

Spearman's Rank Correlation Coefficient

If the data are in ordinal scale then Spearman's rank correlation coefficient is used. It is denoted by the Greek letter ρ (**rho**).

Spearman's correlation can be calculated for the subjectivity data also, like competition scores. The data can be ranked from low to high or high to low by assigning ranks.

Spearman's rank correlation coefficient is given by the formula

$$\rho = 1 - \frac{6 \sum_{i=1}^n D_i^2}{n(n^2 - 1)}$$

where $D_i = R_{1i} - R_{2i}$

R_{1i} = rank of i in the first set of data

R_{2i} = rank of i in the second set of data and

n = number of pairs of observations

Example

Two referees in a flower beauty competition rank the 10 types of flowers as follows:

Referee A	1	6	5	10	3	2	4	9	7	8
Referee B	6	4	9	8	1	2	3	10	5	7

Use the rank correlation coefficient and find out what degree of agreement is between the referees.

Rank by 1 st referee R_{1i}	Rank by 2 nd referee R_{2i}	$D_i = R_{1i} - R_{2i}$	D_i^2
1	6	-5	25
6	4	2	4
5	9	-4	16
10	8	2	4
3	1	2	4
2	2	0	0
4	3	1	1
9	10	-1	1
7	5	2	4
8	7	1	1
			$\sum_{i=1}^n D_i^2 = 60$

Here $n = 10$ and $\sum_{i=1}^n D_i^2 = 60$

$$\rho = 1 - \frac{6 \sum_{i=1}^n D_i^2}{n(n^2 - 1)}$$

$$= 1 - \frac{6 \times 60}{10(10^2 - 1)} = 1 - \frac{360}{10(99)} = 1 - \frac{360}{990} = 0.636$$

Interpretation: Degree of agreement between the referees 'A' and 'B' is 0.636 and they have “strong agreement” in evaluating the competitors.

Example

Calculate the Spearman's rank correlation coefficient for the following data.

Candidates	1	2	3	4	5
Marks in Tamil	75	40	52	65	60
Marks in English	25	42	35	29	33

Tamil		English		$D_i = R_{1i} - R_{2i}$	D_i^2
Marks	Rank (R_{1i})	Marks	Rank (R_{2i})		
75	1	25	5	-4	16
40	5	42	1	4	16
52	4	35	2	2	4
65	2	20	4	-2	4
60	3	33	3	0	0
					40

$$\sum_{i=1}^n D_i^2 = 40 \text{ and } n = 5$$

$$\begin{aligned} \rho &= 1 - \frac{6 \sum_{i=1}^n D_i^2}{n(n^2 - 1)} \\ &= 1 - \frac{6 \times 40}{5(5^2 - 1)} = 1 - \frac{240}{5(24)} = -1 \end{aligned}$$

Interpretation: This perfect negative rank correlation (-1) indicates that scorings in the subjects, totally disagree. Student who is best in Tamil is weakest in English subject and vice-versa.

Repeated ranks

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 $\frac{8+9}{2} = 8.5$ 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 $\frac{8+9+10}{3} = 9$ 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.

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

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

Example

Compute the rank correlation coefficient for the following data of the marks obtained by 8 students in the Commerce and Mathematics.

Marks in Commerce	15	20	28	12	40	60	20	80
Marks in Mathematics	40	30	50	30	20	10	30	60

Marks in Commerce (X)	Rank (R_{1i})	Marks in Mathematics (Y)	Rank (R_{2i})	$D_i = R_{1i} - R_{2i}$	D_i^2
15	2	40	6	-4	16
20	3.5	30	4	-0.5	0.25
28	5	50	7	-2	4
12	1	30	4	-3	9
40	6	20	2	4	16
60	7	10	1	6	36
20	3.5	30	4	-0.5	0.25
80	8	60	8	0	0
				Total	$\sum D^2 = 81.5$

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

Repetitions of ranks

In Commerce (X), 20 is repeated two times corresponding to ranks 3 and 4. Therefore, 3.5 is assigned for rank 2 and 3 with $m_1=2$.

In Mathematics (Y), 30 is repeated three times corresponding to ranks 3, 4 and 5. Therefore, 4 is assigned for ranks 3, 4 and 5 with $m_2=3$.

Therefore,

$$\begin{aligned} \rho &= 1 - 6 \left[\frac{81.5 + \frac{1}{12}(2^3 - 2) + \frac{1}{12}(3^3 - 3)}{8(8^2 - 1)} \right] \\ &= 1 - 6 \frac{[81.5 + 0.5 + 2]}{504} = 1 - \frac{504}{504} = 0 \end{aligned}$$

Interpretation: Marks in Commerce and Mathematics are uncorrelated