TEL502E – Homework 3

Due 22.03.2016

1. Consider a discrete random variable X whose probability mass function (pmf) depends on a parameter θ, where $\theta \in \{0, 1, 2\}$. Suppose that X takes values in $\{0, 1, 2, 3\}$ and its pmf for different values of θ, denoted by $P(x|\theta)$, is as given below.

x	$P(x\|\theta = 0)$	$P(x\|\theta = 1)$	$P(x\|\theta = 2)$
0	1/8	1/4	0
1	1/4	1/2	1/3
2	3/8	1/8	1/3
3	1/4	1/8	1/3

(a) Suppose we are given a realization of X as $x = 1$. Find the maximum likelihood estimate (MLE) for θ.

(b) Suppose we are given two independent realizations of X as $x_1 = 1$, $x_2 = 2$. Find the MLE for θ.

Solution. (a) Note that the likelihood function in this case is $L(\theta) = P(1|\theta)$. According to the table, $L(\theta)$ is maximized for $\hat{\theta} = 1$ (where $L(\hat{\theta}) = 1/2$).

(b) Thanks to independence, the likelihood function is given as $L(\theta) = P(1|\theta)\, P(2|\theta)$. We then have $L(0) = 3/32$, $L(1) = 1/16$, $L(2) = 1/9$. Thus, the ML estimate is $\hat{\theta} = 2$.

2. Suppose $X_1, X_2, \ldots, X_n$ are independent and identically distributed random variables with pdf

$$f_{X_i}(t) = \begin{cases} 0, & \text{if } t < 0, \\ \theta^{-t} \ln(\theta), & \text{if } t \geq 0, \end{cases}$$

where $\theta > 1$ is an unknown constant.

(a) Find the maximum likelihood estimator for θ in terms of $X_1, X_2, \ldots, X_n$.

(b) Specify whether the estimator you found is biased or not.
(Hint : $\int_0^\infty x\, c^{-x}\, dx = \big(\ln(c)\big)^2$, if $c > 1$.)

Solution. (a) Given $X_i = x_i > 0$, the likelihood function is given as,

$$L(\theta) = \big(\ln(\theta)\big)^n \, \theta^{-\left(\sum_i x_i\right)}.$$

Setting the derivative of the log-likelihood function to zero, we find that the maximizer of this expression satisfies,

$$\frac{n}{\ln(\theta)} \frac{1}{\theta} - \left(\sum_i x_i\right) \theta = 0.$$

Solving for θ, we find the ML estimate as $\exp(n/\sum_i x_i)$. Therefore, the ML estimator is,

$$\hat{\theta} = \exp\left(\frac{n}{\sum_i X_i}\right).$$

(b) First notice that, by the provided hint, $\mathbb{E}(X_i) = 1/\ln(\theta)$. Therefore,

$$\mathbb{E}\left(\frac{\sum_{i=1}^n X_i}{n}\right) = \frac{1}{\ln(\theta)}.$$

Recall that Jensen's inequality states that if f is a strictly convex function and X is a continuous random variable, then

$$f\left(\mathbb{E}(X)\right) < \mathbb{E}\big(f(X)\big).$$

Observe that for $t > 0$, $f(t) = \exp(1/t)$ is a strictly convex function. Therefore, it follows that

$$\mathbb{E}(\hat{\theta}) = \mathbb{E}\left(f\left(\frac{\sum_{i=1}^n X_i}{n}\right)\right) > f\left(\mathbb{E}(X)\right) = \theta.$$

Therefore, $\hat{\theta}$ is a biased estimator.

3. Let X_1, X_2 be independent Gaussian random variables with mean θ and variance 1. Also, let θ be a random variable uniformly distributed on $[0,1]$ – that is, the pdf of θ is given by,

$$f_\theta(t) = \begin{cases} 1, & \text{if } t \in [0,1], \\ 0, & \text{if } t \notin [0,1]. \end{cases}$$

(a) Find the joint pdf of θ, X_1, X_2. That is, find $f_{\theta,X_1,X_2}(t, x_1, x_2)$.

(b) Find the maximum a posteriori (MAP) estimate of θ.

(c) Evaluate the estimator you found in part (b) if the data is as given below.

(i) $x_1 = 3/4$, $x_2 = 1$.

(ii) $x_1 = 1/2$, $x_2 = 2$.

Solution. (a) The joint pdf is given as,

$$\begin{aligned} f_{X_1,X_2,\Theta}(x_1,x_2,t) &= f_{X_1,X_2|\Theta}(x_1,x_2|t)\, f_\theta(t) \\ &= \begin{cases} \dfrac{1}{2\pi} \exp\left(-\dfrac{(x_1-t)^2+(x_2-t)^2}{2}\right), & \text{if } 0 \le t \le 1, \\ 0, & \text{otherwise.} \end{cases} \\ &= \begin{cases} \dfrac{1}{2\pi} \exp\left(-\dfrac{x_1^2+x_2^2}{2}\right) \cdot \exp\left(\dfrac{x_1+x_2}{2}t - t^2\right), & \text{if } 0 \le t \le 1, \\ 0, & \text{otherwise.} \end{cases} \end{aligned}$$

(b) Notice that for fixed x_1, x_2, we need to maximize the term $\exp\left(\frac{x_1+x_2}{2}t - t^2\right)$ subject to $t \in [0,1]$. This is equivalent to minimizing $t^2 - \frac{x_1+x_2}{2}t$ with respect to $t \in [0,1]$. But this is a quadratic with a minimum at $(x_1+x_2)/2$. Therefore, the MAP estimate is given as,

$$\hat{t} = \begin{cases} 0, & \text{if } \dfrac{x_1+x_2}{2} < 0, \\ \dfrac{x_1+x_2}{2}, & \text{if } 0 \le \dfrac{x_1+x_2}{2} \le 1, \\ 1, & \text{if } 1 < \dfrac{x_1+x_2}{2}. \end{cases}$$

(c) (i) $\hat{t} = 7/8$. (ii) $\hat{t} = 1$.

4. Suppose we observe $X = \theta + Z$, where θ and Z are independent random variables. Suppose also that θ is uniformly distributed over the unit interval and Z is a standard normal random variable (i.e., $\mathcal{N}(0,1)$). That is, the pdfs of θ and Z are,

$$\begin{aligned} f_\theta(t) &= u(t)\, u(1-t), \\ f_Z(z) &= \frac{1}{\sqrt{2\pi}} e^{-z^2/2}, \end{aligned}$$

where u denotes the step function.

(a) Find the joint pdf of X and θ, that is, $f_{X,\theta}(x,t)$.

(b) Find the maximum a posteriori (MAP) estimator for θ in terms of X.

(c) Evaluate the estimator you found in part (b) if the observation is given as

(c.1) $x = 1/4$,

(c.2) $x = -1$,

(c.3) $x = 2$.

Solution. (a) Notice that $f_{X|\theta}(x|t) = f_Z(x-t)$. Therefore, the joint pdf of X and θ is obtained as,

$$f_{X,\theta}(x,t) = f_{X|\theta}(x|t)\, f_\theta(t) = \begin{cases} \dfrac{1}{\sqrt{2\pi}} \exp\left(-\dfrac{1}{2}(x-t)^2\right), & \text{if } 0 \le t \le 1, \\ 0, & \text{otherwise.} \end{cases}$$

(b) For fixed x, the joint pdf is maximized for

$$\hat{t} = \begin{cases} 0, & \text{if } x < 0, \\ x, & \text{if } 0 \le x \le 1, \\ 1, & \text{if } 1 < x. \end{cases}$$

(c) Evaluating the estimator, we find, (c.1) $\hat{t} = 1/4$, (c.2) $\hat{t} = 0$, (c.3) $\hat{t} = 1$.

5. Suppose X is a Gaussian random variable with mean θ and variance 1. Suppose θ is also a Gaussian random variable with mean 2 and variance 3.

(a) Find the pdf of X.

(b) Find the minimum mean square estimate (MMSE) of θ given X.

Solution. (a) Notice that X can be written as the sum of a standard normal random variable Z and θ, where Z and θ are independent. Since the sum of Gaussian random variables are Gaussian, X is Gaussian. Therefore, it suffices to find the mean and the variance of X. But $\mathbb{E}(X) = \mathbb{E}(Z + \theta) = 2$. Also, since Z and θ are independent, we have, $\text{var}(X) = \text{var}(Z) + \text{var}(\theta) = 4$. Thus,

$$f_X(x) = \frac{1}{2\sqrt{2\pi}} \exp\left(-\frac{1}{8}(x-2)^2\right).$$

(b) Notice that the joint pdf of X and θ is given as,

$$f_{X,\theta}(x,t) = f_{X|\theta}(x|t)\, f_\theta(t) = \frac{1}{\sqrt{2\pi}} \exp\left(-\frac{1}{2}\,(x-t)^2\right) \frac{1}{\sqrt{6\pi}} \exp\left(-\frac{1}{6}\,(t-2)^2\right).$$

We find,

$$\begin{aligned} f_{\theta|X}(t|x) &= \frac{f_{X,\theta}(x,\theta)}{f_X(x)} \\ &= c\, \exp\left(\underbrace{-\frac{1}{2}(x-t)^2 - \frac{1}{6}(t-2)^2 + \frac{1}{8}(x-2)^2}_{h(t)}\right), \end{aligned}$$

where c is a constant. Notice that the form of $f_{\theta|X}(t|x)$, for fixed x and variable t, is the same as that of a Gaussian random variable. To find the mean and variance of this random variable, it's sufficient to find the maximum of $h(t)$ and the factor that multiplies t. But note that, since we are interested in $\mathbb{E}(\theta|X)$, finding the mean is sufficient for our purposes. The mean can be found by setting the derivative of h to zero. This gives the equation,

$$-(t-x) - \frac{1}{3}(t-2) = 0.$$

Solving for t, we find $t = 3/4(x + 2/3)$. Thus,

$$\mathbb{E}(\theta|X) = \frac{3}{4}X + \frac{1}{2} = \frac{3}{4}(X-2) + 2.$$