Solution to Gate ST 2023 Q 26

Mayank Gupta

Question : Consider the following regression model

$$y_k = \alpha_0 + \alpha_1 \log_e k + \epsilon_k, \qquad k = 1, 2, \ldots, n,$$

where ϵ_k's are independent and identically distributed random variables each having probability density function $f(x) = \frac{1}{2}e^{-|x|}, x \in \mathbb{R}$. Then which one of the following statements is true?

(A) The maximum likelihood estimator of α_0 does not exist
(B) The maximum likelihood estimator of α_1 does not exist
(C) The least squares estimator of α_0 exists and is unique
(D) The least squares estimator of α_1 exists, but it is not unique

Solution:

$$\epsilon_k \overset{\text{iid}}{\sim} \mathcal{N}(0, \sigma^2) \tag{1}$$

$$f(\epsilon_k) = \frac{1}{2}e^{-|\epsilon_k|} \tag{2}$$

$$\text{Likelihood function : } f(\epsilon_1\epsilon_2....\epsilon_n) = \prod_{k=1}^{n} f(\epsilon_k) \tag{3}$$

$$L = \prod_{k=1}^{n} \frac{1}{2}e^{-|\epsilon_k|} \tag{4}$$

$$= \left(\frac{1}{2}\right)^n e^{\left(\sum_{k=1}^{n} -|\epsilon_k|\right)} \tag{5}$$

$$= \left(\frac{1}{2}\right)^n e^{\left(\sum_{k=1}^{n}\left(-|y_k - \alpha_0 - \alpha_1 \log_e k|\right)\right)} \tag{6}$$

$$L_1 = \ln L = -n \ln 2 - \sum_{k=1}^{n} \left(|y_k - \alpha_0 - \alpha_1 \log_e k|\right) \tag{7}$$

$$L_1 = \text{function of } \alpha_0, \alpha_1 \tag{8}$$

$\because$ $\exists$ a value of α_0 and α_1 which will maximise the value of L_1 which can be obtained by differentiating L_1 by α_0 and α_1

$\therefore$ Maximum likelihood estimator of α_0 and α_1 exist

$\therefore$ Option (A) and (B) are incorrect

The least square estimator of α_0 and α_1 is $\tilde{\alpha_0}$ and $\tilde{\alpha_1}$ which will minimise

$$Q(\alpha_0, \alpha_1) = \sum_{k=1}^{n} (y_k - \alpha_0 - \alpha_1 \log_e k)^2 \tag{9}$$

$\because$ $\exists$ a value of α_0 and α_1 which will manimise the value of $Q(\alpha_0, \alpha_1)$ which can be obtained by differentiating $Q(\alpha_0, \alpha_1)$ by α_0 and α_1

$\therefore$ Least square estimator of α_0 and α_1 exists and are unique

$\therefore$ Option (C) is correct and (D) is incorrect