CSCE883 Homework 2

Matthew "Mehdi" Hatami Goloujeh

Question 1

To find the maximum likelihood estimate (MLE) of θ given the data $(3, 0, 2, 1, 3, 2, 1, 0, 2, 1)$, we start by considering the probability mass function (PMF) of the discrete random variable X:

$$P(X = x) = \begin{cases} \frac{2\theta}{3} & \text{if } x = 0, \\ \frac{\theta}{3} & \text{if } x = 1, \\ \frac{2(1-\theta)}{3} & \text{if } x = 2, \\ \frac{1-\theta}{3} & \text{if } x = 3. \end{cases}$$

Counting Observations

First, we count the number of times each value appears in the data:

$$\begin{aligned} n_0 &= \text{number of times } X = 0 = 2, \\ n_1 &= \text{number of times } X = 1 = 3, \\ n_2 &= \text{number of times } X = 2 = 3, \\ n_3 &= \text{number of times } X = 3 = 2. \end{aligned}$$

The total number of observations is $N = n_0 + n_1 + n_2 + n_3 = 10$.

Writing the Likelihood Function

The likelihood function $L(\theta)$ is the product of the probabilities of each observation:

$$\begin{aligned} L(\theta) &= \left(\frac{2\theta}{3}\right)^{n_0} \left(\frac{\theta}{3}\right)^{n_1} \left(\frac{2(1-\theta)}{3}\right)^{n_2} \left(\frac{1-\theta}{3}\right)^{n_3} \\ &= \left(\frac{1}{3}\right)^{N} (2\theta)^{n_0} (\theta)^{n_1} (2(1-\theta))^{n_2} (1-\theta)^{n_3}. \end{aligned}$$

Since the constant $\left(\frac{1}{3}\right)^N$ does not affect the maximization, we can ignore it:

$$L(\theta) \propto (2\theta)^{n_0} (\theta)^{n_1} (2(1-\theta))^{n_2} (1-\theta)^{n_3} .$$

Simplifying the Likelihood Function

Combine like terms and exponents:

$$L(\theta) \propto 2^{n_0+n_2}\theta^{n_0+n_1}(1-\theta)^{n_2+n_3}.$$

Computing the Log-Likelihood Function

Taking the natural logarithm simplifies differentiation:

$$\begin{aligned}
\ell(\theta) &= \ln L(\theta) \\
&= (n_0+n_1)\ln\theta + (n_2+n_3)\ln(1-\theta) + (n_0+n_2)\ln 2 + \text{constant}.
\end{aligned}$$

Again, constants can be ignored during maximization:

$$\ell(\theta) = (n_0+n_1)\ln\theta + (n_2+n_3)\ln(1-\theta) + \text{constant}.$$

Differentiating and Setting the Derivative to Zero

Differentiate $\ell(\theta)$ with respect to θ:

$$\frac{d\ell}{d\theta} = \frac{n_0+n_1}{\theta} - \frac{n_2+n_3}{1-\theta}.$$

Set the derivative equal to zero to find the maximum:

$$\frac{n_0+n_1}{\theta} - \frac{n_2+n_3}{1-\theta} = 0.$$

Solving for θ

Multiply both sides by $\theta(1-\theta)$:

$$(n_0+n_1)(1-\theta) - (n_2+n_3)\theta = 0.$$

Simplify the equation:

$$\begin{aligned}
(n_0+n_1)(1-\theta) - (n_2+n_3)\theta &= 0 \\
(n_0+n_1) - (n_0+n_1)\theta - (n_2+n_3)\theta &= 0 \\
(n_0+n_1) - \theta\left[(n_0+n_1) + (n_2+n_3)\right] &= 0.
\end{aligned}$$

Since $(n_0+n_1) + (n_2+n_3) = N$:

$$(n_0 + n_1) - \theta N = 0.$$

Solving for θ:

$$\theta = \frac{n_0 + n_1}{N}.$$

Computing the MLE of θ

Plug in the observed values:

$$\theta_{\text{MLE}} = \frac{n_0 + n_1}{N} = \frac{2 + 3}{10} = \frac{5}{10} = 0.5.$$

Conclusion

The maximum likelihood estimate of θ is $\mathbf{0.5}$.

Question 2

We are given that $X_1, X_2, \ldots, X_n$ are independent and identically distributed (i.i.d.) with the probability density function (pdf):

$$f(x \mid \sigma) = \frac{1}{2\sigma} \exp\left(-\frac{|x|}{\sigma}\right).$$

Our goal is to find the maximum likelihood estimate (MLE) of σ.

Writing the Likelihood Function

The likelihood function $L(\sigma)$ based on the observed data $x_1, x_2, \ldots, x_n$ is:

$$L(\sigma) = \prod_{i=1}^{n} f(x_i \mid \sigma) = \prod_{i=1}^{n} \left[\frac{1}{2\sigma} \exp\left(-\frac{|x_i|}{\sigma}\right)\right].$$

Simplify the likelihood function:

$$L(\sigma) = \left(\frac{1}{2\sigma}\right)^n \exp\left(-\frac{1}{\sigma}\sum_{i=1}^{n} |x_i|\right).$$

Computing the Log-Likelihood Function

Taking the natural logarithm of the likelihood function to simplify differentiation:

$$\begin{aligned}
\ell(\sigma) &= \ln L(\sigma) \\
&= n \ln\left(\frac{1}{2\sigma}\right) - \frac{1}{\sigma}\sum_{i=1}^{n} |x_i| \\
&= -n \ln(2\sigma) - \frac{1}{\sigma}\sum_{i=1}^{n} |x_i|.
\end{aligned}$$

Differentiating and Setting the Derivative to Zero

Differentiate $\ell(\sigma)$ with respect to σ:

$$\frac{d\ell}{d\sigma} = -n \cdot \frac{1}{\sigma} + \frac{1}{\sigma^2}\sum_{i=1}^{n} |x_i| = -\frac{n}{\sigma} + \frac{S}{\sigma^2},$$

where $S = \sum_{i=1}^{n} |x_i|$.

Set the derivative equal to zero to find the maximum:

$$-\frac{n}{\sigma} + \frac{S}{\sigma^2} = 0.$$

Solving for σ

Multiply both sides by σ^2:

$$-n\sigma + S = 0.$$

Solve for σ:

$$n\sigma = S \implies \sigma = \frac{S}{n} = \frac{1}{n}\sum_{i=1}^{n} |x_i|.$$

Conclusion

The maximum likelihood estimate of σ is:

$$\hat{\sigma} = \frac{1}{n}\sum_{i=1}^{n} |x_i|,$$

which is the sample mean of the absolute values of the observations.

Question 3

We are given that $X_1, X_2, \ldots, X_n$ are independent and identically distributed (i.i.d.) with the normal distribution $\mathcal{N}(\mu, \sigma^2)$. Our goal is to find the maximum likelihood estimates (MLE) of μ and σ^2.

Writing the Likelihood Function

The likelihood function $L(\mu, \sigma^2)$ based on the observed data $x_1, x_2, \ldots, x_n$ is:

$$L(\mu, \sigma^2) = \prod_{i=1}^{n} f(x_i \mid \mu, \sigma^2) = \prod_{i=1}^{n} \left[\frac{1}{\sqrt{2\pi\sigma^2}} \exp\left(-\frac{(x_i - \mu)^2}{2\sigma^2} \right) \right].$$

Simplify the likelihood function:

$$L(\mu, \sigma^2) = \left(\frac{1}{\sqrt{2\pi\sigma^2}} \right)^n \exp\left(-\frac{1}{2\sigma^2} \sum_{i=1}^{n} (x_i - \mu)^2 \right).$$

Computing the Log-Likelihood Function

Taking the natural logarithm of the likelihood function to simplify differentiation:

$$\begin{aligned} \ell(\mu, \sigma^2) &= \ln L(\mu, \sigma^2) \\ &= -\frac{n}{2} \ln(2\pi\sigma^2) - \frac{1}{2\sigma^2} \sum_{i=1}^{n} (x_i - \mu)^2. \end{aligned}$$

Estimating μ

Differentiate $\ell(\mu, \sigma^2)$ with respect to μ:

$$\frac{\partial \ell}{\partial \mu} = -\frac{1}{\sigma^2} \sum_{i=1}^{n} (x_i - \mu).$$

Set the derivative equal to zero to find the maximum:

$$-\frac{1}{\sigma^2} \sum_{i=1}^{n} (x_i - \mu) = 0.$$

Simplify:

$$\sum_{i=1}^{n} (x_i - \mu) = 0 \implies \sum_{i=1}^{n} x_i - n\mu = 0.$$

Solve for μ:

$$n\mu = \sum_{i=1}^{n} x_i \implies \hat{\mu} = \frac{1}{n} \sum_{i=1}^{n} x_i.$$

Estimating σ^2

Differentiate $\ell(\mu, \sigma^2)$ with respect to σ^2:

$$\frac{\partial \ell}{\partial \sigma^2} = -\frac{n}{2} \cdot \frac{1}{\sigma^2} + \frac{1}{2\sigma^4} \sum_{i=1}^{n} (x_i - \mu)^2.$$

Set the derivative equal to zero:

$$-\frac{n}{2\sigma^2} + \frac{1}{2\sigma^4} \sum_{i=1}^{n} (x_i - \mu)^2 = 0.$$

Multiply both sides by $2\sigma^4$:

$$-n\sigma^2 + \sum_{i=1}^{n} (x_i - \mu)^2 = 0.$$

Solve for σ^2:

$$n\sigma^2 = \sum_{i=1}^{n} (x_i - \mu)^2 \implies \hat{\sigma}^2 = \frac{1}{n} \sum_{i=1}^{n} (x_i - \hat{\mu})^2.$$

Conclusion

The maximum likelihood estimates are:

$$\hat{\mu} = \frac{1}{n} \sum_{i=1}^{n} x_i,$$

$$\hat{\sigma}^2 = \frac{1}{n} \sum_{i=1}^{n} (x_i - \hat{\mu})^2.$$

Therefore, the MLE of μ is the sample mean, and the MLE of σ^2 is the sample variance (unbiased by n).

Question 4

We are given that $X_1, X_2, \ldots, X_n$ are independent and identically distributed (i.i.d.) Poisson random variables with parameter λ.

Our goal is to find the maximum likelihood estimate (MLE) of λ.

Writing the Likelihood Function

The likelihood function $L(\lambda)$ based on the observed data $x_1, x_2, \ldots, x_n$ is:

$$L(\lambda) = \prod_{i=1}^{n} P(X_i = x_i \mid \lambda) = \prod_{i=1}^{n} \left(\frac{\lambda^{x_i} e^{-\lambda}}{x_i!} \right)$$

Simplify the likelihood function:

$$\begin{aligned} L(\lambda) &= \left(\prod_{i=1}^{n} \lambda^{x_i} \right) e^{-n\lambda} \left(\prod_{i=1}^{n} \frac{1}{x_i!} \right) \\ &= \lambda^{\sum_{i=1}^{n} x_i} e^{-n\lambda} \left(\prod_{i=1}^{n} \frac{1}{x_i!} \right) \end{aligned}$$

Since the term $\prod_{i=1}^{n} \frac{1}{x_i!}$ does not depend on λ, it can be considered a constant with respect to λ during the maximization.

Computing the Log-Likelihood Function

Taking the natural logarithm of the likelihood function to simplify differentiation:

$$\begin{aligned} \ell(\lambda) &= \ln L(\lambda) \\ &= \left(\sum_{i=1}^{n} x_i \right) \ln \lambda - n\lambda + \ln \left(\prod_{i=1}^{n} \frac{1}{x_i!} \right) \\ &= \left(\sum_{i=1}^{n} x_i \right) \ln \lambda - n\lambda + \text{constant} \end{aligned}$$

Again, the constant term $\ln \left(\prod_{i=1}^{n} \frac{1}{x_i!} \right)$ does not depend on λ and can be ignored during maximization.

Let $S = \sum_{i=1}^{n} x_i$.

Differentiating and Setting the Derivative to Zero

Differentiate $\ell(\lambda)$ with respect to λ:

$$\frac{d\ell}{d\lambda} = \frac{S}{\lambda} - n$$

Set the derivative equal to zero to find the maximum:

$$\frac{S}{\lambda} - n = 0$$

Solving for λ

Solving for λ:

$$\frac{S}{\lambda} = n \implies \lambda = \frac{S}{n} = \frac{1}{n}\sum_{i=1}^{n} x_i$$

Conclusion

The maximum likelihood estimate of λ is:

$$\hat{\lambda} = \frac{1}{n}\sum_{i=1}^{n} x_i$$

That is, the MLE of λ is the sample mean of the observed data.

Question 5

Given that $Y \mid X \sim N(\mu(X), \sigma^2(X))$, the probability density function (PDF) of Y for a single observation (X_i, Y_i) is:

$$p(Y_i \mid X_i) = \frac{1}{\sqrt{2\pi\sigma^2(X_i)}} \exp\left(-\frac{(Y_i - \mu(X_i))^2}{2\sigma^2(X_i)}\right).$$

The likelihood for the entire dataset $\{(X_i, Y_i)\}_{i=1}^N$ is:

$$L(\{(X_i, Y_i)\}_{i=1}^N) = \prod_{i=1}^{N} \frac{1}{\sqrt{2\pi\sigma^2(X_i)}} \exp\left(-\frac{(Y_i - \mu(X_i))^2}{2\sigma^2(X_i)}\right).$$

Taking the negative log-likelihood to obtain the loss function:

$$\mathcal{L} = -\log L(\{(X_i, Y_i)\}_{i=1}^N)$$

This results in:

$$\mathcal{L} = -\sum_{i=1}^{N} \left(-\frac{1}{2}\log(2\pi\sigma^2(X_i)) - \frac{(Y_i - \mu(X_i))^2}{2\sigma^2(X_i)}\right)$$

Simplifying:

$$\mathcal{L} = \sum_{i=1}^{N} \left(\frac{\log(2\pi\sigma^2(X_i))}{2} + \frac{(Y_i - \mu(X_i))^2}{2\sigma^2(X_i)}\right)$$