Class 6

Shikhar Saxena

February 06, 2023

Contents

Consistent Estimators

Definition 1. *Let $\hat{\Theta}_1, \hat{\Theta}_2, \dots, \hat{\Theta}_n$ be a sequence of point estimators of θ. We say that $\hat{\Theta}$ is a consistent estimator of θ if*

$$\lim_{n \to \infty} P(|\hat{\Theta} - \theta| \geq \epsilon) = 0 \quad \forall \epsilon > 0$$

An alternate definition:

Definition 2. *Let $\hat{\Theta}_1, \hat{\Theta}_2, \dots, \hat{\Theta}_n$ be a sequence of point estimators of θ. We say that $\hat{\Theta}_n$ is a consistent estimator of θ if*

$$\lim_{n \to \infty} MSE(\hat{\Theta}_n) = 0$$

Proof.

$$\begin{aligned} P(|\hat{\Theta} - \theta| \geq \epsilon) &= P(|\hat{\Theta} - \theta|^2 \geq \epsilon^2) \\ &\leq \frac{E\left[|\hat{\Theta}_n - \theta|^2\right]}{\epsilon^2} \quad \text{By Markov inequality} \\ &= MSE(\hat{\Theta}_n)/\epsilon^2 \end{aligned}$$

□

Maximum Likelihood Estimator

Example 1. *Bag contains 3 balls. Each ball is either red or blue. Let θ be the number of blue balls. Define random variables X_1, X_2, X_3, X_4 (balls are chosen with* *replacement that is why we can have more than three random variables).*

$$X_i = \begin{cases} 1 & \textit{if } i^{th} \textit{ chosen ball is blue} \\ 0 & \textit{otherwise} \end{cases}$$

Then $X_i \sim$ *Bernoulli* $\left(\frac{\theta}{3}\right)$. *By independence,* $P(x_1, x_2, x_3, x_4) = \prod_i P_{X_i}(x_i)$ *where* $P_{X_i}(x;\theta) = \binom{3}{x}\theta^x(1-\theta)^{3-x}$.

Likelihood *function is then defined as:* $L(x_1, x_2, x_3, x_4) = P(x_1, x_2, x_3, x_4; \theta)$.

For jointly continuous random variables, $L(x_1, x_2, x_3, x_4) = f(x_1, x_2, x_3, x_4; \theta)$.

Bayesian Inference

Draw inference of an unknown random variable X by observing random variable Y. The unknown variable is modelled with prior distribution $P_X(x)$.

After observing Y, we find posterior distribution of X, $P_{X|Y}(x|y)$. Usually found using Bayes' formula.

MAP Estimate is then shown by $\hat{x}_{MAP}$. We don't care about the denominator (in Bayes' rule) for MAP estimate because it is just a constant.

So we maximize $f_{Y|X}(y|x)f_X(x)$.