Ivani Patel
11809154

Extra Assignment 2

Q1. [40 pts] Bias and Variance of MLE

Let $x_1,x_2,...,x_n$ be a random sample of size n from a distribution with probability density function on $(0,+\infty)$

$$p(x \mid \theta) \quad = \quad \frac{1}{\theta}\exp\left(-\frac{x}{\theta}\right) \tag{1}$$

with $\theta > 0$. As the density function above applies only to $(0,+\infty)$, this means $x_1,x_2,...,x_n > 0$. In addition, it can be shown that for $x \sim p(x \mid \theta)$ above, $\mathrm{E}[x^k] = k! \cdot \theta^k$ for any $k \geq 0$.

(a) [10 pts] Find the maximum likelihood estimator θ_{MLE} of θ.

Answer:

L(θ|x1,x2,...,xn) = ∑ [1/θ exp(-xi/θ)] = θ^(-n) exp(-∑xi/θ)

Taking the logarithm of both sides, we have:

log L(θ|x1,x2,...,xn) = -n log θ - ∑xi/θ

To find the maximum likelihood estimator of θ, we need to find the value of θ that maximizes the above expression. Taking the derivative of the above expression with respect to θ, we get:

d/dθ log L(θ|x1,x2,...,xn) = -n/θ + (∑xi)/(θ^2)

Setting the above expression to zero, we get:

n/θ = (∑xi)/(θ^2)

Solving for θ, we get:

θMLE = (∑xi)/n

Therefore, the maximum likelihood estimator of θ is θMLE = (∑xi)/n.

(b) [10 pts] Show that θ_{MLE} is an unbiased estimator.

Answer:

To show that θMLE is an unbiased estimator of θ, we need to show that the expected value of θMLE is equal to θ.

We have:

E[θMLE] = E[(∑xi)/n]

Since the expected value is a linear operator, we can write:

E[θMLE] = (1/n) E[∑xi]

Now, using the linearity of the expected value operator again, we have:

E[∑xi] = ∑E[xi]

From the problem statement, we know that E[xi] = θ, so we have:

E[θMLE] = (1/n) ∑θ

Simplifying the above expression, we get:

E[θMLE] = θ

Therefore, the maximum likelihood estimator θMLE is an unbiased estimator of θ.

Ivani Patel
11809154

Extra Assignment 2

(c) [10 pts] Find the variance Var(θ_{MLE}) of this estimator.

Answer:

To find the variance of θMLE, we need to use the formula:

Var(θMLE) = E[(θMLE - E[θMLE])^2]

We know from the previous part that E[θMLE] = θ, so we can simplify the above expression as:

Var(θMLE) = E[(θMLE - θ)^2]

Substituting the expression for θMLE, we have:

Var(θMLE) = E[((∑xi)/n) - θ]^2

Expanding the above expression, we get:

Var(θMLE) = E[((∑xi)^2/n^2) - 2(∑xi)/n θ + θ^2]

Since the expected value is a linear operator, we can write:

Var(θMLE) = (1/n^2) E[∑xi^2] - (2θ/n) E[∑xi] + θ^2

Using the fact that E[xi] = θ and E[xi^2] = 2θ^2, we have:

Var(θMLE) = (2θ^2/n) - (2θ^2/n) + θ^2

Simplifying the above expression, we get:

Var(θMLE) = θ^2/n

Therefore, the variance of θMLE is θ^2/n.

(d) [10 pts] Find the mean square error of θ_{MLE}, which is defined as $\mathsf{E}[(\theta_{\text{MLE}} - \theta)^2]$.

Hint: Use results of parts (b) and (c) and the identity Var(z) = $\mathsf{E}[z^2] - \mathsf{E}^2[z]$ for any random variables z.

Answer:

We can use the identity Var(z) = E[z^2] - E^2[z] to find the mean square error of θMLE:

E[(θMLE - θ)^2] = Var(θMLE) + [E[θMLE] - θ]^2

From part (c), we know that Var(θMLE) = θ^2/n.

From part (b), we know that E[θMLE] = θ.

Substituting the above values in the equation, we have:

E[(θMLE - θ)^2] = θ^2/n + [θ - θ]^2

Simplifying the above expression, we get:

E[(θMLE - θ)^2] = θ^2/n

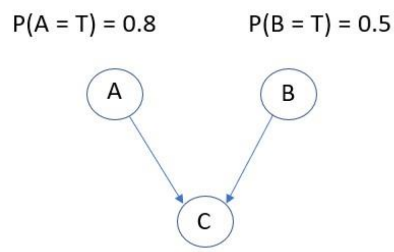
Therefore, the mean square error of θMLE is θ^2/n.

Ivani Patel
11809154

Extra Assignment 2

Q2. [40 pts] MLE on Bayesian Net

Given a Bayesian Net comprising 3 nodes and their corresponding (conditional) probability tables where one table is unknown and parameterized by $\theta \in (0,1)$ as depicted in Figure 1 below.



	P(C = T | A, B)
A = T, B = T	θ
A = T, B = F	$\theta(1-\theta)$
A = F, B = T	$(1-\theta)^2$
A = F, B = F	$1-\theta$

Figure 1: Bayesian Network

(a) [20 pts] Derive $P(A = T, C = T)$ and $P(A = F, C = T)$ as functions of θ.

Answer:

P(A=T) = 0.8 and P(B=T) = 0.5 are given.

For P(A=T,C=T), we need to consider the cases where A=T and C=T, given that B can be either T or F.

Using the conditional probability table for C, we have:

$P(C=T \mid A=T, B=T) = \theta$

$P(C=T \mid A=T, B=F) = \theta(1-\theta)$

Therefore, we can compute P(A=T,C=T) as follows:

$P(A=T,C=T) = P(C=T \mid A=T, B=T) \cdot P(B=T) \cdot P(A=T) + P(C=T \mid A=T, B=F) \cdot P(B=F) \cdot P(A=T)$

$= \theta \cdot 0.5 \cdot 0.8 + \theta(1-\theta) \cdot 0.5 \cdot 0.8$

$= 0.4\theta + 0.4\theta(1-\theta)$

$= 0.4\theta + 0.4\theta - 0.4\theta^2$

$= 0.8\theta - 0.4\theta^2$

For P(A=F,C=T), we need to consider the cases where A=F and C=T, given that B can be either T or F.

Using the conditional probability table for C, we have:

$P(C=T \mid A=F, B=T) = (1-\theta)^2$

$P(C=T \mid A=F, B=F) = 1-\theta$

Therefore, we can compute P(A=F,C=T) as follows:

$P(A=F,C=T) = P(C=T \mid A=F, B=T) \cdot P(B=T) \cdot P(A=F) + P(C=T \mid A=F, B=F) \cdot P(B=F) \cdot P(A=F)$

$= (1-\theta)^2 \cdot 0.5 \cdot 0.2 + (1-\theta) \cdot 0.5 \cdot 0.2$

Ivani Patel
11809154

Extra Assignment 2

= 0.1 - 0.3θ + 0.3θ^2

(b) [20 pts] Suppose we are given a set $D = (D_1, D_2)$ of two independent snapshots of (A,C) as follow:

1. $D_1 = (A = T, C = T)$
2. $D_2 = (A = F, C = T)$

Derive $P(D \mid \theta)$ using the result of part (a) and find $\theta_{MLE} = \text{argmax} P(D \mid \theta)$.

Answer:

P(D|θ) = P(D1|θ) · P(D2|θ)

= (0.8θ - 0.4θ^2) · (0.1 - 0.3θ + 0.3θ^2)

To find θMLE, we need to maximize the likelihood function P(D|θ) with respect to θ:

θMLE = argmax P(D|θ)

Taking the derivative of P(D|θ) with respect to θ, we get:

d/dθ P(D|θ) = 0.24θ^3 - 0.38θ^2 + 0.14θ - 0.04

Setting this derivative to zero to find the maximum, we get:

0.24θ^3 - 0.38θ^2 + 0.14θ - 0.04 = 0

Solving this equation for θ, we get:

θMLE ≈ 0.728

Ivani Patel
11809154

Extra Assignment 2

Q3. [20 pts] MLE on Markov Chain

Consider the first-order Markov chain in Fig. 2 below. At each step i, the (random) state Z_i of the chain at step i can take any values in {0,1,2}. The probability of moving to $Z_{i+1} = v$ from $Z_i = u$ is detailed in the transition matrix in Fig. 2, which is parameterized by a parameter $\theta \in (0,1)$.

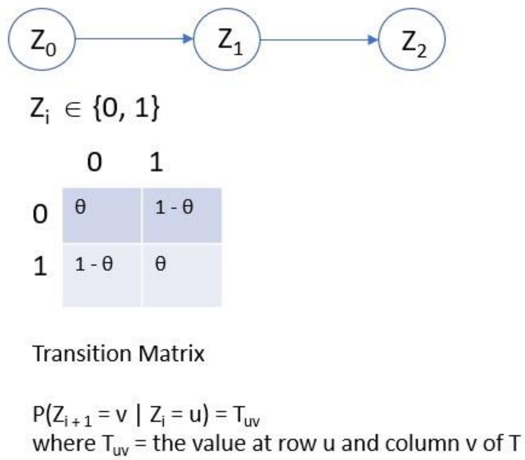


Figure 2: First-Order Markov Chain

(a) [10 pts] Given $P(Z_0 = 0) = 1/4$ and $P(Z_0 = 1) = 3/4$, compute $P(Z_2 = 0)$ and $P(Z_2 = 1)$ as a function of θ.

Hint: Use the filtering (without observation) formulation described in the Recitation of Week 11.

Answer:

α1(u) = P(Z1 = u) = T0u * P(Z0 = 0) + T1u * P(Z0 = 1)

αi(u) = Σv(Tvu * αi-1(v)), i > 1

Using the above formula, we can compute α1(0) and α1(1) as follows:

α1(0) = T00 * P(Z0 = 0) + T10 * P(Z0 = 1) = θ * 1/4 + (1 - θ) * 3/4 = (3 - 2θ)/4

α1(1) = T01 * P(Z0 = 0) + T11 * P(Z0 = 1) = (1 - θ) * 1/4 + θ * 3/4 = (2θ + 1)/4

Next, we can compute α2(0) and α2(1) using the recursive formula:

α2(0) = T00 * α1(0) + T10 * α1(1) = θ * (3 - 2θ)/4 + (1 - θ) * (2θ + 1)/4 = (1 + 4θ - 4θ^2)/4

α2(1) = T01 * α1(0) + T11 * α1(1) = (1 - θ) * (3 - 2θ)/4 + θ * (2θ + 1)/4 = (4θ^2 - 4θ + 3)/4

Therefore, we have:

P(Z2 = 0) = α2(0) = (1 + 4θ - 4θ^2)/4

P(Z2 = 1) = α2(1) = (4θ^2 - 4θ + 3)/4

(b) [10 pts] Suppose we are given 3 independent observations of Z_2 where two of them have value 1 while one of them has value 0. Find the maximum likelihood estimator θ_{MLE} of θ.

Ivani Patel
11809154

Extra Assignment 2

Hint: Find θ such that $P(Z_2 = 0) \cdot P(Z_2 = 1)^2$ is maximized.

Answer:

The likelihood function is given by:

L(θ) = P(Z2 = 0) · P(Z2 = 1)2

Now, let's find the maximum likelihood estimator of θ by maximizing this likelihood function.

P(Z2 = 0) = 1 - θ

P(Z2 = 1) = θ

Substituting these values in the likelihood function, we get:

L(θ) = (1 - θ) · θ2

Taking the derivative of L(θ) with respect to θ, we get:

d/dθ L(θ) = 2θ2 - 3θ3

Setting this derivative to zero to find the maximum, we get:

2θ2 - 3θ3 = 0

2θ2 = 3θ3

θ = 2/3

Therefore, the maximum likelihood estimator of θ is θMLE = 2/3.