UNIT 15
MULTIVARIATE MEASURES OF ASSOCIATION

1. INTRODUCTION

Multivariate Measures of Association (or Multivariate Analysis) refers to a collection of statistical techniques used to analyze data involving multiple variables, either by determining how several independent variables collectively influence a dependent variable (dependence techniques) or by summarizing the structure of relationships among a set of variables (interdependence techniques).

MVA moves beyond simple two-variable relationships (bivariate analysis). We stop looking at how just one factor (like Age) affects an outcome (like Spend) and look at how many factors simultaneously (Age, Income, Location, Satisfaction) combine to influence or structure our data.

2. MULTIPLE REGRESSION

Multiple Regression is an extension of simple regression used to predict the value of a single dependent variable based on the combined linear effect of two or more independent variables. Coefficients (bi) are the weight assigned to each independent variable, indicating its relative importance in the prediction model.

This is used to get a much more accurate prediction than simple regression. Instead of saying, "High satisfaction predicts high spend," we say, "*High satisfaction AND low distance from home AND high income all predict high spend.*" It tells us the unique impact of each predictor while controlling for the others.

Coffee Shop Example: Predicting **Total Monthly Spend (**Y**)** using **Satisfaction Score (**X1**)**, **Distance to Store (**X2**)**, and **Income (**X3**)**. The result might show that Satisfaction is the strongest predictor, but Distance is still significant.

3. LISREL (STRUCTURAL EQUATION MODELING - SEM)

LISREL (LInear Structural RELations) is the original name for a technique now commonly referred to as **Structural Equation Modeling (SEM)**. SEM is a sophisticated statistical methodology that allows the testing of complex, hypothesized causal relationships between multiple variables, including both observed variables and unobserved (latent) variables. Observed variables are the data you can directly measure or record, like a person's age or a test score.Unobserved variables (or latent variables) are theoretical concepts you cannot directly measure, such as intelligence or motivation, which must be inferred from observed data.

SEM is for testing complex theories or models. It tests the entire chain of cause and effect simultaneously. It is often used when key concepts, like 'Loyalty' or 'Service Quality,' cannot be measured directly but are inferred from several survey questions.

Coffee Shop Example: We hypothesize a chain:

(Staff Friendliness → Service Quality) → (Satisfaction) → (Loyalty Intention).

SEM tests if the data collected supports this entire theoretical model, showing the strength of each link in the chain.

4. CONJOINT ANALYSIS

Conjoint Analysis is a dependence technique used to determine the relative importance consumers attach to various product attributes (features) and the utility (value) they derive

from each specific level of those attributes. It requires consumers to evaluate realistic combinations of features.

This technique figures out what customers *really* value in a product. Instead of asking customers, "Is price or quality more important?" (where they usually say both), we force them to make trade-offs between specific features and price points. By analyzing their choices, we can mathematically calculate the hidden value they place on each feature.

Coffee Shop Example: We want to design a new loyalty app. We test scenarios mixing: **1. Price (**5/month **vs.** 10/month**)**, **2. Reward (Free coffee vs.** 50% **off)**, and **3. Access (VIP seating vs. Standard seating)**. Conjoint will reveal if customers value the Free Coffee reward more than the lower Price.

5. FACTOR ANALYSIS

Factor Analysis is an interdependence technique used for data reduction and summarization. It groups a large set of interrelated variables into a smaller, manageable set of underlying, unobserved dimensions called factors (or components) that account for most of the variation in the original data.

Factor analysis helps simplify messy data. If you have 30 survey questions that are all asking similar things about customer experience, Factor Analysis mathematically compresses them into just a few core themes or factors.

Coffee Shop Example: We use 10 separate Likert scale questions to measure the store atmosphere (e.g., "The music is good," "The seating is comfortable," "The lighting is appealing"). Factor Analysis might combine these 10 items into two meaningful factors: **1. Sensory Atmosphere** (music, lighting) and **2. Physical Comfort** (seating, temperature).

6. CLUSTER ANALYSIS

Cluster Analysis is an interdependence technique used to classify objects (e.g., respondents, products, or brands) into a number of relatively homogeneous (similar) groups called clusters. The goal is to maximize the similarity of objects within each cluster while maximizing the dissimilarity between clusters.

This technique is the core method for market segmentation. It lets the data group customers automatically based on how similar they are across all variables (Age, Income, Visit Frequency, etc.), rather than the researcher guessing the segments beforehand.

Coffee Shop Example: Cluster Analysis, using variables like Age, Visit Frequency, and Spend, might identify three segments for the coffee shop:

- **Cluster 1 (The Daily Commuter):** High Visit Frequency, Medium Spend, Medium Age.
- **Cluster 2 (The Weekend Socializer):** Low Visit Frequency, High Spend (on groups), Younger Age.
- **Cluster 3 (The Budget Buyer):** High Visit Frequency, Low Spend (small purchases), Older Age.