

[4]

- OR iii. How common factors extracted? Elaborate **6** 3 1-6,
with some examples. 9-11 3 1-4

Total No. of Questions: 6

Total No. of Printed Pages: 4

Enrollment No.....

Q.6 Attempt any two:

- i. Explain K-Means clustering with an example. **5** 2 1-6,
9-11 4 1-4
- ii. What is hierachal clustering method? Explain
with example. **5** 2 1-6,
9-11 4 1-4
- iii. What is similarity measure? What are the
methods for measure of similarity? **5** 2 1-6,
9-11 4 1-4



Knowledge is Power

Faculty of Engineering
End Sem Examination Dec 2024
EN3BS09 Computational Statistics

Programme: B.Tech.

Branch/Specialisation: CSBS

Duration: 3 Hrs.

Maximum Marks: 60

Note: All questions are compulsory. Internal choices, if any, are indicated. Answers of Q.1 (MCQs) should be written in full instead of only a, b, c or d. Assume suitable data if necessary. Notations and symbols have their usual meaning.

	Marks	BL	PO	CO	PSO
Q.1 i. Which type of analysis involves three or more variables?	1	1	1-6, 8-11	1	1-4
(a) Univariate statistical analysis					
(b) Bivariate statistical analysis					
(c) Multivariate statistical analysis					
(d) All of these					
ii. Descriptive analysis gives-	1	1	1-6, 8-11	1	1-4
(a) Measures of location					
(b) Measures of spread					
(c) Measures of variation					
(d) All of these					
iii. Which analysis is portrayed by the equation: $Y = b_0 + b_1X_1 + b_2X_2 + b_3X_3 \dots + b_nX_n$	1	1	1-6, 8-11	2	1-4
(a) Simple regression					
(b) Multiple regression					
(c) Stepwise regression					
(d) All of these					
iv. Residual is-	1	1	1-6, 8-11	2	1-4
(a) Estimated error					
(b) Difference of observed value and estimated value					
(c) Both (a) and (b)					
(d) None of these					

	[2]					[3]									
v.	Which of the following represent linear combinations of two random variables X_1 and X_2 ?	1	1	1-6, 9-11	3	1-4	Q.2	i.	Explain following terms with examples-	4	1	1-6, 8-11	1	1-4	
	(a) $Y = a_1X_1 + a_2X_2$							(a) Mean	(b) Variance						
	(b) $Y = a_1X_1 * a_2X_2$						ii.	Compute Mean vector and variance-covariance matrix of following set of data-		6	2	1-6, 8-11	1	1-4	
	(c) $Y = a_1X_1 / a_2X_2$							Price of Book (X)	No. of copies sold						
	(d) None of these							42	4						
vi.	What is the primary goal of principal component analysis?	1	1	1-6, 9-11	3	1-4			52	5					
	(a) Classification							48	4						
	(b) Clustering							58	3						
	(c) Dimensionality reduction						OR	iii.	What kind of estimation of parameters needed in multivariate data? Explain with example.	6	2	1-6, 8-11	1	1-4	
	(d) Regression														
vii.	Factor analysis is very useful when number of factors are relatively-	1	1	1-6, 9-11	3	1-4	Q.3	i.	What is the difference between the simple linear regression model and the multiple linear regression model?	4	2	1-6, 8-11	2	1-4	
	(a) Higher than number of variables X							ii.	Calculate the least square estimates and find best fitted equation of following data. Also calculate the residuals and residual sum of square.	6	2	1-6, 8-11	2	1-4	
	(b) Equal to number of variables X							Z : 0	1 2 3 4						
	(c) Zero							y : 1	4 3 8 9						
	(d) Less than number of variables X						OR	iii.	Explain multivariate regression with example.	6	2	1-6, 8-11	2	1-4	
viii.	Which of the following indicates how strongly a measured variable is correlated with a factor?	1	1	1-6, 9-11	3	1-4									
	(a) Factor β	(b) Discriminator					Q.4	i.	Explain discriminant analysis with an example.	4	1	1-6, 9-11	3	1-4	
	(c) Factor link	(d) None of these						ii.	What are objectives of principal component analysis and discriminant analysis? Give example for each.	6	3	1-6, 9-11	3	1-4	
ix.	Which of the following is required by K-means clustering?	1	1	1-6, 9-11	4	1-4		OR	iii.	Explain following in discriminant analysis-	6	3	1-6, 9-11	3	1-4
	(a) Defined distance metric								(a) Two group discriminant analysis						
	(b) Number of clusters								(b) Discriminant function						
	(c) Initial guess as to cluster centroids														
	(d) All of these														
x.	Which of the following is finally produced by Hierarchical Clustering?	1	1	1-6, 9-11	4	1-4	Q.5	i.	What do you mean by factor loading? Explain with an example.	4	1	1-6, 9-11	3	1-4	
	(a) Final estimate of cluster centroids							ii.	Explain following terms which are concerned with factor analysis. Give example for each-	6	3	1-6, 9-11	3	1-4	
	(b) Tree showing how close things are to each other								(a) Observable data						
	(c) Assignment of each point to clusters								(b) Unobservable data						
	(d) All of these														

Scheme of Marking

Computational Statistics-EN3BS09

- | | | |
|-------|---|--------------------------|
| Q.1 | i) Which type of analysis involves three or more variables? | 1 |
| | c) Multivariate statistical analysis | |
| ii) | Descriptive Analysis gives | 1 |
| | d) All of the above | |
| iii) | Which analysis is portrayed by the equation:
$Y = b_0 + b_1X_1 + b_2X_2 + b_3X_3 \dots + b_nX_n$ | 1 |
| | b) Multiple regression | |
| iv) | Residual is | 1 |
| | c) Both a and b. | |
| v) | Which of the following represent linear combinations of two random variables X_1 and X_2 | 1 |
| | a) $Y = a_1X_1 + a_2X_2$ | |
| vi) | c) Dimensionality reduction | 1 |
| vii) | Factor analysis is very useful when number of factors are relatively | 1 |
| | d) Less than number of variables X. | |
| viii) | Which of the following indicates how strongly a measured variable is correlated with a factor? | 1 |
| | d) Factor loading | |
| ix) | Which of the following is required by K-means clustering? | 1 |
| | d) All of these | |
| x) | Which of the following is finally produced by Hierarchical Clustering? | 1 |
| | b) Tree showing how close things are to each other | |
| Q.2 | i. Explain following terms with examples | 4 |
| | a. Mean | |
| | b. Variance | 2-Marks each |
| ii. | Compute Mean vector and variance- covariance matrix of following set of data. | 6 |
| | | 3-Marks each |
| | Price of Book ,X | No of copies sold |
| | 42 | 4 |
| | 52 | 5 |
| | 48 | 4 |
| | 58 | 3 |

- | | | | |
|-----|------|--|---|
| OR | iii. | What kind of estimation of parameters needed in multivariate data? Explain with example.
Explanation 4-Marks
Example 2 Marks | 6 |
| Q.3 | i. | What is the difference between the simple linear regression model and the multiple linear regression model ?
At-least two difference 2 Marks each | 4 |
| | ii. | Calculate the least square estimates and find best fitted equation of following data. Also calculate the residuals and residual sum of square.
Z : 0 1 2 3 4
y : 1 4 3 8 9 | 6 |
| | | Least square estimates 3-Marks
Best fitted equation 1-Marks
Residuals and Residual sum of square 2-Marks each | |
| OR | iii. | Explain multivariate regression with example.

Explanation 4-Marks
Example 2-Marks | 6 |
| Q.4 | i. | Explain discriminant analysis with an example.

Explanation 2-Marks
Example 2-Marks | 4 |
| | ii. | What are objectives of Principal Component Analysis and Discriminant Analysis ? Give example for each.

3-Marks each | 6 |
| OR | iii. | Explain following in discriminant analysis
a. Two group discriminant analysis 3-Marks
b. Discriminant function 3-Marks | 6 |
| Q.5 | i. | What do you mean by factor loading? Explain with an example.
Explanation 2-Marks
Example 2-Marks | 4 |

[2]

- ii. Explain following terms which are concerned with factor analysis. **6**
Give example for each
a. Observable data **3-Marks**
b. Unobservable data **3-Marks**
- OR iii. How common factors extracted? Elaborate with some examples. **6**
Explanation **4-Marks**
Example **2-Marks**

[3]

- Q.6 Attempt any two:
- i. Explain K-Means clustering with an example. **5**
Explanation **3-Marks**
Example **2-Marks**
- ii. What is hierachal clustering method? Explain with example. **5**
Explanation **3-Marks**
Example **2-Marks**
- iii. What is similarity measure ? What are the methods for measure of similarity ?
Similarity measure **3-Marks**
similarity measure methods **2-Marks**
