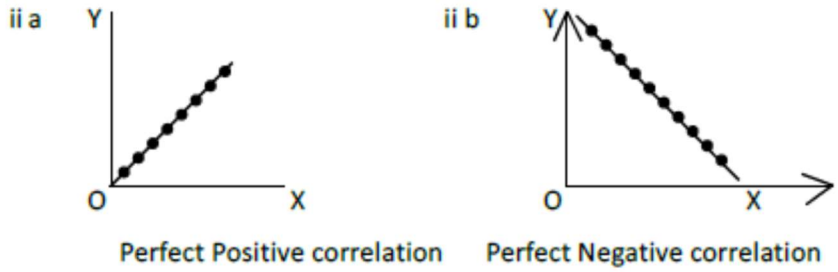
i a



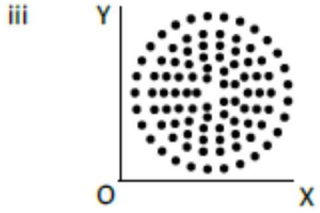
Positive correlation

i b

Negative correlation



Perfect Positive correlation Perfect Negative correlation



No correlation

9 Karl pearson's coefficient of correlation:- Karl person's coefficient of correlation is an important and widely used method of studying correlation. Karl person has measured the degree of relationship between the two variables with help of correlation coefficient. Coefficient of correlation measures the degree of relationship between the two variables.

- Computation of Karl persons coefficient of correlation :- The various formulae used to calculate coefficient of correlation r are : -

I. When deviations are taken from mean

$$r=\frac{\sum XY}{\sqrt{\sum X^2}\sqrt{\sum Y^2}}$$

The formula is also expressed as :

$$r=\frac{\sum X-\overline{X}\ \ Y-\overline{Y}}{\sqrt{\sum X-\overline{X}^{\,2}}\sqrt{\sum Y-\overline{Y}^{\,2}}}$$

ii When deviations are taken from assumed mean

$$r=\frac{N\times\sum dxdy-\sum dx.\sum dy}{\sqrt{N.\sum dx^2-\sum dx^{\,2}}\sqrt{N.\sum dy^2-\sum dy^{\,2}}}$$

where $dx = X - Ax,\ \ dy = Y - Ay.$

iii When step deviations are taken

$$r=\frac{N\times\sum dx'dy'-\sum dx'.\sum dy'}{\sqrt{N.\sum dx'^2-\sum dx'^{\,2}}\sqrt{N.\sum dy'^2-\sum dy'^{\,2}}}$$

Where $dx'=\frac{dx}{ix},\ dy'=\frac{dy}{iy}$

iv When actual data is used product moment correlation-Direct method

$$r=\frac{N\times\sum xy-\sum x.\sum y}{\sqrt{N.\sum x^2-\sum x^{\,2}}\sqrt{N.\sum y^2-\sum y^{\,2}}}$$

10. Properties of coefficient of correlation :- some of the important properties of Karl- person's coefficient of correlation are : -
    - i The correlation coefficient is independent of the units of measurement of the variables:-
    - ii The value of co-relation coefficient lies between +1 and -1.
    - iii The correlation coefficient is independent of the choice of both origin and scale of observations.
    - iv The correlation coefficient of the variables x and y is symmetric, I.e. $r_{xy} = r_{yx}$.

- **Advantages of Karl Pearson's method:-**

  Karl person's method assumes a linear relationship between two variables x and y. If r = 0, it simply means there is no linear correlation between x and y. There may exist quadratic or cubic relationship between x and y. The most important advantage of this method is that it gives an idea about co-variation of the values of two variables and also indicates the direction of such relationships.

11. Rank Correlation :- Charles Edward spearman evolved another method of finding out correlation between different qualitative attributes of a variable. This is known, as rank correlation coefficient. When a group of individuals are arranged according to their degree of possession of a character say, beauty, intelligence etc. they are said to be ranked. Spearman's formula for ranks correlation coefficient in as follows:-

A When Ranks are not repeated :-

$$R=1-\frac{6\sum D^2}{N^3-N}$$

B When Ranks are repeated

$$R=1-\frac{6\left[\sum D^2+\frac{1}{12}M^3-M+\ldots\ldots\right]}{N^3-N}$$

Where D = Difference between the ranks of the two series and

N = number of individuals in each series.

- The most important advantage of rank correlation method is that it can be used when quantitative measurement is not possible.