Maximum Likelihood Estimation

Tuesday, December 6, 2022

Motivating example:

Consider a bag containing 3 balls.
Each ball is either red or blue.

Let θ be the number of blue balls
possible outcomes are $\theta = \{0, 1, 2, 3\}$.

Choose 4 balls with replacement at random.
So, let X_1, X_2, X_3, X_4 be i.i.d. r.v.

$$X_i = \begin{cases} 1 & \text{if blue} \\ 0 & \text{if red} \end{cases}$$

$$X_i \sim \text{Bernoulli}\left(\frac{\theta}{3}\right).$$

Experiment: Taking data.

observations: $x_1 = 1,\ x_2 = 0,\ x_3 = 1,\ x_4 = 1$.
3 blue, 1 red.

1. For each possible values of θ, what is the probability of the observed sample $(x_1, x_2, x_3, x_4) = (1, 0, 1, 1)$.

We know $\theta = \{0, 1, 2, 3\} \longrightarrow$ number of blue balls.

Since $X_i \sim \text{Bernoulli}\left(\frac{\theta}{3}\right)$, then

$$P_{X_i}(x) = \begin{cases} \theta/3 & \text{if } x = 1 \\ 1 - \theta/3 & \text{if } x = 0 \end{cases}$$

Since X_i are independent, the joint PMF for X_1, X_2, X_3, and X_4 is

$$P_{X_1 X_2 X_3 X_4}(x_1, x_2, x_3, x_4) = P_{X_1}(x_1) P_{X_2}(x_2) P_{X_3}(x_3) P_{X_4}(x_4) = L(x_1, x_2, x_3, x_4)$$

So,

$$L(1, 0, 1, 1) = P_{X_1}(1) P_{X_2}(0) P_{X_3}(1) P_{X_4}(1)$$

$$= \left(\frac{\theta}{3}\right)\left(1 - \frac{\theta}{3}\right)\left(\frac{\theta}{3}\right)\left(\frac{\theta}{3}\right)$$

$$L(1,0,1,1;\theta) = \left(\frac{\theta}{3}\right)^3\left(1 - \frac{\theta}{3}\right)$$

$$L(1,0,1,1;\theta) = \left(\frac{\theta}{3}\right)^3\left(1-\frac{\theta}{3}\right)$$

The joint PMF depends on θ. $J(x_1, x_2, x_3, x_4; \theta)$.

Compute the likelihoods.

θ	$J(x_1,x_2,x_3,x_4;\theta)$	
0	0	→ because sample includes both red & blue
1	0.0247	
2	0.0988	→ maximum likelihood
3	0	→ because sample includes both red & blue

Choose $\theta=2$. So our estimate is $\hat{\theta}=2$, this means that the observed data is most likely to occur for $\theta=2$.

Likelihood function Definition

Let $X_1, X_2, X_3, \ldots X_n$ be a random sample from a distribution with parameter θ.

Suppose we observed $X_1 = x_1$, $X_2 = x_2$, $\cdots$, $X_n = x_n$.

1. If X_i's are discrete, then

$$L(x_1, x_2, \cdots, x_n; \theta) = P_{X_1X_2\cdots X_n}(x_1, x_2, \cdots, x_n; \theta)$$

2. If X_i's are jointly continuous, then

$$L(x_1, x_2, \cdots, x_n; \theta) = f_{X_1,X_2\cdots X_n}(x_1, x_2, \cdots, x_n; \theta)$$

In some problems, you can do log likelihood

$$\ln(L(x_1, x_2, \ldots, x_n; \theta)).$$

Example:

1. Find the likelihood function of

$X_i \sim \text{Binomial}(3, \theta)$ and we have observersed $(x_1, x_2, x_3, x_4) = (1, 3, 2, 2)$

n　　p

$$P_{X_i}(x;\theta) = \binom{3}{x}\theta^x(1-\theta)^{3-x}$$

So,

$$L(x_1,x_2,x_3,x_4;\theta) = P_{X_1X_2X_3X_4}(x_1,x_2,x_3,x_4;\theta)$$

$$= P_{X_1}(x_1;\theta)\,P_{X_2}(x_2;\theta)\,P_{X_3}(x_3;\theta)\,P_{X_4}(x_4;\theta)$$

$$= \binom{3}{x_1}\binom{3}{x_2}\binom{3}{x_3}\binom{3}{x_4}\,\theta^{x_1+x_2+x_3+x_4}\,(1-\theta)^{12-(x_1+x_2+x_3+x_4)}$$

plug-in observations:

$$L(1,3,2,2;\theta) = \binom{3}{1}\binom{3}{3}\binom{3}{2}\binom{3}{2}\,\theta^8(1-\theta)^4$$

$$= 27\,\theta^8(1-\theta)^4$$

<u>Maximum Likelihood Estimate (MLE)</u>

Given that we observed $X_1=x_1,\ X_2=x_2,\ \dots,\ X_n=x_n$, a MLE of θ, shown by $\hat{\theta}_{ML}$ is a value of θ that maximizes the likelihood function

$$L(x_1,x_2,x_3,\dots,x_n;\theta).$$

Example:

2. $X_i \sim \text{Binomial}(3,\theta)$

$$L(1,3,2,2;\theta) = 27\,\theta^8(1-\theta)^4.$$

Find MLE.

$$\frac{d}{d\theta}L(1,3,2,2;\theta) = 27\left(8\theta^7(1-\theta)^4 - 4\theta^8(1-\theta)^3\right)$$

$$0 = 27\left(8\theta^7(1-\theta)^4 - 4\theta^8(1-\theta)^3\right)$$

$$\downarrow$$

solutions $\cancel{\theta=0}$, $\theta = 2/3$, $\cancel{\theta=1}$

So, $\hat{\theta}_{ML} = \frac{2}{3}.$