

## 22 March Assignment

May 29, 2023

[ ]: Q1. Pearson correlation coefficient **is** a measure of the linear relationship between two variables. Suppose you have collected data on the amount of time students spend studying **for** an exam **and** their final exam scores. Calculate the Pearson correlation coefficient between these two variables **and** interpret the result.

ANS -

[ ]: The Pearson correlation coefficient **is** a measure of the linear relationship between two variables. It ranges **from** -1 to 1. A value of -1 indicates a perfect negative linear relationship between the two variables, **while** a value of 1 indicates a perfect positive linear relationship between the two variables. A value of 0 indicates no linear relationship between the two variables.

To calculate the Pearson correlation coefficient between the amount of time students spend studying **for** an exam **and** their final exam scores, you can use the following formula:

$$r = (\sum xy - \sum x \sum y) / \sqrt{(\sum x^2 - (\sum x)^2)(\sum y^2 - (\sum y)^2)}$$

where:

**r** **is** the Pearson correlation coefficient  
**x** **is** the amount of time spent studying **for** an exam  
**y** **is** the final exam score  
 **$\Sigma$**  **is** the **sum** of  
**n** **is** the number of observations

The interpretation of the result depends on its value. If **r** **is** close to 1, it means that there **is** a strong positive linear relationship between the two variables. If **r** **is** close to -1, it means that there **is** a strong negative linear relationship between the two variables. If **r** **is** close to 0, it means that there **is** no linear relationship between the two variables.

[ ]:

[ ]: Q2. Spearman's rank correlation is a measure of the monotonic relationship  
↪between two variables.

Suppose you have collected data on the amount of sleep individuals get each  
↪night and their overall job  
satisfaction level on a scale of 1 to 10. Calculate the Spearman's rank  
↪correlation between these two  
variables and interpret the result.

ANS -

[ ]: Spearman's rank correlation is a measure of the monotonic relationship between  
↪two variables. It is used to assess how well the relationship  
between two variables can be described using a monotonic function. In this  
↪case, the two variables are the amount of sleep individuals get  
each night and their overall job satisfaction level on a scale of 1 to 10.

To calculate the Spearman's rank correlation coefficient between these two  
↪variables, you would first rank each variable separately from  
lowest to highest. Then you would compare the ranks of each variable and  
↪calculate the correlation coefficient. The Spearman's rank  
correlation coefficient ranges from -1 to 1, where -1 indicates a perfect  
↪negative correlation, 0 indicates no correlation, and 1 indicates a  
perfect positive correlation.

The Spearman's rank correlation coefficient between sleep and job satisfaction  
↪can be interpreted as follows:

If the coefficient is close to -1, it indicates that there is a strong negative  
↪monotonic relationship between sleep and job satisfaction.

This means that as the amount of sleep decreases, job satisfaction tends to  
↪decrease as well.

If the coefficient is close to 0, it indicates that there is no monotonic  
↪relationship between sleep and job satisfaction.

If the coefficient is close to 1, it indicates that there is a strong positive  
↪monotonic relationship between sleep and job satisfaction.

This means that as the amount of sleep increases, job satisfaction tends to  
↪increase as well.

[ ]:

[ ]: Q3. Suppose you are conducting a study to examine the relationship between the  
↪number of hours of  
exercise per week and body mass index (BMI) in a sample of adults. You  
↪collected data on both variables  
for 50 participants. Calculate the Pearson correlation coefficient and the  
↪Spearman's rank correlation

between these two variables and compare the results.

ANS -

[ ]: To calculate the Pearson correlation coefficient and the Spearman's rank correlation between these two variables, you can use the following formulas:

Pearson correlation coefficient:

$$r = (n\sum xy - \sum x \sum y) / \sqrt{(n\sum x^2 - (\sum x)^2)(n\sum y^2 - (\sum y)^2)}$$

Copy

where  $n$  is the number of observations,  $\sum xy$  is the sum of the products of paired scores,  $\sum x$  is the sum of  $x$  scores,  $\sum y$  is the sum of  $y$  scores,  $\sum x^2$  is the sum of squared  $x$  scores and  $\sum y^2$  is the sum of squared  $y$  scores.

Spearman's rank correlation:

$$r_s = 1 - (6 * \sum d^2) / (n(n^2 - 1))$$

Copy

where  $d$  is the difference between ranks assigned to paired observations and  $n$  is the number of observations.

The Pearson correlation coefficient measures the linear relationship between two continuous variables. It ranges from -1 to 1. A value of 1 indicates a perfect positive linear relationship between two variables. A value of -1 indicates a perfect negative linear relationship between two variables. A value of 0 indicates no linear relationship between two variables.

The Spearman's rank correlation measures the monotonic relationship between two continuous or ordinal variables. It ranges from -1 to 1.

A value of 1 indicates a perfect positive monotonic relationship between two variables. A value of -1 indicates a perfect negative monotonic relationship between two variables. A value of 0 indicates no monotonic relationship between two variables.

In general, if there is a strong linear relationship between two variables, then Pearson's correlation coefficient will be high and Spearman's rank correlation will also be high. However, if there is a non-linear relationship between two variables, then Pearson's correlation coefficient may not be high but Spearman's rank correlation may still be high

[ ]:

[ ]: Q4. A researcher is interested in examining the relationship between the number of hours individuals spend watching television per day and their level of physical activity. The researcher collected data on both variables from a sample of 50 participants. Calculate the Pearson correlation coefficient between these two variables.

AMS -

[ ]: The Pearson correlation coefficient formula is used to measure the strength and direction of the linear relationship between two variables. The formula is:

$$r = \frac{[n(\sum xy) - \sum x \sum y]}{\sqrt{[n(\sum x^2) - (\sum x)^2][n(\sum y^2) - (\sum y)^2]}}$$

where x is the independent variable, y is the dependent variable, n is the sample size, and  $\Sigma$  represents a summation of all values<sup>12</sup>.

In this case, the researcher collected data on both variables from a sample of 50 participants. To calculate the Pearson correlation coefficient between these two variables, we need to know the values of x and y for each participant. Once we have these values, we can use the formula above to calculate r

[ ]:

[ ]: Q5. A survey was conducted to examine the relationship between age and preference for a particular brand of soft drink. The survey results are shown below:

ANS -

[ ]:

[ ]:

[ ]: Q6. A company is interested in examining the relationship between the number of sales calls made per day and the number of sales made per week. The company collected data on both variables from a sample of 30 sales representatives. Calculate the Pearson correlation coefficient between these two variables.

ANS -

[ ]: The Pearson correlation coefficient formula is used to measure the strength and direction of the linear relationship between two variables.

The formula is:

$$r = \frac{[n(\sum xy) - \sum x \sum y]}{\text{Square root of } \sqrt{[n(\sum x^2) - (\sum x)^2][n(\sum y^2) - (\sum y)^2]}}$$

where x is the independent variable, y is the dependent variable, n is the sample size, and  $\Sigma$  represents a summation of all values.

In this case, the independent variable is the number of sales calls made per day and the dependent variable is the number of sales made per week. The Pearson correlation coefficient between these two variables can be calculated using this formula.