

B.E. FULL-TIME DEGREE END SEMESTER EXAMINATIONS, MAY 2012

INDUSTRIAL ENGINEERING

SIXTH SEMESTER- (REGULATION-2008)

IE9351 MULTI-VARIATE STATISTICAL ANALYSIS

14

(Use of Statistical Tables are permitted)

Time: 3 hr

Max. Mark: 100

Answer ALL Questions

Part-A (10 X 2 = 20 Mark)

1. Define inverse relationship. Give an example?
2. What are the two assumptions used in interpreting the standard error of estimate in regression analysis?
3. What are the objectives of multivariate methods?
4. List out the properties of wishart distributions?
5. How will you calculate proportion of total population variance?
6. What are the underlying assumptions in the structure of the data for one way MANOVA?
7. What is average linkage clustering method?
8. How the weights are chosen in the PCA?
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. In a survey on a product, data on three variables 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.

12 (a) Below are some data (table 1.2) on consumption expenditures, CONSUMP, disposable income, INCOME; and sex of the head of household, SEX of 12 randomly chosen families. The variable GENDER has been coded:

$$\text{GENDER} = \begin{cases} 1 & \text{if SEX} = \text{'M'} \text{ (male)} \\ 0 & \text{if SEX} = \text{'F'} \text{ (female)} \end{cases} \quad (16)$$

- (a) Determine the best fitting regression
- (b) If disposable income is held constant, is there a significant difference in consumption between households headed by a male versus those where the head of household is

female? State explicit hypothesis test them at the 0.10 level, and state an explicit conclusion.

- (c) Give an approximate 95 percent confidence interval for consumption for a household with disposable income of \$40,000 headed by a male.

Table 1.1

Respondent	Variable		
	X1	X2	X3
1	5	7	8
2	3	3	5
3	3	1	5
4	1	1	5
5	3	2	4
6	3	3	5
7	2	2	1
8	6	6	5
9	4	2	1
10	4	3	2

Table 1.2

Consump	Income (\$)	Sex	Gender
37,070	45,100	M	1
22,700	28,070	M	1
24,260	35,000	F	0
30,420	35,000	M	1
17,360	18,860	F	0
33,520	41,270	M	1
26,960	32,940	M	1
19,360	21,440	F	0
35,680	44,700	M	1
22,360	24,400	F	0
28,640	33,620	F	0
39,720	46,000	M	1

OR

- 12 (b) (i) Discuss the inferences about population parameters in multiple regression? (6)

(ii) A graduate student trying to purchase a used Neptune car has researched the prices. She believes the year of the car and the number of miles the car has been driven both influence the purchase price. Data are given below for 10 cars with the price in thousands of dollars, year and miles driven in thousands.

(a) Calculate the least square equation that best relates these three variables (7)

(b) The student would like to purchase a 1991 neptune with about 40,000 miles on it
How much do you predict she will pay? (3)

Year	1987	1992	1993	1988	1994	1991	1992	1988	1989	1991
Price (\$ thousands)	2.99	6.02	8.87	3.92	9.55	9.05	9.37	4.2	4.8	5.74
Miles (Thousands)	55.6	18.4	21.3	46.9	11.8	36.4	28.2	44.2	34.9	26.4

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

OR

13 (b) Explain the procedure involved in the computation of Multi-Variate Two Way Fixed Effects Model With Interaction? (16)

14 (a) The potential customers of a computer company rate the product of the company into good or bad based on the time to respond to breakdown calls (X1) & the percentage discount on product price (X2). The ratings by customers are presented in the following table 1.3 (16)

(a) Design the Discriminant function

(b) Compute discriminant ratio, K and identify the variable which is more important in relation to the other variable

(c) Validate the discriminant function using the given data by forming groups based on the critical discriminant score

(d) Test whether the group means are equal in importance at a significance level of 0.05

Table 1.3: Customer data

Customer	Rating	X1 (in hours)	X2 (%)
1	Good	24	5
2	Good	12	8
3	Bad	36	4
4	Bad	12	0
5	Bad	36	3
6	Good	36	10
7	Good	24	3
8	Bad	48	4
9	Bad	96	5
10	Good	36	12

OR

14 (b) Describe the Step-by-Step Method of PCA with illustrations show atleast one model calculation for the convergence of First Principal Component Loadings? (16)

the Q-Q plot for the ten samples of ordered observations: -1.00, 0.10, 0.16, 0.41, 0.62, 0.80, 1.26, 1.54, 1.71 and 2.30. Explain the steps involved in evaluating bivariate normality and constructing a chi-square plot? (16)

OR

15 (b) (i) Describe the procedure followed in ROC for groupings based on Attributes? (8)

(ii) Discuss the interpretation of the standard error of estimate and approximate prediction of intervals? (8)