Total number of printed pages–7

3 (Sem–6/CBCS) STA HC 2

2023

STATISTICS

(Honours)

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 : $1\times7=7$

(a) Write down the probability density function of the bivariate normal distribution.

(b) The characteristic function of the multivariate normal distribution having the mean vector $\underset{\sim}{O}$ and variance covariance matrix I, is _____. *(Fill in the blank)*

(c) Non-parametric tests can be used only if the measurements are

(i) nominal or ordinal

(ii) ratio scale

(iii) interval scale

(iv) interval and ratio scale

(Choose the correct option)

(d) Statement : Non-parametric test does not make any assumption regarding the form of the population. The statement is

(i) True

(ii) False

(Tick the correct answer)

(e) If the correlation coefficient *(e)*, is zero for a bivariate normal distribution, then the variables are

(i) dependent

(ii) independent

(iii) uncorrelated but dependent

(iv) partly dependent

(f) Let $\underset{\sim}{X}$ (with p-components) be distributed according to $N\left(\underset{\sim}{\mu}, \Sigma\right)$. Then $\underset{\sim}{Y} = C\underset{\sim}{X}$ is distributed according to

(i) $N\left(\underset{\sim}{\mu}, C\Sigma C'\right)$ for singular C

(ii) $N\left(C\underset{\sim}{\mu}, \Sigma\right)$ for C non-singular

(iii) $N\left(C\underset{\sim}{\mu}, C\Sigma C'\right)$ for C non-singular

(iv) $N(O, I)$

(g) Define partial correlation coefficient.

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

(a) Explain the test for randomness.

(b) Find $Cov\left(A\underset{\sim}{X}, B\underset{\sim}{Y}\right)$ where A, B are matrices of constant elements.

(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) Describe the advantages and drawbacks of the non-parameteric methods over parametric methods.

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

(a) For a bivariate distribution

$$f(x,y) = \frac{1}{2\pi\sqrt{(1-p^2)}} exp\left[-\frac{1}{2(1-p^2)}\left(x^2 - 2pxy + y^2\right)\right],$$

$-\infty < (x,y) < \infty$.

Find the conditional distribution of Y given X.

(b) Let (X, Y) be a bivariate normal random variable with $E(X) = E(Y) = 0$, $V(X) = V(Y) = 1$ and $Cov(X,Y) = P$. Find the probability density function (pdf) of $Z = Y/X$.

(c) For a multivariate normal distribution $N\left(\underset{\sim}{\mu}, \Sigma\right)$, if $\underset{\sim}{\mu} = \underset{\sim}{0}$ and

$$\Sigma = \begin{pmatrix} 1 & 0.80 & -0.40 \\ 0.80 & 1 & -0.56 \\ -0.40 & -0.56 & 1 \end{pmatrix}$$

Find the partial correlation between X_1 and X_3 given X_2.

(d) Write a short note on principal component analysis.

(e) Describe the sign test for one sample.

4. Answer **either** (a) **or** (b) : 10

(a) Write a note on Hotelling T^2 mentioning its applications. Prove that Hotelling T^2 is invariant under a non-singular linear transformation. 4+6=10

Or

(b) For a bivariate normal distribution

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

find —

(i) the value of K;

(ii) marginal distribution of Y;

(iii) expectation of the conditional distribution of Y given X.

5. Answer **either** *(a)* **or** *(b)* : 10

(a) Describe explicitly the Kruskal-Wallis test with example.

Or

(b) Write a note on Wilcoxon-Mann-Whitney test for non-parametric methods.

6. Answer **either** *(a)* **or** *(b)* : 10

(a) Prove that if $X_1, X_2, \ldots X_p$ have a joint normal distribution, a necessary and sufficient condition for one subset of the random variables and the subset consisting of the remaining variables to be independent is that each covariance of a variable from one set and a variable from the other set is 0 (zero).

Or

(b) Derive the pdf of bivariate normal distribution as a particular case of multivariate normal distribution.
