**Correlation**

Correlation refers to the statistical relationship between two entities. In other words, it's how two variables move in relation to one another. Correlation can be used for various data sets, as well.

**Meaning of Correlation:**

To measure the degree of association or relationship between two variables quantitatively, an index of relationship is used and is termed as co-efficient of correlation.

**Correlation Examples in Real Life**

1. The more time you spend running on a treadmill, the more calories you will burn.
2. Taller people have larger shoe sizes and shorter people have smaller shoe sizes.
3. The longer your hair grows, the more shampoo you will need.
4. The less time I spend marketing my business, the fewer new customers I will have.
5. The more hours you spend in direct sunlight, the more severe your sunburn.
6. The more money she saves, the more financially secure she feels.
7. As the temperature goes up, ice cream sales also go up.
8. When an employee works more hours his paycheck increases proportionately.
9. The more gasoline you put in your car, the farther it can go.
10. The longer someone invests, the more compound interest he will earn.
11. The longer amount of time you spend in the bath, the more wrinkly your skin becomes.
12. As it snows more, the sales for deicers go up.
13. As you drink more coffee, the number of hours you stay awake increases.
14. As a child grows, so does his clothing size.
15. As her salary increased, so did her spending.
16. The more it rains, the more sales for umbrellas go up.
17. Taller people tend to weigh more and shorter people tend to weigh less.
18. As a person’s level of happiness decreases, so does his level of helpfulness.
19. People who suffer from depression have higher rates of suicide than those who do not.
20. As more people go to the movies, the amount of money spent on tickets increases.
21. When workers get a raise, morale improves.
22. As tread on your car’s tires decreases, traction with the road decreases.
23. As the level of water lowers in a fish tank, the volume of the habitat for the fish decreases.
24. The faster a jet pilot flies, the higher the G-forces are.
25. The more you exercise your muscles, the stronger they get.
26. When employees make a high salary, efficiency increases.
27. As the wattage of light bulbs increases, the light output increases.

**Different degree of correlation**

The **degree** of intensity of relationship between two variables is measured with the coefficient of correlation. They are as follows:

**1. Perfect correlation:** If two variables change in the same proportion (increase or decrease), then the correlation between them is perfect correlation. Here, perfect correlation can be a positive or negative correlation.

~~Coefficient of correlation (r) = 1: If there is perfect positive relationship between two variables, then the value of correlation will be +1.~~

~~Coefficient of correlation (r) = −1: If there is perfect negative relationship between two variables, then the value of correlation will be −1.~~

**2. Zero correlation:** If two variables have no relationship between them, then the correlation is zero. It implies that a change in the value of one variable has no effect on the change in the value of the other variable.

~~Coefficient of correlation (r) = 0: If there is no relationship between the two variables, then the value of correlation will be zero. However, it does not imply that these two variables are independent. It only indicates non-existence of linear relation between the two variables.~~

**3. Limited degree of correlation:** A limited degree of correlation exists between perfect correlation and zero correlation, i.e. the value of the coefficient of correlation lies between +1 and −1. This limited degree of correlation may be high, moderate or low.

~~High degree of correlation: Correlation of two series of data is closer to one.~~

~~Medium degree of correlation: Correlation of two series of data is neither large nor small.~~

~~Low degree of correlation: Correlation of two series of data is small.~~

**Uses of Correlations**

**Prediction**

If there is a relationship between two variables, we can make predictions about one from another.

**Validity**

Concurrent validity (correlation between a new measure and an established measure).

**Reliability**

Test-retest reliability (are measures consistent).

Inter-rater reliability (are observers consistent).

**Theory verification**

Predictive validity.

**Different properties of correlation.**

1. Coefficient of Correlation lies between -1 and +1:

The coefficient of correlation cannot take value less than -1 or more than one +1. Symbolically,

-1<=r<= + 1 or | r | <1.

2. Coefficients of Correlation are independent of Change of Origin:

This property reveals that if we subtract any constant from all the values of X and Y, it will not affect the coefficient of correlation.

3. Coefficients of Correlation possess the property of symmetry:

The degree of relationship between two variables is symmetric as shown below:

4. Coefficient of Correlation is independent of Change of Scale: This property reveals that if we divide or multiply all the values of X and Y, it will not affect the coefficient of correlation.

5. Co-efficient of correlation measures only linear correlation between X and Y.

6. If two variables X and Y are independent, coefficient of correlation between them will be zero.

**Different Methods of Computing the Co-Efficient of Correlation:**

In ease of ungrouped data of bivariate distribution, the following three methods are used to compute the value of co-efficient of correlation:

1. Scatter diagram method.

2. Karl Pearson’s coefficient of correlation.

3. Spearman’s Rank-correlation coefficient.

4. Method of Least Squares.

**Scatter Diagram Method**

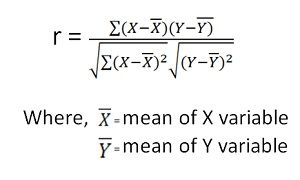
The Scatter Diagram Method is the simplest method to study the correlation between two variables wherein the values for each pair of a variable is plotted on a graph in the form of dots thereby obtaining as many points as the number of observations. Then by looking at the scatter of several points, the degree of correlation is ascertained.

The degree to which the variables are related to each other depends on the manner in which the points are scattered over the chart. The more the points plotted are scattered over the chart, the lesser is the degree of correlation between the variables. The more the points plotted are closer to the line, the higher is the degree of correlation. The degree of correlation is denoted by “r”.

**Karl Pearson’s Coefficient of Correlation**

Karl Pearson’s Coefficient of Correlation is widely used mathematical method wherein the numerical expression is used to calculate the degree and direction of the relationship between linear related variables.

Pearson’s method, popularly known as a Pearsonian Coefficient of Correlation, is the most extensively used quantitative methods in practice. The coefficient of correlation is denoted by “r”.

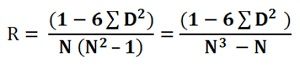
If the relationship between two variables X and Y is to be ascertained, then the following formula is used:

**Spearman’s Rank Correlation Coefficient**

The Spearman’s Rank Correlation Coefficient is the non-parametric statistical measure used to study the strength of association between the two ranked variables. This method is applied to the ordinal set of numbers, which can be arranged in order, i.e. one after the other so that ranks can be given to each.

In the rank correlation coefficient method, the ranks are given to each individual on the basis of its quality or quantity, such as ranking starts from position 1st and goes till Nth position for the one ranked last in the group.

The formula to calculate the rank correlation coefficient is:



**Method of Least Squares**

The Method of Least Squares is another mathematical method that tells the degree of correlation between the variables by using the square root of the product of two regression coefficient that of x on y and y on x.

The numerical notation of the formula to calculate the correlation by the coefficient method of least squares is given below:

