INSTITUTE OF INFORMATION TECHNOLOGY

JAHANGIRNAGAR UNIVERSITY

Final Assignment

Submission Date : 30/07/2021

Course Tittle : Probability and Statistics for Engineers.

Course Code : ICT - 2107

Submitted To

Dr. Mohd. Muzibur Rahman

Associate Professor

IIT – JU

Submitted By

Md. Shakil Hossain

Class Roll – 2023

Exam Roll – 192340

IIT – JU

Correlation

Correlation refers to the stastistical relationship between two entities. In other words it's how to variable move in relation to one othother.

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 it rains, the more sales for umbrellas go up.
2. The faster a jet Pilot flies the higher the G-forces are.
3. As the temperature goes up, ice cream sales also go up.

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 then the correlation between them is Perfect Correlation. Here Perfect correlations can be a Positive or negative correlation.

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.

3. Limited degree of Correlation: A limited degree of correlation exists between Perfect Correlation and zero correlation. The value of the coefficient of correlation lies between +1 and -1. This limited degree of correlation may be high moderate or low.

Uses of Correlations:

Prediction:

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

Validity:

Concurrent validity

Reliability:

Test-retest reliability

Inter-rater reliability

Theory Verification:

Predictive validity.

Different Properties of Correlation:

1. Coefficient of correlation lies between –1 and +1:
2. The coefficient of correlation cannot take value less than –1 or more then 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 constants 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: $r_{xy} = r_{yx}$.

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 case of ungrouped data of bivariate distribution the following th methods are used to compute the value of co-efficient of correlation.

1. Scatter diagram method.
2. Karl Pearson's coefficient of correlation
3. Sperman'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 ofon 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 ~~points are scattered~~ 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 variable.

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.

$$r = \frac{\sum (x-\bar{x})(Y-\bar{Y})}{\sqrt{\sum (x-\bar{x})^2}\sqrt{(Y-\bar{Y})^2}}$$

where,

$\bar{x}$ = mean of x variable

$\bar{Y}$ = mean of y variable

Spearman's Rank Correlation Coefficient:

The Spearman 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. 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

$$R = \frac{(1 - 6\Sigma D^2)}{N(N^2 - 1)}$$

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:

$$r = \sqrt{b_{xy} \times b_{yx}}$$

Let us Consider the Sales is X and Expenses is Y.

By using calculator we get

$\Sigma x^2 = 120088 \quad \Sigma x = 1094 \quad n = 10$

$\Sigma y^2 = 40000 \quad \Sigma y = 630 \quad \Sigma xy = 69240$

we know that. Correlation of Coefficient

$$r = \frac{\Sigma XY - \frac{\Sigma X \Sigma Y}{n}}{\sqrt{\left[\left\{\Sigma X^2 - \frac{(\Sigma X)^2}{n}\right\}\left\{\Sigma Y^2 - \frac{(\Sigma Y)^2}{n}\right\}\right]}}$$

$$r = \frac{69240 - \frac{1094 \times 630}{10}}{\sqrt{\left(120088 - \frac{(1094)^2}{10}\right)\left(40000 - \frac{(630)^2}{10}\right)}}$$

$= 0.898$ Ans.

THE END