13/2/23

Reg. No. | | | | | | | | | | | | |

Question Paper Code	11710

B.E. / B.Tech. - DEGREE EXAMINATIONS, NOV/DEC 2022

Third Semester

Computer Science and Business Systems

20BSMA305 - COMPUTATIONAL STATISTICS

(Regulations 2020)

(Usage of Statistical Table Permitted)

Duration: 3 Hours | Max. Marks: 100

PART - A (10 × 2 = 20 Marks)

Answer ALL Questions

		Marks, K-Level,CO
1.	Find $\mu_x, \mu_y, \sigma_x, \sigma_y$ and ρ_{xy} for the following $\frac{1}{2\pi}e^{-\frac{1[x^2+y^2+4x-6y+13]}{2}}$.	2,K3,CO1
2.	Find the correlation co-efficient of the random vector X = (X_1, X_2) following BND with mean $(\mu_1, \mu_2) = (7,6)$ and $\Sigma = \begin{pmatrix} 8 & 2 \\ 2 & 5 \end{pmatrix}$.	2,K3,CO1
3.	Write the formula to find the residual of multivariate linear regression.	2,K1,CO2
4.	Write the properties of a matrix $\overline{H}$.	2,K1,CO2
5.	Define discriminant analysis and write its application.	2,K1,CO3
6.	What is the test statistic for the variable p= 2 and population g ≥ 2 in MANOVA ?	2,K3,CO3
7.	Define scree plot.	2,K1,CO4
8.	What is the covariance structure for the orthogonal factor model?	2,K1,CO4
9.	Define k-mean clustering.	2,K1,CO5
10.	Write the four properties of a metric or distance measure.	2,K1,CO5

PART - B (5 × 16 = 80 Marks)

Answer ALL Questions

11. a) (i) The life of a tube X_1 and X_2 are distributed as BND (2000, 0.1, 2500, 0.01, 0.87). If a filament diameter is 0.098, what is the probability that the tube will last 1950 hrs? — 6,K3,CO1

(ii) The co-variance matrix of a 3-dimensional vector X= (X_1, X_2, X_3) is given by $\Sigma = \begin{pmatrix} 25 & -2 & 4 \\ -2 & 4 & 1 \\ 4 & 1 & 9 \end{pmatrix}$. Determine the correlation matrix and correlation between X_1 and $\frac{X_2}{3} + \frac{X_3}{3}$. — 10,K3,CO1

OR

b) (i) Consider a bivariate normal distribution with $\mu_1 = 1, \mu_2 = 3,$ $\sigma_{11} = 2, \sigma_{22} = 1$ and $\rho_{12} = -0.8$ — 8,K3,CO1

1) Write bivariate normal density.
2) Write the squared statistical distance expression.

(ii) Find the maximum likelihood estimates of 2x1 mean vector μ and 2x2 covariance matrix Σ based on the random sample $X = \begin{pmatrix} 3 & 6 \\ 4 & 4 \\ 5 & 7 \\ 7 & 7 \end{pmatrix}$ from a bivariate normal population. — 8,K3,CO1

12. a) (i) If the response variables take the value $Y = \begin{pmatrix} 10 & 100 \\ 12 & 110 \\ 11 & 105 \end{pmatrix}$ and the design matrix $X = \begin{pmatrix} 1 & 9 & 62 \\ 1 & 8 & 58 \\ 1 & 7 & 64 \end{pmatrix}$ takes the value. — 8,K3,CO2

1) Calculate $\hat{\beta}$.
2) Calculate $\hat{\varepsilon}$.
3) Find a linear regression model.

(ii) Find the regression co-efficients for the following — 8,K3,CO2

Y_1	10	12	11
Y_2	100	110	105
X_1	9	8	7
X_2	62	58	64

OR

b) Given the mean vector and covariance matrix of Y, X_1, X_2 — 16,K3,CO2

$$\mu = \begin{pmatrix} \mu_y \\ \mu_x \end{pmatrix} = \begin{pmatrix} 5 \\ 2 \\ 0 \end{pmatrix}, \Sigma = \begin{pmatrix} 10 & 1 & -1 \\ 1 & 7 & 3 \\ -1 & 3 & 2 \end{pmatrix}.$$

Determine the following.

1) The best linear predictor $\beta_0 + \beta_1 X_1 + \beta_1 X_2$.
2) Mean square Error.
3) Multiple correlation coefficient.
4) Also verify that mean square error equals $\sigma_{yy}(1-\rho_y^2(x))$.

13. a) Construct a MANOVA table for the following — 16,K3,CO3

$$\begin{pmatrix} \begin{pmatrix} 9 \\ 3 \end{pmatrix} & \begin{pmatrix} 6 \\ 2 \end{pmatrix} & \begin{pmatrix} 9 \\ 7 \end{pmatrix} \\ \begin{pmatrix} 0 \\ 4 \end{pmatrix} & \begin{pmatrix} 2 \\ 0 \end{pmatrix} & \\ \begin{pmatrix} 3 \\ 8 \end{pmatrix} & \begin{pmatrix} 1 \\ 9 \end{pmatrix} & \begin{pmatrix} 2 \\ 7 \end{pmatrix} \end{pmatrix}.$$

OR

b) (i) Consider the two data sets $X_1 = \begin{pmatrix} 6 & 9 \\ 5 & 7 \\ 4 & 8 \end{pmatrix}$ for which $\overline{X_1} = \begin{pmatrix} 3 \\ 6 \end{pmatrix}$, $\overline{X_2} = \begin{pmatrix} 5 \\ 8 \end{pmatrix}$ & spooled = $\begin{pmatrix} 1 & 1 \\ 1 & 2 \end{pmatrix}$. *8,K3,CO3*

1) Calculate the linear discriminant function.

2) Classify the observation $x_0{}' = (2,7)$ as population π_1 or π_2 using rule with equal prior and equal cost.

(ii) Compute the linear discriminant projection for the following set of data: *8,K3,CO3*

$X_1 = \{ (4, 1) (2, 4) (2, 3) (3, 6) (4, 4) \}$

$X_2 = \{ (9, 10) (6, 8) (9, 5) (8, 7) (10, 8) \}$.

14. a) Consider an m = 1 factor model for the population with the covariance matrix $\Sigma = \begin{pmatrix} 1 & 0.9 & 0.7 \\ 0.9 & 1 & 0.4 \\ 0.7 & 0.4 & 1 \end{pmatrix}$. Show that there is a unique choice of L and Ψ with $\Sigma = LL' + \Psi$, but $\Psi_1 < 0$, so the choice is not admissable. *16,K3,CO4*

OR

b) Convert the covariance matrix $\Sigma = \begin{pmatrix} 1 & 4 \\ 4 & 100 \end{pmatrix}$ to correlation matrix ρ. Determine the principle components Y_1 and Y_2 from the ρ matrix. Also compute ρ_{Y_1, Z_1}, ρ_{Y_1, Z_2}, ρ_{Y_2, Z_1} and ρ_{Y_2, Z_2}. *16,K3,CO4*

15. a) Suppose we measure two variables X_1,X_2 for each of four items A,B,C,D. *16,K3,CO5*

	X_1	X_2
A	2	4
B	3	6
C	2	5
D	3	2

Use k-means clustering technique to divide the items into k = 2 clusters.

OR

b) Find the clusters using Single Linkage procedure. Use Euclidean distance and draw the dendogram. *16,K3,CO5*

Points	A	B	C	D	E
X	2	6	2	2	5
Y	5	5	4	2	4