Computation of Coefficient of Correlation

Basuki Nath Jha (B.Ed)

Sem– 2, TC- 204

Godda College Godda

There are different methods which helps us to find out whether the variables are related or not. But generally two methods are used:

1. Karl Pearson's Coefficient of correlation.
2. Rank Method.

Rank Correlation

There are many problems of business and industry when it is not possible to measure the variable under consideration quantitatively or the statistical series is composed of items which can not be exactly measured. For instance, it may be possible for the two judges to rank six different brands of cigarettes in terms of taste, whereas it may be difficult to give them a numerical grade in terms of taste. In such problems. Spearman's coefficient of rank correlation is used. The formula for rank correlation is :

$$\rho = 1 - \frac{6 \sum D^2}{n\,(n^2-1)}$$

where, **ρ** stands for rank coefficient of correlation.

D refers to the difference of ranks between paired items.

n refers to the number of paired observations.

The value of rank correlation coefficient varies between +1 and –1. When the value of ρ = +1, there is complete agreement in the order of ranks and the ranks will be in the same order. When ρ = – l, the ranks will be in opposite direction showing complete disagreement in the order of ranks.

Procedure for calculating Coefficient of Correlation

- Rank the two data sets. Ranking is achieved by giving the ranking '1' to the biggest number in a column, '2' to the second biggest value and so on. The smallest value in the column will get the lowest ranking. This should be done for both sets of measurements.
- Tied scores are given the mean (average) rank. For example, the three tied scores are ranked fifth in order, but occupy three positions (fifth, sixth and seventh) in a ranking hierarchy of ten. The mean rank in this case is calculated as (5+6+7) ÷ 3 = 6.
- Find the difference in the ranks (D): This is the difference between the ranks of the two values on each row of the table. The rank of the second value is subtracted from the rank of the first.

- Square the differences (D^2) to remove negative values and then sum them ($\sum D^2$).

Illustration: The scores for nine students in physics and math are as follows:

Physics: 35, 23, 47, 17, 10, 43, 9, 6, 28

Mathematics: 30, 33, 45, 23, 8, 49, 12, 4, 31

Compute the student's ranks in the two subjects and compute the Spearman rank correlation.

Solution:

Physics	Rank (R_1)	Mathematics	Rank (R_2)	Difference D = (R_1 - R_2)	D^2
35	3	30	5	2	4
23	5	33	3	2	4
47	1	45	2	1	1
17	6	23	6	0	0
10	7	08	8	1	1
43	2	49	1	1	1
09	8	12	7	1	1
06	9	04	9	0	0
28	4	31	4	0	0
					$\sum D^2 = 12$

Here **n** = 9, $\mathbf{D^2 =}$ 12

So,

$$\rho = 1 - \frac{6 \sum D^2}{n\,(n^2-1)}$$

$$\rho = 1 - \frac{6 \times 12}{9\,(9^2-1)}$$

$$\rho = 1 - \frac{72}{9 \times 80}$$

$$\rho = 1 - \frac{1}{10}$$

$$\rho = \frac{9}{10}$$

$$\rho = 0.9$$

The Spearman Rank Coefficient of Correlation for this set of data is 0.9. It means the correlation is very high.

The End