Total number of printed pages–7

3 (Sem–6) STS M2

2021

STATISTICS

(Major)

Paper : 6·2

(Design of Experiments)

Full Marks : 60

Time : Three hours

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

1. Answer the following : 1×5=5

   (a) What are the assumptions made about the error component in the analysis of variance techniques ?

   (b) Name the Simplest design making use of all the three basic principles of design of experiments.

Contd.

(c) What will be the error degrees of freedom (d.f.) in an RBD with 4 blocks comparing 5 treatments, having one missing observation ?

(d) What is contrast ?

(e) Name *one* technique which is used to reduce the error in design of experiment.

2. Answer the following : 2×5=10

(a) Explain the situation in which a randomized block design is considered as an improvement over a completely randomized design.

(b) In a RBD, the yield of the plot for first treatment and first block is 50 *kg*. Mean of the first treatment is 25 *kg*, mean of the first block is 12 *kg* and grand mean is 10 *kg*. Find the estimate of error component for the corresponding plot.

(c) Show that in a 2^3 factorial design with the factors A, B and C each at two levels *viz.*, 0 and 1, the main effect A and the interaction effect AB are orthogonal contrast.

(d) State why a particular block in each replication of confounded factorial design, whether completely or partially is known as the 'Key block' or the 'Control block' .

(e) When will you call a confounded factorial design to be balanced ? Explain with 2^3 factorial experiment for three replications.

3. Answer the following : 5×3=15

(a) What do you mean by control of error in Design of experiments ? Describe *two* methods for control of error.

(b) In a LSD, five treatments are used. The treatments are tabulated as below :

Treatment	*A*	*B*	*C*	*D*	*E*
Mean	48.48	46.94	41.04	43.00	51.40

The treatment SS and error SS are respectively 348.24 and 304.08.

Examine at 5% level whether there exists real difference between means of treatment *B* and *E* .

Given,

F □.□ □ ; □, □ □ = 3.26

t □.□ □ □, □ □ = 2.18

(c) Describe the layout of a 2^3 experiment where all the interactions are partially confounded. In such a case, indicate degrees of freedoms (d.f.s) and sum of squares (SS) for all the components of treatment SS.

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

(a) Suppose you have k treatments to be compared in k^2 plots. How will you carry out the experiment under each of the following situations ?

(i) There is no fertility difference among k^2 plots

(ii) The fertility changes along one direction only

(iii) The fertility changes along two perpendicular directions

Give the appropriate analysis of variance table for each case.

(b) Show that when missing plot technique is applied in a randomized block design (RBD).
[adjusted treatment SS] ≥ [SS_t]
where SS_t = treatment SS obtained from original data.

(c) Suppose you wish to set up an experiment to test the effectiveness of 2 levels of Nitrogen, 2 of phosphate and 2 of potash on the yield of potatoes and have enough land to plant 40 plots. Show how you will set up this experiment and how you will analyse the results obtained.

(d) What are the main considerations in the use of confounded factorial design ? Suppose three factors A, B and C (all parametric) are to be studied, each at two levels. In carrying out the experiment it is necessary to run it in 2 blocks of 4 plots. Two replicates are planned. Set up the formulas for the sum of squares and degrees of freedom for each effect, if the first replicate has blocks confounded with ABC and second replicate has blocks confounded with BC.

(e) If Y_{ij} is the observation from plot in the i^{th} block to which j^{th} treatment is applied and $E(Y_{ij}) = \mu + a_i + t_j$; where μ is the average effect, a_i and t_j are fixed effects of i^{th} block and j^{th} treatment respectively ($i = 1, 2, \ldots\ldots b$; $j = 1, 2, \ldots\ldots r$). Obtain estimate of a_i and t_j.

Analyse the above design and compare it with completely randomized design.

Total number of printed pages–4

3 (Sem–6) STS M 3

2021

STATISTICS

(Major)

Paper : 6·3

(Applied Statistics – 2)

Full Marks : 60

Time : Three hours

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

GROUP–A

1. Answer the following as directed : 1×5=5

   (a) Statistical control charts help in controlling the quality during manufacturing process.

   (Write true ***or*** *false)*

Contd.

(b) In a C-chart, if average number of defects $\bar{c}$ is 9, Then lower control limit is :

(i) 9

(ii) 0

(iii) 18

(iv) None of these

(Choose the correct option)

(c) General fertility rates mainly depend on —

(i) total female population

(ii) total population

(iii) female population of child bearing age

(iv) None of these

(Choose the correct option)

(d) In a life table $d_x = 1 - \frac{l_{x+1}}{l_x}$.

(Write true ***or*** *false)*

(e) Infant mortality rate is computed for children —

(i) above 5 years age

(ii) between the age of 1 to 5 years

(iii) under the age of 1 year

(iv) None of these.

(Choose the correct option)

2. Answer ***all*** questions : 2×5=10

(a) What is meant by process control and product control in industrial statistics ?

(b) Why is an abridged life table prepared ?

(c) Explain the usefulness of R-chart.

(d) What do you understand by acceptance quality level (A.Q.L) ?

(e) Write a note on expectation of life.

3. Answer ***all*** questions : 5×3=15

(a) What is meant by Sampling inspection plan ? Describe the single sampling inspection plan.

(b) What is C-chart ? How are the control limits for C-chart obtained ? Justify the use of Poisson distribution in their computation.

(c) Define central mortality rate. Show that with the usual notations :

$$m_x = \frac{2q_x}{2 - q_x}$$

GROUP–B

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

(a) What do you understand by a crude birth rate ? Is it an accurate measure of the population growth of a country ? If not, how can it be improved to give better results ?

(b) Describe in detail, the construction of a complete life table and its uses. Give also the assumptions regarding the population under which a life table is constructed.

(c) Distinguish between crude and specific death rates. Explain clearly the purpose and procedure of standardised death rates.

(d) Explain in detail, $\overline{X}$ and R-charts. What purposes do they serve ? What are their advantages over the P-chart ?

(e) What are the advantages of statistical quality control ? Also explain the justificaiton for using the three sigma (3σ) limits in the control charts.

———

Total number of printed pages–7

3 (Sem–6) STS M 4

2021

STATISTICS

(Major)

Paper : 6·4

(Computer Programming and Multivariate Analysis)

Full Marks : 60

Time : Three hours

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

GROUP–A

(30 marks)

1. Answer the following questions as directed: 1×5=5

   (a) Let (X, Y) follows bivariate normal distribution with parameters $\left(\mu_1, \mu_2, \sigma_1^2, \sigma_2^2, \rho\right)$. Then the conditional variance of $Y/X = x$ is _____.
   (Fill in the blank)

Contd.

(b) State whether the following is acceptable as Fortran 77 statement.

$$X + Y = Z$$

(c) Let $(X, Y) \sim \text{BVND}(0, 0, \sigma_1^2, \sigma_2^2, \rho)$. Then the pdf will be —

(i) $$f(x) = \frac{1}{2\pi\sigma_1\sigma_2\sqrt{(1-\rho^2)}} e^{-\frac{1}{2(1-\rho^1)}\left[\left(\frac{x}{\sigma_1}\right)^2 + \left(\frac{y}{\sigma_2}\right)^2 - 2\rho\frac{xy}{\sigma_1\sigma_2}\right]}$$

(ii) $$f(x) = \frac{1}{2\pi\sigma_1\sigma_2\sqrt{(1-\rho^2)}} e^{-\frac{1}{2(1-\rho^1)}\left[\left(\frac{x}{\sigma_1}\right)^2 + \left(\frac{y}{\sigma_2}\right)^2 + 2\rho\frac{xy}{\sigma_1\sigma_2}\right]}$$

(iii) $$f(x) = \frac{1}{2\pi\sigma_1\sigma_2\sqrt{(1-\rho^2)}} e^{-\frac{1}{2(1-\rho^1)}\left[\left(\frac{x}{\sigma_1}\right)^2 + \left(\frac{y}{\sigma_2}\right)^2 + 2\frac{xy}{\sigma_1\sigma_2}\right]}$$

(iv) None of the above

(Choose the correct option)

(d) Define Hotelling T^2 statistic.

(e) A flowchart is :

(i) Programming language

(ii) A graphical representation of an algorithm

(iii) Step by step procedure of a programme written in English

(iv) None of the above

(Choose the correct option)

2. Answer the following questions : 2×5=10

(a) Write equivalent Fortran 77 statements of the following expressions:

(i) $e^{-x}x^{kx}$

(ii) $e^{-\left(\frac{x-a}{b}\right)^2}$

(b) Let $(X, Y) \sim \text{BVND}\left(\mu_1, \mu_2, \sigma_1^2, \sigma_2^2, \rho\right)$. If $\rho = 0$, then prove that $X \sim N\left(\mu_1, \sigma_1^2\right)$ and $Y \sim N\left(\mu_2, \sigma_2^2\right)$.

(c) Let $\underset{\sim}{X} \sim N_3\left(\underset{\sim}{\mu}, \Sigma\right)$ where

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

then find ρ_{13}.

(d) Given the sides of a triangle a, b, c. Write an algorithm to find its area

$[\, Area = \sqrt{s(s-a)(s-b)(s-c)}, \quad where \quad 2s = a+b+c \,]$

(e) Find $var\left(C \underset{\sim}{X}\right)$, where C is a $p \times p$ matrix of constant elements and X is a $p \times 1$ vector with variance-covariance matrix Σ.

3. Answer the following questions : 3×5=15

(a) Let $(X, Y) \sim$ BVND (0, 0, 1, 1, ρ), then prove (or disprove) that $X+Y$ and X–Y are independently distributed.

(b) Examine whether Hotelling T^2 is invariant under changes in the units of measurements.

(c) Derive bivariate normal density as a particular case of multivariate normal.

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

Find the distribution of $\underset{\sim}{Y} = \begin{pmatrix} X_1 - X_2 \\ X_2 - X_3 \end{pmatrix}$

(e) Write a note on flowchart symbols and their uses.

GROUP–B

(30 marks)

Answer ***any three*** *questions from this Group :* 10×3=30

4. Derive the Characteristic function $\Phi(t)$ of the bivariate normal distribution (with usual parameters) and hence deduce the expression $\Phi(t)$ when variables are independent. 8+2=10

5. *(a)* Let $\underset{\sim}{X} \sim N_3\left(\underset{\sim}{\mu}, \Sigma\right)$ where

$$\Sigma = \begin{pmatrix} 1 & \rho & 0 \\ \rho & 1 & \rho \\ 0 & \rho & 1 \end{pmatrix}.$$

For what value of ρ, are $X_1+X_2+X_3$ and $X_1-X_2-X_3$ are statistically independent? 5

(b) Let $\underset{\sim}{X} = N_p\left(\underset{\sim}{\mu}, \Sigma\right)$. Consider the transformation $\underset{\sim}{X} - \underset{\sim}{\mu} = C\underset{\sim}{Y}$, where C is $p \times p$ non-singular matrix and $\underset{\sim}{Y} = \left(y_1\ y_2 \cdots y_p\right)'$. Prove that the corresponding Jacobian of transformation is $|J| = |\Sigma|^{1/2}$. 5

6. *(a)* If $\underset{\sim}{X} \sim N_2(0, \Sigma)$, where

$\Sigma = (\sigma_{ij})$, $i, j = 1, 2$, then prove that

$$\left(\underset{\sim}{X}' \Sigma^{-1} \underset{\sim}{X} - \frac{X_1^2}{\sigma_{11}} \right) \sim \chi_1^2$$ 8

(b) Assuming that $k = 1$, $l = 45$, $m = 7$, $n = 5$, evaluate the following Fortran 77 expressions : 2

(i) $(l+m)/n+k$

(ii) $l+m/n+k$

7. Let $\underset{\sim}{X} \sim N_3\left(\underset{\sim}{\mu}, \Sigma\right)$, where $\underset{\sim}{\mu} = (2, -3, 1)'$ and

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

Find the following :

(i) Conditional pdf of X_1 given X_2 and X_3.

(ii) Regression equation of X_1 on X_2 and X_3.

(iii) Conditional variance of X_1 given X_2 and X_3. 6+2+2=10

8. *(a)* Draw a flowchart indicating the steps to find the regression equation of Y on X. 5

(b) Given a random sample drawn from a population $\underset{\sim}{X} \sim N_p\left(\underset{\sim}{\mu}, \Sigma\right)$. Discuss the procedure to test the following hypothesis —

$\underset{\sim}{\mu} = \underset{\sim}{\mu}_0$ (*specified*), when Σ is known. 5