|  |  |  |
| --- | --- | --- |
| Related image | **KONERU LAKSHMAIAH EDUCATION FOUNDATION**  (Deemed to be University estd, u/s, 3 of the UGC Act, 1956) (NAAC Accredited “A++” Grade University)  Green Fields, Guntur District, A.P., India – 522502  **Department of Computer Science and Engineering**  (DST - FIST Sponsored Department) |  |

**B.Tech. II CSE(H) PROGRAM**

**A.Y. 2023-24 ODD, Semester-II**

**Course Code: 22MT2005**

**PROBABILITY, STATISTICS AND QUEUING THEORY**

**Course Outcome-2**

**Session 17:** **RankCorrelations**

1. **Course Description (Description about the subject)**

Rank correlation is a statistical measure that quantifies the degree of association between two variables. It is used to determine whether two sets of rankings or scores are related to each other, and if so, how strongly.

1. **Aim**

To explain identify the linear relationship between two variables using different measures of correlation and regression.

1. **Instructional** **Objectives (Course Objectives)**

To Calculate the linear relationship between two variables using different measures of correlation and regression

1. **Learning** **Outcomes (Course Outcome)**

**CO2**: Students will be able to Explain different measures of central tendency, and dispersion.

1. **Module** **Description** **(CO-2 Description)**

Types of Scatter diagrams, Correlation and Regression

1. **Session** **Introduction**

Scatter diagrams are a useful tool for visualizing the relationship between two variables. They can be used to identify the type of relationship, to identify outliers, and to make predictions.

1. **Session description**

**Rank correlation coefficient for untied ranks:**

A group of n individuals may be arranged in order of merit with respect to some characteristic. The same group would give different orders for different characteristics. Considering the orders corresponding to two characteristics A and B, the correlation between these n pairs of ranks is called the rank correlation in the characteristics A and B for that group of individuals.

Let xi, yi be the ranks of the ith individuals in A and B respectively. Assuming that no two individuals are bracketed equal in either case, each of the variables taking the values 1, 2, 3, ..., n.

Hence the correlation between the variables is

**Rank correlation coefficient for tied ranks:**

Ranks of some individuals are same.

If some individuals receive the same rank, the ranks of them are said to be tied. Let us suppose that ranks of m individuals are tied after K ranks, hence they receive the ranks K+1, K+2, ..., K+m. Since ranks are tied for m individuals, they should assign a common rank for all m individuals as the average of K+1, K+2... K+m.

Hence the Spearman’s rank correlation coefficient formulae for tie ranks is given by

where

1. **Activities/ Case studies/related to the session.**

**NA**

1. **Examples & contemporary extracts of articles/ practices to convey the idea of the Session**

**Example 1:** Compute Spearman’s rank correlation coefficient for the following data

X: 20 14 36 29 5 11

Y: 19 9 25 10 2 6

**Solution:**

|  |  |  |  |  |  |
| --- | --- | --- | --- | --- | --- |
| **X** | **Y** | **Ranks of X** | **Ranks of Y** | **di** | **di2** |
| **xi** | **yi** |
| 20  14  36  29  5  11 | 19  9  25  10  2  6 | 3  4  1  2  6  5 | 2  4  1  3  6  5 | 1  0  0  -1  0  0 | 1  0  0  1  0  0 |
|  |  |  |  |  | **=2** |

**Example 2:**  Marks of 11 students in two subjects A and B are given below. Obtain rank correlation coefficient.

Student No: 1 2 3 4 5 6 7 8 9 10 11

Marks in A: 25 36 20 36 48 52 25 65 35 45 60

Marks in B: 35 42 30 42 56 68 45 50 42 55 68

**Solution:**  Let us denote ranks of two subjects A and B be x and y.

|  |  |  |  |  |  |  |
| --- | --- | --- | --- | --- | --- | --- |
| Student No. | Marks in A | Marks in B | Ranks of A=xi | Ranks of B=yi | di=xi-yi | di2 |
| 1  2  3  4  5  6  7  8  9  10  11 | 25  36  20  36  48  52  25  65  35  45  60 | 35  42  30  42  56  68  45  50  42  55  68 | 9.5  6.5  11  6.5  4  3  9.5  1  8  5  2 | 10  8  11  8  3  1.5  6  5  8  4  1.5 | 0.5  -1.5  0  -1.5  1  0.5  3.5  -4  0  1  0.5 | 0.25  2.25  0  2.25  1  0.25  12.25  16  0  1  0.25 |
|  |  |  |  |  |  |  |

n=11,

In subject A, 25 repeated twice and 36 repeated twice i.e. m1=2, m2=2

In subject B, 68 repeated twice and 42 repeated thrice i.e. m1=2, m2=3

Rank correlation coefficient with tied ranks is

1. **SAQ's-Self Assessment Questions**

1: Which of the following statements is true regarding rank correlation?

A) It measures the strength of a linear relationship between two variables.

B) It measures the strength of a non-linear relationship between two variables.

C) It measures the similarity of rankings between two variables.

D) It measures the difference of rankings between two variables.

Answer: C) It measures the similarity of rankings between two variables.

2: The value of rank correlation coefficient (Spearman's rho) lies between:

A) -1 and 1

B) 0 and 1

C) 0 and 100

D) -100 and 100

Answer: A) -1 and 1

3: In which scenario would you prefer to use Kendall's Tau rank correlation coefficient over Spearman's rho?

A) When dealing with small sample sizes.

B) When the data contains tied ranks.

C) When the relationship between variables is assumed to be linear.

D) When the variables have different units of measurement.

Answer: B) When the data contains tied ranks.

4: If the rank correlation coefficient (Spearman's rho) is close to -1, it indicates:

A) A strong positive correlation.

B) A strong negative correlation.

C) A weak positive correlation.

D) No correlation between the variables.

Answer: B) A strong negative correlation.

5: Which rank correlation coefficient is more sensitive to outliers?

A) Spearman's rho

B) Kendall's Tau

C) Both are equally sensitive

D) Rank correlation coefficients are not affected by outliers.

Answer: A) Spearman's rho

1. **Summary**

Rank correlation measures the degree of association between ranked variables, making it useful when dealing with non-parametric data or when the relationship between variables is not assumed to be linear. It is particularly helpful in situations where the data contains tied ranks or is not normally distributed.

1. **Terminal Questions**

1. The value of Karlpearon’s correlation (r) for the following data is **0.636**.

x: 0.05 0.14 0.24 0.30 0.47 0.52 0.57 0.61 0.67 0.72

y: 1.08 1.15 1.27 1.33 1.41 1.46 1.54 2.72 4.01 9.63

1. Calculate the Spearman’s rank correlation for this data.
2. What advantage of ρ brought out in this problem?

2. Twelve recruits were subjected to a selection test to ascertain their suitability for a certain course of training. At the end of training there were given a proficiency test.

The marks secure by recruits in the selection test (x) and the proficiency test (Y) are given below

x: 65 63 67 64 68 62 70 66 68 67 69 71

y: 68 66 68 65 69 66 68 65 71 67 68 70

Calculate coefficient of rank correlation.

3. Ten competitors in a beauty contest were ranked by the three judges A, B, and C in the following order:

**Competitors**

**Judges**: 1 2 3 4 5 6 7 8 9 10

**A** : 6 5 3 10 2 4 9 7 8 1

**B** : 5 8 4 7 10 2 1 6 9 3

**C**  : 4 9 8 1 2 3 10 5 7 6

Discuss which pair of judges has the nearest approach to common tastes of beauty.

4. Determine than Spearman’s rank correlation

|  |  |  |  |  |  |
| --- | --- | --- | --- | --- | --- |
| x | 2 | 3 | 8 | 9 | 6 |
| y | 2.5 | 3.5 | 2.1 | 3 | 4 |

1. **Case Studies (CO Wise)**

**NA**

1. **Answer Key**

**NA**

1. **Glossary**

**NA**

1. **References of books, sites, links Textbooks:**

**Textbooks:**

1. Probability and Statistics Rukmangad Achari E. and E. Keshava Reddy
2. Probability and Statistics for Engineers and Scientists” Ronald E. Walpole, Sharon L. Myers and Keying Ye 8th Edition Pearson pub
3. Probability & Statistics for Engineers Dr. J. Ravichandran first Edition Wiley-India

**Reference books:**

1. Hossein Pishro-Nik, Introduction to Probability, Statistics, and Random Processes, 2014, by Kappa Research LLC; ISBN-13: 978-0990637202

**Web Resources**

1. https://ncert.nic.in/textbook.php?kemh1=0- 16
2. https://ncert.nic.in/textbook.php?jemh1=ps-15
3. **Keywords**

Rank correlation, Spearman’s rank correlation, tied and untied ranks.