

Q1. Define Covariance and explain how it differs from Correlation in terms of scale and interpretation.

Answer:

Covariance measures the direction of the linear relationship between two variables. It shows whether they increase or decrease together.

Difference from Correlation:

Covariance values depend on the scale of measurement, so they are difficult to interpret.

Correlation is a standardized measure and always lies between -1 and $+1$, making it easier to interpret.

Covariance only shows direction, while correlation shows both direction and strength.

Q2. What does a positive, negative, and zero covariance indicate about the relationship between two variables?

Answer:

Positive covariance: Both variables move in the same direction.

Negative covariance: One variable increases while the other decreases.

Zero covariance: No linear relationship exists between the variables.

Q3. Discuss the limitations of covariance as a measure of relationship between two variables. Why is correlation preferred in many cases?

Answer:

Limitations of covariance:

No fixed range \rightarrow difficult to compare.

Units depend on variables' scales.

Does not indicate strength clearly.

Why correlation is preferred:

Scale-free and standardized.

Easy interpretation (-1 to $+1$).

Widely comparable across datasets.

Q4. Explain the difference between Pearson's correlation coefficient and Spearman's rank correlation coefficient. When would you prefer to use Spearman's correlation?

Answer:

Pearson's Correlation:

Measures linear relationship

Uses actual values

Sensitive to outliers

Spearman's Correlation:

Measures monotonic relationship

Uses ranks

Robust to outliers

Use Spearman's when:

Data is ordinal

Relationship is non-linear

Outliers are present

Q5. If the correlation coefficient between two variables X and Y is 0.85, interpret this value in context. Can y causation from this value? Why or why not?□

Answer:

A value of 0.85 indicates a strong positive relationship.

As X increases, Y also increases strongly.

Causation cannot be inferred because correlation does not prove cause-and-effect; other factors m both variables.

Q6. Using the dataset below, calculate the covariance between X and Y.

| | | | | |
|---|---|---|---|----|
| X | 2 | 4 | 6 | 8 |
| Y | 3 | 7 | 5 | 10 |

Answer:

4.75

Q7. Compute the Pearson correlation coefficient between variables A and B:

| | | | | | |
|---|----|----|----|----|----|
| A | 10 | 20 | 30 | 40 | 50 |
| B | 8 | 14 | 18 | 24 | 28 |

Answer:

0.99760861

Q8. The following table shows heights (in cm) and weights (in kg) of 5 students.

Find the correlation coefficient between Height and Weight.

| | | | | | |
|--------|-----|-----|-----|-----|-----|
| Height | 150 | 160 | 165 | 170 | 180 |
| Weight | 50 | 55 | 58 | 62 | 70 |

Answer:

0.99218437

Q9. Given the dataset below, determine whether there is a positive or negative correlation between X and Y (No need for exact calculation, just reasoning.)

| | | | | | |
|---|----|----|---|---|---|
| X | 1 | 2 | 3 | 4 | 5 |
| Y | 15 | 12 | 9 | 7 | 3 |

Answer:

As X increases from 1 to 5, Y consistently decreases from 15 to 3.

This means:

The variables move in opposite directions

When one goes up, the other goes down

Answer :

-0.9958641

There is a negative correlation between X and Y.

Q10. Two investment portfolios have the following returns (%) over 5 years. Compute the covariance and correlation coefficient, and interpret whether the portfolios move together.

| Year | Portfolio A | Portfolio B |
|------|-------------|-------------|
| 1 | 8 | 6 |
| 2 | 10 | 9 |
| 3 | 12 | 11 |
| 4 | 9 | 8 |
| 5 | 11 | 10 |

Answer:

Covariance:

2.4

Correlation:

0.98639392

on.

er variables

ret.

les?

lation

on

you infer

may influence

4.

