Invariance and Equivariance in Estimation

Statistics · Module 3 — Point Estimation

Node 3.6

Position in Knowledge Graph

System:
Statistics

Structural Domain:
Point Estimation

Node:
3.6

Core Principle

Estimation procedures should behave naturally under transformations of the parameter.

If we reparameterize the model, the estimator should transform accordingly.

This is called **invariance** (or more precisely, equivariance).

Definition: Invariance of the MLE

Let $\hat{\theta}_{\text{MLE}}$ be the MLE of θ.

If $\phi = g(\theta)$ is a one-to-one transformation, then:

$$\hat{\phi}_{\text{MLE}} = g(\hat{\theta}_{\text{MLE}}).$$

That is:

> The MLE of a transformed parameter equals the transformation of the MLE.

This is called the **invariance property of the MLE**.

Why It Holds

The likelihood for ϕ is:

$$L_\phi(\phi) = L_\theta(g^{-1}(\phi)).$$

Maximizing over ϕ is equivalent to maximizing over θ.

Thus the argmax transforms accordingly.

No recalculation needed.

Intuition

If you estimate:

- variance σ^2,
- then want standard deviation σ,

you do not re-estimate. You transform the estimate.

MLE respects the structure of the parameter space.

Other estimators (e.g., unbiased estimators) generally do **not** satisfy this property.

Structural Role

This explains why MLE is "natural":

- It respects reparameterization.
- It behaves coherently under model transformations.
- It is defined through the likelihood geometry rather than moment constraints.

This prepares for:

- Asymptotic normality
- Information geometry
- Delta method
- Likelihood-based testing

Worked Example

Example: Normal Variance

Suppose $X_i \sim \mathcal{N}(\mu, \sigma^2)$ with known μ.

MLE for variance:

$$\hat{\sigma}^2 = \frac{1}{n} \sum (X_i - \mu)^2.$$

Now consider $\sigma = \sqrt{\sigma^2}$.

By invariance:

$$\hat{\sigma} = \sqrt{\hat{\sigma}^2}.$$

No separate optimization needed.

Contrast: Unbiased Estimators

Unbiased estimators do **not** generally preserve invariance.

Example:

If $\hat{\theta}$ is unbiased for θ,

$$\mathbb{E}[\hat{\theta}] = \theta,$$

it does **not** imply:

$$\mathbb{E}[g(\hat{\theta})] = g(\theta).$$

Thus unbiasedness is not invariant under nonlinear transformations.

This is one reason unbiasedness is not always the dominant principle in modern inference.

Equivariance Under Group Transformations

More generally, if a model is invariant under a transformation group G, estimators can be required to satisfy:

$$\delta(g \cdot x) = g \cdot \delta(x).$$

Such estimators are called **equivariant**.

This becomes important in:

- Location-scale families
- Decision theory
- Invariant testing

Cross-Links

- **Linear Algebra:** Coordinate changes preserve geometric objects.
- **AI:** Parameterization choice affects optimization but not the optimum distribution.
- **Economics:** Reparameterizations in structural models.
- **Quantum:** Different parameterizations of states must yield consistent estimators.

Structural Compression

Invariant:
MLE commutes with smooth one-to-one transformations.

Mechanism:
Argmax of likelihood preserved under reparameterization.

Cross-System Echo:
In geometry, intrinsic properties do not depend on coordinates.

Exercises

1. Prove the invariance property formally.
2. Give an example where unbiasedness fails under transformation.
3. Show that MLE for $\log \theta$ equals $\log(\hat{\theta}_{\text{MLE}})$.
4. Consider a location family $f(x - \theta)$. What form must an equivariant estimator take?

Statistics · Module 3 — Point Estimation · Node 3.6
Concepts: maximum-likelihood-estimation
Prerequisites: 3.2
Enables: 7.1, 8.1

hbar.university — own.your.cognition