Practice Tasks for Maximum Likelihood Estimation (MLE)

1. Suppose you observe ten coin flips out of which seven lands on heads. If the probability of heads is p, write down the likelihood function and derive the Maximum Likelihood Estimator for p.
2. Suppose you have observed the sample values $X = \{3,7,5,9\}$ from a uniform distribution $U(0, \theta)$. Find the MLE for θ.
3. Given data $X = \{x_1, \ldots, x_n\}$ assumed to follow a normal distribution N(μ,1), write down the likelihood function and derive the MLE for μ.
4. Assume you collect count data $X = \{x_1, \ldots, x_n\}$ from a Poisson distribution with parameter λ. Write down the likelihood function and find the MLE for λ.

Solutions

Task 1.

The probability mass function (PMF) of a Bernoulli distributed random variable X is:

$$P(X = a) = p^a(1-p)^{1-a}, a \in \{0,1\}$$

Since we have ten independent observations $x_1, x_2, \dots, x_{10}$, the likelihood function with $n = 10$ is:

$$L(p) = \prod_{i=1}^{n} p^{x_i}(1-p)^{1-x_i}$$

Taking the logarithm to produce the log-likelihood function:

$$l(p) = \ln L(p) = \sum_{i=1}^{n} x_i \ln p + \sum_{i=1}^{n} (1-x_i) \ln(1-p)$$

$$l(p) = \left(\sum x_i\right) \ln p + \left(n - \sum x_i\right) \ln(1-p)$$

where $n = 10$ (total observations) and $\sum x_i = 7$ (number of heads).

To maximize $l(p)$, differentiate with respect to p:

$$\frac{\partial l(p)}{\partial p} = \frac{\sum x_i}{p} - \frac{n - \sum x_i}{1-p}$$

Setting it to zero:

$$\frac{\sum x_i}{p} = \frac{n - \sum x_i}{1-p}$$

Rearrange to express p:

$$(1-p) \sum x_i = p\left(n - \sum x_i\right)$$

$$\sum x_i - p \sum x_i = pn - p \sum x_i$$

$$\sum x_i = pn$$

$$\frac{\sum x_i}{n} = \hat{p}_{ML}$$

Substituting for our specific sample $n = 10$ and $\sum x_i = 7$, we have that $\hat{p}_{ML} = \frac{7}{10} = 0.7$

Task 2.

We have observed sample values $X = \{3,7,5,9\}$ and we know that they follow $U(0, \theta)$, so the probability density function (PDF) to work with is:

$$P(X = a) = \begin{cases} \frac{1}{\theta}, & 0 \leq a \leq \theta \\ 0, & \text{otherwise} \end{cases}$$

So, the likelihood function is with sample size $n = 4$:

$$L(\theta) = \prod_{i=1}^{n} \frac{1}{\theta} = \left(\frac{1}{\theta}\right)^n = \left(\frac{1}{\theta}\right)^4$$

The likelihood is valid only if θ is at least as large as the maximum observation. That is, $\theta \geq \max(X)$. MLE is found by maximizing $L(\theta)$. Since $L(\theta)$ is a decreasing function of θ, it is maximized for the smallest possible θ, which is:

$$\hat{\theta}_{ML} = \max(X) = 9$$

Task 3.

The probability density function (PDF) of a $X \sim N(\mu, 1)$ random variable is

$$P(X = a) = \frac{1}{\sqrt{2\pi}\sigma} e^{-\frac{(a-\mu)^2}{2\sigma^2}} = \frac{1}{\sqrt{2\pi}} e^{-\frac{(a-\mu)^2}{2}}$$

Since we have n independent observations $x_1, x_2, \dots, x_n$, the likelihood function becomes:

$$L(\mu) = \prod_{i=1}^{n} \frac{1}{\sqrt{2\pi}} e^{-\frac{(x_i-\mu)^2}{2}}$$

Taking the logarithm, we have:

$$l(\mu) = \ln L(\mu) = -\frac{n}{2}\ln(2\pi) - \sum_{i=1}^{n} \frac{(x_i - \mu)^2}{2}$$

To find the MLE, differentiate with respect to μ:

$$\frac{\partial l(\mu)}{\partial \mu} = \sum_{i=1}^{n} (x_i - \mu)$$

Setting it to zero and solve for μ:

$$\sum (x_i - \mu) = 0$$

$$\sum x_i - n\mu = 0$$

$$\hat{\mu}_{ML} = \frac{\sum x_i}{n}$$

Task 4.

The probability mass function (PMF) of a Poisson distributed random variable X with parameter λ is:

$$P(X = a) = \frac{\lambda^a e^{-\lambda}}{a!}, a = 0,1,2, \ldots$$

Since we have n independent observations $x_1, x_2, \ldots, x_n$, the likelihood function becomes:

$$L(\lambda) = \prod_{i=1}^{n} \frac{\lambda^{x_i} e^{-\lambda}}{x_i!}$$

Taking the logarithm, we have:

$$l(\lambda) = \ln L(\lambda) = \sum_{i=1}^{n} (x_i \ln \lambda - \lambda - \ln x_i!)$$

To find the MLE, differentiate with respect to λ:

$$\frac{\partial l(\lambda)}{\partial \lambda} = \sum_{i=1}^{n} \left(\frac{x_i}{\lambda} - 1 - 0\right) = \sum \frac{x_i}{\lambda} - n$$

Setting it to zero and solve for λ:

$$\sum \frac{x_i}{\lambda} - n = 0$$

$$\sum x_i = n\lambda$$

$$\frac{\sum x_i}{n} = \hat{\lambda}_{ML}$$