



B.Com Honours

Semester IV

Calicut University

BUSINESS STATISTICS

Course Code: COM4CJ205 • Module 2 Notes

Module II: Correlation and Regression Analysis

Covers bivariate statistical association using Karl Pearson's correlation, Spearman's rank correlation, and simple linear regression modeling for business forecasting.

Unit 3: Correlation Analysis – Karl Pearson's (r) & Rank Correlation

1. Karl Pearson's Coefficient of Correlation ((r)):

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

Properties: Limits are $(-1 \leq r \leq +1)$; independent of change of origin and scale.

Coefficient of Determination ((r²)): Proportion of total variance in dependent variable explained by independent variable. Coefficient of Non-determination = $(1 - r^2)$.

2. Spearman's Rank Correlation ((r_s)): Applicable for qualitative data (beauty, intelligence, ethics):

$$r_s = 1 - \frac{6 \sum D^2}{N(N^2 - 1)} \quad ((D = R_1 - R_2))$$

Unit 4: Regression Analysis & Simple Linear Regression Equations

1. Regression Equation of (Y) on (X) (estimating Sales (Y) from Ad Spend (X)):

$$(Y - \bar{Y}) = b_{yx} (X - \bar{X}) \quad | \quad (b_{yx} = \frac{N \sum XY - (\sum X)(\sum Y)}{N \sum X^2 - (\sum X)^2})$$

2. Regression Equation of (X) on (Y) (estimating Ad Spend (X) from Sales (Y)):

$$(X - \bar{X}) = b_{xy} (Y - \bar{Y}) \quad | \quad (b_{xy} = \frac{N \sum XY - (\sum X)(\sum Y)}{N \sum Y^2 - (\sum Y)^2})$$

3. Interrelationship: $(r = \pm \sqrt{b_{yx} \cdot b_{xy}})$. Standard Error of Estimate: $(S_{ey} = S_y \sqrt{1 - r^2})$.

Practical Problem 2 (Karl Pearson Correlation & Linear Regression Solved Computation): Advertising

Expenditure ((X), in Lakhs) and Sales Revenue ((Y), in Lakhs) for 5 quarters:

(X) values: 2, 4, 6, 8, 10 (Total $(\sum X = 30)$); (Y) values: 5, 7, 11, 13, 14 (Total $(\sum Y = 50)$).

Calculating products: $(\sum X^2 = 220)$, $(\sum Y^2 = 560)$, $(\sum XY = 354)$. Calculate (a) Karl Pearson's Correlation Coefficient ((r)), (b) Coefficient of Determination ((r²)), (c) Regression Coefficient (b_{yx}), and (d) Estimate Sales ((Y)) when Advertising Expenditure ($X = 12$ ext{ Lakhs}).

Step-by-Step Solved Correlation & Regression Calculation:

- **1. Means Computation:** $(\bar{X} = 30 / 5 = 6.0)$; $(\bar{Y} = 50 / 5 = 10.0)$.
- **2. Karl Pearson Correlation Coefficient ((r)):**
Numerator = $(5(354) - (30)(50) = 1770 - 1500 = 270)$.
Denominator = $(\sqrt{[5(220) - 900][5(560) - 2500]} = \sqrt{[200][300]} = \sqrt{60,000} = 244.949)$.
 $(r = 270 / 244.949 = \mathbf{+0.9839})$ (Strong positive correlation!)
Coefficient of Determination $(r^2 = (0.9839)^2 = \mathbf{0.9681}$ ext{ (or 96.81% variance explained)}).
- **3. Regression Slope Coefficient ((b_{yx})):**

$$(b_{yx}) = 270 / 200 = 1.35$$

• **4. Linear Regression Equation of (Y) on (X):**

$(Y - 10 = 1.35 (X - 6))$ implies $Y - 10 = 1.35 X - 8.10$ implies $\mathbf{Y = 1.35 X + 1.90}$.

• **5. Sales Prediction for (X = 12 ext{ Lakhs}):**

Estimated Sales $(Y = 1.35(12) + 1.90 = 16.20 + 1.90 = \mathbf{ext{Rs. 18.10 Lakhs}})$

Practical Problem 3 (Spearman's Rank Correlation Solved Computation): Two judges rank 5 job candidates for executive selection:

Judge 1 Ranks ((R₁)): 1, 2, 3, 4, 5; Judge 2 Ranks ((R₂)): 2, 1, 4, 3, 5. Calculate Spearman's Rank Correlation Coefficient ((r_s)).

Step-by-Step Rank Correlation Table:

Candidate	Judge 1 Rank ((R ₁))	Judge 2 Rank ((R ₂))	Difference (D = R ₁ - R ₂)	Squared Difference (D ²)
A	1	2	-1	1
B	2	1	+1	1
C	3	4	-1	1
D	4	3	+1	1
E	5	5	0	0
TOTAL	-	-	0	(sum D² = 4)

$(r_s = 1 - \frac{6 \times 4}{5(25 - 1)}) = 1 - \frac{24}{120} = 1 - 0.20 = \mathbf{+0.80})$ (High agreement between judges!)

Acing your exams is just a click away!

Visit www.degreeelive.in to download the next module for free.

DegreeLive