6 SEM TDC STSH (CBCS) C 14

2 0 2 4

(May)

STATISTICS

(Core)

Paper : C–14

(Multivariate Analysis and Non-Parametric 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 alternatives : $1 \times 5 = 5$

   (a) If (X, Y) are bivariate $N(0, 0, 1, 1, \rho)$, then the variables $(x + y)$ and $(x - y)$ are

   (i) correlated with $\rho = \dfrac{1}{2}$

   (ii) independently distributed

   (iii) negatively correlated

   (iv) None of the above

(Turn Over)

(b)   The quantity $(x - \mu)' \Sigma^{-1} (x - \mu)$ involved in multivariate normal density function represents

    (i) multivariate normal density

    (ii) dispersion matrix

    (iii) exponential series

    (iv) Mahalanobis squared distance

(c)   The partial correlation coefficient $r_{13.2}$ is called

    (i) first-order partial correlation

    (ii) zero-order partial correlation

    (iii) second-order partial correlation

    (iv) None of the above

(d)   Principal component analysis is meant for

    (i) reducing the number of observed variables

    (ii) increasing the number of variables

    (iii) reducing the correlation between variables

    (iv) reducing the variance among variables

(Continued)

(e) Kolmogorov-Smirnov test is useful as

    (i) a test of goodness of fit

    (ii) a test of identicalness of two populations

    (iii) a measure of confidence band

    (iv) All of the above

2. Answer the following questions in brief :

$$2 \times 5 = 10$$

(a) State the properties of random vector.

(b) Define partial correlation coefficient with examples.

(c) How to obtain discriminant scores?

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

(e) What do you mean by empirical distribution function?

3. (a) State the properties of bivariate normal distribution.

Let the joint p.d.f. of X and Y be

$$f(x, y) = \frac{1}{2\pi\sigma_1\sigma_2\sqrt{1-\rho^2}} \exp\left\{-\frac{1}{2(1-\rho^2)}\left[\frac{(x-\mu_1)^2}{\sigma_1^2}\right.\right.$$

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

where

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

Find—

(i) the marginal distribution of X;

(ii) $P(3 < y < 8 | X = 7)$ given that $\mu_1 = 3$, $\mu_2 = 1$, $\sigma_1^2 = 16$, $\sigma_2^2 = 25$, $\rho = 0.6$.

2+3+2=7

Or

(b) How do you express and organize multivariate data?

Let

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

Calculate the mean vector and variance-covariance matrix.

3+4=7

4. (a) (i) Define multivariate normal distribution. If $X \sim N_P (\mu, \Sigma)$, then show that the components of X will be jointly independent if and only if the covariance of X_j and X_k $(\dot{y} \neq k = 1, 2, \cdots, p)$ is zero.   2+6=8

        (ii) Define Wishart distribution and its properties.   3

Or

(b) (i) Clarify the concept of Hotelling's T^2 distribution.   4

        (ii) Discuss the properties of multiple correlation coefficient. Show that the relationship between regression coefficient and partial regression coefficient is

$$1 - R^2_{1 \cdot 23} = (1 - r^2_{12})(1 - r^2_{13 \cdot 2})$$

    3+4=7

5. (a) What does discriminant analysis signify? What are the assumptions underlying discriminant analysis? Usually, what test statistics are involved in discriminant analysis technique?

    2+3+2=7

Or

(b) How is principal component analysis used for dimensionality reduction? What are the major uses of principal component analysis? What is the purpose of factor analysis?   2+3+2=7

6. (a) What do you mean by empirical distribution function? Define run in a sequence of symbols. Explicate Mann-Whitney U-test for testing the identicalness of two populations.

2+2+6=10

Or

(b) Write some applications of Kolmogorov-Smirnov test. How to resolve the problem of zero differences in sign test? Discuss the Kruskal-Wallis method of analysis for one-way classification of data.   2+2+6=10

★ ★ ★