



Ramakrishna Mission Vivekananda Educational and Research Institute

PO Belur Math, Howrah, West Bengal 711 202

School of Mathematical Sciences

Department of Computer Science

MSc Big Data Analytics : Batch 2022-24, Semester III, End-Semester

DA310: Multivariate Statistics

Dr. Sudipta Das

Student Name (in block letters):

Date: 04 December 2023

Student Roll No:

Max Marks: 100

Signature:

Time: 3hrs

Answer all questions

1. (20 points)

Consider $f(x_1, x_2, x_3) = k(x_1 + x_2x_3)$; $0 < x_1, x_2, x_3 < 1$.

(a) (6 points) Determine k so that f is a valid pdf of $X = (X_1, X_2, X_3)^T$.

(b) (14 points) Hence, compute the conditional variance of $(X_2, X_3)^T$ given $X_1 = x_1$.

2. (30 points)

Consider an *iid* sample of size n from a bivariate population with pdf

$$f(x_1, x_2) = \frac{1}{\theta_1\theta_2} \exp \left\{ - \left(\frac{x_1}{\theta_1} + \frac{x_2}{\theta_2} \right) \right\}, x_1, x_2 > 0.$$

(a) (15 points) Compute the MLE of $\theta = (\theta_1, \theta_2)^T$.

(b) (15 points) Find the Cramer-Rao lower bound. Is it possible to derive a minimum variance unbiased estimator of θ ?

3. (30 points)

A p -dimensional Wishart distribution with parameter Σ and n , [i.e., $W_p(\Sigma, n)$] is defined as the distribution of $\sum_{i=1}^n Z_i Z_i^T$, where Z_i s are i.i.d and $Z_i \sim N_p(0, \Sigma)$.

(a) (15 points) Let $\mathcal{X}_{n \times p}$ be a data matrix from a $N_p(0, \Sigma)$ distribution and let $C_{n \times n}$ be a symmetric matrix. Prove that

$$\mathcal{X}^T C \mathcal{X} = \sum_{i=1}^n \lambda_i W_p(\Sigma, 1),$$

where $\lambda_i, i = 1, \dots, n$ are the eigenvalues of C .

(b) (15 points) Hence, prove that $nS = \mathcal{X}^T H \mathcal{X}$ is distributed as $W_p(\Sigma, n-1)$, where $H = I_n - \frac{1}{n} \mathbf{1}_n \mathbf{1}_n^T$, (note that S is the sample covariance matrix).

4. (20 points)

Apply a PCA to $\Sigma = \begin{bmatrix} 1 & \rho \\ \rho & 1 \end{bmatrix}$, where $0 < \rho < 1$. Now change the scale of X_1 , i.e., consider the covariance of cX_1 and X_2 , where $c > 1$. How do the PC directions change with the screeplot?

This exam has total 4 questions, for a total of 100 points and 0 bonus points.