AUA CS 108, Statistics, Fall 2019

Lecture 25

Michael Poghosyan
YSU, AUA
michael@ysu.am, mpoghosyan@aua.am

21 Oct 2019

Contents

Last Lecture ReCap

- Give the definition of a Consistent Estimator

Last Lecture ReCap

- Give the definition of a Consistent Estimator
- In which case and in which way we can use Unbiased Estimator?

Last Lecture ReCap

- Give the definition of a Consistent Estimator
- In which case and in which way we can use Unbiased Estimator?
- In which case and in which way we can use Consistent Estimator?

Fisher Information, Refresher

Assume we have a parametric family of distributions $\mathcal{F}_\theta$, $\theta \in \Theta$, and $f(x|\theta)$ is the PD(M)F of $\mathcal{F}_\theta$.

Fisher Information, Refresher

Assume we have a parametric family of distributions $\mathcal{F}_\theta$, $\theta \in \Theta$, and $f(x|\theta)$ is the PD(M)F of $\mathcal{F}_\theta$.

Definition: The following quantity is called **the Fisher Information** of the parametric family $\mathcal{F}_\theta$:

$$I(\theta) = -\mathbb{E}\left(\frac{\partial^2}{\partial\theta^2} \ln f(X|\theta)\right) = \mathbb{E}\left[\left(\frac{\partial}{\partial\theta} \ln f(X|\theta)\right)^2\right],$$

where $X \sim \mathcal{F}_\theta$.

Example

Example: Calculate the Fisher Information for the *Bernoulli*(p) family

Solution: OTB

Example

Example: Calculate the Fisher Information for the *Bernoulli*(p) family

Solution: OTB

Example: Calculate the Fisher Information for the $\mathcal{N}(\mu, \sigma^2)$ family (separately for the Parameter μ and σ^2)

Solution: OTB

Cramer-Rao Inequality, C-R Lower Bound

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

Cramer-Rao Inequality, C-R Lower Bound

This 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

This 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): 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

This 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): 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,

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

Note: There is a version of C-R Inequality for the Biased case.

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 $> \dfrac{1}{n \cdot I(\theta)}$).

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.