(f) For a simple regression model $y_i = \beta_0 + \beta_1 X_i + \varepsilon_i$; $i = 1, 2, ..., \gamma$, find the $100(1-\alpha)\%$ confidence interval for mean response at a particular value of the regression variable X.

Total number of printed pages–8

3 (Sem–4/CBCS) STA HC 2

2024

STATISTICS

(Honours Core)

Paper : STA–HC–4026

(Linear Models)

Full Marks : 60

Time : Three hours

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

1. Answer the following questions as directed : $1\times7=7$

   (a) As variability due to chance decreases, the value of F will

   (i) increase

   (ii) decrease

   (iii) stay the same

   (iv) Can't tell from the given information

   (Choose the correct option)

(b) The degree of freedom associated with error mean of squares in two-way classification (with one observation per cell) is :

(i) $n-1$

(ii) $k-1$

(iii) $(n-1)(k-1)$

(iv) $nk-1$

(Choose the correct option)

(c) The term $\sum_i \sum_j \left(y_{ij} - \bar{y}_{io}\right)^2$ in one-way ANOVA is called

(i) Variance

(ii) Total sum of squares

(iii) Sum of squares due to treatments

(iv) Error sum of squares

(Choose the correct option)

(d) The equation that describes how the response variable (Y) is related to the explanatory variable (X) is

(i) the correlation model

(ii) the regression model

(iii) used to compare the correlation coefficient

(iv) None of the above

(Choose the correct option)

(e) SSE cannot be

(i) larger than SST

(ii) smaller than SST

(iii) equal to none

(iv) equal to zero

(Choose the correct option)

(f) If the coefficient of determination is equal to 1, then the correlation coefficient

(i) must also be equal to 1

(ii) can be either -1 or $+1$

(iii) can be any value between -1 and 1

(iv) must be -1

(Choose the correct option)

(g) If the coefficient of determination is 0·64, the correlation coefficient

(i) is 0·529

(ii) could be either 0·80 or $-0·80$

(iii) must be positive

(iv) must be negative

(Choose the correct option)

2. Answer the following questions briefly : 2×4=8

(a) State some applications of analysis of variance.

(b) What is the difference between R^2 and adj R^2.

(c) Define linear model.

(d) Consider the model

$$Y_i = i\beta + \varepsilon_i;\ i = 1, 2, 3$$

where ε_1, ε_2 and ε_3 are independent with zero mean and variance σ^2, $2\sigma^2$ and $3\sigma^2$ respectively. Find the mean and variance of

$$T = \frac{Y_1 + Y_2 + Y_3}{6}$$

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

(a) There are 3 observations independently drawn such that

$$y_1 = \theta_1 + \varepsilon_1$$
$$y_2 = \theta_1 + \theta_2 + \varepsilon_2$$
$$y_3 = \theta_2 + \varepsilon_3$$

where θ_1 and θ_2 are unknown parameters and ε_1, ε_2 and ε_3 are i.i.d. $N(0, \sigma^2)$. Find the best linear unbiased estimator (BLUE) of θ_1 and θ_2.

(b) Let $y_i = \alpha_1 + \alpha_2 x_i + \varepsilon_i;\ i = 1, 2, \ldots, 10$ where x_i's are fixed covariates and ε_i's are i.i.d. $N(0, \sigma^2)$ random variables. Suppose $\hat{\alpha}_1$ and $\hat{\alpha}_2$ are least square estimators of α_1 and α_2 respectively. Given the following data :

$$\sum_{i=1}^{10} x_i = 0,\ \sum_{i=1}^{10} x_i^2 = 21,\ \sum_{i=1}^{10} x_i y_i = 21,$$

$$\sum_{i=1}^{10} y_i = 12,\ \sum_{i=1}^{10} y_i^2 = 58.$$

Show that correlation coefficient between $\hat{\alpha}_1$ and $\hat{\alpha}_2$ is zero.

(c) Define analysis of variance (AoV). Write down the assumptions involved in AoV. 2+3=5

(d) Using the following data :

Y :	65	57	57	54	66
X :	26	13	16	−7	27

Estimate the regression line $Y = \alpha + \beta X$.

(e) Write short notes on : 2½+2½=5

(i) Coefficient of determination

(ii) Parametric function

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

(a) Define estimable function. Let Y_1, Y_2, Y_3 be uncorrelated observations with common variance σ^2 and $E(Y_1) = \theta_0 + \theta_1$, $E(Y_2) = \theta_0 + \theta_2$, $E(Y_3) = \theta_0 + \theta_3$, where θ_i's; $i = 1, 2, 3$ are unknown parameters. Show that $\theta_1 - \theta_2$, $\theta_1 - \theta_3$ and $\theta_2 - \theta_3$ are each estimable.

(b) (i) State the Gauss-Markov theorem.

(ii) Let y_1, y_2, y_3 be uncorrelated observations with common variance σ^2 and $E(y_1) = \alpha_1$, $E(y_2) = \alpha_2$ and $E(y_3) = \alpha_1 + \alpha_2$, where α_1 and α_2 are unknown parameters. Show that $T = \frac{y_1 + y_2 + y_3}{3}$ is the Best Linear Unbiased Estimator of $\alpha_1 + \alpha_2$.

(c) Derive the analysis of variance for two-way classification data with one observation per cell under fixed effect model.

(d) There are 4 observations independently drawn such that

$$y_1 = \beta_1 + \varepsilon_1$$
$$y_2 = \beta_2 + \varepsilon_2$$
$$y_3 = \beta_3 + \varepsilon_3$$
$$y_4 = \beta_1 + \beta_2 + \beta_3 + \beta_4$$

where $\beta_1, \beta_2, \beta_3$ and β_4 are unknown parameters and ε_i's, $i = 1, 2, 3, 4$ are independently and identically distributed normal variables with mean zero and known variance σ^2. Find the variance of the unbiased estimator of $\beta_1, \beta_2, \beta_3$ and β_4.

(e) Define simple linear regression model. Write the basic assumptions of simple linear regression model. Estimate the parameters of the model. 2+3+5=10