LECTURE - 16

Agenda:

1. Definition of likelihood
2. Alternative definition of sufficiency
3. Examples

In the context of estimation, "likelihood of the data" refers to the joint probability density (mass function) of the observed sample.

Definition: Suppose $Y_1, Y_2, \ldots, Y_n$ is a sample from a population with parameter θ (unknown). Suppose $y_1, y_2, \ldots, y_n$ are the observed values of $Y_1, Y_2, \ldots, Y_n$ in a specific instance.

1. If $Y_1, Y_2, \ldots, Y_n$ are all continuous random variables, the likelihood (denoted by $L(y_1, y_2, \ldots, y_n \mid \theta)$) is defined to be the joint density of $Y_1, Y_2, \ldots, Y_n$ evaluated at $y_1, y_2, \ldots, y_n$.

2. If $Y_1, Y_2, \ldots, Y_n$ are all discrete random variables, the likelihood (again denoted by $L(y_1, y_2, \ldots, y_n \mid \theta)$) is defined to be the joint probability mass function of $Y_1, Y_2, \ldots, Y_n$ evaluated at $y_1, y_2, \ldots, y_n$.

Example: Let $Y_1, Y_2, ---, Y_n$ be a random sample from a ~~exponential~~ population which is Exponential with mean θ. What is the likelihood of observing $y_1, y_2, ---, y_n$ in a specific instance?

Note that

(1) $Y_1, Y_2, ---, Y_n$ are continuous random variables.

(2) $Y_1, Y_2, ---, Y_n$ are independent random variables, hence their joint density is the product of their marginal densities.

Hence,

$$L(y_1, y_2, ---, y_n \mid \theta)$$

$$= f_{Y_1}(y_1) \times f_{Y_2}(y_2) \times --- \times f_{Y_n}(y_n)$$

$$= \begin{cases} \dfrac{e^{-y_1/\theta}}{\theta} \times \dfrac{e^{-y_2/\theta}}{\theta} \times --- \times \dfrac{e^{-y_n}}{\theta} & \text{if } y_1, y_2, ---, y_n > 0, \\ 0 & \text{otherwise.} \end{cases}$$

$$= \begin{cases} \dfrac{e^{-\sum_{i=1}^{n} y_i/\theta}}{\theta^n} & \text{if } y_1, y_2, ---, y_n > 0, \\ 0 & \text{otherwise.} \end{cases}$$

An alternative definition of sufficiency can be provided in terms of likelihood.

<u>RESULT:</u> Let W be an estimator based on a sample $Y_1, Y_2, ---, Y_n$ from a population θ. Then W is a sufficient statistic (or a sufficient estimator) for θ if and only if the likelihood can be factored as

$$L(y_1, y_2, ---, y_n | \theta) = g(w, \theta) \times h(y_1, y_2, ---, y_n),$$

where $g(w, \theta)$ is a function only of w and θ and $h(y_1, y_2, ---, y_n)$ does not depend on θ.

<u>IMPORTANT:</u> The above result should hold for <u>ALL</u> possible observed values $y_1, y_2, ---, y_n$ of the sample and the corresponding observed value w of the estimator W.

<u>Example:</u> Let us continue to discuss the Exponential example. We derived the likelihood as

$$L(y_1, y_2, ---, y_n | \theta) = \begin{cases} \dfrac{e^{-\Sigma y_i/\theta}}{\theta^n} & \text{if } y_1, y_2, ---, y_n > 0, \\ 0 & \text{otherwise.} \end{cases}$$

Now, since θ is the population mean, an intuitive estimator for θ is the sample mean . $\bar{Y}$. The question is whether $\bar{Y}$ is a sufficient statistic?

Note that,

$$L(y_1, y_2, \ldots, y_n \mid \theta) = \frac{e^{-\sum_{i=1}^{n} y_i / \theta}}{\theta^n} \mathbb{1}_{\{y_1, y_2, \ldots, y_n > 0\}}$$

<u>NOTATION:</u> For any set S, the function

$\mathbb{1}_S$ is defined as

$$\mathbb{1}_S(y) = \begin{cases} 1 & \text{if } y \in S, \\ 0 & \text{otherwise} \end{cases}$$

Hence, we can observe that indeed factorise the likelihood as

$$L(y_1, y_2, \ldots, y_n \mid \theta) = \underbrace{\frac{e^{-n\bar{y}/\theta}}{\theta^n}}_{g(\bar{y}, \theta)} \times \underbrace{\mathbb{1}_{\{y_1, y_2, \ldots, y_n > 0\}}}_{h(y_1, y_2, \ldots, y_n)}$$

Hence $\bar{Y}$ is a sufficient statistic for θ.