CORRELATION vs REGRESSION

Subject Expert: Prof. Karunesh Saxena

Faculty of Management Studies

Director IQAC Cell

Mohanlal Sukhadia University

Udaipur

Flow of Presentation

- ➢ **CORRELATION** – Introduction to correlation
- ▪ Types of correlation
- ▪ Methods of correlation
- ▪ Merits and Demerits of correlation
- ➢ **REGRESSION** - Introduction to Regression
- ▪ Types of Regression
- ▪ Models of Regression
- ▪ Merits and Demerits
- ➢ **COMPARISION TABLE OF CORRELATION AND REGRESSION.**

CORRELATION

Introduction

- **Correlation** is a statistical measure (expressed as a number) that describes the size and direction of a relationship between two or more variables. A correlation between variables, however, does not automatically mean that the change in one variable is the cause of the change in the values of the other variables

CORRELATION

Introduction

- Correlation is the tool that helps to study the relationship between two variables.

- The measure of correlation is called the **Coefficient Of Correlation** (denoted by symbol r) summarizes in one figure the direction and degree of correlation.

- A Correlation Coefficient can be calculated when there are two (or more) sets of scores for the *same* individuals or matched groups.

- It lies between +1 to -1 and zero shows no relation between variables

- The higher the correlation the stronger the relationship between variables and vice versa.

TYPES OF CORRELATION

POSITIVE CORRELATION

➢ When the increase in one variable (X) is followed by a corresponding increase in the other variable (Y); it is said to be **Positive Correlation**.

➢ For example 'Heat'and 'Temperature' have a perfect positive correlation.

NEGATIVE CORRELATION

➢ The increase in one variable (X) results in a corresponding decrease in the other variable (Y), the correlation is said to be **Negative Correlation**.

➢ For example: A student who has many absences has a decrease in grades.

TYPES OF CORRELATION

SIMPLE CORRELATION

▪ Under simple correlation problem there are only two variables are studied.

MULTIPLE CORRELATION

▪ Under Multiple Correlation three or more than three variables are studied. Ex. Q_d = f (P,P_C, P_S, t, y)

PARTITAL CORRELATION

▪ analysis recognizes more than two variables but considers only two variables keeping the other constant.

TYPES OF CORRELATION

LINEAR CORRELATION

- Correlation is said to be linear when the amount of change in one variable tends to bear a constant ratio to the amount of change in the other. The graph of the variables having a linear relationship will form a straight line.

NON – LINEAR CORRELATION

- The correlation would be non linear if the amount of change in one variable does not bear a constant ratio to the amount of change in the other variable.

❖Scatter Diagram Method

METHODS OF CORRELATION

The scatter diagram graphs pairs of numerical data, with one variable on each axis, to look for a relationship between them. If the variables are correlated, the points will fall along a line or curve.

METHODS OF CORRELATION

- Scatter Diagram is a graph of observed plotted points where each points represents the values of X & Y as a coordinate. It portrays the relationship between these two variables graphically.

- Simple & Non Mathematical method

- Not influenced by the size of extreme item

- First step in investing the relationship between two variables

METHODS OF CORRELATION

❖KARL PEARSON'S COEFFICIENT OF CORRELATION

- Most widely used measure of Correlation.

- This method is also known as Pearson's Product Moment Method and Product Moment Method.

- The Coefficient of Correlation relation is denoted by symbol r.

METHODS OF CORRELATION

- The standard formula used in the computation of Pearson's Product Moment Correlation Coefficient is as follows:

$$\frac{N\sum XY - \sum X \sum Y}{\sqrt{N\sum X^2 - (\sum X)^2 X\, N\sum Y^2 - (\sum Y)^2}}$$

METHODS OF CORRELATION

❖ Spearman's Rank Coefficient of Correlation

- When statistical series in which the variables under study are not capable of quantitative measurement but can be arranged in serial order, in such situation Pearson's correlation coefficient can not be used in such case Spearman Rank correlation can be used.

- $R = 1 - \{6\sum D^2 / N(N^2 - 1)\}$

- R = Rank correlation coefficient

- D = Difference of rank between paired item in two series.

- N = Total number of observation.

Advantages & Disadvantages of Correlation

ADVANTAGES

- Show the amount (strength) of relationship present
- Can be used to make predictions about the variables under study.
- Can be used in many places, including natural settings, libraries, etc.
- Easier to collect co relational data

DISADVANTAGES

- Can't assume that a cause-effect relationship exists
- Little or no control (experimental manipulation) of the variables is possible
- Relationships may be accidental or due to a third, unmeasured factor common to the 2 variables that are measured

REGRESSION

Introduction

- Regression Analysis is a very powerful tool in the field of statistical analysis in predicting the value of one variable, given the value of another variable, when those variables are related to each other.

- Regression Analysis is mathematical measure of average relationship between two or more variables.

- Regression analysis is a statistical tool used in prediction of value of unknown variable from known variable.

TYPES OF REGRESSION

- SIMPLE AND MULTIPLE REGRESSION

- LINEAR AND CURVI LINEAR REGRESSION

- PARTIAL AND TOTAL REGRESSION

- STEPWISE REGRESSION

METHODS OF REGRESSION

SIMPLE REGRESSION MODEL

y = a+bx+e

- Is calculated with the one independent and one dependent variable.

MULTIPLE REGRESSION MODEL

y =a+b1x1+b2x2+bnxn+e

- Is calculated with more than two variables.
- One is dependent variables and others are independent variables

LEAST SQUARE METHOD

- Here mathematical equation is formed to predict the value of unknown variable from given known variables
- Least square method is mostly used model and method of regression analysis

Advantages & Disadvantages of Regression

ADVANTAGES

- Accuracy of Results
- Assessment Tools
- Use of Multi-Variables
- Input for New Management Trends

DISADVANTAGES

- Very lengthy and complicated procedure of calculations
- Cannot be used in case of qualitative phenomenon

COMPARISION TABLE OF CORRELATION & REGRESSION.

	CORRELATION	REGRESSION
PURPOSE	The main purpose of correlation analysis is to predict which are the most dependable forecasts.	The main purpose of regression analysis is to predict or estimate the unknown variable with the help of known variable.
SCOPE	Correlation analysis has limited applications	Regression analysis has wider applications.
USAGE	To represent linear relationship between two variables.	To fit a best line and estimate one variable on the basis of another variable.
NATURE OF VARIABLES	Both the variables are mutually dependent.	One variable is dependent and other variable is independent.

COMPARISION TABLE OF CORRELATION & REGRESSION.

RESPONDING NATURE	Independent of the change of Origin or change of Scale.	Independent of the change of Origin but dependent on the change of Scale.
MEASURES	Measures the degree to which two variables move together,	Describes the fundamental level the nature of any linear relationship between two variables.
RELATIONSHIP	Correlation is confined to the linear relationship between variables only.	Regression studies linear and Non-linear relationships also.
Association	Correlation coefficient measures the extent and direction of a linear association between two variables.	Linear regression allows us to describe one variable as a linear function of another variable.

THANK YOU