INDIAN STATISTICAL INSTITUTE

Linear Statistical Models

(B.Stat. 3rd year)

Semester 1, 2025-26

Mid-Semester Examination

Date: September 9,2025

Total time: 2 hours

(Total score is 40 points. One sheet (both sided) of notes is allowed.)

1. Suppose $\mathbf{Y} \sim N_n(\mathbf{X}\boldsymbol{\beta}, \sigma^2 I_n)$, where $\beta \in \mathbb{R}^p$, and $\text{rank}(\mathbf{X}) = p$. Further, suppose that $\mathbf{X} = [\mathbf{Z} : \mathbf{W}]$ where $\mathbf{Z}$ is an $n \times q$ matrix with $1 \leq q < p$. Define

$$S_1 = \min_{\beta \in \mathbb{R}^p} \|\mathbf{Y} - \mathbf{X}\boldsymbol{\beta}\|_2^2, \qquad \text{and} \qquad S_0 = \min_{\gamma \in \mathbb{R}^q} \|\mathbf{Y} - \mathbf{Z}\gamma\|_2^2$$

Then show that, S_1/S_0 is distributed as $\text{Beta}(k_1, k_2)$ where $k_1 = (n-p)/2$ and $k_2 = (n-q)/2$, provided $\mathbb{E}(\mathbf{Y}) \in \mathcal{C}(\mathbf{Z})$.

[5 points]

2. Consider the model

$$Y_{ijk} = \alpha_i + \beta_j + \varepsilon_{ijk}, \qquad 1 \leq i, j \leq 2; \qquad k = 1, \ldots, m$$

where ε_{ijk} are assumed to be i.i.d. $N(0, \sigma^2)$ and $m \geq 2$.

(a) Show that the corresponding design matrix $\mathbf{X}$ has rank 3.

(b) Find the BLUE of $\boldsymbol{\theta} = \mathbf{X}^T\boldsymbol{\mu}$, where the mean parameter $\boldsymbol{\mu} = (\mu_{11}, \mu_{12}, \mu_{21}, \mu_{22})^T$ with $\mu_{ij} = \mathbb{E}(Y_{ijk})$, for $1 \leq i, j \leq 2$, and $k \geq 1$.

(c) In each of the following cases, check whether the parameter is estimable, and if so, find its BLUE. Justify your answer.

(i) α_1

(ii) $\alpha_1 - \alpha_2$

(iii) $\alpha_1 + \alpha_2 + \beta_1 + \beta_2$

(iv) $\alpha_1 - \alpha_2 + \beta_1 - \beta_2$

[15 points = 3 + 4 + 8]

3. Suppose that we obtain experimental data from the model

$$Y_{ij} = \mu_i + \beta_1 z_{ij} + \beta_2 w_{ij} + \varepsilon_{ij}, \qquad j = 1, \ldots, n_i, \quad i = 1, \ldots, r. \tag{1}$$

where $\varepsilon_{ij} \overset{i.i.d.}{\sim} N(0, \sigma^2)$ and index i corresponds to the i-th treatment. Here Z and W are two continuous variables (covariates), and z_{ij} and w_{ij} denote the measurements on these variables for the j-th subject corresponding to the i-th treatment. Define

$$S_{zz} = \sum_{i=1}^{r}\sum_{j=1}^{n_i}(z_{ij}-\bar{z}_{i\cdot})^2 \qquad S_{zw} = \sum_{i=1}^{r}\sum_{j=1}^{n_i}(z_{ij}-\bar{z}_{i\cdot})(w_{ij}-\overline{w}_{i\cdot}) \qquad S_{ww} = \sum_{i=1}^{r}\sum_{j=1}^{n_i}(w_{ij}-\overline{w}_{i\cdot})^2$$

$$S_{zy} = \sum_{i=1}^{r}\sum_{j=1}^{n_i}(z_{ij}-\bar{z}_{i\cdot})(Y_{ij}-\overline{Y}_{i\cdot}) \qquad S_{wy} = \sum_{i=1}^{r}\sum_{j=1}^{n_i}(w_{ij}-\overline{w}_{i\cdot})(Y_{ij}-\overline{Y}_{i\cdot})$$

and $\bar{z}_{i\cdot} = \frac{1}{n_i}\sum_{j=1}^{n_i} z_{ij}$, $\overline{w}_{i\cdot} = \frac{1}{n_i}\sum_{j=1}^{n_i} w_{ij}$, $\overline{Y}_{i\cdot} = \frac{1}{n_i}\sum_{j=1}^{n_i} Y_{ij}$, and $\widehat{\sigma}^2 = MSE$ for model (1).

(a) Show that the least squares estimator of $\boldsymbol{\beta} = (\beta_1, \beta_2)^T$ is given by

$$\widehat{\boldsymbol{\beta}} = \begin{bmatrix} S_{zz} & S_{zw} \\ S_{zw} & S_{ww} \end{bmatrix}^{-1} \begin{bmatrix} S_{zy} \\ S_{wy} \end{bmatrix}$$

(b) Show that the likelihood ratio test for $H_0 : \beta_2 = 0$ against $H_A : \beta_2 \neq 0$, rejects H_0 for large values of the statistic U, where

$$U = (S_{ww} - S_{zw}^2/S_{zz})\,\frac{\widehat{\beta}_2^2}{\widehat{\sigma}^2}\ .$$

(c) Show that the likelihood ratio test for $H_0 : \beta_1 = \beta_2 = 0$ against $H_A : \beta_1 \neq 0$ or $\beta_2 \neq 0$, rejects H_0 for large values of the statistic V, where

$$V = \frac{1}{\widehat{\sigma}^2}\widehat{\boldsymbol{\beta}}^T \begin{bmatrix} S_{zz} & S_{zw} \\ S_{zw} & S_{ww} \end{bmatrix} \widehat{\boldsymbol{\beta}}$$

[**12 points = 4 + 4 + 4**]

4. Suppose $\mathbf{A}$ and $\mathbf{B}$ are symmetric $n \times n$ matrices. Further, suppose that $\mathbf{B}$ is idempotent, and $\mathbf{A}$ and $I_n - \mathbf{A} - \mathbf{B}$ are non-negative definite matrices. Show that $\mathbf{AB} = \mathbf{BA} = \mathbf{0}_n$, where $\mathbf{0}_n$ is the $n \times n$ matrix of zero entries.

[**Hint:** Show that $\mathbf{y}^T\mathbf{A}\mathbf{y} = 0$ for all $\mathbf{y} \in \mathcal{C}(\mathbf{B})$.]

[**8 points**]