Department of BBA **SEM -II**
Dr. M.K.Singh

Correlation

Correlation is the measure of the degree of relationship between two and more than two variables in which the change in the values of one variable brings changes in the values of other variable also.

"The correlation between variables is a measure of the nature and degree of association between the variables".

The degree of relationship is measured by a coefficient known as correlation coefficient and is denoted by `r.The values of `r' lies between $\pm$ 1.

Types of correlation

On the basis of the nature of the relationship correlation may be :

(1) Positive and Negative Correlation
(2) Linear and Curvilinear Correlation
(3) Simple Multiple and Partial Correlation

(1) Positive and Negative Correlation :When the two variables change in the same direction, that is, both increase or both decrease the relationship is said to be ***positive or direct***.But when the change is in opposite direction ,that is, one is increasing the other is decreasing the correlation is ***negative or inverse.***

Examples :(Positive Correlations)

1) Age of husband and age of wife
2) Price of a commodity and the quantity supplied
3) Income of a consumer and expenditure
4) Rainfall and agricultural yield

Examples:(Negative correlation)

1) Price of a commodity and Quantity demanded
2) Expenditure and Saving
3) Yield of crops and price

(2) Linear and Curvilinear Correlation:

If the ratio of change between two variables is uniform then there exist linear correlation between them.Linear and non-linear correlation is based upon the consistency of the ratio of change between the

variables.The linear relationship is described by a straight line.In case of non-linear relationship ,the amount of change in one variable does not bear a constant ratio to the amount of change in the other variable.Such relationship is described by a curve.

(3) Simple,Multiple and Partial Correlation:

Simple ,multiple and partial correlation is based upon the number of variables under study.When only two variables are studied ,the analysis of the relationship between them is called simple correlation.When three or more variables are studied, the relationship can be either multiple or partial.In multiple correlation, three or more variables are studied simultaneously. But in partial correlation we consider only two variables influencing each other while the effect of other variable(s) is held constant.

Degree of Correlation

The relationship between two variables can be determined by the quantitative value of coefficient of correlation.

Perfect Correlation : When changes in two variables are exactly proportional correlation is said to be perfect.If equal proportional changes are in the same direction then there is perfect positive correlation and is describe as +1; and if equal proportional changes are in opposite direction ,there is perfect negative correlation and is described as -1.

Zero correlation : If the two variables are uncorrelated then there is zero correlation.

Limited Degree of correlation : Correlation is said to be limited positive when there is unequal changes in the two variables in the same direction and correlation is limited negative when there are unequal changes in the opposite direction.The limited degree of correlation can be high(between ± 0.75 to 1),moderate (±0.25 to 0.75) or low (between ± 0 to 0.25).

Degree of Correlation	Positive	Negative
Perfect Correlation	+1	-1
Very high degree of Correlation	+0.9	-0.9
Moderate degree of Correlation	+.75 to +0.9	-.75 to-.9
Only the possibility of correlation	+.5 to+.75	-.5 to -.75
Low Correlation	+.3to +.5	-.3 to-.5
Zero Correlation	0	0

Methods of finding Correlation:

A.Karl Pearson's Coefficient of Correlation

A mathematical method for measuring the intensity or the magnitude of *linear relationship* between two variables was suggested by Karl Pearson (1867-1936), a great British Biometrician and Statistician and, it is by far the most widely used method in practice. Karl Pearson's measure, known as Pearsonian correlation coefficient between two variables *X* and *Y*, usually denoted by *r(X,Y)* or *rxy* or simply *r* is a numerical measure of linear relationship between them and is defined as the ratio of the covariance between *X* and *Y,* to the product of the standard deviations of *X* and *Y*.

Symbolically,

$$r = \text{Covariance}(XY) / SD_x * SD_y$$

$$r = cov.xy / s_x * s_y$$

Where r= coefficient of correlation

Cov.xy =covariance xy $\qquad$ $covxy = \sum(X-\bar{X})(Y-\bar{Y}) / N$

S_x=Standard deviation of X

S_y=Standard deviation of Y

Short-Cut Method

$$r_{xy} = \frac{\sum d_x d_y - \sum d_x * \sum d_y / N}{\sqrt{[\sum d_x^2 - (\sum d_x)^2/N] * [\sum d_y - (\sum d_y)^2/N]}}$$

Example

		d_x	dy			
X	Y	X-A	Y-A	dx^2	dy^2	dxdy
10	5	0	0	0	0	0
10	6	0	1	0	1	0
11	4	1	-1	1	1	-1
12	3	2	-2	4	4	-4
12	2	2	-3	4	9	-6
		5	-5	9	15	-11

$\sum dx = 5$, $\sum dx^2 = 9$ $\sum dxdy = -11$

$\sum dy = -5$ $\sum dy^2 = 15$

$$rxy = \frac{-11-(5*-5)/5}{\sqrt{(9-(5)^2/5)*(15-(-5)^2/5}}$$

$$= \frac{-6}{\sqrt{(9-5)*(15-5)}}$$

$$= \frac{-6}{\sqrt{40}} = \frac{-6}{6.342} = -0.95$$

Product Moment Method

$$r_{xy} = \frac{\sum XY - N\bar{X}\bar{Y}}{\sqrt{(\sum X^2 - N(\bar{X})^2 * (\sum Y^2 - N(\bar{Y})^2}} = \frac{214-(5*11*4)}{\sqrt{(609-5*11^2)(90-5*4^2)}}$$

X	Y	X^2	Y^2	XY
10	5	100	25	50
10	6	100	36	60
11	4	121	16	44
12	3	144	9	36
12	2	144	4	24
55	20	609	90	214

$$= \frac{214-220}{\sqrt{4*10}} = \frac{-6}{6.324}$$

$$=$$

Mathematical properties of coefficient of correlation

1. The coefficient of correlation is independent of change of origin and change of scale.If all the values of X and Y series are multiplied or divided by some constant then scale will change but the value of correlation coefficient will not change.Similarly if a constant is added to or

subtracted from all the values of X and Y series the the origin will change but the value of correlation coefficient will not change.

2. The value of coefficient of correlation lies between ± 1.Symbolically,

$$-1 \leq r \leq +1$$