- Fisher Information
- Cramer-Rao Lower Bound (Cramer-Rao Inequality)
- MVUE
- The Method of Moments

Example

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

Solution: OTB

Example: Calculate the Fisher Information for the *Exp*(λ) 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

Fisher Information, cont'd

Another interpretation of the Fisher Information is the following.

It is easy to see that (under the regularity conditions)

$$\mathbb{E}\left(\frac{\partial}{\partial\theta}\ln f(X|\theta)\right) = 0.$$

Then,

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

So the Fisher Information is the Variance of the Score function

$$\frac{\partial}{\partial\theta}\ln f(X|\theta).$$

Fisher Information in the Multidimensional case

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

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

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. 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.

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 $\hat{\theta}$,

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

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.