STAT 713 sp 2023 Lec 07 slides

Fisher information, Cramér-Rao lower bound

Karl B. Gregory

University of South Carolina

These slides are an instructional aid; their sole purpose is to display, during the lecture, definitions, plots, results, etc. which take too much time to write by hand on the blackboard. They are not intended to explain or expound on any material.

Table of Contents

Throughout let $\mathbf{X} = (X_1, \ldots, X_n)$ and $\mathbf{x} = (x_1, \ldots, x_n)$.

Another framework for evaluating estimators: Cramér-Rao.

Cramér-Rao condition

Suppose $\mathbf{X} \sim f(\mathbf{x}; \theta)$, where $f(\mathbf{x}; \theta)$ is differentiable in θ. In addition, suppose the true parameter value lies in an open subset of Θ. We will assume

$$\frac{\partial}{\partial \theta} \mathbb{E}_\theta h(\mathbf{X}) = \mathbb{E}_\theta \left(h(\mathbf{X}) \frac{\partial}{\partial \theta} \log f(\mathbf{X}; \theta) \right) \quad \text{whenever } \mathbb{E}_\theta |h(\mathbf{X})| < \infty$$

for all θ. We will call this the *Cramér-Rao (CR)* condition.

For now assume $\Theta \subset \mathbb{R}$, so that θ is a scalar.

Under this condition, we can find a lower bound for the variance of any estimator.

The CR condition requires that the support of $f(\mathbf{x}; \theta)$ not depend on θ.

Score function, Fisher information

Under the CR condition, define the

1. *score function* as $S(\theta; \mathbf{X}) = \frac{\partial}{\partial\theta} \log f(\mathbf{x}; \theta)$.
2. *Fisher information* as $I(\theta) = \mathbb{E}_\theta[S^2(\theta; \mathbf{X})]$.

Theorem (Properties of the score and Fisher information)

Under the CR condition we have

1. $\mathbb{E}_\theta S(\theta; \mathbf{X}) = 0$ *for all* θ.
2. $\mathbb{E}_\theta |h(\mathbf{X})| < \infty$ *for all* $\theta \implies \frac{\partial}{\partial\theta}\mathbb{E}_\theta h(\mathbf{X}) = \text{Cov}_\theta(h(\mathbf{X}), S(\theta; \mathbf{X}))$ *for all* θ.
3. $I(\theta) = \text{Var}_\theta(S(\theta; \mathbf{X}))$.
4. *If* $\frac{\partial^2}{\partial\theta^2} f(\mathbf{x}; \theta)$ *exists and* $\frac{\partial}{\partial\theta}\mathbb{E}_\theta S(\theta; \mathbf{X}) = \int_{\mathcal{X}} \frac{\partial}{\partial\theta}[S(\theta; \mathbf{x}) f(\mathbf{x}; \theta)] d\mathbf{x}$, *then*

$$I(\theta) = -\mathbb{E}_\theta \left(\frac{\partial^2}{\partial\theta^2} \log f(\mathbf{X}; \theta) \right).$$

Exercise: Prove the above.

Theorem (The Cramér-Rao lower bound)

Let $\tilde{\tau}(\mathbf{X})$ be an estimator with finite mean and suppose $\mathbb{E}_\theta \tilde{\tau}(\mathbf{X})$ is differentiable wrt θ. Then under the CR condition,

$$\operatorname{Var} \tilde{\tau}(\mathbf{X}) \geq \frac{[\frac{\partial}{\partial \theta} \mathbb{E}_\theta \tilde{\tau}(\mathbf{X})]^2}{I(\theta)}.$$

The right hand side is called the *Cramér-Rao Lower Bound (CRLB)*.

No estimator can have a variance smaller than the CRLB.

Exercise: Prove the above using the Cauchy-Schwarz inequality.

Corollary (CRLB for unbiased estimator with iid sample)

Let $\mathbf{X}$ *be an iid sample and let* $\tilde{\tau}(\mathbf{X})$ *be an unbiased estimator of* $\tau = \tau(\theta)$, *where* τ *is differentiable with derivative* τ'. *Then under the CR condition,*

$$\operatorname{Var} \tilde{\tau}(\mathbf{X}) \geq \frac{[\tau'(\theta)]^2}{n \operatorname{Var}_\theta(\frac{\partial}{\partial\theta} \log f(X_1;\theta))}.$$

No unbiased estimator of τ can have a variance smaller than the right hand side.

Exercise: Show above.

Exercise: Let $\mathbf{X} \overset{\text{ind}}{\sim}$ Geometric(θ), $\theta \in (0, 1)$.

1. Find the Fisher Information $I(\theta)$ as $\mathbb{E}[S^2(\theta; \mathbf{X})]$.
2. Verify that $I(\theta) = -\mathbb{E}[\frac{\partial^2}{\partial\theta^2} \log f(\mathbf{X}; \theta)]$.
3. Give the CRLB for unbiased estimators of θ.
4. What can we say about the variance of the UMVUE of θ?
5. Give the CRLB for unbiased estimators of $1/\theta$.
6. Give the UMVUE for $1/\theta$.

Exercise: Let $\mathbf{X}$ be iid with cdf $F(x;\mu) = 1/(1+e^{-(x-\mu)})$, for $\mu \in \mathbb{R}$.

1. Find the Fisher Information $I(\mu)$.
2. Give the CRLB for unbiased estimators of μ.
3. Check whether $\bar{X}_n$ achieves the CRLB.

Exercise: Let $\mathbf{X} \overset{\text{ind}}{\sim} \text{Uniform}(0, \theta)$, for $\theta > 0$.

1. Naïvely find $S(\theta; \mathbf{X})$, ignoring the fact that the support depends on θ.
2. Obtain the corresponding Fisher information.
3. Discuss applicability of the CRLB to this setting.

Theorem (Attainment of CRLB)

An unbiased estimator $\tilde{\tau}(\mathbf{X})$ of $\tau(\theta)$ achieves its CRLB iff

$$S(\theta; \mathbf{X}) = a(\theta)[\tilde{\tau}(\mathbf{X}) - \tau(\theta)]$$

for some $a(\theta)$, i.e. iff $\tilde{\tau}(\mathbf{X})$ is perfectly correlated with $S(\theta; \mathbf{X})$.

Exercise: Prove the result.

Exercise: Let $\mathbf{X} \overset{\text{ind}}{\sim} \text{Poisson}(\lambda)$, $\lambda > 0$.

1. Find the CRLB for unbiased estimators of λ.
2. Verify that $\bar{X}_n$ is unbiased and is perfectly correlated with the score function.

Table of Contents

When $\theta = (\theta_1, \ldots, \theta_d)^T$, let

$$\frac{\partial}{\partial\theta} a(\theta) = \left[\frac{\partial}{\partial\theta_1} a(\theta), \ldots, \frac{\partial}{\partial\theta_d} a(\theta)\right]^T \quad \text{and} \quad \frac{\partial^2}{\partial\theta\partial\theta^T} = \left[\frac{\partial^2}{\partial\theta_i\partial\theta_j} a(\theta)\right]_{1\leq i\leq j\leq d}.$$

In this case the Fisher information is given by the matrix

$$I(\theta) = \mathbb{E}_\theta[S(\theta;\mathbf{X})S(\theta;\mathbf{X})^T] = -\mathbb{E}_\theta\left(\frac{\partial^2}{\partial\theta\partial\theta^T}\log f(\mathbf{X};\theta)\right),$$

where 2nd eq. holds if $\frac{\partial^2}{\partial\theta\partial\theta^T} f(\mathbf{x};\theta)$ exists and $\frac{\partial}{\partial\theta}\mathbb{E}_\theta S(\theta;\mathbf{X})^T = \mathbb{E}_\theta \frac{\partial}{\partial\theta} S(\theta;\mathbf{X})^T$.

Theorem (Cramér-Rao lower bound when parameter multidimensional)

Let $\tilde{\tau}(\mathbf{X})$ be an estimator of (scalar) $\tau = \tau(\theta)$ with finite mean and suppose $\mathbb{E}_\theta\tilde{\tau}(\mathbf{X})$ is differentiable wrt θ. Then under the CR condition,

$$\operatorname{Var}\tilde{\tau}(\mathbf{X}) \geq [\tfrac{\partial}{\partial\theta}\mathbb{E}_\theta\tilde{\tau}(\mathbf{X})]^T[I(\theta)]^{-1}[\tfrac{\partial}{\partial\theta}\mathbb{E}_\theta\tilde{\tau}(\mathbf{X})].$$

Exercise: Let $\mathbf{X} \overset{\text{ind}}{\sim} \text{Normal}(\mu, \gamma)$, where $\gamma > 0$ is the variance.

1. Find the Fisher information $I(\mu, \gamma)$.
2. Find the CRLB for unbiased estimators of
   - $\tau(\mu, \gamma) = \mu + 3\sqrt{\gamma}$.
   - $\tau(\mu, \gamma) = \gamma$.