

The Cramér-Rao Lower Bound

Let X_1, X_2, \dots, X_n be a random sample from some distribution with pdf $f(x; \theta)$.

Consider estimating some $\tau(\theta)$.

Suppose that $\hat{\tau}(\theta)$ is any unbiased estimator of $\tau(\theta)$.

Then

$$\text{Var}[\hat{\tau}(\theta)] \geq \frac{[\tau'(\theta)]^2}{I_n(\theta)}$$

the CRLB

The Cramér-Rao Lower Bound

$$\text{Var}[\hat{\tau}(\theta)] \geq \frac{[\tau'(\theta)]^2}{I_n(\theta)}$$

Fisher information:

$$I_n(\theta) := E \left[\left(\frac{\partial}{\partial \theta} \ln f(\vec{X}; \theta) \right)^2 \right]$$

Notation:

$$\text{Var}[\hat{\tau}(\theta)] \geq \frac{[\tau'(\theta)]^2}{I_n(\theta)}$$

The estimator: $\hat{\tau}(\theta) = T = t(\vec{X})$

so

$$\text{Var}[T] \geq \frac{[\tau'(\theta)]^2}{I_n(\theta)}$$

The proof of the Cramér-Rao Lower Bound depends on the **Cauchy-Schwartz inequality**:

$$\left(\int g(x) h(x) dx \right)^2 \leq \left(\int g^2(x) dx \right) \left(\int h^2(x) dx \right)$$

Proof: Consider

$$\iint \left(g(x)h(y) - g(y)h(x) \right)^2 dx dy \geq 0$$

Square the left-hand side out and run the integrals through.

Get

$$0 \leq \iint g^2(x)h^2(y) \, dx \, dy$$

$$- 2 \iint g(x)h(y)g(y)h(x) \, dx \, dy$$

$$+ \iint g^2(y)h^2(x) \, dx \, dy$$

these
are
the
same



Get

$$0 \leq 2 \iint g^2(x) h^2(y) \, dx \, dy$$

$$- 2 \iint g(x) h(y) g(y) h(x) \, dx \, dy$$

$$\iint g(x) h(y) g(y) h(x) \, dx \, dy$$

$$= \int g(y) h(y) \left(\int g(x) h(x) \, dx \right) dy$$

$$= \left(\int g(x) h(x) \, dx \right) \left(\int g(y) h(y) \, dy \right)$$

$$= \left(\int g(x) h(x) \, dx \right)^2$$

Get

$$0 \leq 2 \iint g^2(x)h^2(y) \, dx \, dy \\ - 2 \left(\int g(x)h(x) \, dx \right)^2$$

$$\left(\int g(x)h(x) \, dx \right)^2 \leq \iint g^2(x)h^2(y) \, dx \, dy$$

$$\leq \left(\int g^2(x) \, dx \right) \left(\int h^2(y) \, dy \right)$$

$$\leq \left(\int g^2(x) \, dx \right) \left(\int h^2(x) \, dx \right)$$

$$\left(\int g(x) h(x) dx \right)^2 \leq \left(\int g^2(x) dx \right) \left(\int h^2(x) dx \right)$$

- This holds with sums. (discrete version)
- This holds with dx replaced by $f(x)dx$ where f is a pdf:

$$(E[g(X)h(X)])^2 \leq E[g^2(X)] E[h^2(X)]$$

$$\text{Var}[\hat{\tau}(\theta)] \geq \frac{[\tau'(\theta)]^2}{I_n(\theta)}$$

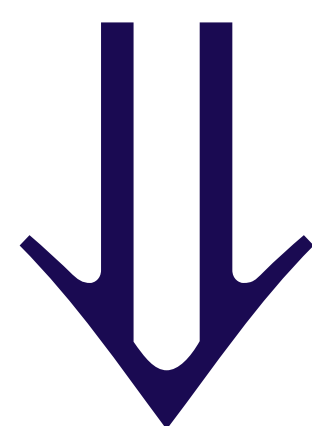
$$I_n(\theta) := E \left[\left(\frac{\partial}{\partial \theta} \ln f(\vec{X}; \theta) \right)^2 \right]$$

Will be done if we can show that:

$$\tau'(\theta) = \underbrace{(T - \tau(\theta))}_{g(\vec{X})} \underbrace{\left(\frac{\partial}{\partial \theta} \ln f(\vec{X}; \theta) \right)}_{h(\vec{X})}$$

where $T = t(\vec{X})$.

$$\tau'(\theta) = (\mathsf{T} - \tau(\theta)) \left(\frac{\partial}{\partial \theta} \ln f(\vec{X}; \theta) \right)$$



$$\mathbb{E}[\tau'(\theta)] = \mathbb{E} \left[(\mathsf{T} - \tau(\theta)) \left(\frac{\partial}{\partial \theta} \ln f(\vec{X}; \theta) \right) \right]$$

\parallel

$$\tau'(\theta)$$

$$[\tau'(\theta)]^2 = \left(\mathbb{E} \left[(\mathsf{T} - \tau(\theta)) \left(\frac{\partial}{\partial \theta} \ln f(\vec{X}; \theta) \right) \right] \right)^2$$

$$[\tau'(\theta)]^2 = \left(\mathbb{E} \left[(\mathsf{T} - \tau(\theta)) \left(\frac{\partial}{\partial \theta} \ln f(\vec{X}; \theta) \right) \right] \right)^2$$

$$\leq \mathbb{E} \left[(\mathsf{T} - \tau(\theta))^2 \right] \cdot \mathbb{E} \left[\left(\frac{\partial}{\partial \theta} \ln f(\vec{X}; \theta) \right)^2 \right]$$

$$= \text{Var}[\mathsf{T}] \cdot \mathsf{I}_n(\theta)$$

$$\Rightarrow \text{Var}[\mathsf{T}] \leq \frac{[\tau'(\theta)]^2}{\mathsf{I}_n(\theta)}$$

$$\tau'(\theta) = (\mathsf{T} - \tau(\theta)) \left(\frac{\partial}{\partial \theta} \ln f(\vec{\mathsf{X}}; \theta) \right)$$


$$\tau'(\theta) = \frac{\partial}{\partial \theta} \tau(\theta) = \frac{\partial}{\partial \theta} \mathsf{E}[\mathsf{T}]$$

$$= \frac{\partial}{\partial \theta} \int \mathsf{t}(\vec{\mathsf{x}}) f(\vec{\mathsf{x}}; \theta) d\vec{\mathsf{x}}$$

$$= \frac{\partial}{\partial \theta} \int \mathsf{t}(\vec{\mathsf{x}}) f(\vec{\mathsf{x}}; \theta) d\vec{\mathsf{x}} - \tau(\theta) \underbrace{\frac{\partial}{\partial \theta} \int f(\vec{\mathsf{x}}; \theta) d\vec{\mathsf{x}}}_{\text{1}}$$

$$= \int (\mathsf{t}(\vec{\mathsf{x}}) - \tau(\theta)) \frac{\partial}{\partial \theta} f(\vec{\mathsf{x}}; \theta) d\vec{\mathsf{x}}$$




$$\tau'(\theta) = \int (\mathbf{t}(\vec{\mathbf{x}}) - \tau(\theta)) \frac{\partial}{\partial \theta} f(\vec{\mathbf{x}}; \theta) d\vec{\mathbf{x}}$$


Want to see



- $E \left[(\mathbf{t}(\vec{\mathbf{X}}) - \tau(\theta)) \left(\frac{\partial}{\partial \theta} \ln f(\vec{\mathbf{X}}; \theta) \right) \right]$

- Note that

$$\frac{\partial}{\partial \theta} f(\vec{\mathbf{x}}; \theta) = \frac{\partial}{\partial \theta} \ln f(\vec{\mathbf{x}}; \theta) f(\vec{\mathbf{x}}; \theta)$$


The CRLB is valid if

- $\frac{\partial}{\partial \theta} \int f(\vec{x}; \theta) dx = \int \frac{\partial}{\partial \theta} f(\vec{x}; \theta) dx$

- $\frac{\partial}{\partial \theta} \ln f(\vec{x}; \theta)$ **exists**

- $0 < E \left[\left(\frac{\partial}{\partial \theta} \ln f(\vec{X}; \theta) \right)^2 \right] < \infty$

The CRLB is valid if

- $$\frac{\partial}{\partial \theta} \int f(\vec{x}; \theta) dx = \int \frac{\partial}{\partial \theta} f(\vec{x}; \theta) dx$$

This one doesn't hold whenever the parameter is in the indicator or support of the distribution.

Example:

The CRLB doesn't hold for the $\text{unif}(0, \theta)$ distribution!

$$\text{Var}[\hat{\tau}(\theta)] \geq \frac{[\tau'(\theta)]^2}{I_n(\theta)}$$
$$I_n(\theta) := E \left[\left(\frac{\partial}{\partial \theta} \ln f(\vec{X}; \theta) \right)^2 \right]$$

Example:

$$X_1, X_2, \dots, X_n \stackrel{\text{iid}}{\sim} \text{Bernoulli}(p)$$

Find the Cramér-Rao lower bound of the variance of all unbiased estimators of p .

Here, $\theta=p$ and $\tau(p)=p$.

The Fisher Information:

$$I_n(p) := E \left[\left(\frac{\partial}{\partial p} \ln f(\vec{X}; p) \right)^2 \right]$$

pdf: $f(x; p) = p^x(1 - p)^{1-x} I_{\{0,1\}}(x)$

joint pdf: $f(\vec{X}; p)$

$$= p^{\sum_{i=1}^n x_i} (1 - p)^{n - \sum_{i=1}^n x_i} \prod_{i=1}^n I_{\{0,1\}}(x_i)$$

Take the log:

$$\ln f(\vec{x}; p)$$

$$= \left(\sum_{i=1}^n x_i \right) \ln p + \left(n - \sum_{i=1}^n x_i \right) \ln(1 - p)$$

Take the derivative:

$$\frac{\partial}{\partial p} \ln f(\vec{x}; p) = \frac{\sum_{i=1}^n x_i}{p} - \frac{n - \sum_{i=1}^n x_i}{1 - p}$$

Simplify:

$$\frac{\partial}{\partial p} \ln f(\vec{x}; p) = \frac{(1 - p) \sum_{i=1}^n x_i - p \left(n - \sum_{i=1}^n x_i \right)}{p(1 - p)}$$

$$= \frac{\sum_{i=1}^n x_i - np}{p(1 - p)}$$

Put the random variables in, square, and take the expectation.

Note that $Y = \sum_{i=1}^n X_i \sim \text{binomial}(n, p)$

$$I_n(p) = E \left[\left(\frac{\partial}{\partial p} \ln f(\vec{X}; p) \right)^2 \right]$$

$$= E \left[\left(\frac{Y - np}{p(1 - p)} \right)^2 \right]$$

$$= \frac{1}{p^2(1 - p)^2} \underbrace{E \left[(Y - np)^2 \right]}$$

variance of
binomial

Note that $Y = \sum_{i=1}^n X_i \sim \text{binomial}(n, p)$

$$I_n(p) = \frac{1}{p^2(1-p)^2} E \left[(Y - np)^2 \right]$$

$$= \frac{1}{p^2(1-p)^2} \text{Var}[Y]$$

$$= \frac{np(1-p)}{p^2(1-p)^2} = \frac{n}{p(1-p)}$$

Example:

$$X_1, X_2, \dots, X_n \stackrel{\text{iid}}{\sim} \text{Bernoulli}(p)$$

$$\begin{aligned} \text{Var}[\hat{p}] &= \frac{[\tau'(p)]^2}{I_n(p)} \\ &= \frac{1^2}{n/[p(1-p)]} \\ &= \frac{p(1-p)}{n} \end{aligned}$$

Example:

$$X_1, X_2, \dots, X_n \stackrel{\text{iid}}{\sim} \text{Bernoulli}(p)$$

mean: p

variance: $p(1 - p)$

$$E[\bar{X}] = E[X_1] = p$$

$$\text{Var}[\bar{X}] = \frac{\text{Var}[X_1]}{n} = \frac{p(1 - p)}{n}$$

So \bar{X} is an unbiased estimator of p with the smallest possible variance! Woot!

This estimator for p for the Bernoulli distribution is actually a

Uniformly
Minimum
Variance
Unbiased
Estimator.

