Total number of printed pages–8

3 (Sem–6/CBCS) STA HC 2

2024

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) Let X and Y be two normally correlated variables. In derving the bivariate normal distribution we make the following asumptions :

(i) Regression of Y on X is linear.

(ii) The arrays are homoscedastic.

(iii) The distribution of Y in different arrays is normal.

Contd.

Choose the correct answer from the following options :

(I) Only *(i)* and *(ii)* are the basic assumptions required

(II) Only *(ii)* is the basic assumption deriving bivariate normal distribution

(III) No assumptions are required to derive bivariate normal distribution

(IV) All three assumptions *(i)*, *(ii)* and *(iii)* are required to derive bivariate normal distribution

(b) The moment generating function of the bivariate normal distribution is

(I) $M_{XY}(t_1, t_2) = exp\left[\frac{1}{2}\left(t_1^2 + t_2^2 + 2Pt_1t_2\right)\right]$

(II) $M_{XY}(t_1, t_2) = exp\left[\frac{1}{2}\left(t_1 + t_2 + 2Pt_1t_2\right)\right]$

(III) $M_{XY}(t_1, t_2) = exp\left[\left(t_1^2 + t_2^2 + 2Pt_1t_2\right)\right]$

(IV) $M_{XY}(t_1, t_2) = exp\left[2\left(t_1^2 + t_2^2 + \frac{1}{2}Pt_1t_2\right)\right]$

(Choose the correct option)

(c) If the marginal distribution of X and Y are normal (Gaussian)

(I) it does not necessarily imply that the joint distribution of (X, Y) is bivariate normal

(II) it implies that the joint distribution of (X, Y) is bivariate normal

(III) it implies that the variables X and Y are independently distributed

(IV) the joint distribution is bivariate normal, but the random variables X and Y are independently distributed

Choose the appropriate answer from the above options.

(d) There are certain assumptions associated with non-parametric tests, which are

(i) sample observations are independent

(ii) the variables under study is continuous

(iii) the distributions are continuous

(iv) lower order moments exists

Choose the appropriate answer from the following options :

(I) No assumptions are required for non-parametric tests

(II) The non-parametric tests are valid for the discrete variables

(III) All above assumptions *(i)*, *(ii)*, *(iii)* and *(iv)* are required for non-parametric tests

(IV) The non-parametric tests can be done only for discrete distributions.

State which of the following statements is True **or** *False* :

(e) To study the correlation coefficient of two variables in the qualitative data, Pearson correlation being used to find the correlation coefficient of the two varriables.

(I) True

(II) False

Choose the correct answer.

(f) If $R_{1\cdot 23} = 1$, the multiple linear regression of X_1 on X_2 and X_3 is considered as perfect for all predictions.

(I) True

(II) False

Choose the correct answer.

(g) Let $\underset{\sim}{X} = \left(X_1, X_2, ..., X_p\right)'$ be the data vector be divided as $\underset{\sim}{X}^{(1)} = \left(X_1, ..., X_q\right)'$ and $\underset{\sim}{X}^{(2)} = \left(X_{q+1}, ..., X_p\right)'$, then for $i = 1, 2, ..., q$, the residual variate $X_{i\cdot q+1, ..., p}$ of X_1 is independent of the regressor $\underset{\sim}{X}^{(2)}$.

(I) True

(II) False

Choose the correct answer.

Answer the following questions : 2×4=8

(a) State *two* assumptions required for a non-parametric test.

(b) Let X be a p-dimensional random vector and if A is an $(n \times p)$ matrix of known constants, then AX is the product of a vector X by a matrix A, then find $V(AX)$.

(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

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

(d) Write a note on test for randomness.

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

(a) Write a note on ordinary sign test.

(b) Let $\underset{\sim}{X}$ (with p-components) be distributed according to $N\left(\underset{\sim}{\mu}, \Sigma\right)$, prove that $\underset{\sim}{Y} = C\underset{\sim}{X}$ is distributed according to $N\left(C\underset{\sim}{\mu}, C\Sigma C'\right)$ for C non-singular.

(c) If the bivariate random variables $X = (X_1, X_2)'$ follow bivariate normal distribution with mean vector 0 and dispersion matrix $\sum = \begin{pmatrix} 1 & \rho \\ \rho & 1 \end{pmatrix}$. Find the pdf of X_1/X_2.

(d) If (X, Y) is $\text{BVN}\left(\mu_1, \mu_2, \sigma_1^2, \sigma_2^2, \rho\right)$. Compute the correlation coefficient of e^X and e^Y.

(e) If X and Y are standard normal variates with coefficient of correlation ρ, show that regression of Y on X is linear.

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

(a) If $X_1, \ldots, X_p$ have a joint normal distribution, a necessary and sufficient condition for one subset of 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. Prove it.

(b) Let $(X, Y) \sim \text{BVN}$, prove that if and only if every linear combination of X and Y is a normal variate.

(c) Discuss Mann-Whitney test. What is the mean and variance of the test statistic ? How is the test carried out for large sample ? 6+2+2=10

Contd.

(d) Describe the Kolmogorov-Smirnov test for one sample, explaining assumptions and hypothesis.

(e) (i) Write a note on principal component analysis. 5

(ii) Write a note on Hotelling's T^2. 5

(f) Suppose $\underset{\sim}{X} \sim N_p\left(\underset{\sim}{\mu}, \Sigma\right)$ and let $\underset{\sim}{X}$ be partitioned as $\underset{\sim}{X}' = \left(\underset{\sim}{X}^{(1)'}, X^{(2)'}\right)$, where $\underset{\sim}{X}^{(1)}$ has k-components and $\underset{\sim}{X}^{(2)}(p-k)$ components. Prove that the regression for the mean of $\underset{\sim}{X}^{(1)}$ on $\underset{\sim}{X}^{(2)}$ coincides with the linear mean square regression of $\underset{\sim}{X}^{(1)}$ on $\underset{\sim}{X}^{(2)}$.