

## INDUSTRIAL ENGINEERING

SIXTH SEMESTER- (REGULATION-2008)

**IE9351 MULTI VARIATE STATISTICAL ANALYSIS**

(Use of Statistical Tables are permitted)

Time: 3 hr

Max. Mark: 100

Answer ALL Questions

Part-A (10 X 2 = 20 Mark)

1. Draw and Define inverse relationship.
2. What are the advantages of coefficient of multiple determination?
3. State few examples for the usage of hypothesis in multivariate analysis?
4. What are the objectives of multivariate methods?
5. List out the properties of wishart distributions?
6. What is the formula used for proportion of total population variance?
7. What are the applications of communality?
8. List out the applications of discriminant analysis?
9. Define Association coefficient.
10. State the sampling distribution for multivariate normal data when no of variables is one and no of groups is greater than or equal two.

Part-B (5 X 16 = 80 Marks)

11. How will you design the Discriminant function? Explain the Computation of discriminant ratio, K and identify the variable which is more important in relation to the other variable? How will you validate the discriminant function using the set of data by forming groups based on the critical discriminant score? How will you frame the hypothesis?

12 (a) Explain the calculation of sample coefficient of determination and the sample coefficient of correlation? (8)

(ii) Explain the Multicollinearity in Multiple Regression with suitable illustrations? (8)

OR

12 (b)(i) In finance, it is of interest to look at the relationship between Y, a stock's average return, and X, the overall market return. The slope coefficient computed by linear regression is

called the stock's beta by investment analysts. A beta greater than 1 indicates that the stock is relatively sensitive to changes in the market; a beta less than 1 indicates that the stock is relatively insensitive. For the following data, compute the beta and test to see whether it is significantly less than 1. Use alpha equal to 0.05. (10)

Y (%)	10	12	8	15	9	11	8	10	13	11
X (%)	11	15	3	18	10	12	6	7	18	13

- (ii) In a regression problem with a sample size of 17, the slope was found to be 3.73 and the standard error of estimate 28.654. The quantity  $(\sum X^2 - n\bar{X}^2) = 871.56$ . Find the standard error of the regression slope coefficient (6)

- 13 (a) (i) How will you estimate mean and variance of multivariate normal distributions? (3)  
 (ii) How do you assess the assumptions of normality? (3)  
 (iii) Explain the construction of Q-Q plot, checking bivariate normality and a chi-square plot? (10)

OR

- 13 (b) Derive an MANOVA model for comparing g population mean vectors? (16)

- 14 (a) In a survey on a product, data on three variable have been collected, as shown in the following table 1.1. Perform factor analysis using centroid method to identify two factors which can represent the variables of the study. (16)

OR

- 14 (b) (i) In a survey, the number of years of experience of employee in two different skills are summarized in table 1.2. Perform cluster analysis using the agglomerative hierarchical clustering method. (16)

- 15 (a) Describe the Principal Component Analysis in detail? (16)

OR

- 15 (b) Explain the Rank ordered cluster algorithm in detail with a suitable illustrations? (16)

y data

Variable		
X1	X2	X3
5	7	8
3	3	5
3	1	5
1	1	5
3	2	4
3	3	5
2	2	1
6	6	5
4	2	1
4	3	2

Table 1.2 Survey of Years of Experience on Skills

Skill-X	Skill-Y
2	8
8	15
3	6
6	9
8	7
10	10