ACTL2131 2.3 - Evaluating Point Estimators

Tadhg Xu-Glassop

2025T1

Motivation

So far, we've found ways of estimating unknown parameters θ using estimators, $\hat{\theta}$.

But we can find multiple estimators for the one quantity, (e.g. by finding MME or MLE or guessing your favourite number), so we need a way of assessing how 'good' and estimator is, and comparing different estimators.

Estimators are random - they are based off a sample, which is itself random. So, we can use the same methods of analysing estimators as we would random variables...

Bias, Variance and MSE

We define the bias of an estimator as

$$\text{Bias}(\hat{\theta}) = \mathbb{E}[\hat{\theta}] - \theta,$$

variance as usual, and Mean Squared Error (MSE) as

$$\mathrm{MSE} = \mathbb{E}[(\theta - \hat{\theta})^2] = \text{Var}(\hat{\theta}) + \text{Bias}^2(\hat{\theta}).$$

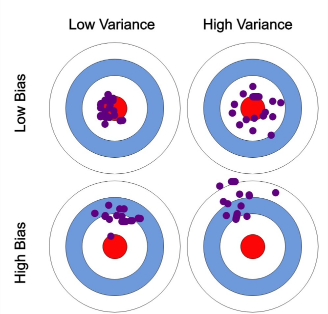
We say an estimator is unbiased if $\text{Bias}(\hat{\theta}) = 0$ or $\mathbb{E}[\hat{\theta}] = \theta$.

Intuitively,

- Bias is how wrong we expect our estimator to be;
- Variance is how variable - 'random' our estimator is. The square root gives **standard error**, which is the standard deviatin of an estimator;
- MSE is a measure of the distance between our estimator and the true value.

Bias, Variance and MSE (cont.)

Below is a visualisation of different levels of bias and variance. Try to understand the intuition behind each target.



Cramer-Rao Lower Bound

Definition

Define $I_f(\theta)$ the Fisher information of the parameter θ wrt a RV with pdf f, where

$$I_f(\theta) = -\mathbb{E}\left[\frac{\partial^2 \log(f_X(x|\theta))}{\partial\theta^2}\right].$$

Then, the Cramer-Rao Lower Bound for the variance of unbiased estimators for θ is given by

$$\mathrm{CRLB}(\theta) = \frac{1}{n \cdot I_f(\theta)}$$

So, if I have an unbiased estimator $\hat{\theta}$, the smallest possible variance it could have is $\mathrm{CRLB}(\theta)$. If $\mathrm{Var}(\theta) = \mathrm{CRLB}(\theta)$, then this estimator has minimal variance.

Properties of MLE's

Theorem (Asymptotic distribution of MLE's)

Let $\hat{\theta}$ be the MLE for θ. Then

$$\mathbb{E}[\hat{\theta}] \xrightarrow{p} \theta, \qquad \text{and} \qquad \mathsf{Var}(\hat{\theta}) \xrightarrow{p} \mathrm{CRLB}(\theta).$$

Further,

$$\hat{\theta} \xrightarrow{d} \mathcal{N}\left(\theta, \mathrm{CRLB}(\theta)\right).$$

This means we can approximate

$$\frac{\hat{\theta} - \theta}{\sqrt{\mathrm{CRLB}(\theta)}} \sim \mathcal{N}(0, 1),$$

and we can use this to make confidence intervals for θ.