YSU ASDS, Statistics, Fall 2019

Lecture 18

Michael Poghosyan

23 Oct 2019

Contents

Last Lecture ReCap

- Give the definition of Consistency.

Last Lecture ReCap

- Give the definition of Consistency.
- Give the definition of the Fisher Information.

Fisher Information in the Multidimensional case

Now assume that the parameter θ is d-dimensional. Then the Fisher Information Matrix is defined as

$$I(\theta) = \mathbb{E}\left[\Big(\nabla_\theta \ln f(X|\theta)\Big) \cdot \Big(\nabla_\theta \ln f(X|\theta)\Big)^T\right],$$

where $\nabla_\theta g(\theta)$ denotes the Gradient of $g(\theta)$ w.r.t θ.

Cramer-Rao Inequality, C-R Lower Bound

CR LB is a remarkable inequality, giving a lower bound for the Variance of an Unbiased Estimator.

Cramer-Rao Inequality, C-R Lower Bound

CR LB is a remarkable inequality, giving a lower bound for the Variance of an Unbiased Estimator. Under some regularity conditions on the family of Distributions $\mathcal{F}_\theta$, the following holds:

Cramer-Rao Inequality, C-R Lower Bound

CR LB is a remarkable inequality, giving a lower bound for the Variance of an Unbiased Estimator. Under some regularity conditions on the family of Distributions $\mathcal{F}_\theta$, the following holds:

Theorem (Cramer-Rao, Unbiased Case): Assume we have a Random Sample

$$X_1, X_2, ..., X_n \overset{IID}{\sim} \mathcal{F}_\theta$$

and the Fisher Information for the family $\mathcal{F}_\theta$ is $I(\theta)$. Assume also that $\hat{\theta}$ is an unbiased estimator for θ obtained from our Random Sample. Then, under the above mentioned regularity conditions,

$$Var(\hat{\theta}) \geq \frac{1}{n \cdot I(\theta)}.$$

Cramer-Rao Inequality, C-R Lower Bound, Biased Case

There is a version of C-R Inequality for the general (not necessarily UnBiased) case.

Cramer-Rao Inequality, C-R Lower Bound, Biased Case

There is a version of C-R Inequality for the general (not necessarily UnBiased) case.

Theorem (Cramer-Rao, General Case): Assume we have a Random Sample

$$X_1, X_2, ..., X_n \overset{IID}{\sim} \mathcal{F}_\theta,$$

and we are using an Estimator $\hat{\theta}$ with the Expectation $k(\theta) = \mathbb{E}(\hat{\theta})$. Then

$$Var(\hat{\theta}) \geq \frac{[k'(\theta)]^2}{n \cdot I(\theta)}.$$

Cramer-Rao Inequality, C-R Lower Bound, Biased Case

There is a version of C-R Inequality for the general (not necessarily UnBiased) case.

Theorem (Cramer-Rao, General Case): Assume we have a Random Sample

$$X_1, X_2, ..., X_n \overset{IID}{\sim} \mathcal{F}_\theta,$$

and we are using an Estimator $\hat{\theta}$ with the Expectation $k(\theta) = \mathbb{E}(\hat{\theta})$. Then

$$Var(\hat{\theta}) \geq \frac{[k'(\theta)]^2}{n \cdot I(\theta)}.$$

In particular, if $\hat{\theta}$ is unbiased, then $k(\theta) = \theta$, so we will obtain the previous C-R Inequality.

Consequences of the C-R LB

Recall that for an Unbiased Estimator $\hat{\theta}$,

$$MSE(\hat{\theta}, \theta) = Var(\hat{\theta}).$$

Consequences of the C-R LB

Recall that for an Unbiased Estimator $\hat{\theta}$,

$$MSE(\hat{\theta}, \theta) = Var(\hat{\theta}).$$

So the C-R LB gives us

$$MSE(\hat{\theta}, \theta) \geq \frac{1}{n \cdot I(\theta)}.$$

Consequences of the C-R LB

Recall that for an Unbiased Estimator $\hat{\theta}$,

$$MSE(\hat{\theta}, \theta) = Var(\hat{\theta}).$$

So the C-R LB gives us

$$MSE(\hat{\theta}, \theta) \geq \frac{1}{n \cdot I(\theta)}.$$

And this is a fundamental restriction on the MSE: *you cannot do better when estimating θ than the Estimator with*

$$MSE(\hat{\theta}, \theta) = \frac{1}{n \cdot I(\theta)},$$

you cannot obtain smaller MSE using Unbiased Estimators!

Consequences of the C-R LB

Recall that for an Unbiased Estimator $\hat{\theta}$,

$$MSE(\hat{\theta}, \theta) = Var(\hat{\theta}).$$

So the C-R LB gives us

$$MSE(\hat{\theta}, \theta) \geq \frac{1}{n \cdot I(\theta)}.$$

And this is a fundamental restriction on the MSE: *you cannot do better when estimating θ than the Estimator with*

$$MSE(\hat{\theta}, \theta) = \frac{1}{n \cdot I(\theta)},$$

you cannot obtain smaller MSE using Unbiased Estimators!

And if there exists an Unbiased Estimator $\hat{\theta}$ with

$$MSE(\hat{\theta}, \theta) = \frac{1}{n \cdot I(\theta)},$$

we call $\hat{\theta}$ an **Efficient Estimator** for θ, and that Estimator is a MVUE for θ.

Notes

Note: Not always there exists an Unbiased Estimator with

$$MSE(\hat{\theta}, \theta) = Var(\hat{\theta}) = \frac{1}{n \cdot I(\theta)}.$$

Notes

Note: Not always there exists an Unbiased Estimator with

$$MSE(\hat{\theta}, \theta) = Var(\hat{\theta}) = \frac{1}{n \cdot I(\theta)}.$$

But even in that cases we can have MVUE (and, in that case, the minimum of the Variance will be $> \frac{1}{n \cdot I(\theta)}$).

Notes

Note: Not always there exists an Unbiased Estimator with

$$MSE(\hat{\theta}, \theta) = Var(\hat{\theta}) = \frac{1}{n \cdot I(\theta)}.$$

But even in that cases we can have MVUE (and, in that case, the minimum of the Variance will be $> \dfrac{1}{n \cdot I(\theta)}$).

Note: Sometimes, in different Textbooks, an Unbiased Estimator with Minimum Variance (not necessarily with $Var(\hat{\theta}) = \dfrac{1}{n \cdot I(\theta)}$) is called an **Efficient Estimator** for θ.

Example

Example: Show that in the Bernoulli Model, with a Random Sample

$$X_1, X_2, ..., X_n \sim Bernoulli(p), \qquad p \in [0, 1],$$

the Estimator

$$\hat{p} = \overline{X}$$

is the MVUE of p.

Example

Example: Show that in the Bernoulli Model, with a Random Sample

$$X_1, X_2, ..., X_n \sim Bernoulli(p), \qquad p \in [0, 1],$$

the Estimator

$$\hat{p} = \overline{X}$$

is the MVUE of p.

Example: Show that in the Poisson Model, with a Random Sample

$$X_1, X_2, ..., X_n \sim Pois(\lambda), \qquad \lambda > 0,$$

the Estimator

$$\hat{\lambda} = \overline{X}$$

is the MVUE of λ.

Methods to find (good) Estimators

The Problem

Here, in this Section, our Problem will be the following:

The Problem

Here, in this Section, our Problem will be the following:

We are given a Random Sample from a Parametric Family of Distributions,

$$X_1, X_2, ..., X_n \sim \mathcal{F}_\theta, \qquad \theta \in \Theta.$$

The Problem

Here, in this Section, our Problem will be the following:

We are given a Random Sample from a Parametric Family of Distributions,

$$X_1, X_2, ..., X_n \sim \mathcal{F}_\theta, \qquad \theta \in \Theta.$$

Problem: The Problem is to find/construct a good Estimator for θ, using our Random Sample.

The Method of Moments

Theoretical Moments

First we define the Theoretical and Sample (Empirical) Moments:

Theoretical Moments

First we define the Theoretical and Sample (Empirical) Moments:

Assume we consider the Parametric Family of Distributions $\mathcal{F}_\theta$, $\theta \in \Theta$.

Theoretical Moments

First we define the Theoretical and Sample (Empirical) Moments:

Assume we consider the Parametric Family of Distributions $\mathcal{F}_\theta$, $\theta \in \Theta$. Let $X \sim \mathcal{F}_\theta$.

Theoretical Moments

First we define the Theoretical and Sample (Empirical) Moments:

Assume we consider the Parametric Family of Distributions $\mathcal{F}_\theta$, $\theta \in \Theta$. Let $X \sim \mathcal{F}_\theta$.

Definition: The **Theoretical Moment** of order $k \in \mathbb{N}$ of the Distribution $\mathcal{F}_\theta$ (or the r.v. X) is $\mathbb{E}(X^k)$.

Theoretical Moments

First we define the Theoretical and Sample (Empirical) Moments:

Assume we consider the Parametric Family of Distributions $\mathcal{F}_\theta$, $\theta \in \Theta$. Let $X \sim \mathcal{F}_\theta$.

Definition: The **Theoretical Moment** of order $k \in \mathbb{N}$ of the Distribution $\mathcal{F}_\theta$ (or the r.v. X) is $\mathbb{E}(X^k)$.

Example: Say, the first Moment of $\mathcal{F}_\theta$ is just the Mean of X, $\mathbb{E}(X)$.

Theoretical Moments

First we define the Theoretical and Sample (Empirical) Moments:

Assume we consider the Parametric Family of Distributions $\mathcal{F}_\theta$, $\theta \in \Theta$. Let $X \sim \mathcal{F}_\theta$.

Definition: The **Theoretical Moment** of order $k \in \mathbb{N}$ of the Distribution $\mathcal{F}_\theta$ (or the r.v. X) is $\mathbb{E}(X^k)$.

Example: Say, the first Moment of $\mathcal{F}_\theta$ is just the Mean of X, $\mathbb{E}(X)$. The second Moment of $\mathcal{F}_\theta$ is $\mathbb{E}(X^2)$ etc.

Theoretical Moments

First we define the Theoretical and Sample (Empirical) Moments:

Assume we consider the Parametric Family of Distributions $\mathcal{F}_\theta$, $\theta \in \Theta$. Let $X \sim \mathcal{F}_\theta$.

Definition: The **Theoretical Moment** of order $k \in \mathbb{N}$ of the Distribution $\mathcal{F}_\theta$ (or the r.v. X) is $\mathbb{E}(X^k)$.

Example: Say, the first Moment of $\mathcal{F}_\theta$ is just the Mean of X, $\mathbb{E}(X)$. The second Moment of $\mathcal{F}_\theta$ is $\mathbb{E}(X^2)$ etc.

Note: Note that, in general, the Theoretical Moments of $\mathcal{F}_\theta$ are functions of θ.

Empirical (Sample) Moments

Now assume we have a Random Sample from $\mathcal{F}_\theta$:

$$X_1, X_2, ..., X_n \sim \mathcal{F}_\theta, \qquad \theta \in \Theta.$$

Empirical (Sample) Moments

Now assume we have a Random Sample from $\mathcal{F}_\theta$:

$$X_1, X_2, ..., X_n \sim \mathcal{F}_\theta, \qquad \theta \in \Theta.$$

Definition: The **Empirical (Sample) Moment** of order $k \in \mathbb{N}$ of the Distribution $\mathcal{F}_\theta$ is

$$\frac{(X_1)^k + (X_2)^k + ... + (X_n)^k}{n}.$$

Empirical (Sample) Moments

Now assume we have a Random Sample from $\mathcal{F}_\theta$:

$$X_1, X_2, ..., X_n \sim \mathcal{F}_\theta, \qquad \theta \in \Theta.$$

Definition: The **Empirical (Sample) Moment** of order $k \in \mathbb{N}$ of the Distribution $\mathcal{F}_\theta$ is

$$\frac{(X_1)^k + (X_2)^k + ... + (X_n)^k}{n}.$$

Example: Say, the second order Empirical Moment is

$$\frac{(X_1)^2 + (X_2)^2 + ... + (X_n)^2}{n}.$$

Empirical (Sample) Moments

Now assume we have a Random Sample from $\mathcal{F}_\theta$:

$$X_1, X_2, ..., X_n \sim \mathcal{F}_\theta, \qquad \theta \in \Theta.$$

Definition: The **Empirical (Sample) Moment** of order $k \in \mathbb{N}$ of the Distribution $\mathcal{F}_\theta$ is

$$\frac{(X_1)^k + (X_2)^k + ... + (X_n)^k}{n}.$$

Example: Say, the second order Empirical Moment is

$$\frac{(X_1)^2 + (X_2)^2 + ... + (X_n)^2}{n}.$$

Note: The Empirical Moment is independent of the Parameter θ, it is just a Statistics, it is a function of $X_1, X_2, ..., X_n$.

The Method of Moments, 1D case

Now, the idea of the Method of Moments (MoM) is to match the corresponding Theoretical and Empirical Moments.

The Method of Moments, 1D case

Now, the idea of the Method of Moments (MoM) is to match the corresponding Theoretical and Empirical Moments.

First we consider the case when our Parameter θ is 1D, i.e., $\theta \in \Theta \subset \mathbb{R}$.

The Method of Moments, 1D case

Now, the idea of the Method of Moments (MoM) is to match the corresponding Theoretical and Empirical Moments.

First we consider the case when our Parameter θ is 1D, i.e., $\theta \in \Theta \subset \mathbb{R}$. In that case, we need just one equation to find θ.

The Method of Moments, 1D case

Now, the idea of the Method of Moments (MoM) is to match the corresponding Theoretical and Empirical Moments.

First we consider the case when our Parameter θ is 1D, i.e., $\theta \in \Theta \subset \mathbb{R}$. In that case, we need just one equation to find θ.

MoM suggests: Solve the Equation

1-st order Theoretical Moment = 1-st order Empirical Moment,

to find (express) θ.

The Method of Moments, 1D case

Now, the idea of the Method of Moments (MoM) is to match the corresponding Theoretical and Empirical Moments.

First we consider the case when our Parameter θ is 1D, i.e., $\theta \in \Theta \subset \mathbb{R}$. In that case, we need just one equation to find θ.

MoM suggests: Solve the Equation

1-st order Theoretical Moment = 1-st order Empirical Moment,

to find (express) θ. The value of θ, expressed by the Random Sample, is called **the MoM Estimator**.

The Method of Moments, 1D case

Now, the idea of the Method of Moments (MoM) is to match the corresponding Theoretical and Empirical Moments.

First we consider the case when our Parameter θ is 1D, i.e., $\theta \in \Theta \subset \mathbb{R}$. In that case, we need just one equation to find θ.

MoM suggests: Solve the Equation

1-st order Theoretical Moment = 1-st order Empirical Moment,

to find (express) θ. The value of θ, expressed by the Random Sample, is called **the MoM Estimator**. And if this will not give a value for θ, solve instead

2-nd order Theoretical Moment = 2-nd order Empirical Moment.

The Method of Moments, 1D case

Now, the idea of the Method of Moments (MoM) is to match the corresponding Theoretical and Empirical Moments.

First we consider the case when our Parameter θ is 1D, i.e., $\theta \in \Theta \subset \mathbb{R}$. In that case, we need just one equation to find θ.

MoM suggests: Solve the Equation

1-st order Theoretical Moment = 1-st order Empirical Moment,

to find (express) θ. The value of θ, expressed by the Random Sample, is called **the MoM Estimator**. And if this will not give a value for θ, solve instead

2-nd order Theoretical Moment = 2-nd order Empirical Moment.

If this is giving the value for θ, then you have the MoM Estimator.

The Method of Moments, 1D case

Now, the idea of the Method of Moments (MoM) is to match the corresponding Theoretical and Empirical Moments.

First we consider the case when our Parameter θ is 1D, i.e., $\theta \in \Theta \subset \mathbb{R}$. In that case, we need just one equation to find θ.

MoM suggests: Solve the Equation

1-st order Theoretical Moment = 1-st order Empirical Moment,

to find (express) θ. The value of θ, expressed by the Random Sample, is called **the MoM Estimator**. And if this will not give a value for θ, solve instead

2-nd order Theoretical Moment = 2-nd order Empirical Moment.

If this is giving the value for θ, then you have the MoM Estimator. Otherwise, continue with the 3-rd Moments, then with 4-th Moments, etc.

Examples, MoM

Example: Find the MoM Estimator for p in the *Bernoulli*(p) Model.

Solution: OTB

Examples, MoM

Example: Find the MoM Estimator for p in the *Bernoulli*(p) 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, MoM

Example: Find the MoM Estimator for p in the *Bernoulli*(p) 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 MoM Estimator and MoM Estimate for θ.

Solution: OTB

Examples, MoM

Example: Find the MoM Estimator for p in the *Bernoulli*(p) 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 MoM Estimator and MoM Estimate for θ.

Solution: OTB

Example: Find the MoM Estimator for λ in the *Exp*(λ) Model.

Solution: OTB

Examples, MoM

Example: Find the MoM Estimator for θ in the $\mathit{Unif}\,[0, \theta]$ Model.

Solution: OTB

Examples, MoM

Example: Find the MoM Estimator for θ in the $\mathit{Unif}\,[0, \theta]$ Model.

Solution: OTB

Example: Find the MoM Estimator for θ in the $\mathit{Unif}\,[-\theta, \theta]$ Model.

Solution: OTB