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:

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

$$f(\epsilon_k) = \frac{1}{\sigma\sqrt{2\pi}} e^{\frac{-1}{2}\frac{(\epsilon_k - 0)^2}{\sigma^2}} \tag{2}$$

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

$$L = \prod_{k=1}^{n} \frac{1}{\sigma\sqrt{2\pi}} e^{\frac{-1}{2}\frac{(\epsilon_k)^2}{\sigma^2}} \tag{4}$$

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

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

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

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

$$\frac{\partial L_1}{\partial \alpha_0} = -\frac{1}{\sigma^2}\sum_{k=1}^{n}\left(y_k - \alpha_0 - \alpha_1 \log_e k\right) = 0 \tag{9}$$

$$\sum_{k=1}^{n}\left(y_k - \alpha_0 - \alpha_1 \log_e k\right) = 0 \tag{10}$$

$$n\bar{y} - n\alpha_0 - \alpha_1 n\bar{x} = 0, \quad \text{where } \bar{x} = \frac{1}{n}\sum_{k=1}^{n}\log_e k, \quad \bar{y} = \frac{1}{n}\sum_{k=1}^{n} y_k \tag{11}$$

$$\implies \tilde{\alpha_0} = \bar{y} - \tilde{\alpha_1}\bar{x} \tag{12}$$

$$\frac{\partial L_1}{\partial \alpha_1} = -\frac{1}{\sigma^2}\sum_{k=1}^{n}\left(y_k - \alpha_0 - \alpha_1 \log_e k\right)\log_e k = 0 \tag{13}$$

$$\implies \tilde{\alpha_1} = \frac{\sum_{k=1}^{n}\left(\log_e k - \bar{x}\right)\left(y_k - \bar{y}\right)}{\sum_{k=1}^{n}\left(\log_e k - \bar{x}\right)^2} \tag{14}$$

$\because$ Maximum likelihood estimator of α_0 and α_1 exist
$\therefore$ Option (A) and (B) are incorrect
The least square estimator of α_0 is $\tilde{\alpha}_0$ and is unique
The least square estimator of α_1 is $\tilde{\alpha}_1$ and is unique
$\therefore$ Option (D) is incorrect
$\therefore$ Option (C) is correct