The Cramér–Rao Lower Bound

Statistics · Module 3 — Point Estimation

Node 3.4

Position in Knowledge Graph

System:
Statistics

Structural Domain:
Point Estimation

Node:
3.4

Regularity Conditions (Important)

Assume:

1. The model $f_\theta(x)$ is differentiable in θ.
2. Differentiation under the integral sign is valid.
3. The support of f_θ does not depend on θ.

These are required for the identities below.

Theorem (Cramér–Rao Inequality)

Let $\hat{\theta}$ be an unbiased estimator of θ.

Then:

$$\mathrm{Var}_\theta(\hat{\theta}) \geq \frac{1}{\mathcal{I}_n(\theta)},$$

where:

$$\mathcal{I}_n(\theta) = n\mathcal{I}(\theta) = n\mathbb{E}_\theta\left[\left(\frac{\partial}{\partial\theta}\log f_\theta(X)\right)^2\right].$$

Interpretation

No unbiased estimator can have variance smaller than the reciprocal of Fisher information.

Fisher Information sets a fundamental precision limit.

More curvature → smaller lower bound → more precise estimation possible.

Proof Sketch (Core Geometry)

Let:

$$S(\theta) = \frac{\partial}{\partial\theta} \log L(\theta).$$

Then:

$$\mathbb{E}_\theta[S(\theta)] = 0.$$

Differentiate unbiasedness:

$$\mathbb{E}_\theta[\hat{\theta}] = \theta.$$

Differentiating both sides w.r.t. θ:

$$\frac{d}{d\theta}\mathbb{E}_\theta[\hat{\theta}] = 1.$$

Using interchange of derivative and expectation:

$$\mathbb{E}_\theta\left[\hat{\theta}S(\theta)\right] = 1.$$

Now apply Cauchy–Schwarz:

$$1^2 \leq \mathrm{Var}_\theta(\hat{\theta}) \cdot \mathrm{Var}_\theta(S(\theta)).$$

But:

$$\mathrm{Var}_\theta(S(\theta)) = \mathcal{I}_n(\theta).$$

Thus:

$$\mathrm{Var}_\theta(\hat{\theta}) \geq \frac{1}{\mathcal{I}_n(\theta)}.$$

Equality Condition

Equality holds iff:

$$\hat{\theta} = \theta + c(\theta)S(\theta).$$

In exponential families, the MLE often achieves the bound.

Example: Normal Mean (Known Variance)

If $X_i \sim \mathcal{N}(\mu, \sigma^2)$,

$$\mathcal{I}_n(\mu) = \frac{n}{\sigma^2}.$$

Thus:

$$\mathrm{Var}(\hat{\mu}) \geq \frac{\sigma^2}{n}.$$

The sample mean satisfies:

$$\mathrm{Var}(\bar{X}) = \frac{\sigma^2}{n}.$$

It achieves the bound → efficient.

Cross-Links

- **Linear Algebra:** Bound derived from Cauchy–Schwarz in L^2 space.
- **AI:** Fisher Information underlies natural gradient methods.
- **Economics:** Efficiency bounds in structural estimation.
- **Quantum:** Quantum Cramér–Rao bound generalizes this inequality.

Structural Compression

Invariant:
Variance ≥ inverse information.

Mechanism:
Geometry of score function via Cauchy–Schwarz.

Cross-System Echo:
In physics, uncertainty principles bound precision.

Exercises

1. Prove the unbiasedness differentiation step carefully.
2. Compute Fisher information for Poisson and derive CRLB.
3. Show sample mean achieves CRLB for Normal mean.
4. Explain why biased estimators can beat the bound.

Statistics · Module 3 — Point Estimation · Node 3.4
Concepts: fisher-information, variance-and-covariance
Prerequisites: 2.5, 3.3, 2.4
Enables: 3.5, 7.1

hbar.university — own.your.cognition