Total number of printed pages–7

3 (Sem–6/CBCS) STA HC 2

2025

STATISTICS

(Honours Core)

Paper : STA–HC–6026

(Multivariate Analysis and Non-parametric Methods)

Full Marks : 60

Time : Three hours

The figures in the margin indicate full marks for the questions.

1. Answer the following questions as directed : $1 \times 7 = 7$

   (a) In a bivariate normal distribution, the correlation coefficient (ρ) is said to measure :

   (A) Variance

   (B) Covariance

Contd.

(C) Linear relationship between variables

(D) Independence of variables

(b) Define partial correlation coefficient.

(c) For a multivariate normal distribution, if the covarience matrix is diagonal, it implies –

(A) the variables are independent

(B) the variables are dependent

(C) the mean vector is zero

(D) the distribution is uniform

(d) The MGF of a bivariate normal distribution can be used to compute –

(A) moments of individual variables

(B) joint moments of the variables

(C) both individual and joint moments

(D) None of the above

(e) Which of the following is a characteristic of non-parametric tests ?

(A) They assume a normal distribution of the data

(B) They are based on ranking rather than actual data values

(C) They require knowledge of population parameters

(D) They are limited to large sample sizes only

(f) Which of the following tests is a non-parametric ?

(A) t-test

(B) ANOVA

(C) z-test

(D) Mann-Whitney U-test

(g) In which scenario would you prefer a non-parametric test over a parametric test ?

(A) When the data follow a normal distribution

(B) When the sample size is very large

(C) When the data are ordinal or not normally distributed

(D) When the mean is the parameter of interest

2. Answer the following questions : 2×4=8

(a) State the marginal distributions of a bivariate normal distribution.

(b) Explain the role of the covariance matrix in a multivariate normal distribution.

(c) Let $\underset{\sim}{X} = \begin{pmatrix} X_1 \\ X_2 \end{pmatrix}$. If X_1 and X_2 are independent and

$g\left(\underset{\sim}{x}\right) = g^{(1)}(X_1)g^{(2)}(X_2)$, prove that

$$E[g(X)] = E\left[g^{(1)}(X_1)\right]E\left[g^{(2)}(X_2)\right]$$

(d) Write a note on test for randomness.

3. Answer **any three** questions from the following : 5×3=15

(a) Derive the conditional mean and variance of Y given X in bivariate normal distribution, where the variables (X, Y) follow bivariate normal distribution.

(b) Define a multivariate normal distribution and explain the role of the mean vector and covariance matrix.

(c) Explain the significance of the correlation coefficient (ρ) in the bivariate normal distribution and how it affects the shape of the joint pdf.

(d) Derive the steps for conducting a Mann-Whitney U-test and describe its applications.

(e) Describe the sign test for *one* sample.

(f) Discuss the purpose and assumptions of the Kolmogorov-Smirnov test for comparing distributions.

4. Answer ***any three*** questions from the following : 10×3=30

(a) (i) Discuss the assumptions underlying Hotelling's T^2 test and their implications in multivariate analysis. 3+3=6

(ii) Explain how Hotelling's T^2 test is used to compare the means of two multivariate samples. 4

(b) Provide a step-by-step explanation of the PCA procedure and describe the significance of principal components in data reduction and feature reduction. Discuss its limitations in real-world applications. 4+3+3=10

(c) Derive the joint probability density function for a bivariate normal distribution and explain its components. What is the effect of the correlation coefficient (ρ) on the distribution's shape ? 5+3+2=10

(d) Discuss how linear transformations of a multivariate normal distribution affect its mean vector and covariance matrix. Describe the positive semi-definiteness of the covariance matrix and its importance in a multivariate normal distribution. 6+2+2=10

(e) For a bivariate normal distribution

$$dF = K\exp\left[-\frac{2}{3}\left(x^2 - xy + y^2 - 3x + 3y + 3\right)\right]dxdy,$$

find –

(i) the value of K

(ii) marginal distribution of Y

(iii) expectation of the conditional distribution of Y given X. 2+4+4=10

(f) (i) Explain how the Chi-square test of independence can be applied to analyse categorical data, and interpret its test statistic. 5

(ii) Discuss Kruskal-Wallis test. 5

2000