Spearman's Rank Correlation Coefficient

Data Science and A.I. Lecture Series

Bindeshwar Singh Kushwaha

PostNetwork Academy

Need for Spearman's Rank Correlation Coefficient

- In many cases, the relationship between variables is not linear, making Pearson's correlation coefficient unsuitable.

Need for Spearman's Rank Correlation Coefficient

- In many cases, the relationship between variables is not linear, making Pearson's correlation coefficient unsuitable.
- Spearman's Rank Correlation measures the strength and direction of a monotonic relationship between two variables.

Need for Spearman's Rank Correlation Coefficient

- In many cases, the relationship between variables is not linear, making Pearson's correlation coefficient unsuitable.
- Spearman's Rank Correlation measures the strength and direction of a monotonic relationship between two variables.
- It is particularly useful when:

Need for Spearman's Rank Correlation Coefficient

- In many cases, the relationship between variables is not linear, making Pearson's correlation coefficient unsuitable.
- Spearman's Rank Correlation measures the strength and direction of a monotonic relationship between two variables.
- It is particularly useful when:
  - The data is ordinal or ranked.

Need for Spearman's Rank Correlation Coefficient

- In many cases, the relationship between variables is not linear, making Pearson's correlation coefficient unsuitable.
- Spearman's Rank Correlation measures the strength and direction of a monotonic relationship between two variables.
- It is particularly useful when:
  - The data is ordinal or ranked.
  - The relationship is not linear.

Need for Spearman's Rank Correlation Coefficient

- In many cases, the relationship between variables is not linear, making Pearson's correlation coefficient unsuitable.
- Spearman's Rank Correlation measures the strength and direction of a monotonic relationship between two variables.
- It is particularly useful when:
  - The data is ordinal or ranked.
  - The relationship is not linear.
- Spearman's method uses ranks instead of raw data values, making it robust to outliers and non-normal distributions.

Formula for Spearman's Rank Correlation

Formula

$$r_s = 1 - \frac{6 \sum d_i^2}{n(n^2 - 1)}$$

- d_i: Difference between ranks of corresponding values.

Formula for Spearman's Rank Correlation

Formula

$$r_s = 1 - \frac{6 \sum d_i^2}{n(n^2 - 1)}$$

- d_i: Difference between ranks of corresponding values.
- n: Number of data pairs.

Formula for Spearman's Rank Correlation

Formula

$$r_s = 1 - \frac{6 \sum d_i^2}{n(n^2 - 1)}$$

- d_i: Difference between ranks of corresponding values.
- n: Number of data pairs.

Illustrative Example: Ranks

- Two ladies rank 7 different types of lipsticks.

Illustrative Example: Ranks

- Two ladies rank 7 different types of lipsticks.
- Calculate Spearman's rank correlation coefficient for the ranks provided:

Illustrative Example: Ranks

- Two ladies rank 7 different types of lipsticks.
- Calculate Spearman's rank correlation coefficient for the ranks provided:

Lipstick	x_i	y_i	$d_i = x_i - y_i$	d_i^2
A	1	2	-1	1
B	4	3	1	1
C	2	1	1	1
D	5	4	1	1
E	3	5	-2	4
F	6	6	0	0
G	7	7	0	0

Solution: Step-by-Step Calculation

- Compute $\sum d_i^2 = 12$.

Solution: Step-by-Step Calculation

- Compute $\sum d_i^2 = 12$.
- Substitute values into the formula:

Solution: Step-by-Step Calculation

- Compute $\sum d_i^2 = 12$.
- Substitute values into the formula:

$$r_s = 1 - \frac{6 \cdot \sum d_i^2}{n(n^2 - 1)}$$

Solution: Step-by-Step Calculation

- Compute $\sum d_i^2 = 12$.
- Substitute values into the formula:

$$r_s = 1 - \frac{6 \cdot \sum d_i^2}{n(n^2 - 1)}$$

- Substitute $n = 7$ and $\sum d_i^2 = 12$:

Solution: Step-by-Step Calculation

- Compute $\sum d_i^2 = 12$.
- Substitute values into the formula:

$$r_s = 1 - \frac{6 \cdot \sum d_i^2}{n(n^2 - 1)}$$

- Substitute $n = 7$ and $\sum d_i^2 = 12$:

$$r_s = 1 - \frac{6 \cdot 12}{7(7^2 - 1)}$$

Solution: Step-by-Step Calculation

- Compute $\sum d_i^2 = 12$.
- Substitute values into the formula:

$$r_s = 1 - \frac{6 \cdot \sum d_i^2}{n(n^2 - 1)}$$

- Substitute $n = 7$ and $\sum d_i^2 = 12$:

$$r_s = 1 - \frac{6 \cdot 12}{7(7^2 - 1)}$$

- Simplify:

$$r_s = 1 - \frac{72}{336} = 1 - 0.2143 = 0.7857$$

Final Result

Spearman's Rank Correlation Coefficient

The Spearman's Rank Correlation Coefficient is:

$$r_s = 0.786$$

- This indicates a strong positive correlation between the two rankings.

Final Result

Spearman's Rank Correlation Coefficient

The Spearman's Rank Correlation Coefficient is:

$$r_s = 0.786$$

- This indicates a strong positive correlation between the two rankings.
- The closer r_s is to 1, the stronger the positive monotonic relationship.