MATH 7241 Fall 2020: Problem Set #9

Due date: Thursday December 10

Reading: Ross Section 1.6; "Information Theory, Inference and Learning Algorithms" by David J.C. MacKay (available online), Chapter 3; Ross Chapter 5; additional notes "Poisson-notes.pdf".

$\boxed{\textbf{Exercise 1}}$ A continuous random variable X is uniformly distributed over the interval $[0, \theta]$. However the parameter θ is unknown, and initially it is modeled as the value of a continuous random variable which is uniformly distributed on the interval $[0, 1]$.

a). Write down the prior distribution f_0 for θ.

b). Write down the likelihood function $f(x \mid \theta)$ (this is the pdf for X conditioned on the value θ). Be careful to specify the intervals where $f(x \mid \theta)$ is zero and where it is nonzero.

c). The random variable X is measured and gives the value $x_1 \in [0, 1]$. Let $D = \{x_1\}$ denote the data from this trial. The posterior pdf of θ is given by the Bayesian update rule

$$f_1(\theta \mid D) = \frac{f(x_1 \mid \theta)\, f_0(\theta)}{Z}$$

where the evidence is

$$Z = \int_0^1 f(x_1 \mid \theta)\, f_0(\theta)\, d\theta$$

Sketch a graph of $f(x_1 \mid \theta)\, f_0(\theta)$ as a function of θ. Again be careful to account for the intervals where $f(x_1 \mid \theta)$ is zero and where it is nonzero.

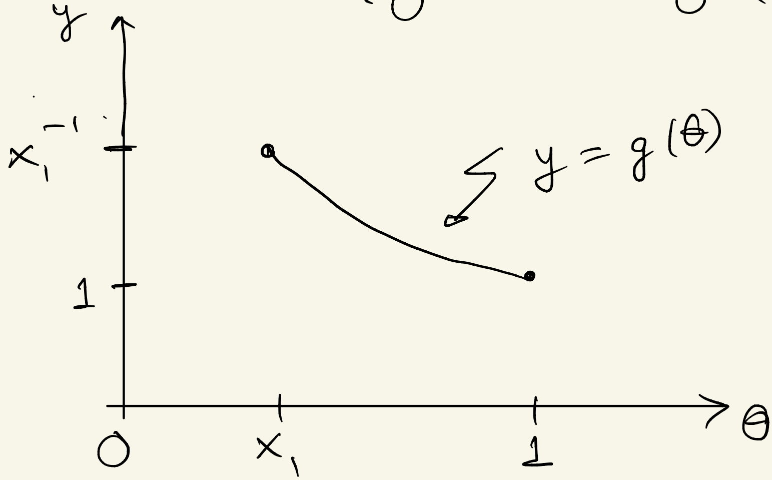
d). Evaluate the evidence Z and hence determine the posterior pdf f_1.

a) $$f_0(\theta) = \begin{cases} 1 & 0 \le \theta \le 1 \\ 0 & \text{else} \end{cases}$$

b) $f(x|\theta) = \begin{cases} \frac{1}{\theta} & 0 \le x \le \theta \\ 0 & x < 0 \text{ or } x > \theta \end{cases}$

c) let $g(\theta) = f(x_1|\theta)\, f_0(\theta)$

$$= \begin{cases} \frac{1}{\theta} & 1 \ge \theta \ge x_1 \\ 0 & \theta < x_1 \text{ or } \theta > 1 \end{cases}$$



d) $$Z = \int_0^1 f(x_1|\theta)\, f_0(\theta)\, d\theta$$

$$= \int_{x_1}^{1} \frac{1}{\theta}\, d\theta$$

$$= \ln\theta \,\Big|_{x_1}^{1}$$

$$= -\ln x_1$$

$$= \ln\left(\frac{1}{x_1}\right)$$

$$\Rightarrow \quad f_1(\theta \mid D) = \begin{cases} \dfrac{1}{\theta \ln\left(\frac{1}{x_1}\right)} & 1 \geq \theta \geq x_1 \\ 0 & \theta < x_1 \text{ or } \theta > 1. \end{cases}$$

Exercise 2 Consider the same setup as the previous problem, where X is uniform on $[0, \theta]$, and initially θ is modeled as uniform on $[0, 1]$. Suppose now that the random variable X is measured twice, giving values x_1 and x_2, so the data is $D_2 = \{x_1, x_2\}$. Following the same method, compute the posterior pdf $f_1(\theta \mid D_2)$.

$$D_2 = \{x_1, x_2\}$$

$$\Rightarrow f(D_2 \mid \theta) = f(x_1 \mid \theta)\, f(x_2 \mid \theta)$$ (b/c independent measurements)

$$= \begin{cases} \left(\frac{1}{\theta}\right)\left(\frac{1}{\theta}\right) & 0 \le x_1 \le \theta,\ 0 \le x_2 \le \theta \\ 0 & \text{else} \end{cases}$$

$$= \begin{cases} \frac{1}{\theta^2} & \max\{x_1, x_2\} \le \theta \\ 0 & \text{else} \end{cases}$$

evidence

$$Z = \int_0^1 f(D \mid \theta)\, f_0(\theta)\, d\theta$$

$$= \int_{x_{max}}^1 \frac{1}{\theta^2}\, 1.\ d\theta$$

$$= \frac{1}{x_{max}} - 1$$

$\Rightarrow$ posterior is

$$f_1(\theta \mid D) = \begin{cases} \frac{1}{\theta^2}\left(\frac{1}{x_{max}} - 1\right)^{-1} & x_{max} \leq \theta \leq 1 \\ 0 & \text{else.} \end{cases}$$

Exercise 3 Events occur according to a Poisson process with rate $\lambda = 2$ per hour.
(a) What is the probability that no event occurs between 8pm and 9pm?
(b) Starting at noon, what is the expected time at which the fourth event occurs?
(c) What is the probability that two or more events occur between 6pm and 8pm?

a) $\mathbb{P}(N(9) - N(8) = 0)$

$= \mathbb{P}(N(1) = 0)$ (memoryless)

$= e^{-\lambda(1)} = e^{-2}$

b) $\mathbb{E}[T_4] = 4\,\mathbb{E}[T_1] = \frac{4}{\lambda} = \frac{4}{2} = 2$

c) $\mathbb{P}(N(8) - N(6) \geq 2)$

$= \mathbb{P}(N(2) \geq 2)$ (memoryless)

$= 1 - \mathbb{P}(N(2) = 0) - \mathbb{P}(N(2) = 1)$

$= 1 - e^{-\lambda(2)} - \lambda(2)e^{-\lambda(2)}$

$= 1 - 5e^{-4}$

Exercise 4 Cars pass a certain point in a highway in accordance with a Poisson process with rate $\lambda = 10$ per minute. The number of passengers in the cars are independent and identically distributed, with the following distribution: if Y is the number of passengers in a car, then $P(Y = 1) = 0.4$, $P(Y = 2) = 0.3$, $P(Y = 3) = 0.2$, $P(Y = 4) = 0.1$. A car is full if it has four passengers.
a) Find the probability that the next car that passes is full.
b) Find the probability that two full cars pass in the next minute.
c) Find the expected number of passengers that pass in the next minute.
d) Find the probability that at least two passengers pass in the next ten seconds.
[You may want to use the result about thinning of Poisson processes].

a) # passenges is independent and IID

$$\Rightarrow \mathbb{P}(\text{next car is full}) = \mathbb{P}(Y=4) = 0.1$$

b) $N_4(t) =$ # full cars in $[0, t]$

Thinning result $\Rightarrow N_4 \sim$ P.P. rate $\lambda(0.1) = 1 = \lambda_4$

$$\Rightarrow \mathbb{P}(N_4(1) = 2) = \frac{(\lambda_4 t)^2}{2!} e^{-\lambda_4 t} = \frac{1}{2} e^{-1} = \frac{1}{2e}$$

c) N_1, N_2, N_3, N_4 independent P.P.'s.

$\Rightarrow \mathbb{E}$[# passenges in next minute]

$$= \mathbb{E}[N_1(1)] + 2\,\mathbb{E}[N_2(1)] + 3\,\mathbb{E}[N_3(1)] + 4\,\mathbb{E}[N_4(1)]$$

$$= \lambda_1(1) + 2\lambda_2(1) + 3\lambda_3(1) + 4\lambda_4(1)$$

$$= [(0.4) + 2(0.3) + 3(0.2) + 4(0.1)]10 \quad (\lambda = 10)$$

$$= 20$$

d) $\mathbb{P}($at least 2 passengers in 10 sec$)$

$= 1 - \mathbb{P}($zero passengers in 10 sec.$)$

$\quad - \mathbb{P}($1 passenger in 10 sec.$)$

$$= 1 - \mathbb{P}(N(\tfrac{1}{6}) = 0)$$

$$- \mathbb{P}(Y=1)\,\mathbb{P}(N(\tfrac{1}{6}) = 1)$$

$$= 1 - e^{-\lambda(\frac{1}{6})} - (0.4)\,(\lambda\tfrac{1}{6})\,e^{-\lambda(\frac{1}{6})}$$

$$= 1 - e^{-5/3} - (0.4)(\tfrac{5}{3})\,e^{-5/3}$$

$$= 1 - \tfrac{5}{3}\,e^{-5/3}$$

Exercise 5 Men and women enter a bank according to independent Poisson processes at rates $\mu = 3$ and $\lambda = 2$ per minute respectively. Starting at an arbitrary time, find the probability that at least two men arrive before the next two women arrive. [You may want to use the result about superposition of Poisson processes].

$\mathbb{P}$(at least 2 Men before next 2 Women)

$$= \mathbb{P}(MM) + \mathbb{P}(MWM) + \mathbb{P}(WMM)$$

$$\mathbb{P}(M) = \frac{\mu}{\mu+\lambda} = \frac{3}{5} \qquad P(W) = \frac{2}{5}$$

$\Rightarrow \mathbb{P}$(at least 2M before 2W)

$$= \left(\frac{3}{5}\right)^2 + \left(\frac{3}{5}\right)^2\left(\frac{2}{5}\right) + \left(\frac{3}{5}\right)^2\left(\frac{2}{5}\right)$$

$$= \left(\frac{3}{5}\right)^2\left(\frac{9}{5}\right) = \frac{81}{125}$$

Exercise 6 Customers arrive at a bank according to a Poisson process with rate λ. Suppose two customers arrived during the first hour. What is the probability that (a) both arrived during the first 20 minutes? (b) at least one arrived during the first 20 minutes? [You may want to use the result about conditional distribution of arrival times].

Given $N(1) = 2$

$\Rightarrow$ arrival times are

$$T_1 = U_{(1)}, \quad T_2 = U_{(2)}$$

where U_1, U_2 are IID uniform on $[0,1]$.

a) $P(\text{both arrived in 20 mins} \mid N(1)=2)$

$$= P(T_1 \le \tfrac{1}{3}, T_2 \le \tfrac{1}{3} \mid N(1)=2)$$

$$= P(U_1 \le \tfrac{1}{3}, U_2 \le \tfrac{1}{3})$$

$$= P(U \le \tfrac{1}{3})^2 = \frac{1}{9}.$$

b) $P(\text{at least arrived in 20 mins} \mid N(1)=2)$

$$= 1 - P(\text{none arrived in 20 mins} \mid N(1)=2)$$

$$= 1 - P(T_1 > \tfrac{1}{3}, T_2 > \tfrac{1}{3} \mid N(1) = 2)$$

$$= 1 - \left(\tfrac{2}{3}\right)^2 = \frac{5}{9}.$$