EE 503

Homework 14 Solution

(Not to be handed in for grading)

Work all 3 problems.

Problem 1. Let $X = (X_1, X_2, \ldots, X_n)$ be i.i.d. where each $X_i \sim N(\mu, \sigma^2)$ where both μ and σ^2 are unknown. Find the MLE for μ and σ^2.

Solution: We compute the likelihood function as

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

and then

$$\log L(\mu, \sigma^2 | x) = -\frac{n}{2}\log 2\pi - \frac{n}{2}\log \sigma^2 - \frac{1}{2\sigma^2}\sum_{i=1}^{n}(x_i - \mu)^2.$$

The partial derivatives w.r.t. the unknowns are

$$\frac{\partial}{\partial \mu}\log L(\mu, \sigma^2 | x) = \frac{1}{\sigma^2}\sum_{i=1}^{n}(x_i - \mu)$$

and

$$\frac{\partial}{\partial \sigma^2}\log L(\mu, \sigma^2 | x) = -\frac{n}{2\sigma^2} + \frac{1}{2\sigma^4}\sum_{i=1}^{n}(x_i - \mu)^2.$$

Setting these partial derivatives equal to zero and solving yields the solutions

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

To verify that we have found a global maximum we can use the fact that if $\bar{\mu} \neq \bar{X}$ then $\sum(x_i - \theta)^2 > \sum(x_i - \bar{X})^2$. Hence, for any value of σ^2

$$\frac{1}{(2\pi\sigma^2)^{n/2}} \exp\left(-\frac{1}{2\sigma^2}\sum_{i=1}^{n}(x_i - \bar{X})^2\right) \geq \frac{1}{(2\pi\sigma^2)^{n/2}} \exp\left(-\frac{1}{2\sigma^2}\sum_{i=1}^{n}(x_i - \bar{\mu})^2\right).$$

Therefore, verifying we have found the MLE reduces to a one-dimensional problem in verifying that $(\sigma^2)^{-n/2}\exp\left(-\frac{1}{2\sigma^2}\sum(x_i - \bar{X})^2\right)$ achieves its global maximum at $\sigma^2 = n^{-1}\sum(x_i - \bar{X})^2$ which is straightforward using univariate calculus.

Problem 2. Text 11.10 (modified). An information source generates i.i.d. bits X_n for which $P(X_n = 0) = P(X_n = 1) = 1/2$.

a. Suppose X_n is transmitted over a binary symmetric channel (BSC) with probability of error $= p$ (a BSC just means it is just as likely a 0 is received as a 1 as it is a 1 is received as a 0). Find the probabilities of the possible outputs of the channel.

   Solution: Let O denote output and I denote input. Then

$$\begin{aligned} P(O=0) &= P(O=0|I=0)P(I=0) + P(O=0|I=1)P(I=1) \\ &= (1-p)\frac{1}{2} + p\frac{1}{2} = \frac{1}{2} \end{aligned}$$

   and

$$\begin{aligned} P(O=1) &= P(O=1|I=0)P(I=0) + P(O=1|I=1)P(I=1) \\ &= p\frac{1}{2} + (1-p)\frac{1}{2} = \frac{1}{2}. \end{aligned}$$

b. Suppose X_n is transmitted over K consecutive identical and independent BSCs. Does the sequence of channel outputs form a Markov chain?

   Solution: The one-step transition probability matrix is

$$P = \begin{bmatrix} 1-p & p \\ p & 1-p \end{bmatrix}.$$

   Note that in a system of K independent BSCs

$$P(O_{n+1} = o_{n+1}|O_n = o_n, \ldots, O_0 = o_0) = P(O_{n+1} = o_{n+1}|O_n = o_n)$$

   thus we have a Markov chain.

c. Find the K-step transition probabilities that relate the input bits from the source to the output bits of the Kth channel.

Solution: We find

$$P^K = \begin{bmatrix} 1-p & p \\ p & 1-p \end{bmatrix}^K.$$

To compute this we use $P^K = E\Lambda^K E^{-1}$ where E is the matrix of eigenvectors and Λ is the matrix of eigenvalues. We find

$$P^K = \begin{bmatrix} \frac{1}{2} + \frac{1}{2}(1-2p)^K & \frac{1}{2} - \frac{1}{2}(1-2p)^K \\ \frac{1}{2} - \frac{1}{2}(1-2p)^K & \frac{1}{2} + \frac{1}{2}(1-2p)^K \end{bmatrix}.$$

d. What are the probabilities of the possible outputs of the Kth channel as $K \to \infty$? We get

- $$P^K \to \begin{bmatrix} \frac{1}{2} & \frac{1}{2} \\ \frac{1}{2} & \frac{1}{2} \end{bmatrix}$$

as $K \to \infty$ for $|1-2p| < 1$. Hence, the limiting value of the outputs (0 and 1) each occur with probability 1/2.

Problem 3. Here we consider a Heuristic estimator. Suppose you have available to you outputs of an algorithm that produces a random variable that has a noncentral chi-square distribution with two degrees of freedom with mean and variance

$$\begin{aligned} E[X] &= 2\sigma^2 + s^2 \\ Var[X] &= 4\sigma^4 + 4\sigma^2 s^2. \end{aligned}$$

You have available to you n samples of this distribution.You wish to estimate s^2. Can you think of a way to do this from observing the mean and variance expressions?

Solution: Observe

$$(E[X])^2 - Var[x] = s^4$$

so we can estimate the moments empirically using the n samples and combine these statistics this way to get an estimate of s^4 and then take the square root to get an estimate of s^2.