

Correlation and Regression

Analysis

Unit 3

1. What do you mean by correlation?

Correlation refers to the statistical relationship between two or more variables. It measures the degree to which changes in one variable correspond to changes in another. The correlation coefficient quantifies this relationship, ranging from -1 to +1.

2. What are the properties of the correlation coefficient?

The properties of the correlation coefficient are:

- **Range:** It lies between -1 and +1.
- **Significance:** A positive value indicates a direct relationship, while a negative value indicates an inverse relationship.
- **Symmetry:** The correlation between X and Y is the same as between Y and X.
- **Unit-free measure:** It is a pure number independent of units of measurement.
- **Value interpretation:**
 - $r = +1$: Perfect positive correlation.
 - $r = -1$: Perfect negative correlation.
 - $r = 0$: No correlation.

3. What are the importance of correlation?

The importance of correlation includes:

- Helps in predicting one variable based on another.
- Useful in finance, economics, and business decision-

making.

- Assists in identifying relationships for research and analysis.
- Helps in regression analysis for forecasting trends.

4. What are the limitations of correlation?

Some limitations of correlation are:

- It does not imply causation.
- Affected by extreme values.
- Can be misleading if variables are non-linearly related.
- Does not provide a detailed explanation of the relationship.

5. What are the different types of correlation?

- **Positive and Negative Correlation:** If two variables move in the same direction, it's positive; otherwise, it's negative.
- **Simple, Multiple, and Partial Correlation:** Simple deals with two variables, multiple involves more than two, and partial examines the effect of one variable while keeping others constant.
- **Linear and Non-Linear Correlation:** Linear follows a straight-line pattern, while non-linear does not.

6. What do you mean by a scatter diagram? Write its merits and demerits.

A scatter diagram is a graphical representation of data points that show the relationship between two variables.

- **Merits:**
- Simple and easy to understand.
- Helps visualize relationships between variables.
- **Demerits:**
- Does not quantify correlation.

- Cannot determine causation.

7. What do you mean by Karl Pearson's correlation coefficient? Write its merits and demerits.

Karl Pearson's correlation coefficient (r) is a statistical measure that calculates the strength and direction of a linear relationship between two variables.

- **Merits:**
- Measures the degree of linear correlation.
- Provides precise numerical value.
- **Demerits:**
- Sensitive to outliers.
- Assumes a linear relationship.

8. What do you mean by Rank correlation? Write its merits and demerits.

Rank correlation measures the degree of association between two ranked variables (Spearman's Rank Correlation is commonly used).

- **Merits:**
- Suitable for ordinal data.
- Less affected by extreme values.
- **Demerits:**
- Not suitable for exact measurement.
- Less accurate than Pearson's correlation.

9. Write a note on multiple correlation.

Multiple correlation examines the relationship between one dependent variable and two or more independent variables. It helps in understanding how multiple factors contribute to a variable.

10. What do you mean by regression analysis?

Regression analysis is a statistical technique used to model

and analyze relationships between dependent and independent variables. It helps in forecasting and decision-making.

11. What do you mean by regression lines?

Regression lines represent the best-fit line through a set of data points, showing the relationship between dependent and independent variables.

12. What are the properties of regression coefficients?

- They remain constant for a given relationship.
- They can be positive or negative.
- If one coefficient is greater than 1, the other must be less than 1.
- The two regression coefficients are reciprocals of each other.

13. Explain the relationship between the correlation coefficient and regression coefficient.

- Correlation measures the strength of a relationship, while regression quantifies it.
- Regression helps predict values, while correlation does not.
- The square of the correlation coefficient represents the proportion of explained variance in regression.

14. Distinguish between correlation and regression .

- **Definition:** Correlation measures the strength and direction of the relationship between two variables, while regression determines the dependency of one variable on another.
- **Direction:** Correlation is symmetric (X and Y are interchangeable), whereas regression is asymmetric

(one is dependent, and the other is independent).

- **Purpose:** Correlation helps in understanding the degree of association, whereas regression is used for prediction and forecasting.

15. Write a note on multiple regression.

Multiple regression is an extension of simple regression where multiple independent variables are used to predict the dependent variable. It is widely used in business, economics, and social sciences for better decision-making.

