Estimating parameters from data

2026-09-09

Summary

Going from data to parameters instead of from parameters to simulations, and three ways to do it.

- The inverse problem: weeks 1 and 2 started from a known distribution and simulated forward; now we start from data and recover the parameters.
- A hat marks an estimate from data: $\hat{\theta}$ is computed, θ is the unknown truth.
- The likelihood $p(y \mid \theta)$ measures how consistent the data are with a proposed θ.
- $p(\theta \mid y) \neq p(y \mid \theta)$: the probability of the parameters given the data is not the probability of the data given the parameters.
- Maximum likelihood finds the peak of the likelihood; for the Normal, $\hat{\mu} = \bar{y}$.
- Method of moments matches sample moments to distribution moments; for the Normal, the same answer.
- Bayes' theorem reverses the direction: prior $\times$ likelihood $\propto$ posterior, giving a full distribution over θ rather than a point estimate.
- With enough data, all three converge and the prior stops mattering.

Key results.

$$p(y \mid \theta) = \prod_{i=1}^{n} p(y_i \mid \theta) \tag{1}$$

$$\hat{\theta}_{\text{MLE}} = \arg\max_{\theta} \log p(y \mid \theta) \tag{2}$$

$$p(\theta \mid y) \propto p(y \mid \theta)\, p(\theta) \tag{3}$$

Bibliography