1. In question 1 of homework 1, Romeo figured out that girls are Exponential(\(\Theta\)) hours late to a date, where \(\Theta\) is a Uniform(0, 1) random variable. He then updated his prior after Juliet showed up 10 minutes late and then 30 minutes late on the first two dates. Find Romeo’s CE and MAP posterior estimates for \(\Theta\) after the first and second dates, respectively.

**Solution:** Let \(X_1\) be the time of Juliet’s arrival in hours in the first date. From homework 1 we know that:

\[
f_{\Theta|X_1}(\theta|X_1 = 1/6) = \frac{\theta e^{-\theta/6}}{\int_0^1 \theta e^{-\theta/6}} \approx 2.232 \theta e^{-\theta/6}, 0 \leq \theta \leq 1
\]

Therefore the CE estimate is given by:

\[
CE_1 = E[\Theta|X_1 = 1/6] = \int_0^1 \theta \cdot 2.232 \theta e^{-\theta/6} d\theta \approx 0.657
\]

For the MAP estimate we need to find a value \(0 \leq \theta \leq 1\) that maximizes \(\theta e^{-\theta/6}\). We can find the critical points by taking derivatives: \(d(\theta e^{-\theta/6})/d\theta = (1 - \theta/6)e^{-\theta/6}\), so the only critical point is \(\theta = 6\), which is outside the range of interest. So the MAP estimate must be one of the endpoints of the range, in this case \(MAP_1 = 1\).

After the second date,

\[
f_{\Theta|X_1,X_2}(\theta|1/6, 1/2) \approx 4.904 \theta^2 e^{-2\theta/3}
\]

The CE estimate is given by:

\[
CE_{1,2} = E[\Theta|X_1 = 1/6, X_2 = 1/2] = \int_0^1 \theta \cdot 4.904 \theta^2 e^{-2\theta/3} d\theta \approx 0.724
\]

The MAP estimate is given by \(0 \leq \theta \leq 1\) that maximizes \(\theta^2 e^{-2\theta/3}\). The critical points are 0 and 3, so by the same reasoning as before \(MAP_{1,2} = 1\).

2. In a group of ten people, including Alice and Bob, each pair is friends with probability \(P\), independently of the other pairs.

(a) Let \(A\) be the number of Alice’s friends and \(B\) be the number of Bob’s friends in the group. Conditioned on \(P = p\), what kind of random variables are \(A\) and \(B\)? Are they independent?

**Solution:** The PMF of \(A\) conditioned on \(P\) is:

\[
P_{A|P}(A = a|P = p) = \binom{10 - 1}{a} p^a (1 - p)^{10-1-a}
\]

Therefore \(A\) is a binomial random variable and so is \(B\). However they are not independent conditioned on \(P\): \(P_{A,B|P}(A = 9, B = 0|P = p) = 0\) (Alice is a friend of all other nine people, contradicting the fact that Bob has no friend) while \(P(A = 9|P = p) \cdot P(B = 0|P = p) > 0\).

---

1. Recall that an Exponential(\(\theta\)) random variable has mean \(1/\theta\), so Romeo’s prior posits that on average, even the most punctual girl will be at least an hour late. After seeing Juliet show up only 10 minutes late, Romeo’s MAP estimate is as optimistic as it can be.

2. This is the Erdős-Rényi random graph model.
(b) Now suppose $P$ is unknown. Alice counts five friends in the group. What is her MAP estimate of $P$ assuming a Uniform(0, 1) prior?

**Solution:** $P$ given $A = 5$ is a Beta($5 + 1, 4 + 1$) random variable. The MAP estimate for such a random variable is $5/(5 + 4) = 5/9$.

(c) Bob, who is one of Alice’s friends, tells her that he has only one other friend in the group. How does this information affect Alice’s MAP estimate of $P$?

**Solution:** Even though $A$ and $B$ are dependent, the random variable $C$ that counts the total number of friendships that involve Alice or Bob is a Binomial($17, P$) random variable, so $P$ conditioned on $C = 5 + 1 = 6$ is a Beta($6 + 1, 11 + 1$) random variable. The MAP estimate is $6/17$.

3. The TAs of ENGG2780A want to estimate the hardness of the problem that they prepared for an upcoming quiz. They know from experience that the conditional PDF of the time $X$ it takes a TA to solve the problem (in minutes) given the hardness $\Theta$ is

$$f_{X|\Theta}(x|\Theta) \propto \begin{cases} e^{-\lambda(\Theta)x}, & \text{if } 5 \leq x \leq 60 \\ 0, & \text{otherwise,} \end{cases}$$

where when $\Theta = 1$ (hard problem) $\lambda(1) = 0.04$, and when $\Theta = 2$ (easy problem) $\lambda(2) = 0.16$.

(a) Assume that the prior probability that the problem is hard is 0.3. Given that a TA’s solution time was 20 minutes, which hypothesis will they accept and what will be the probability of error?

**Solution:** The posterior PMF of $\Theta$ given $X = 20$ is:

$$P(\Theta = 1|X = 20) \propto f_{X|\Theta}(X = 20|\Theta = 1)P(\Theta = 1)$$

$$\propto e^{-0.04\cdot20} \cdot 0.3 = 0.3e^{-0.8} \approx 0.135$$

$$P(\Theta = 2|X = 20) \propto f_{X|\Theta}(X = 20|\Theta = 2)P(\Theta = 2)$$

$$\propto e^{-0.16\cdot20} \cdot (1 - 0.3) = 0.7e^{-3.2} \approx 0.029$$

Therefore they accept that it is a hard problem ($MAP = 1$). The probability of error is

$$P(MAP \neq \Theta) = P(\Theta = 2) = 0.029/(0.135 + 0.029) \approx 0.177.$$ 

(b) All five TAs solve the problem and the recorded solution times are 10, 25, 15, and 35 minutes. Which hypothesis will they accept now and what will be the probability of error? [Adapted from textbook problem 8.2.6]

**Solution:** Let $X_1, X_2, \ldots, X_5$ be the solution time of the five TAs.

$$P(\Theta = 1|X_1 = 20, X_2 = 10, \ldots, X_5 = 35) \propto f_{X|\Theta}(20|1) \cdots f_{X|\Theta}(35|1)P(\Theta = 1)$$

$$\propto e^{-0.04\cdot(20+10+25+15+35)} \cdot 0.3 \approx 4.5 \cdot 10^{-3}$$

Similarly,

$$f_{\Theta|X_1, X_2, \ldots, X_5}(\Theta = 2|x_1, x_2, \ldots, x_5) \propto e^{-0.16\cdot(20+10+25+15+35)} \cdot (1 - 0.3) \approx 3.5 \cdot 10^{-8}$$

Again $MAP = 1$. As $4.5 \cdot 10^{-3}$ is much bigger than $3.5 \cdot 10^{-8}$, the probability of error is around $3.5 \cdot 10^{-8}/4.5 \cdot 10^{-3} \approx 7.8 \cdot 10^{-6}$. The TAs are now much more confident that the problem is hard.
4. Let $X$ be a uniform $\text{Uniform}(0, 2^\Theta)$ random variable with unknown $\Theta$. Your prior on $\Theta$ is $\text{Geometric}(1/2)$.

(a) What is your MAP estimate of $\Theta$ given that you observed $X = 6.18$? What is the estimation error?

**Solution:** From observation $X = 6.18$ we know that $\Theta$ must be at least 3, so

$$f_{\Theta|X}(\theta|6.18) \propto f_{X|\Theta}(6.18|\theta) f_\Theta(\theta) = \frac{1}{2^\theta} \cdot \frac{1}{2^\Theta} = 4^{-\theta}$$

for every $\theta \geq 3$. (The first $1/2^\theta$ represents the PDF value of a $\text{Uniform}(0, 2^\Theta)$ random variable at 6.18, while the second one represents the PMF value of a $\text{Geometric}(1/2)$ random variable at $\theta$.) By the total probability theorem,

$$f_X(6.18) = \sum_{\theta=3}^{\infty} f_{X|\Theta}(6.18|\theta) f_\Theta(\theta) = \sum_{\theta=3}^{\infty} 4^{-\theta}.$$  

Using the hint in part (b) (which should have been given earlier),

$$\sum_{\theta=3}^{\infty} 4^{-\theta} = 4^{-3} \sum_{\theta=0}^{\infty} 4^{-\theta} = 4^{-3} \cdot \frac{1}{1 - 1/4} = \frac{1}{48},$$

so $f_{\Theta|X}(\theta|6.18) = 48 \cdot 4^{-\theta}$. The most likely posterior value is MAP = 3. The estimation error is

$$P(\Theta \neq 3|X = 6.18) = 1 - P(\Theta = 3|X = 6.18) = 1 - 48 \cdot 4^{-3} = \frac{1}{4}.$$  

(b) What is the average MAP estimation error $P(\text{MAP} \neq \Theta)$ given a single sample $X$? (Hint: $1 + x + x^2 + \cdots = 1/(1 - x)$ when $|x| < 1$.)

**Solution:** If the observation is $X = x$ then $\Theta$ must be an integer that is at least as large as $\log_2 x$. The smallest such integer is usually denoted as $[\log_2 x]$. By the same calculation as in part (a), $f_{\Theta|X}(\theta|x)$ has value proportional to $4^{-\theta}$ for every $\theta \geq [\log_2 x]$. The MAP estimate picks out the $\theta$ that maximizes this expression, namely $\theta = [\log_2 x]$. The estimation error given $X = x$ is then

$$P(\Theta \neq [\log_2 x]|X = x) = 1 - P(\Theta = [\log_2 x]|X = x)$$

$$= 1 - \frac{4^{-[\log_2 x]}}{\sum_{\theta=[\log_2 x]}^{\infty} 4^{-\theta}}$$

$$= 1 - \frac{1}{\sum_{\theta=0}^{\infty} 4^{-\theta}}$$

$$= \frac{1}{4}.$$  

This conditional error has the unusual feature of being independent of $X$, so the average estimation error is also $P(\Theta \neq [\log_2 X]) = 1/4$.

(c) (Optional) What is the average MAP estimation error given $n$ independent samples?

**Solution:** The smallest possible value that $\Theta$ can take given $X_1 = x_1, \ldots, X_n = x_n$ is now $t = \max ([\log_2 x_1], [\log_2 x_2], \ldots, [\log_2 x_n])$. The posterior is now

$$f_{\Theta|X_1,X_2,\ldots,X_n}(\theta|x_1, x_2, \ldots, x_n) \propto f_{X_1|\Theta}(x_1|\theta) \cdots f_{X_n|\Theta}(x_n|\theta) f_{\Theta}(\theta)$$

$$= 2^{-\theta} \cdots 2^{-\theta} \cdot 2^{-\theta}$$

$$= 2^{-(n+1)\theta},$$

for every $\theta \geq t$. By the total probability theorem,

$$f_{X_1,X_2,\ldots,X_n}(x_1, x_2, \ldots, x_n) = \sum_{\theta=t}^{\infty} f_{X_1,X_2,\ldots,X_n|\Theta}(x_1, x_2, \ldots, x_n|\theta) f_{\Theta}(\theta)$$

$$= \sum_{\theta=t}^{\infty} 2^{-(n+1)\theta}.$$
which is again maximized at $\theta = t$, so the MAP estimate is $t$. By a similar calculation as in part (b), the estimation error given $X_1 = x_1, \ldots, X_n = x_n$ is

$$1 - \frac{2^{-(n+1)t}}{\sum_{\theta=t}^{\infty} 2^{-(n+1)\theta}} = 1 - \frac{1}{\sum_{\theta=0}^{\infty} 2^{-(n+1)\theta}} = 2^{-(n+1)}$$

and so is the average error.