SIMPLE AND LINEAR CORRELATION ANALYSIS: MEANING, MEASUREMENT

By

Dr. Santosh Kumar Lal

Dept. of Commerce

Sariya College, Suriya

Classified as Microsoft Confidential

SIMPLE AND LINEAR CORRELATION ANALYSIS: MEANING, MEASUREMENT

Correlation is a statistical technique used to measure the **direction and degree of relationship between two or more variables**.

For example:

- Income and expenditure
- Price and demand
- Advertising expenditure and sales
- Production and cost
- Marks obtained and study hours
- If a change in one variable is associated with a change in another variable, the variables are said to be **correlated**.
- **Correlation analysis does not necessarily establish cause-and-effect relationship.** It only tells us how closely two variables move together.

सहसंबंध विश्लेषण का अर्थ

सहसंबंध (Correlation) एक सांख्यिकीय तकनीक है जिसके द्वारा दो या दो से अधिक चरों के बीच पाए जाने वाले **संबंध की दिशा और मात्रा/डिग्री** को मापा जाता है।

उदाहरण:

- आय और व्यय
- कीमत और माँग
- विज्ञापन व्यय और बिक्री
- उत्पादन और लागत
- अध्ययन के घंटे और प्राप्त अंक
- यदि एक चर में परिवर्तन होने पर दूसरे चर में भी व्यवस्थित रूप से परिवर्तन होता है, तो दोनों चरों को **सहसंबंधित** (Correlated) कहा जाता है।

ध्यान देने योग्य बात यह है कि **सहसंबंध से कारण-परिणाम संबंध** (Cause and Effect Relationship) **सिद्ध नहीं होता।** यह केवल यह बताता है कि दो चर किस सीमा तक साथ-साथ परिवर्तित होते हैं।

SIMPLE CORRELATION

When the relationship between only two variables is studied, it is called Simple Correlation.

For example, studying the relationship between:

- Sales and advertising expenditure
- Demand and price
- Income and consumption

Thus, simple correlation is concerned with the relationship between two variables, generally represented by X and Y.

सरल सहसंबंध

जब **केवल दो चरों** के बीच संबंध का अध्ययन किया जाता है, तो उसे **सरल सहसंबंध** (Simple Correlation) कहते हैं।

उदाहरण के लिए:

- बिक्री और विज्ञापन व्यय
- माँग और कीमत
- आय और उपभोग

इसमें सामान्यतः दो चरों को X **और** Y द्वारा प्रदर्शित किया जाता है।

LINEAR CORRELATION

- When the relationship between two variables can be represented approximately by a straight line, it is called Linear Correlation.
- In linear correlation, a change in one variable is associated with a relatively constant change in the other variable.
- For example, if an increase in advertising expenditure is generally accompanied by a proportionate increase in sales, the relationship may be considered approximately linear

- जब दो चरों के बीच संबंध को लगभग **एक सीधी रेखा** (Straight Line) द्वारा प्रदर्शित किया जा सकता है, तो उसे **रैखिक सहसंबंध** (Linear Correlation) कहते हैं।
- रैखिक सहसंबंध में एक चर में परिवर्तन के साथ दूसरे चर में भी लगभग नियमित दर से परिवर्तन होता है।
- उदाहरण के लिए, विज्ञापन व्यय में वृद्धि के साथ बिक्री में भी लगभग समान दिशा में वृद्धि हो रही हो, तो दोनों के बीच रैखिक सहसंबंध माना जा सकता है।

FEATURES OF SIMPLE AND LINEAR CORRELATION

- It studies the relationship between two variables.
- It indicates the direction of relationship.
- It measures the degree or strength of relationship.
- Linear correlation assumes that the relationship can be represented approximately by a straight line.
- The correlation coefficient is generally represented by r.
- The value of r lies between -1 and +1.
- It is useful in business forecasting and decision-making.
- Correlation does not by itself prove causation.

सरल एवं रैखिक सहसंबंध की विशेषताएँ

- इसमें मुख्यतः **दो चरों** के बीच संबंध का अध्ययन किया जाता है।
- यह संबंध की **दिशा** बताता है।
- यह संबंध की **तीव्रता या मात्रा** को मापता है।
- रैखिक सहसंबंध में संबंध को लगभग **सीधी रेखा** से प्रदर्शित किया जा सकता है।
- सहसंबंध गुणांक को सामान्यतः r से प्रदर्शित किया जाता है।
- r **का मान** -1 **से** +1 **के बीच** होता है।
- इसका उपयोग व्यावसायिक पूर्वानुमान एवं निर्णय लेने में किया जाता है।
- केवल सहसंबंध के आधार पर कारण-परिणाम संबंध सिद्ध नहीं किया जा सकता।

DIRECTION OF CORRELATION

A. Positive Correlation

- When both variables move in the **same direction**, the correlation is positive.
- If X increases and Y also increases, or X decreases and Y also decreases, there is positive correlation.

Examples:

- Income and consumption
- Advertising expenditure and sales
- Production and total cost

सहसंबंध की दिशा

A. **धनात्मक सहसंबंध**

- जब दोनों चर **एक ही दिशा में** परिवर्तन करते हैं, तो उसे **धनात्मक सहसंबंध** (Positive Correlation) कहते हैं।
- अर्थात् X बढ़ने पर Y भी बढ़े अथवा X घटने पर Y भी घटे, तो दोनों के बीच धनात्मक सहसंबंध होता है।

उदाहरण: आय और उपभोग, विज्ञापन और बिक्री, उत्पादन और कुल लागत।

NEGATIVE CORRELATION

- When the variables move in opposite directions, the correlation is negative.
- If X increases and Y decreases, or vice versa, the correlation is negative.

Examples:

- Price and demand
- Interest rate and investment, under appropriate conditions
- Price and quantity demanded

ऋणात्मक सहसंबंध

जब दो चर **विपरीत दिशाओं में** परिवर्तन करते हैं, तो उसे **ऋणात्मक सहसंबंध** (Negative Correlation) कहते हैं।

यदि X बढ़ता है और Y घटता है अथवा X घटता है और Y बढ़ता है, तो दोनों के बीच ऋणात्मक सहसंबंध होता है।

उदाहरण: कीमत और माँग।

ZERO CORRELATION

When there is no systematic linear relationship between two variables, the linear correlation coefficient is zero or approximately zero.

शून्य सहसंबंध

जब दो चरों के बीच कोई व्यवस्थित **रैखिक संबंध** नहीं पाया जाता, तो रैखिक सहसंबंध गुणांक का मान **शून्य या शून्य के निकट** हो सकता है।

MEASUREMENT OF CORRELATION

Simple and linear correlation can be measured through different methods. The most important method in Business Statistics is:

Karl Pearson's Coefficient of Correlation

- It is represented by **r**.

KARL PEARSON’S COEFFICIENT OF CORRELATION

Karl Pearson's coefficient of correlation is one of the most widely used statistical measures for determining the degree and direction of linear relationship between two variables.

- The value of **r** ranges from:
- $-1 \leq r \leq +1$

कार्ल पियर्सन का सहसंबंध गुणांक दो चरों के बीच **रैखिक संबंध की दिशा तथा मात्रा** को मापने की सबसे महत्वपूर्ण सांख्यिकीय विधियों में से एक है।

- इसका मान:
- $-1 \leq r \leq +1$

के बीच होता है।

DIRECT METHOD – प्रत्यक्ष विधि

- **X** = First variable / **पहला चर**
- **Y** = Second variable / **दूसरा चर**
- **X̄** = Mean of X / X **का माध्य**
- **Ȳ** = Mean of Y / Y **का माध्य**
- **r** = Coefficient of correlation / **सहसंबंध गुणांक**

इस विधि में पहले दोनों चरों का arithmetic mean निकाला जाता है। इसके बाद प्रत्येक observation के mean से deviations निकाले जाते हैं। इन deviations के आधार पर सूत्र से r की गणना की जाती है।

INTERPRETATION OF THE VALUE OF R
R के मान की व्याख्या

Value of r	English Interpretation	हिंदी व्याख्या
+1	Perfect positive correlation	पूर्ण धनात्मक सहसंबंध
+0.75 to +0.99	Very high positive correlation	बहुत उच्च धनात्मक सहसंबंध
+0.50 to +0.74	High/Moderate positive correlation	उच्च/मध्यम धनात्मक सहसंबंध
+0.25 to +0.49	Low positive correlation	निम्न धनात्मक सहसंबंध
0	No linear correlation	कोई रैखिक सहसंबंध नहीं
-0.25 to -0.49	Low negative correlation	निम्न ऋणात्मक सहसंबंध
-0.50 to -0.74	Moderate/High negative correlation	मध्यम/उच्च ऋणात्मक सहसंबंध
-0.75 to -0.99	Very high negative correlation	बहुत उच्च ऋणात्मक सहसंबंध
-1	Perfect negative correlation	पूर्ण ऋणात्मक सहसंबंध

CHARACTERISTICS OF CORRELATION COEFFICIENT

- The value of r lies between -1 and +1.
- The sign of r indicates the direction of relationship.
- The absolute value of r indicates the strength of the linear relationship.
- If $r = +1$, there is perfect positive correlation.
- If $r = -1$, there is perfect negative correlation.
- If $r = 0$, there is no linear correlation.
- The coefficient is unit-free.
- Interchanging X and Y does not change the value of r.
- Correlation does not necessarily imply causation.

सहसंबंध गुणांक की विशेषताएँ

- r **का मान** -1 **और** +1 **के बीच** होता है।
- r **का चिह्न** (+ **या** -) संबंध की दिशा बताता है।
- |r| **का मान** रैखिक संबंध की तीव्रता बताता है।
- r = +1 होने पर पूर्ण धनात्मक सहसंबंध होता है।
- r = -1 होने पर पूर्ण ऋणात्मक सहसंबंध होता है।
- r = 0 होने पर कोई रैखिक सहसंबंध नहीं होता।
- सहसंबंध गुणांक **इकाई-रहित** (Unit-free) होता है।
- X और Y को आपस में बदलने पर भी r का मान नहीं बदलता।
- सहसंबंध का अर्थ अनिवार्य रूप से कारण-परिणाम संबंध नहीं है।

IMPORTANCE IN BUSINESS STATISTICS

Correlation analysis is particularly useful in business because it helps managers understand relationships among economic and business variables.

Major applications include:

- **Sales forecasting:** Relationship between advertising and sales.
- **Demand analysis:** Relationship between price and demand.
- **Cost analysis:** Relationship between production and cost.
- **Financial analysis:** Relationship among financial variables.
- **Market research:** Studying relationships between consumer characteristics and purchasing behaviour.
- **Business planning:** Understanding relationships helps in managerial decision-making.

BUSINESS STATISTICS में सहसंबंध का महत्व

व्यवसाय में अनेक आर्थिक एवं व्यावसायिक चर एक-दूसरे से जुड़े होते हैं। इसलिए सहसंबंध विश्लेषण प्रबंधकों के लिए महत्वपूर्ण है।

मुख्य उपयोग:

- **बिक्री पूर्वानुमान:** विज्ञापन और बिक्री के बीच संबंध।
- **माँग विश्लेषण:** कीमत और माँग के बीच संबंध।
- **लागत विश्लेषण:** उत्पादन और लागत के बीच संबंध।
- **वित्तीय विश्लेषण:** विभिन्न वित्तीय चरों के बीच संबंध।
- **बाजार अनुसंधान:** उपभोक्ता विशेषताओं और खरीद व्यवहार के बीच संबंध।
- **व्यावसायिक योजना:** विभिन्न चरों के बीच संबंधों की जानकारी निर्णय लेने में सहायता करती है।

LIMITATIONS OF CORRELATION ANALYSIS

- Correlation does not prove **cause and effect**.
- A high correlation may sometimes arise due to a **third variable**.
- Pearson's correlation measures **linear association**; it may not adequately describe a nonlinear relationship.
- Extreme observations or **outliers** can substantially affect the coefficient.
- Correlation alone does not provide a complete explanation of the business relationship.
- The reliability of the result depends on the quality and representativeness of the data.

सहसंबंध विश्लेषण की सीमाएँ

- सहसंबंध से **कारण-परिणाम संबंध** सिद्ध नहीं होता।
- उच्च सहसंबंध किसी **तीसरे चर** के प्रभाव के कारण भी हो सकता है।
- पियर्सन का गुणांक मुख्यतः **रैखिक संबंध** को मापता है।
- असामान्य या **अत्यधिक मान** (Outliers) सहसंबंध गुणांक को प्रभावित कर सकते हैं।
- केवल सहसंबंध से व्यावसायिक संबंध की पूरी व्याख्या नहीं की जा सकती।
- परिणाम की विश्वसनीयता data की गुणवत्ता एवं प्रतिनिधित्व पर निर्भर करती है।

सरल एवं रैखिक सहसंबंध विश्लेषण : अर्थ एवं मापन

Simple and Linear Correlation Analysis: Meaning and Measurement

1. अर्थ (Meaning)

- सहसंबंध विश्लेषण दो चरों के मध्य संबंध की दिशा और उसकी डिग्री को मापने की एक सांख्यिकीय विधि है।
- सरल सहसंबंध में केवल दो चरों (X और Y) के बीच संबंध का अध्ययन किया जाता है।
- रैखिक सहसंबंध का अर्थ है कि दोनों चर परिवर्तन एक सीधी रेखा के रूप में एक-दूसरे से संबंधित होते हैं।

सरल एवं रैखिक सहसंबंध की विशेषताएँ

1. केवल दो चरों के बीच संबंध का अध्ययन करता है।
2. संबंध की दिशा (धनात्मक या ऋणात्मक) बताता है।
3. संबंध की डिग्री (कमजोर, मध्यम या मजबूत) बताता है।
4. यह विश्लेषण केवल रैखिक संबंध को मापता है, अर्थात संबंध को एक सीधी रेखा से निरूपित किया जा सकता है।

2. मापन (Measurement)

सरल एवं रैखिक सहसंबंध को मापने के लिए सहसंबंध गुणांक (Coefficient of Correlation) का उपयोग किया जाता है, जिसे 'r' द्वारा निरूपित किया जाता है। यह गुणांक −1 से +1 के बीच मान ग्रहण करता है।

(क) पियसन उत्पाद क्षण सहसंबंध गुणांक (Karl Pearson's Coefficient of Correlation)

(i) प्रत्यक्ष विधि (Direct Method)

$$r = \frac{\sum (X-\overline{X})(Y-\overline{Y})}{\sqrt{\sum (X-\overline{X})^2 \times \sum (Y-\overline{Y})^2}}$$

जहाँ, X = स्वतंत्र चर, Y = आश्रित चर

$\overline{X}$ = X का माध्य, $\overline{Y}$ = Y का माध्य

(ii) संक्षिप्त विधि (Short-cut Method)

$$r = \frac{n\sum XY - (\sum X)(\sum Y)}{\sqrt{[n\sum X^2 - (\sum X)^2][n\sum Y^2 - (\sum Y)^2]}}$$

जहाँ, n = प्रेक्षणों की संख्या

(ख) सहसंबंध गुणांक की व्याख्या (Interpretation of 'r')

r का मान	संबंध की डिग्री	संबंध की प्रकृति
+ 1.00	पूर्ण सहसंबंध	पूर्ण धनात्मक (Positive Perfect Correlation)
+0.75 से +0.99	मजबूत सहसंबंध	धनात्मक (Positive Correlation)
+0.25 से +0.74	मध्यम सहसंबंध	धनात्मक (Positive Correlation)
+0.00 से +0.24	कमजोर सहसंबंध	धनात्मक (Positive Correlation)
0.00	कोई सहसंबंध नहीं	कोई रैखिक संबंध नहीं (No Linear Correlation)
−0.00 से −0.24	कमजोर सहसंबंध	ऋणात्मक (Negative Correlation)
−0.25 से −0.74	मध्यम सहसंबंध	ऋणात्मक (Negative Correlation)
−0.75 से −0.99	मजबूत सहसंबंध	ऋणात्मक (Negative Correlation)
−1.00	पूर्ण सहसंबंध	पूर्ण ऋणात्मक (Negative Perfect Correlation)

सहसंबंध गुणांक की विशेषताएँ

- r का मान −1 और +1 के बीच होता है।
- r का चिह्न (+ या −) संबंध की दिशा को दर्शाता है।
- |r| का मान संबंध की शक्ति (डिग्री) को दर्शाता है।
- r इकाई रहित (Unitless) होता है, इसलिए विभिन्न श्रृंखलाओं की तुलना की जा सकती है।
- r कारण-कार्य संबंध को सिद्ध नहीं करता, यह केवल सहसंबंध को मापता है।

निष्कर्ष

सरल एवं रैखिक सहसंबंध विश्लेषण दो चरों के बीच संबंध की दिशा और डिग्री को मापने का सरल और उपयोगी उपकरण है। व्यवसायिक निर्णय, पूर्वानुमान, नियंत्रण तथा योजना-निर्माण में इसका व्यापक उपयोग किया जाता है।

1. Meaning

- Correlation analysis is a statistical method used to measure the direction and degree of relationship between two variables.
- Simple correlation involves the study of relationship between only two variables (X and Y).
- Linear correlation means that changes in the two variables are related to each other in the form of a straight line.

Features of Simple and Linear Correlation

1. It studies the relationship between only two variables.
2. It indicates the direction of relationship (positive or negative).
3. It measures the degree of relationship (weak, moderate or strong).
4. It measures only linear relationship, i.e., the relationship can be represented by a straight line.

2. Measurement

Simple and linear correlation is measured by the Coefficient of Correlation represented by 'r'. Its value lies between -1 and +1.

(a) Karl Pearson's Coefficient of Correlation

(i) Direct Method

$$r = \frac{\sum (X-\overline{X})(Y-\overline{Y})}{\sqrt{\sum (X-\overline{X})^2 \times \sum (Y-\overline{Y})^2}}$$

Where, X = Independent variable, Y = Dependent variable

$\overline{X}$ = Mean of X, $\overline{Y}$ = Mean of Y

(ii) Short-cut Method

$$r = \frac{n\sum XY - (\sum X)(\sum Y)}{\sqrt{[n\sum X^2 - (\sum X)^2][n\sum Y^2 - (\sum Y)^2]}}$$

Where, n = Number of observations

(b) Interpretation of 'r'

Value of r	Degree of Relationship	Nature of Relationship
+ 1.00	Perfect correlation	Positive Perfect Correlation
+0.75 to +0.99	High correlation	Positive Correlation
+0.25 to +0.74	Moderate correlation	Positive Correlation
+0.00 to +0.24	Low correlation	Positive Correlation
0.00	No correlation	No Linear Correlation
−0.00 to −0.24	Low correlation	Negative Correlation
−0.25 to −0.74	Moderate correlation	Negative Correlation
−0.75 to −0.99	High correlation	Negative Correlation
−1.00	Perfect correlation	Negative Perfect Correlation

Characteristics of Coefficient of Correlation

- The value of r lies between −1 and +1.
- The sign of r (+ or −) indicates the direction of relationship.
- The absolute value |r| indicates the strength (degree) of relationship.
- r is unitless, so it can be used to compare different series.
- r does not establish cause and effect relationship; it only measures association.

Conclusion

Simple and linear correlation analysis is a simple and useful tool to measure the direction and degree of relationship between two variables. It is widely used in business decisions, forecasting, control and planning.

CONCLUSION

Simple and linear correlation analysis is an important part of Business Statistics. It provides a quantitative measure of the **direction and strength of the linear relationship between two variables**. Karl Pearson's coefficient of correlation, represented by **r**, is the most commonly used measure. Its value ranges from **-1 to +1**, where the sign indicates direction and the magnitude indicates strength.

निष्कर्ष

सरल एवं रैखिक सहसंबंध विश्लेषण Business Statistics का एक महत्वपूर्ण विषय है। इसके द्वारा दो चरों के बीच संबंध की **दिशा और तीव्रता** को मात्रात्मक रूप से मापा जाता है। कार्ल पियर्सन का सहसंबंध गुणांक r इसका प्रमुख माप है, जिसका मान -1 **से** +1 के बीच होता है। इसका चिह्न संबंध की दिशा तथा इसका परिमाण संबंध की तीव्रता को दर्शाता है।