STA 6351, Report.2.11

Carson Slater *Baylor University*

2.11

(a) Show that the likelihood for $\boldsymbol{X} = (X_1, \ldots, X_n)$ is, for $\boldsymbol{X} = \boldsymbol{x}$,

$$L(\theta; \boldsymbol{x}) = \theta^{-n} \mathbb{I}\left(\theta \geq \max_i x_i\right).$$

The likelihood function for the vector $\boldsymbol{X} = (X_1, \ldots, X_n)$ at the observed values $\boldsymbol{X} = \boldsymbol{x}$ is the joint density:

$$\begin{aligned} L(\theta; \boldsymbol{x}) &= \prod_{i=1}^{n} f(x_i; \theta) \\ &= \prod_{i=1}^{n} \left[\frac{1}{\theta} \mathbb{I}(0 \leq x_i \leq \theta)\right] \\ &= \theta^{-n} \prod_{i=1}^{n} \mathbb{I}(0 \leq x_i \leq \theta). \end{aligned}$$

The product term $\prod_{i=1}^{n} \mathbb{I}(0 \leq x_i \leq \theta)$ is equal to 1 if and only if $0 \leq x_i \leq \theta$ for all $i = 1, \ldots, n$. Assuming the observed data are non-negative ($x_i \geq 0$), this condition is strictly equivalent to requiring that the maximum value in the sample does not exceed θ. Therefore:

$$\prod_{i=1}^{n} \mathbb{I}(x_i \leq \theta) = \mathbb{I}\left(\max_i x_i \leq \theta\right) = \mathbb{I}\left(\theta \geq \max_i x_i\right).$$

Substituting this indicator back into the joint density gives the final likelihood:

$$L(\theta; \boldsymbol{x}) = \theta^{-n} \mathbb{I}\left(\theta \geq \max_i x_i\right).$$

(b) Verify that the minimal sufficient statistic is

$$T(\boldsymbol{X}) = \max_i X_i.$$

We use the Factorization Theorem for sufficiency. The likelihood function is:

$$L(\theta; \boldsymbol{x}) = \theta^{-n} \mathbb{I}\left(\theta \geq \max_i x_i\right).$$

We factor this likelihood into $g(T(\boldsymbol{x}), \theta) \cdot h(\boldsymbol{x})$ where $T(\boldsymbol{x}) = \max_i x_i$:

- $g(T(\boldsymbol{x}), \theta) = \theta^{-n} \mathbb{I}(\theta \geq T(\boldsymbol{x}))$.
- $h(\boldsymbol{x}) = 1$.

By the Factorization Theorem, $T(\boldsymbol{X}) = \max_i X_i$ is a sufficient statistic. To show it is minimal sufficient, we consider the ratio of the likelihoods for any two sample points $\boldsymbol{x}$ and $\boldsymbol{y}$:

$$\frac{L(\theta; \boldsymbol{x})}{L(\theta; \boldsymbol{y})} = \frac{\mathbb{I}(\theta \geq \max_i x_i)}{\mathbb{I}(\theta \geq \max_i y_i)}.$$

This ratio is independent of θ if and only if $\max_i x_i = \max_i y_i$, or $T(\boldsymbol{x}) = T(\boldsymbol{y})$. Thus, $T(\boldsymbol{X})$ is the minimal sufficient statistic.

(c) Compute the score function formally:

$$S(\theta) = \frac{\partial}{\partial\theta} \log L(\theta; \boldsymbol{x}) = -\frac{n}{\theta} + \frac{\partial}{\partial\theta} \log \mathbb{I}\left(\theta \geq \max_i x_i\right).$$

$$S(\theta) = \frac{\partial}{\partial\theta} \log L(\theta; \boldsymbol{x}) = -\frac{n}{\theta} + \frac{\partial}{\partial\theta} \log \mathbb{I}\left(\theta \geq \max_i x_i\right).$$

The log-likelihood is:

$$\begin{aligned}
\log L(\theta; \boldsymbol{x}) &= \log\left[\theta^{-n} \mathbb{I}\left(\theta \geq \max_i x_i\right)\right] \\
&= -n \log\theta + \log \mathbb{I}\left(\theta \geq \max_i x_i\right).
\end{aligned}$$

The score function $S(\theta)$ is the derivative of the log-likelihood with respect to θ:

$$\begin{aligned}
S(\theta) &= \frac{\partial}{\partial\theta}\left[-n \log\theta + \log \mathbb{I}\left(\theta \geq \max_i x_i\right)\right] \\
&= \frac{\partial}{\partial\theta}(-n \log\theta) + \frac{\partial}{\partial\theta} \log \mathbb{I}\left(\theta \geq \max_i x_i\right) \\
&= -\frac{n}{\theta} + \frac{\partial}{\partial\theta} \log \mathbb{I}\left(\theta \geq \max_i x_i\right).
\end{aligned}$$

This matches the expression provided in the problem statement.

(d) Conclude that the score function is not well defined at the boundary $\theta = \max_i x_i$. Hence the Taylor expansion argument breaks down, illustrating why the approximate result requires smoothness.

Let $x_{(n)} = \max_i x_i$. The second term in the score function is $\frac{\partial}{\partial\theta} \log \mathbb{I}(\theta \geq x_{(n)})$.

For $\theta > x_{(n)}$, the indicator function $\mathbb{I}(\theta \geq x_{(n)})$ is constant and equal to 1. Therefore, $\log \mathbb{I}(\theta \geq x_{(n)}) = \log(1) = 0$. The derivative is:

$$\frac{\partial}{\partial \theta} \log \mathbb{I}(\theta \geq x_{(n)}) = 0, \quad for \theta > x_{(n)}.$$

For $\theta < x_{(n)}$, the indicator function $\mathbb{I}(\theta \geq x_{(n)})$ is constant and equal to 0. The log-likelihood is then $\log(0)$, which is undefined (tends to $-\infty$). This region is outside the support of the likelihood function.

At the boundary $\theta = x_{(n)}$, the indicator function (and thus the log-likelihood) jumps discontinuously from 0 to 1 (or is undefined, depending on conventions for $\theta = x_{(n)}$). The derivative of the log-likelihood, $S(\theta)$, is therefore discontinuous (or singular) at $\theta = x_{(n)}$ and is not well-defined in the classic sense of a continuous derivative.

The total score function for $\theta > x_{(n)}$ is $S(\theta) = -n/\theta$. As θ approaches the boundary $x_{(n)}$ from the right, the score function approaches $S(x_{(n)}^+) = -n/x_{(n)}$. For any $\theta < x_{(n)}$, the likelihood is zero, and the score is undefined.

The breakdown occurs because the standard results of Maximum Likelihood Estimation, such as the asymptotic normality of the MLE and the reliance on the Fisher Information (often derived from the Taylor expansion of the score function around the true parameter θ_0), rely on the **regularity condition** that the log-likelihood function is differentiable with respect to θ for *all* x and that the support of the density does not depend on θ. Since the support *does* depend on θ here, this condition is violated, and the Taylor expansion argument used to establish the large-sample properties of the MLE breaks down.