Total number of printed pages–7

3 (Sem–6/CBCS) STA HC 1

2022

STATISTICS

(Honours)

Paper : STA–HC–6016

(Design of Experiments)

Full Marks : 60

Time : Three hours

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

1. Answer ***any seven*** of the following as directed : $1 \times 7 = 7$

   (a) ________ is the simplest design making use of all the three basic principles of design. *(Fill in the blank)*

   (b) The error degrees of freedom in a $m \times m$ L.S.D. is ________ . *(Fill in the blank)*

Contd.

(c) The error d.f. in an RBD with 4 blocks comparing 6 treatments is ______ .
(Fill in the blank)

(d) The error d.f. for a $p \times p$ L.S.D. with one missing observation is ______ .
(Fill in the blank)

(e) In a split plot design ______ effect is confounded. *(Fill in the blank)*

(f) In the linear model considered in analysis of variance the error term is distributed as ______ .
(Fill in the blank)

(g) In a 2^4 factorial experiment with the four factors A, B, C, D, each at two levels, the interaction effects ABC and ABD is confounded. Name the other factor which is also confounded.

(h) Define the term 'contract'.

(i) Write down the main effects and interaction effects for a 3^2 design with two factors A and B each at three levels 0, 1, 2.

(j) The concept of confounding is not deliberately introduced in a factorial experiment. *(State True or False)*

2. Answer ***any four*** questions from the following : $2 \times 4 = 8$

(a) Give the layout of a 4×4 Latin square design.

(b) Explain why there cannot be a 2×2 L.S.D.

(c) Write a note on the assumptions made in a linear model in analysis of variance.

(d) Explain the use of local control in Latin square design.

(e) In a 5×5 LSD, the following results were obtained :

Row mean square = 11·66

Column mean square = 3·5

Treatment mean square = 49·15

Total sums of square = 285·34

Complete the ANOVA table.

(f) A 2^3 experiment was conducted with three factors N, P and k, each at two levels. The central blocks for the replications are

np, npk, (1), k

(1), npk, nk, p

pk, nk, (1), np

respectively. Find the effect confounded in each replication.

(g) Define balanced incomplete block design.

(h) What do you mean by the term 'efficiency' in a design of experiment?

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

(a) Obtain the estimate of the missing plot in a randomised block design.

(b) What is confounding in a factorial experiment? Explain the difference between complete and partial confounding in case of a 2^4 factorial experiment.

(c) Write a note on the advantages and disadvantages of confounding.

(d) Obtain a balanced confounded 2^4 design in a number of replications having four blocks in each.

(e) Write an introductory note on balanced incomplete block design.

(f) What is factorial experiment ? What are the advantages of a factorial experiment over single factor experiment ?

(g) Describe the layout of a 2^3 experiment where the 2nd order interaction is confounded in all the four replications. Give the structure of the AOV table in this case.

(h) What is a split plot design ? Why is it said that in a split plot design main effect is unfounded ?

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

(a) Give the outline of the analysis of variance of a randomised block design. Obtain the expression for standard error of the difference between two treatment means, when one of them has a missing observation in a randomised block design.

(b) Discuss the analysis of a Latin square design.

(c) The elements of control block of each of six replications of a 2^4 design are (1), ab, acd, bcd. Identify the confounding subgroup and give an outline of the analysis of the data obtained from the experiment.

(d) In a 2^3 factorial experiment conducted with three factors A, B, C, each at two levels, all the interactions effects are confounded in one of the four replications. Give an outline of the analysis of the data.

(e) Describe the layout and give an outline of the analysis of a split plot design.

(f) Find the standard error of the difference between two treatments mean when one of them has a missing observation in a Latin square design. Also write the expression of standard error when there is no missing observation under any of the treatments.

(g) *(i)* Write a note on uniformity trials. 5

(ii) Give an idea of 3^2 factorial experiment. 5

(h) Discuss briefly ***any two*** of the following :

(i) Basic principles of design of experiment.

(ii) Bio-arrays

(iii) Relative efficiency of LSD and RBD

Total number of printed pages–8

3 (Sem–6/CBCS) STA HC 2

2022

STATISTICS

(Honours)

Paper : STA–HC–6026

(Multivariate Analysis and Nonparametric Analysis)

Full Marks : 60

Time : Three hours

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

1. Answer ***any seven*** of the following questions as directed : $1 \times 7 = 7$

(a) The moment generating function of bivariate normal distribution with parameters $\left(0, 0, \sigma_1^2, \sigma_2^2, \rho\right)$ is ______.

(Fill in the blank)

Contd.

(b) Let $\underset{\sim}{X} \sim N_P\left(\underset{\sim}{\mu}, \Sigma\right)$. Then the characteristic of $\underset{\sim}{X}$ is given by

(i) $e^{i\underset{\sim}{t}\underset{\sim}{\mu}+\frac{1}{2}\underset{\sim}{t}'\Sigma\underset{\sim}{t}}$

(ii) $e^{i\underset{\sim}{t}'\underset{\sim}{\mu}-\frac{1}{2}\underset{\sim}{t}'\Sigma\underset{\sim}{t}}$

(iii) $e^{i\underset{\sim}{t}'\underset{\sim}{\mu}+\frac{1}{2}\underset{\sim}{t}'\Sigma\underset{\sim}{t}}$

(iv) None of the above

(Choose the correct option)

(c) Ordinary sign test considers the difference of observed values from the hypothetical median value in terms of :

(i) signs only

(ii) magnitudes only

(iii) sign and magnitude both

(iv) None of the above

(Choose the correct option)

(d) What is dispersion matrix in Multivariate data analysis?

(e) Let $(X, Y) \sim \text{BVND}\left(\mu_1, \mu_2, \sigma_1^2, \sigma_2^2, \rho\right)$. Then state the conditional pdf of Y given $X = x$.

(f) What is run in non-parametric method?

(g) Define Multiple correlation coefficient.

(h) Let $\underset{\sim}{X} \sim N_3\left(\underset{\sim}{\mu}, \Sigma\right)$. Given that

$$\Sigma = \begin{pmatrix} 1 & 2 & 3 \\ 2 & 3 & 0 \\ 3 & 0 & 4 \end{pmatrix}$$

Are X_2 and X_3 independent?

(i) The marginal distribution of a Bivariate normal distribution follows univariate normal distribution. *(State True or False)*

(j) The Kruskal-Wallis test is meant for :

(i) one way classification

(ii) two way classification

(iii) non classified data

(iv) None of the above

(Choose the correct option)

2. Answer ***any four*** of the following questions briefly : 2×4=8

(a) Define mean vector and dispersion matrix for multivariate data analysis.

(b) State the marginal pdfs of X and Y in case of $(X, Y) \sim \text{BVND}\left(\mu_1, \mu_2, \sigma_1^2, \sigma_2^2, \rho\right)$.

(c) What assumptions are generally made for a non-parametric test?

(d) Let $\underset{\frown}{X} = (X_1 \;\; X_2 \;\; X_3)'$ have variance covariance matrix

$$\Sigma = \begin{pmatrix} 25 & -2 & 4 \\ -2 & 4 & 1 \\ 4 & 1 & 9 \end{pmatrix}$$

Find ρ_{12}.

(e) Define marginal distribution of $X_1, X_2, \cdots, X_k \; (k < p)$ in a p-variate multivariate analysis. Also define the conditional distribution of $X_{k+1}, X_{k+2}, \cdots, X_p$ given $X_1, X_2, \cdots, X_k$.

(f) What indication can one get from the number of runs?

(g) Give a brief idea of Principal component analysis.

(h) The pdf of bivariate normal distribution is

$$f(x, y) = k \exp\left[-\frac{1}{2\left(1-\rho^2\right)}\left(x^2 - 2\rho xy + y^2\right)\right],$$
$$-\infty < (x, y) < \infty$$

Find the constant k.

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

(a) If $(X, Y) \sim \text{BVND}\left(\mu_1, \mu_2, \sigma_1^2, \sigma_2^2, \rho\right)$, then show that X and Y are independent if and only if $\rho = 0$.

(b) Describe Kolmogorov-Smirnov one sample test stating its assumptions and hypotheses.

(c) Let $(X, Y) \sim \text{BVND}(0, 0, 1, 1, \rho)$. Then show that

$$Q = \frac{X^2 - 2\rho XY + Y^2}{\left(1-\rho^2\right)}$$

is distributed as chi-square with 2d.f.

(d) Let $\underset{\frown}{X} \sim N_p\left(\underset{\frown}{\mu}, \Sigma\right)$. Then find the distribution of $C\underset{\frown}{X}$ where C is a $p \times p$ non-singular matrix of constant elements.

(e) Write an explanatory note on test of randomness.

(f) With usual notations, prove that

$$r_{12\cdot3} = \frac{r_{12} - r_{13}r_{23}}{\sqrt{\left(1 - r_{13}^2\right)\left(1 - r_{23}^2\right)}}$$

(g) Examine if Hotelling's T^2 is invariant under changes in the units of measurement.

(h) Describe one sample sign test for testing the null hypothesis that the population median is a given value.

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

(a) (i) State *any two* applications of multivariate analysis. 2

(ii) Let $(X, Y) \sim \text{BVND}\left(\mu_1, \mu_2, \sigma_1^2, \sigma_2^2, \rho\right)$. Find the conditional distributions of $X/Y=y$ and $Y/X=x$. 8

(b) Derive the probability density function of p-variate normal distribution.

(c) (i) Describe the Wilcoxon Mann-Whitney U test. 5

(ii) Let $(X, Y) \sim \text{BVND}$ with parameters $\mu_x = 60$, $\mu_y = 75$, $\sigma_x = 5$, $\sigma_y = 12$ and $\rho = 0\cdot55$. Then find $P\{65 \le X \le 75\}$ 5

(d) Let $\underset{\frown}{X}_\alpha \ (\alpha = 1, 2, \cdots, N)$ be a random sample from $N_P\left(\underset{\frown}{\mu}, \Sigma\right)$ and let $\overline{\underset{\frown}{X}} = \frac{1}{N}\sum_{\alpha=1}^{N} \underset{\frown}{X}_\alpha$ be the sample mean vector.

Then prove that $\overline{\underset{\frown}{X}}$ is distributed as $N_P\left(\underset{\frown}{\mu}, \frac{\Sigma}{N}\right)$.

(e) (i) Let $\underset{\frown}{X}_\alpha{}^{(1)}(\alpha = 1, 2, \cdots N_1)$ be a random sample from $N_P\left(\underset{\frown}{\mu}^{(1)}, \Sigma\right)$ and let $\underset{\frown}{X}_\alpha{}^{(2)}(\alpha = 1, 2, \cdots N_2)$ be another random sample from $N_P\left(\underset{\frown}{\mu}^{(2)}, \Sigma\right)$ where the common dispersion matrix Σ is unknown. Discuss the procedure to test the hypothesis $H_0 : \underset{\frown}{\mu}^{(1)} = \underset{\frown}{\mu}^{(2)}$ using Hotelling's T^2 statistic. 5

(ii) In what way the ordinary sign test can be performed for paired samples? Explain. 5

(f) (i) State *any two* properties of multivariate normal distribution. 2

(ii) Derive the bivariate normal density as a particular case of multivariate normal distribution. 8

(g) (i) Let $\underset{\frown}{X} \sim N_3\left(\underset{\frown}{\mu}, \Sigma\right)$. Find the distribution of 5

$$\begin{pmatrix} X_1 - X_2 \\ X_2 - X_3 \end{pmatrix}$$

(ii) Derive the formula for Multiple correlation coefficient for a trivariate distribution. 5

(h) (i) Explain the distribution free method. 3

(ii) Derive the moment generating function of a bivariate normal distribution with usual parameters. 7
