Maximum Likelihood Estimation and Bayesian Statistics

Rebecca C. Steorts

Agenda

- Maximum Likelihood Esimation
- Unbiased Estimators

Traditional inference

You are given **data** X and there is an **unknown parameter** you wish to estimate θ

How would you estimate θ?

- Find an unbiased estimator of θ.
- Find the maximum likelihood estimate (MLE) of θ by looking at the likelihood of the data.
- Suppose that $\hat{\theta}$ estimates θ. Note: $\hat{\theta}$ may depend on the data $x_{1:n} = x_1, \ldots x_n$.

Unbiased Estimator

Recall that $\hat{\theta}$ is an unbiased estimator of θ if

$$E[\hat{\theta}] = \theta. \tag{1}$$

.

Maximum Likelihood Estimation

For each sample point $x_{1:n}$, let $\hat{\theta}$ be a parameter value at which $p(x_{1:n} \mid \theta)$ attains it's maximum as a function of θ, with $x_{1:n}$ held fixed. A *maximum likelihood esimator (MLE)* of the parameter θ based on a sample $x_{1:n}$ is $\hat{\theta}$.

Find the MLE

The solution to the MLE are the possible candidates (θ) that solve

$$\frac{\partial p(x_{1:n} \mid \theta)}{\partial \theta} = 0. \tag{2}$$

Solution to the above equation are only possible candidates for the MLE since the first derivative being 0 is a necessary condition for a maximum (but not a sufficient one).

Our job is to find a global maximum. Thus, we need to ensure that we haven't found a local one.

MLE of Normal distribution

Consider

$$X_1, \ldots, X_n \stackrel{iid}{\sim} \text{Normal}(\theta, 1).$$

Show that the MLE is $\hat{\theta} = \bar{x}$.

MLE of Normal distribution

$$p(x_{1:n} \mid \theta) = (2\pi)^{-n/2} \times \exp\{\frac{-1}{2}\sum_i (x_i - \theta)^2\} \tag{3}$$

Consider

$$\log p(x_{1:n}) = -n/2 \log(2\pi) - \frac{1}{2}\sum_i (x_i - \theta)^2 \tag{4}$$

MLE of Normal distribution

$$\frac{\partial p(x_{1:n} \mid \theta)}{\partial \theta} = \sum_i (x_i - \theta) \tag{5}$$

This implies that

$$\sum_i (x_i - \theta) = 0 \implies \hat{\theta} = \bar{x}.$$

MLE of Normal distribution

Consider

$$\frac{\partial^2 p(x_{1:n} \mid \theta)}{\partial \theta^2} = -n < 0.$$

Thus, our solution is unique.

Exercise

Show that

$$\hat{\theta} = \bar{x}$$

is an unbiased estimator for θ.

Proof.

$$E[\hat{\theta}] = E[\bar{x}] = \frac{1}{n}\sum_i E[x_i] = \frac{1}{n}\sum_i \theta = \theta.$$

Thus, we have showed that the MLE is an unbiased estimator for θ.

Normal-Normal model

Suppose that

$$X_1, \ldots, X_n \overset{iid}{\sim} \text{Normal}(\theta, 1),$$

where we now consider

$$\theta \overset{ind}{\sim} \text{Normal}(\mu, \tau^2).$$

Let $\lambda = 1$ and $\lambda_o = 1/\tau^2$.

Recall that from module 3,

$$\theta \mid x_{1:n} \sim N(M, L^{-1}),$$

where

$$L = n\lambda + \lambda_o$$

and

$$M = \frac{n\lambda\bar{x} + \lambda_o\mu}{n\lambda + \lambda_o}.$$

Normal-Normal model

Observe that

$$M = \frac{n\lambda\bar{x} + \lambda_o\mu}{n\lambda + \lambda_o} = \frac{n\lambda\hat{\theta} + \lambda_o\mu}{n\lambda + \lambda_o} = \frac{n\lambda}{n\lambda + \lambda_o}\hat{\theta} + \frac{\lambda_o}{n\lambda + \lambda_o}\mu.$$

Thus, we can write the posterior mean as a function of the MLE and the prior mean μ.