Properties of Maximum Likelihood estimators

Session 12

09/04/2021

Properties of ML estimators

- MLE need not be necessarily unbiased.
- MLE is NOT necessarily unique.
- MLEs are consistent.
- MLEs are most efficient.

Remark. *MLE's* are always consistent estimators but need not be unbiased. For example in sampling from $N(\mu, \sigma^2)$ population, [c.f. Example 15·31],

$\text{MLE}(\mu) = \bar{x}$ (sample mean), which is both unbiased and consistent estimator of μ.

$\text{MLE}(\sigma^2) = s^2$ (sample variance), which is consistent but not unbiased estimator of σ^2.

MLE is NOT necessarily unique.

- Example 1 :Find the M.L.E. of the parameter θ of the distribution

$$f(x;\theta) = \frac{1}{2}e^{-|x-\theta|} \quad , \qquad -\infty < x < \infty, -\infty < \theta < \infty,$$

Solution :$L(\theta)$= $\prod_{i=1}^{n}\left[\frac{1}{2}e^{-|x-\theta|}\right]$=$\left(\frac{1}{2}\right)^{n} e^{-\sum|x_i-\theta|}$

Log $L(\theta)$= $-nlog2 - \sum|x_i - \theta|$.

But log L is not differentiable w.r.t θ. But log L is maximum when $\sum|x_i - \theta|$ is minimum. Also we know that $\frac{1}{n}\sum|x_i - \theta|$ or MAD is minimum when θ=median(x).

When n is odd, $\hat{\theta} = med(x_1, x_2, \dots x_n)$.

When n is even, $\hat{\theta} = Y_{\frac{n}{2}} \leq \hat{\theta} \leq Y_{\frac{n}{2}}$+1, where $Y_1, Y_2, \dots Y_n$ are taken to be order statistics

- Example 2 : Let $x_1, x_2, \ldots x_n$ denote a random sample of size n from uniform population with p.d.f.

$$f(x, \theta) = 1 \; ; \; \theta - \frac{1}{2} \le x \le \theta + \frac{1}{2}, -\infty < \theta < \infty$$

- Obtain MLE of θ.
- Solution :

$$L = L(\theta; x_1, x_2, \ldots, x_n) = 1, \; \theta - \frac{1}{2} \le x_i \le \theta + \frac{1}{2}$$
$$= 0, \text{ elsewhere}$$

If $x_{(1)}, x_{(2)}, \ldots, x_{(n)}$ is the ordered sample then

$$\theta - \frac{1}{2} \le x_{(1)} \le x_{(2)} \le \ldots \le x_{(n)} \le \theta + \frac{1}{2}$$

Thus L attains the maximum if

$$\theta - \frac{1}{2} \leq x_{(1)} \qquad \wedge \qquad x_{(n)} \leq \theta + \frac{1}{2}$$

$$\Rightarrow \qquad \theta \leq x_{(1)} + \frac{1}{2} \quad \wedge \quad x_{(n)} - \frac{1}{2} \leq \theta$$

Hence *every statistic* $t = t(x_1, x_2, \ldots, x_n)$ *such that*

$$x_{(n)} - \frac{1}{2} \leq t\,(x_1, x_2, \ldots, x_n) \leq x_{(1)} + \frac{1}{2}$$

provides an M.L.E. for θ.

Remark. This example illustrates that M.L.E. for a parameter need not be unique.

- Example 3:Let $x_1, x_2, \dots x_n$ denote a random sample of size n from uniform population with p.d.f.

$$f(x, \theta) = \frac{1}{\theta}, 0 < x < \infty, \theta > 0$$
$$= 0, \quad elsewhere$$

- Obtain MLE of θ.

- Solution : Here

$$L = \prod_{i=1}^{n} f(x_i, \theta) = \frac{1}{\theta} \cdot \frac{1}{\theta} \dots \frac{1}{\theta} = \left(\frac{1}{\theta}\right)^n \qquad \dots(*)$$

Likelihood equation, *viz.*, $\frac{\partial}{\partial\theta} \log L = 0$, gives

$$\frac{\partial}{\partial\theta}(-n \log \theta) = 0 \quad \Rightarrow \quad \frac{-n}{\theta} = 0 \quad \Rightarrow \quad \hat{\theta} = \infty,$$

obviously an absurd result.

In this case we locate M.L.E. as follows :

We have to choose θ so that L in (*) is maximum. Now L is maximum if θ is minimum.

Let $x_{(1)}, x_{(2)}, \ldots, x_{(n)}$ be the *ordered* sample of n independent observations from the given population so that

$$0 \leq x_{(1)} \leq x_{(2)} \leq \ldots \leq x_{(n)} \leq \theta \quad \Rightarrow \quad \theta \geq x_{(n)}$$

Since the minimum value of θ consistent with the sample is $x_{(n)}$, the largest sample observation, $\hat{\theta} = x_{(n)}$.

- **In the last two examples, it is obvious that whenever the range of the variable involves the parameter(s) to be estimated, the likelihood equations fail to give us valid estimates and in this case M.L.Es are obtained by adopting some other approach of maximizing *L* or log *L* directly.**

Theorem : If a sufficient estimator exists, it is a function of the maximum likelihood estimator

Example : Obtain the maximum likelihood estimate of θ in

$$f(x, \theta) = (1 + \theta)\, x^{\theta}, 0 < x < 1,$$

based on an independent sample of size n. Examine whether this estimate is sufficient for θ.

Solution :

$$L(x, \theta) = \prod_{i=1}^{n} f(x_i, \theta) = (1 + \theta)^n \cdot \left(\prod_{i=1}^{n} x_i\right)^{\theta}$$

$$\log L = n \log (1 + \theta) + \theta . \sum_{i=1}^{n} \log x_i$$

$$\frac{\partial}{\partial\theta}\log L = \frac{n}{1+\theta} + \sum_{i=1}^{n} \log x_i = 0$$

$$\Rightarrow \qquad n + \theta \sum_i \log x_i + \sum_i \log x_i = 0$$

$$\therefore \qquad \hat{\theta} = \frac{-n}{\sum_{i=1}^{n} \log x_i} - 1 = \frac{-n}{\log\left(\prod_{i=1}^{n} x_i\right)} - 1 \qquad \ldots(*)$$

Also $$L(x, \theta) = \left\{(1+\theta)^n \cdot \left(\prod_{i=1}^{n} x_i\right)^{\theta-1}\right\} \cdot \left(\prod_{i=1}^{n} x_i\right)$$

Hence by Factorisation theorem, $T = \left(\prod_{i=1}^{n} x_i\right)$ is a sufficient statistic for θ, and $\hat{\theta}$ being a one to one function of sufficient statistic $\left(\prod_{i=1}^{n} x_i\right)$, is also sufficient for θ.