3/41

2211E779

Candidate's Seat No :________

M.Sc Sem-3 Examination
503
Statistics
November-2024

Time : 2-30 Hours] **[Max. Marks : 70**

Note: Attempt all questions.

Q.1

(i) Define multivariate normal distribution and obtain its characteristic function. [7]

(ii) Let $\mathbf{X}=\begin{bmatrix}\mathbf{X}_1\\ \hline \mathbf{X}_2\end{bmatrix}$ be distributed as $N_p(\boldsymbol{\mu},\boldsymbol{\Sigma})$ with $\boldsymbol{\mu}=\begin{bmatrix}\boldsymbol{\mu}_1\\ \boldsymbol{\mu}_2\end{bmatrix}$, $\boldsymbol{\Sigma}=\begin{bmatrix}\boldsymbol{\Sigma}_{11} & \boldsymbol{\Sigma}_{12}\\ \boldsymbol{\Sigma}_{21} & \boldsymbol{\Sigma}_{22}\end{bmatrix}$, and $|\Sigma_{22}|>0$. Then show that the conditional distribution of $\mathbf{X}_1$, given $\mathbf{X}_2=\mathbf{x}_2$, is normal with

Mean = $\boldsymbol{\mu}_1+\boldsymbol{\Sigma}_{12}\boldsymbol{\Sigma}_{22}^{-1}(\mathbf{x}_2-\boldsymbol{\mu}_2)$ and Covariance = $\boldsymbol{\Sigma}_{11}-\boldsymbol{\Sigma}_{12}\boldsymbol{\Sigma}_{22}^{-1}\boldsymbol{\Sigma}_{21}$. [7]

OR

(i) If $\mathbf{X}\sim N_p(\boldsymbol{\mu},\boldsymbol{\Sigma})$ and $\mathbf{C}$: $m\times p$ be any arbitrary matrix then show that [7]

$\mathbf{Y}_{m\times 1}=\mathbf{C}_{m\times p}\mathbf{X}_{p\times 1}\sim N_m(\mathbf{C}\boldsymbol{\mu},\mathbf{C}\boldsymbol{\Sigma}\mathbf{C}')$

(ii) Show that if $\mathbf{X}\sim N_p(\boldsymbol{\mu},\boldsymbol{\Sigma})$ then any sub-vector of $\mathbf{X}$ is also normally distributed with mean equal to the corresponding sub-vector of $\boldsymbol{\mu}$ and with variance-covariance matrix equal to the corresponding sub-matrix of $\boldsymbol{\Sigma}$. [7]

Q.2

(i) Define multiple correlation coefficients. In usual notation obtain the expression in terms of elements of $\Sigma^{-1}=(\sigma_{ij})$ for multiple correlation coefficient $R_{1.23\cdots p}$. [7]

(ii) Describe canonical correlation coefficients and canonical variates. In usual notation show that canonical correlation are solution of the determinant equation [7]

$$\begin{vmatrix}-\lambda\Sigma_{11} & \Sigma_{12}\\ \Sigma_{12} & -\lambda\Sigma_{22}\end{vmatrix}=0$$

OR

(i) Show that (a) $1-R_{1.23}^2=(1-r_{12}^2)(1-r_{13.2}^2)$ (b) $R_{1.23}^2=r_{12}^2+r_{13}^2$, if $r_{23}=0$. [7]

(ii) Define canonical correlation and write it's application. [7]

(P.T.O)

E 779-2

Q.3

(i) If $V_1 \sim W_p(V_1 \mid n_1 \mid \Sigma)$ and $V_2 \sim W_p(V_2 \mid n_2 \mid \Sigma)$ are independently distributed then show that $V_1 + V_2 \sim W_p(V_1 + V_2 \mid n_1 + n_2 \mid \Sigma)$. [7]

(ii) If $V \sim W_p(V \mid n \mid I_p)$ then obtain (i) $E(V^k)$ (ii) $E(V^{-k})$ [7]

OR

(i) Define Wishart distribution and give its applications. [7]

(ii) If $D \sim W_p(D, n, \Sigma)$, then obtain the distribution of $|D|/|\Sigma|$. Also obtain $E|D|^h$. [7]

Q.4

(i) Let Σ be the covariance matrix associated with the random vector $X' = [X_1, X_2, \cdots, X_p]$. Let Σ have the eigenvalue-eigenvector pairs $(\lambda_1, e_1), (\lambda_2, e_2), \cdots, (\lambda_p, e_p)$ where $\lambda_1 \ge \lambda_2 \ge \cdots \ge \lambda_p \ge 0$. The ith principal component is given by $Y_i = e_i' X = e_{1i} X_1 + e_{2i} X_2 + \cdots + e_{pi} X_p$, $i = 1, 2, \cdots, p$. Then show that Var $(Y_i) = e_i' \Sigma e_i = \lambda_i$, and Cov $(Y_i, Y_k) = e_i' \Sigma e_k = 0$, $i \ne k$. [7]

(ii) What do you understand by discriminant and classification? Explain their goals. [7]

OR

(i) Define principal components. Write their important applications. [7]

(ii) Explain how Hotelling T^2 is used for testing equality of two mean vectors from two different multivariate normal populations with unknown but equal covariance matrices. [7]

Q.5 Answer any seven: [14]

(i) For a multivariate normal random vector, the variance-covariance matrix is always________.
(A) square matrix (B) non-negative definite (C) symmetric (D) all of these

(ii) Which of the following is true for a random vector **X** having a multivariate normal distribution?
(A) Linear combinations of the components of **X** are normally distributed
(B) Zero covariance implies that the corresponding components are independently distributed
(C) The conditional distributions of the components are (multivariate) normal
(D) all of these

E779-3

(iii) Multivariate normal distribution is most commonly used in multiple statistical methods because:
(A) it is easier than other distributions
(B) it's assumptions holds due to multiple central limit theorem in all cases
(C) this distribution holds for all types of variables
(D) it converges to unity

(iv) The quantity $(\mathbf{x}-\boldsymbol{\mu})^T \Sigma^{-1}(\mathbf{x}-\boldsymbol{\mu})$ involved in multivariate normal density function represents:
(A) multivariate normal density
(B) dispersion matrix
(C) exponential series
(D) Mahalanobis squared distance

(v) Give one limitation of canonical correlation.

(vi) Define population correlation matrix.

(vii) The _______ distribution is a multivariate generalization of chi-square distribution.
(A) Multivariate Normal
(B) Hotelling's T^2
(C) Wishart distribution
(D) none of these

(viii) The Wishart distribution is a family of distributions for _______ matrices
(A) symmetric positive definite
(B) asymmetric positive definite
(C) symmetric negative definite
(D) asymmetric negative definite

(ix) Give one limitation of Wishart distribution.

(x) What are the general objectives of principal component analysis?

(xi) What is the purpose of eigenvalues in principal component analysis?
(A) They represent the variance explained by each principal component
(B) They determine the number of principal components to retain
(C) They are used for data visualization
(D) They measure the similarity between data points

(xii) In principal component analysis, what is the relationship between the first principal component and the second principal component?
(A) They are orthogonal (uncorrelated) to each other
(B) They are positively correlated
(C) They are negatively correlated
(D) There is no defined relationship between them

***