6 SEM TDC STSH (CBCS) C14

2022

(June/July)

STATISTICS

(Core)

Paper : C–14

(Multivariate Analysis and Nonparametric Methods)

Full Marks : 50
Pass Marks : 20

Time : 2 hours

The figures in the margin indicate full marks for the questions

1. Choose the correct answer from the following : 1×5=5

(a) The variance-covariance matrix for the three variables $\underset{\sim}{X} = (x_1, x_2, x_3)'$ is given by

$$\Sigma = \begin{pmatrix} 1 & -2 & 5 \\ 3 & 9 & 2 \\ -2 & 4 & 16 \end{pmatrix}$$

then the s.d. of x_3 is

(i) 4

(ii) 16

(iii) 5

(iv) 9

(b) Let $\underset{\sim}{Z}_1, \underset{\sim}{Z}_2, \cdots, \underset{\sim}{Z}_n$ be independent $N_P(\underset{\sim}{O}, \Sigma)$. Then $A = \sum_{i=1}^{n} Z_i Z_i'$ is said to have

(i) multivariate normal distribution

(ii) Wishart distribution

(iii) exponential distribution

(iv) Hotelling T^2 statistic

(c) The range of multiple correlation coefficient is

(i) 0 to ∞

(ii) 0 to 1

(iii) -1 to 1

(iv) $-\infty$ to $+\infty$

(d) In discriminant analysis, the criterion for scale of dependent variable is _____ and the predictor or independent variables are _____ in nature.

(i) interval; categorical

(ii) ordinal; interval

(iii) categorical; interval

(iv) ordinal; categorical

(e) While performing Kruskal-Wallis test, the ranks are assigned

(i) independently to the observation for each treatment

(ii) for observations in each block independently

(iii) by pooling all the observations

(iv) None of the above

2. Answer the following questions in brief : $2\times5=10$

(a) For a bivariate normal distribution

$$(X, Y) \sim BvN\left(1, 2, 16, 25, \frac{12}{13}\right)$$

find $P(X > 2)$.

(b) State the properties of multiple correlation coefficient.

(c) Distinguish between principal component and factor analysis.

(d) Distinguish between non-parametric methods and distribution-free methods.

(e) How to resolve the problem of zero differences in sign test?

3. (a) Let the joint p.d.f. of X and Y be

$$f(x, y) = \frac{1}{2\pi\sigma_1\sigma_2\sqrt{1-\rho^2}}$$

$$\exp\left\{-\frac{1}{2(1-\rho^2)}\left[\frac{(x-\mu_1)^2}{\sigma_1^2} - 2\rho\frac{(x-\mu_1)}{\sigma_1}\frac{(y-\mu_2)}{\sigma_2} + \frac{(y-\mu_2)^2}{\sigma_2^2}\right]\right\}$$

where $-\infty < x < \infty;\ -\infty < y < \infty,\ -1 < \rho < 1.$

(i) Find the marginal distribution of X; 4

(ii) Find the conditional distribution of Y given $X = x$. 3

Or

(b) Suppose

$$X = \begin{bmatrix} 10 & 100 \\ 12 & 110 \\ 11 & 105 \end{bmatrix}$$

then—

(i) find the mean vector;

(ii) construct the variance-covariance matrix.

3+4=7

4. *(a)* *(i)* Let $\underset{\sim}{X}$ (with P components) be distributed as $N_P(\underset{\sim}{\mu}, \Sigma)$. Show that $Y = CX$ is distributed according to $Y = CX \sim N_P(C\underset{\sim}{\mu}, C\Sigma C')$, for C is non-singular. 7

(ii) Let X be $N_3(\underset{\sim}{\mu}, \Sigma)$ with $\underset{\sim}{\mu}' = [2, -3, 1]'$ and

$$\Sigma = \begin{bmatrix} 1 & 1 & 1 \\ 1 & 3 & 2 \\ 1 & 2 & 2 \end{bmatrix}$$

Find the distribution of

$$3X_1 - 2X_2 + X_3$$ 4

Or

(b) *(i)* Explain the concept of partial correlation coefficient with examples. Show that partial correlation coefficient is the geometric mean between the regression coefficients. 2+3=5

(ii) Derive the coefficient of multiple correlation. 6

5. *(a)* How is principal component analysis used for dimensionality reduction? Determine the population principal

components Y_1, Y_2 for the covariance matrix

$$\Sigma = \begin{bmatrix} 5 & 2 \\ 2 & 2 \end{bmatrix}$$

Also calculate the proportion of the total population variance explained by the first principal component. 2+5=7

Or

(b) State the assumptions of discriminant analysis. Explicate Fisher's linear discriminate function. 2+5=7

6. *(a)* What do you mean by empirical distribution function? Describe briefly Kolmogorov-Smirnov test of goodness of fit in case of one sample. Write some applications of Kolmogorov-Smirnov test. 2+6+2=10

Or

(b) *(i)* Following is a sequence of heads (H) and tails (T) in tossing of a coin 14 times : 3

H T T H H H T H T T H H T H

Test whether the heads (H) and tails (T) occur in a random order. [Given $\alpha = 0 \cdot 05$, $\pi_L = 2$, $\pi_u = 12$]

(ii) Explicate Mann-Whitney *U*-test for testing the identicalness of two populations. 7

★★★