YSU ASDS, Statistics, Fall 2019

Lecture 20

Michael Poghosyan

04 Nov 2019

Contents

- Maximum Likelihood Estimation

Last Lecture ReCap

- Describe the MLE.

Examples

Example: Find the MLE Estimator for θ in the $Unif\,[0, \theta]$ Model.

Solution: OTB

Examples

Example: Find the MLE Estimator for θ in the $Unif[0, \theta]$ Model.

Solution: OTB

Example: Find the MLE Estimator for (μ, σ^2) in the $\mathcal{N}(\mu, \sigma^2)$ Model.

Solution: OTB

Examples

Example: Find the MLE Estimator for θ in the *Unif* $[0, \theta]$ Model.

Solution: OTB

Example: Find the MLE Estimator for (μ, σ^2) in the $\mathcal{N}(\mu, \sigma^2)$ Model.

Solution: OTB

Example: Assume we have an observation

$$0, 1, 1, 2, 1, 0, 0, 1, 1$$

from the following Model:

X	0	1	2
$\mathbb{P}(X = x)$	$\frac{\theta}{10}$	$\frac{\theta}{5}$	$1 - \frac{3\theta}{10}$,

where $\theta \in [0, \frac{10}{3}]$.

Examples

Example: Find the MLE Estimator for θ in the $Unif[0, \theta]$ Model.

Solution: OTB

Example: Find the MLE Estimator for (μ, σ^2) in the $\mathcal{N}(\mu, \sigma^2)$ Model.

Solution: OTB

Example: Assume we have an observation

$$0, 1, 1, 2, 1, 0, 0, 1, 1$$

from the following Model:

X	0	1	2
$\mathbb{P}(X = x)$	$\frac{\theta}{10}$	$\frac{\theta}{5}$	$1 - \frac{3\theta}{10}$,

where $\theta \in [0, \frac{10}{3}]$. Find the MLE Estimator and MLE Estimate for θ.

Solution: OTB

Examples

Example: Find the MLE Estimator for θ in the *Unif* $[0, \theta]$ Model.

Solution: OTB

Example: Find the MLE Estimator for (μ, σ^2) in the $\mathcal{N}(\mu, \sigma^2)$ Model.

Solution: OTB

Example: Assume we have an observation

$$0, 1, 1, 2, 1, 0, 0, 1, 1$$

from the following Model:

X	0	1	2
$\mathbb{P}(X = x)$	$\frac{\theta}{10}$	$\frac{\theta}{5}$	$1 - \frac{3\theta}{10}$,

where $\theta \in [0, \frac{10}{3}]$. Find the MLE Estimator and MLE Estimate for θ.

Solution: OTB

Example: Construct, for the previous problem, the Estimator.

Examples

Example: Find the MLE Estimator for θ in the *Unif* $[0, \theta]$ Model.

Solution: OTB

Example: Find the MLE Estimator for (μ, σ^2) in the $\mathcal{N}(\mu, \sigma^2)$ Model.

Solution: OTB

Example: Assume we have an observation

$$0, 1, 1, 2, 1, 0, 0, 1, 1$$

from the following Model:

X	0	1	2
$\mathbb{P}(X = x)$	$\frac{\theta}{10}$	$\frac{\theta}{5}$	$1 - \frac{3\theta}{10}$,

where $\theta \in [0, \frac{10}{3}]$. Find the MLE Estimator and MLE Estimate for θ.

Solution: OTB

Example: Construct, for the previous problem, the Estimator.

Examples, MLE

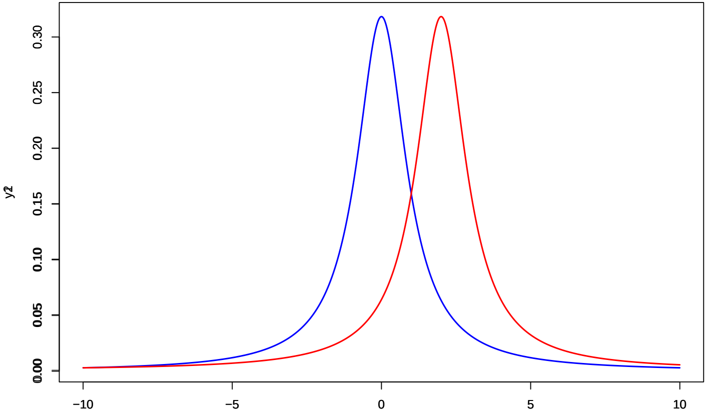
Example: Find the MLE Estimator for θ in the one-Parametric Cauchy Distribution $\textit{Cauchy}(\theta)$ Model. Here, the PDF of $X \sim \textit{Cauchy}(\theta)$ is given by

$$f(x|\theta) = \frac{1}{\pi(1 + (x - \theta)^2)}, \qquad x \in \mathbb{R},$$

and $\theta \in \mathbb{R}$ is called the *location parameter*.

PDF of Cauchy(0) and Cauchy(2)

```
x <- seq(from = -10, to = 10, by = 0.01)
y1 <- dcauchy(x); y2 <- dcauchy(x, location = 2);
plot(x, y1, type = "l", lwd = 2, xlim = c(-10,10), col = "blue")
par(new = T)
plot(x, y2, type = "l", lwd = 2, xlim = c(-10,10), col = "red")
```



Examples, MLE

Example: Find the MLE Estimator for θ in the one-Parametric Cauchy Distribution $Cauchy(\theta)$ Model. Here, the PDF of $X \sim Cauchy(\theta)$ is given by

$$f(x|\theta) = \frac{1}{\pi(1 + (x - \theta)^2)}, \qquad x \in \mathbb{R},$$

and $\theta \in \mathbb{R}$ is called the *location parameter*.

Solution: OTB

Examples, MLE

Example: Find the MLE Estimator for θ in the one-Parametric Cauchy Distribution *Cauchy*(θ) Model. Here, the PDF of $X \sim \mathit{Cauchy}(\theta)$ is given by

$$f(x|\theta) = \frac{1}{\pi(1 + (x-\theta)^2)}, \qquad x \in \mathbb{R},$$

and $\theta \in \mathbb{R}$ is called the *location parameter*.

Solution: OTB

Note: The MoM will not work in this case, since Moments of this Distribution do not exist.

Examples, MLE

Example: Find the MLE Estimator for θ in the one-Parametric Cauchy Distribution *Cauchy*(θ) Model. Here, the PDF of $X \sim$ *Cauchy*(θ) is given by

$$f(x|\theta) = \frac{1}{\pi(1 + (x - \theta)^2)}, \qquad x \in \mathbb{R},$$

and $\theta \in \mathbb{R}$ is called the *location parameter*.

Solution: OTB

Note: The MoM will not work in this case, since Moments of this Distribution do not exist. Btw, the Median of this Distribution is θ.

Examples, MLE

Example: Find the MLE Estimator for θ in the one-Parametric Cauchy Distribution *Cauchy*(θ) Model. Here, the PDF of $X \sim \mathit{Cauchy}(\theta)$ is given by

$$f(x|\theta) = \frac{1}{\pi(1 + (x - \theta)^2)}, \qquad x \in \mathbb{R},$$

and $\theta \in \mathbb{R}$ is called the *location parameter*.

Solution: OTB

Note: The MoM will not work in this case, since Moments of this Distribution do not exist. Btw, the Median of this Distribution is θ.

Note: In **R**, the name of Cauchy Distribution is *cauchy*, so you can use, say, `rcauchy` or `dcauchy` .

Examples, MLE

Example: Find the MLE Estimator for θ in the one-Parametric Cauchy Distribution *Cauchy*(θ) Model. Here, the PDF of $X \sim Cauchy(\theta)$ is given by

$$f(x|\theta) = \frac{1}{\pi(1 + (x - \theta)^2)}, \qquad x \in \mathbb{R},$$

and $\theta \in \mathbb{R}$ is called the *location parameter*.

Solution: OTB

Note: The MoM will not work in this case, since Moments of this Distribution do not exist. Btw, the Median of this Distribution is θ.

Note: In **R**, the name of Cauchy Distribution is *cauchy*, so you can use, say, `rcauchy` or `dcauchy` .

Note: Do you know what is the Distribution of the ratio $\frac{X}{Y}$, when X and Y are Independent Standard Normal Rvs?

Examples, MLE

Example: Find the MLE Estimator for θ in the one-Parametric Cauchy Distribution $Cauchy(\theta)$ Model. Here, the PDF of $X \sim Cauchy(\theta)$ is given by

$$f(x|\theta) = \frac{1}{\pi(1+(x-\theta)^2)}, \qquad x \in \mathbb{R},$$

and $\theta \in \mathbb{R}$ is called the *location parameter*.

Solution: OTB

Note: The MoM will not work in this case, since Moments of this Distribution do not exist. Btw, the Median of this Distribution is θ.

Note: In **R**, the name of Cauchy Distribution is *cauchy*, so you can use, say, `rcauchy` or `dcauchy` .

Note: Do you know what is the Distribution of the ratio $\dfrac{X}{Y}$, when X and Y are Independent Standard Normal Rvs? Cauchy(0) ! $\ddot\smile$

Some Notes about MLE

Note: The MLE does not always exist;

Some Notes about MLE

Note: The MLE does not always exist;

Example: Consider the Model $Unif(a, b)$, i.e., Uniform over the open inteval (a, b).

Some Notes about MLE

Note: The MLE does not always exist;

Example: Consider the Model $Unif(a, b)$, i.e., Uniform over the open inteval (a, b).

Example: Consider the Model with PDF

$$f(x|\theta) = \begin{cases} e^{\theta - x}, & x > \theta \\ 0, & x \leq \theta. \end{cases}$$

Some Notes about MLE

Note: The MLE does not always exist;

Example: Consider the Model $Unif(a, b)$, i.e., Uniform over the open inteval (a, b).

Example: Consider the Model with PDF

$$f(x|\theta) = \begin{cases} e^{\theta - x}, & x > \theta \\ 0, & x \leq \theta. \end{cases}$$

Note: The MLE is not always unique;

Some Notes about MLE

Note: The MLE does not always exist;

Example: Consider the Model $Unif(a, b)$, i.e., Uniform over the open inteval (a, b).

Example: Consider the Model with PDF

$$f(x|\theta) = \begin{cases} e^{\theta - x}, & x > \theta \\ 0, & x \leq \theta. \end{cases}$$

Note: The MLE is not always unique;

Example: Consider the Model $Unif[\theta, \theta + 1]$.

Some Notes about MLE

Note: The MLE does not always exist;

Example: Consider the Model $\mathit{Unif}(a, b)$, i.e., Uniform over the open inteval (a, b).

Example: Consider the Model with PDF

$$f(x|\theta) = \begin{cases} e^{\theta - x}, & x > \theta \\ 0, & x \leq \theta. \end{cases}$$

Note: The MLE is not always unique;

Example: Consider the Model $\mathit{Unif}[\theta, \theta + 1]$.

Note: Sometimes it is necessary to find it numerically, say, using the NR Method

Some Notes about MLE

Note: The MLE does not always exist;

Example: Consider the Model $Unif(a, b)$, i.e., Uniform over the open inteval (a, b).

Example: Consider the Model with PDF

$$f(x|\theta) = \begin{cases} e^{\theta - x}, & x > \theta \\ 0, & x \leq \theta. \end{cases}$$

Note: The MLE is not always unique;

Example: Consider the Model $Unif[\theta, \theta + 1]$.

Note: Sometimes it is necessary to find it numerically, say, using the NR Method

Note: It is remarkable, that ML Estimators, under some conditions (if they exist, of course $\ddot\smile$), possess some nice properties. These properties make MLE one of the widely used methods of Estimation.

Properties of the MLE

It can be proven that, under some regularity conditions on the Parametric Family $\mathcal{F}_\theta$,

- The MLE $\hat{\theta}_n^{MLE}$ is Consistent, i.e.,

Properties of the MLE

It can be proven that, under some regularity conditions on the Parametric Family $\mathcal{F}_\theta$,

- The MLE $\hat{\theta}_n^{MLE}$ is Consistent, i.e.,

$$\hat{\theta}_n^{MLE} \xrightarrow{\mathbb{P}} \theta$$

Properties of the MLE

It can be proven that, under some regularity conditions on the Parametric Family $\mathcal{F}_\theta$,

- The MLE $\hat{\theta}_n^{MLE}$ is Consistent, i.e.,

$$\hat{\theta}_n^{MLE} \xrightarrow{\mathbb{P}} \theta$$

- The MLE $\hat{\theta}_n^{MLE}$ is Asymptotically Normal and Efficient:

$$\frac{\hat{\theta}_n^{MLE} - \theta}{\sqrt{\frac{1}{n \cdot \mathcal{I}(\theta)}}} \xrightarrow{D} \mathcal{N}(0,1)$$

Properties of the MLE

It can be proven that, under some regularity conditions on the Parametric Family $\mathcal{F}_\theta$,

- The MLE $\hat{\theta}_n^{MLE}$ is Consistent, i.e.,

$$\hat{\theta}_n^{MLE} \xrightarrow{\mathbb{P}} \theta$$

- The MLE $\hat{\theta}_n^{MLE}$ is Asymptotically Normal and Efficient:

$$\frac{\hat{\theta}_n^{MLE} - \theta}{\sqrt{\frac{1}{n \cdot \mathcal{I}(\theta)}}} \xrightarrow{D} \mathcal{N}\left(0, 1\right)$$

  or, as people usually use:

$$\sqrt{n} \cdot \left(\hat{\theta}_n^{MLE} - \theta\right) \xrightarrow{D} \mathcal{N}\left(0, \frac{1}{\mathcal{I}(\theta)}\right),$$

Properties of the MLE

It can be proven that, under some regularity conditions on the Parametric Family $\mathcal{F}_\theta$,

- The MLE $\hat{\theta}_n^{MLE}$ is Consistent, i.e.,

$$\hat{\theta}_n^{MLE} \xrightarrow{\mathbb{P}} \theta$$

- The MLE $\hat{\theta}_n^{MLE}$ is Asymptotically Normal and Efficient:

$$\frac{\hat{\theta}_n^{MLE} - \theta}{\sqrt{\frac{1}{n \cdot \mathcal{I}(\theta)}}} \xrightarrow{D} \mathcal{N}(0, 1)$$

  or, as people usually use:

$$\sqrt{n} \cdot \left(\hat{\theta}_n^{MLE} - \theta\right) \xrightarrow{D} \mathcal{N}\left(0, \frac{1}{\mathcal{I}(\theta)}\right),$$

  or, put in other way,

$$\hat{\theta}_n^{MLE} \overset{D}{\approx} \mathcal{N}\left(\theta, \frac{1}{n \cdot \mathcal{I}(\theta)}\right)$$

Properties of the MLE, Cont'd

Let us give an interpretation to this last property:

$$\hat{\theta}_n^{MLE} \overset{D}{\approx} \mathcal{N}\left(\theta, \frac{1}{n \cdot \mathcal{I}(\theta)}\right)$$

Properties of the MLE, Cont'd

Let us give an interpretation to this last property:

$$\hat{\theta}_n^{MLE} \overset{D}{\approx} \mathcal{N}\left(\theta, \frac{1}{n \cdot \mathcal{I}(\theta)}\right)$$

We can interpret this as:

- $\hat{\theta}_n^{MLE}$ is approximately (Asymptotically) Normal;

Properties of the MLE, Cont'd

Let us give an interpretation to this last property:

$$\hat{\theta}_n^{MLE} \stackrel{D}{\approx} \mathcal{N}\left(\theta, \frac{1}{n \cdot \mathcal{I}(\theta)}\right)$$

We can interpret this as:

- $\hat{\theta}_n^{MLE}$ is approximately (Asymptotically) Normal;
- The Mean of $\hat{\theta}_n^{MLE}$ is approximately θ, so it is Asymptotically Unbiased;

Properties of the MLE, Cont'd

Let us give an interpretation to this last property:

$$\hat{\theta}_n^{MLE} \overset{D}{\approx} \mathcal{N}\left(\theta, \frac{1}{n \cdot \mathcal{I}(\theta)}\right)$$

We can interpret this as:

- $\hat{\theta}_n^{MLE}$ is approximately (Asymptotically) Normal;
- The Mean of $\hat{\theta}_n^{MLE}$ is approximately θ, so it is Asymptotically Unbiased;
- The Variance of $\hat{\theta}_n^{MLE}$ is approximately $\dfrac{1}{n \cdot \mathcal{I}(\theta)}$. And we know that, by CR LB, this is the possible smallest Variance we can have with Unbiased Estimators.

Properties of the MLE, Cont'd

Let us give an interpretation to this last property:

$$\hat{\theta}_n^{MLE} \overset{D}{\approx} \mathcal{N}\left(\theta, \frac{1}{n \cdot \mathcal{I}(\theta)}\right)$$

We can interpret this as:

- $\hat{\theta}_n^{MLE}$ is approximately (Asymptotically) Normal;
- The Mean of $\hat{\theta}_n^{MLE}$ is approximately θ, so it is Asymptotically Unbiased;
- The Variance of $\hat{\theta}_n^{MLE}$ is approximately $\dfrac{1}{n \cdot \mathcal{I}(\theta)}$. And we know that, by CR LB, this is the possible smallest Variance we can have with Unbiased Estimators. So we interpret this as **MLE is Asymptotically Efficient**.

Properties of the MLE, Cont'd

Let us give an interpretation to this last property:

$$\hat{\theta}_n^{MLE} \overset{D}{\approx} \mathcal{N}\left(\theta, \frac{1}{n \cdot \mathcal{I}(\theta)}\right)$$

We can interpret this as:

- $\hat{\theta}_n^{MLE}$ is approximately (Asymptotically) Normal;
- The Mean of $\hat{\theta}_n^{MLE}$ is approximately θ, so it is Asymptotically Unbiased;
- The Variance of $\hat{\theta}_n^{MLE}$ is approximately $\dfrac{1}{n \cdot \mathcal{I}(\theta)}$. And we know that, by CR LB, this is the possible smallest Variance we can have with Unbiased Estimators. So we interpret this as **MLE is Asymptotically Efficient**.

So, MLE is **Consistent** and **Asymptotocally Efficient**.

Properties of the MLE, Cont'd

Let us give an interpretation to this last property:

$$\hat{\theta}_n^{MLE} \overset{D}{\approx} \mathcal{N}\left(\theta, \frac{1}{n \cdot \mathcal{I}(\theta)}\right)$$

We can interpret this as:

- $\hat{\theta}_n^{MLE}$ is approximately (Asymptotically) Normal;
- The Mean of $\hat{\theta}_n^{MLE}$ is approximately θ, so it is Asymptotically Unbiased;
- The Variance of $\hat{\theta}_n^{MLE}$ is approximately $\dfrac{1}{n \cdot \mathcal{I}(\theta)}$. And we know that, by CR LB, this is the possible smallest Variance we can have with Unbiased Estimators. So we interpret this as **MLE is Asymptotically Efficient**.

So, MLE is **Consistent** and **Asymptotocally Efficient**. And this is why, for large Sample Size n, MLE is the Top 1 Choice, is (almost) unbeatable.

Properties of the MLE, Cont'd

- Also,

$$\frac{\hat{\theta}_n^{MLE} - \theta}{\sqrt{\frac{1}{n \cdot \mathcal{I}\left(\hat{\theta}_n^{MLE}\right)}}} \xrightarrow{D} \mathcal{N}\left(0, 1\right)$$

Properties of the MLE, Cont'd

- Also,

$$\frac{\hat{\theta}_n^{MLE} - \theta}{\sqrt{\frac{1}{n \cdot \mathcal{I}\left(\hat{\theta}_n^{MLE}\right)}}} \xrightarrow{D} \mathcal{N}(0, 1)$$

or, in more non-rigorous terms,

$$\hat{\theta}_n^{MLE} \overset{D}{\approx} \mathcal{N}\left(\theta, \frac{1}{n \cdot \mathcal{I}\left(\hat{\theta}_n^{MLE}\right)}\right)$$

Properties of the MLE, Cont'd

- Also,

$$\frac{\hat{\theta}_n^{MLE} - \theta}{\sqrt{\frac{1}{n \cdot \mathcal{I}\left(\hat{\theta}_n^{MLE}\right)}}} \xrightarrow{D} \mathcal{N}(0,1)$$

  or, in more non-rigorous terms,

$$\hat{\theta}_n^{MLE} \overset{D}{\approx} \mathcal{N}\left(\theta, \frac{1}{n \cdot \mathcal{I}\left(\hat{\theta}_n^{MLE}\right)}\right)$$

Note: This is almost the above Property,

$$\hat{\theta}_n^{MLE} \overset{D}{\approx} \mathcal{N}\left(\theta, \frac{1}{n \cdot \mathcal{I}(\theta)}\right),$$

but, instead of $\mathcal{I}(\theta)$ we have $\mathcal{I}\left(\hat{\theta}_n^{MLE}\right)$.

Properties of the MLE, Cont'd

- Also,

$$\frac{\hat{\theta}_n^{MLE} - \theta}{\sqrt{\frac{1}{n \cdot \mathcal{I}\left(\hat{\theta}_n^{MLE}\right)}}} \xrightarrow{D} \mathcal{N}(0, 1)$$

  or, in more non-rigorous terms,

$$\hat{\theta}_n^{MLE} \overset{D}{\approx} \mathcal{N}\left(\theta, \frac{1}{n \cdot \mathcal{I}\left(\hat{\theta}_n^{MLE}\right)}\right)$$

Note: This is almost the above Property,

$$\hat{\theta}_n^{MLE} \overset{D}{\approx} \mathcal{N}\left(\theta, \frac{1}{n \cdot \mathcal{I}(\theta)}\right),$$

but, instead of $\mathcal{I}(\theta)$ we have $\mathcal{I}\left(\hat{\theta}_n^{MLE}\right)$.

Note: We will use this later, to construct an (approximate) Confidence Interval for θ and for testing Hypotheses about θ.