05.05.2020

B.COM PROG SEMESTER II-
BUSINESS MATHEMATICS AND STATISTICS

Spearman Rank Correlation Coefficient

It is a non-parametric measure of correlation, using ranks to calculate the correlation. Whenever we are interested to know if two variables are related to each other, we use a statistical technique known as correlation.

The Spearman Rank Correlation Coefficient is its analogue when the data is in terms of ranks. One can therefore also call it correlation coefficient between the ranks. The correlation coefficient is sometimes denoted by rs.

Interpretation of Numerical Values

The numerical value of the correlation coefficient, rs, ranges between -1 and +1. The correlation coefficient is the number indicating the how the scores are relating.

r_s = correlation coefficient

In general,

- $r_s > 0$ implies positive agreement among ranks
- $r_s < 0$ implies negative agreement (or agreement in the reverse direction)
- $r_s = 0$ implies no agreement

Closer r_s is to 1, better is the agreement while rs closer to -1 indicates strong agreement in the reverse direction.

Formula for calculating Spearman's correlation coefficient

The following formula can be used to calculate this coefficient, it is formula where Sd2 is the sum of the squared differences between the pairs of ranks, and n is the number of pairs.

$$r_s = 1 - \frac{6\Sigma d^2}{n(n^2-1)}$$

The advantages of this coefficient are that, if calculation is to be done by hand, it is easier to calculate, and can be used for any data that can be ranked - which includes quantitative data.

Computing the rank Correlation coefficient

CASE I: When Actual ranks are given:

Following steps are involved:

- Compute D, The difference between two ranks given to each individual.
- Compute D^2 and obtain the sum of $\sum D^2$.
- Apply the formula:

$$r_s = 1 - \frac{6\Sigma d^2}{n(n^2-1)}$$

  Where n is the no. of observations.

CASE II: When ranks are not given.

Sometimes we are given the actual bivariate data on two variables and not the ranks. In such situations, it is necessary to assign the ranks. Ranks can be assigned by taking either the highest value as 1 or the lowest value as 1. The next highest or the next lowest value is given rank 2 and so on. But

whether we start with the, lowest value or the highest value we must follow the same method in case of both the variables.

CASE III: When ranks are equal

When two or more items have equal values (i.e., a tie) it is difficult to give ranks to them. In such cases the items are given the average of the ranks they would have received. For example, if two individuals are placed in the 8th place, they are given the rank [8+9] / 2 = 8.5 each, which is common rank to be assigned and the next will be 10; and if three ranked equal at the 8th place, they are given the rank [8 + 9 +10] /3 = 9 which is the common rank to be assigned to each; and the next rank will be 11.

In this case, a different formula is used when there is more than one item having the same value.

$$\rho = 1 - 6\left[\frac{\sum D_i^2 + \frac{1}{12}\left(m_1^3 - m_1\right) + \frac{1}{12}\left(m_2^3 - m_2\right) + \ldots}{n\left(n^2 - 1\right)}\right]$$

where m_i is the number of repetitions of i^{th} rank