Practice questions

1. You flip a coin with unknown probability of heads Θ. Assume a Uniform$(0,1)$ prior on Θ.

   (a) You observe the outcome `HTH` after three flips. What is the posterior probability that the next flip will be a head?

   Solution: The posterior Θ given two heads and a tail is a Beta$(3,2)$ random variable with PDF $f_{\Theta|X_1X_2X_3}(\theta|\mathtt{HTH}) = 12\theta^2(1-\theta)$. The probability that the next toss is a head given $\Theta = \theta$ is $\mathrm{P}(X_4 = \mathtt{H}|\Theta = \theta) = \theta$. By the total probability theorem

   $$\mathrm{P}(X_4 = \mathtt{H}) = \int_0^1 \mathrm{P}(X_4 = \mathtt{H}|\Theta = \theta) f_{\Theta|X_1X_2X_3}(\theta|\mathtt{HTH})d\theta = \int_0^1 \theta \cdot 12\theta^2(1-\theta)d\theta = \frac{3}{5}.$$

   (b) (**Optional**) More generally, you observe h heads and $t = n - h$ tails after n flips. What is the posterior probability that the next flip will be a head?

   Solution: By the same reasoning, the posterior on Θ is Beta$(h+1, t+1)$, and the probability that the next toss is a head given $\Theta = \theta$ is θ. The posterior probability that the next flip is a head is

   $$\begin{aligned}\mathrm{P}(X_4 = \mathtt{H}) &= \int_0^1 \mathrm{P}(X_{n+1} = \mathtt{H}|\Theta = \theta) f_{\Theta|X_1\cdots X_n}(\theta|h \text{ heads, } t \text{ tails})d\theta \\ &= \int_0^1 \frac{1}{B(h+1,t+1)} \cdot \theta^h(1-\theta)^t \cdot \theta d\theta \\ &= \frac{B(h+2,t+1)}{B(h+1,t+1)} \cdot \int_0^1 \frac{1}{B(h+2,t+1)} \cdot \theta^{h+1}(1-\theta)^t \cdot \theta d\theta \\ &= \frac{B(h+2,t+1)}{B(h+1,t+1)}.\end{aligned}$$

   The last equality is true because the integrand is the PDF of a Beta$(h+2, t+1)$ random variable so the integral evaluates to one. Using the formula for the Beta function we get that

   $$\mathrm{P}(X_4 = \mathtt{H}) = \frac{(h+1)!t!/(h+t+2)!}{h!t!/(h+t+1)!} = \frac{h+1}{h+t+2} = \frac{h+1}{n+2}.$$

2. Romeo has figured out that girls have a hidden parameter Θ that determines how late they tend to show up to dates. As usual his prior on Θ is Uniform$(0,1)$.

   (a) Assume that that likelihood $f_{X|\Theta}$ of Juliet being X hours late on a given date is Uniform$(0,\Theta)$. Calculate the probability that Juliet shows up at least 10 minutes late at her first date.

   Solution: The lihelihood of X is $f_{X|\Theta}(x|\theta) = 1/\theta$ when $0 \le x \le \theta$ and zero otherwise. The conditional probability of the event $X \ge 1/6$ given $\Theta = \theta$ is

   $$\mathrm{P}(X \ge 1/6|\Theta = \theta) = \int_{1/6}^{\infty} f_{X|\Theta}(x|\theta)dx = \int_{1/6}^{\theta} \frac{1}{\theta}dx = \begin{cases}(1/\theta)(\theta - 1/6), & \text{if } \theta \ge 1/6 \\ 0, & \text{otherwise.}\end{cases}$$

   The total probability is

   $$\mathrm{P}(X \ge 1/6) = \int_{-\infty}^{\infty} \mathrm{P}(X \ge 1/6|\Theta = \theta) f_\Theta(\theta)d\theta = \int_{1/6}^{1} \frac{1}{\theta}(\theta - 1/6) \cdot 1 d\theta = \frac{5 - \ln 6}{6},$$

which is about 0.535.

(b) Juliet shows up exactly 10 minutes late at her first date. What is the probability that she shows up at least 10 minutes late at her second date?

Solution: Let X' be Juliet's arrival time at her second date. The likelihood of X' is the same as the likelihood of X, but the PDF of Θ is now the posterior PDF after the first date, namely

$$f_{\Theta|X}(\theta|1/6) \propto f_{X|\Theta}(1/6|\theta)f_\Theta(\theta) = \begin{cases} 1/\theta, & \text{if } 1/6 \leq \theta \leq 1 \\ 0, & \text{otherwise.} \end{cases}$$

The proportionality constant must be $1/\int_{1/6}^{1}(1/\theta)d\theta = 1/\ln 6$. The total probability is therefore

$$\begin{aligned} \mathrm{P}(X' \geq 1/6|X = 1/6) &= \int_{-\infty}^{\infty} \mathrm{P}(X' \geq 1/6|\Theta = \theta)f_{\Theta|X}(\theta|1/6) \\ &= \int_{1/6}^{1} \frac{1}{\theta}(\theta - 1/6) \cdot \frac{1}{(\ln 6)\theta}d\theta \\ &= 1 - \frac{5}{6\ln 6}, \end{aligned}$$

which is about 0.535. The probability didn't change up to third decimal (but it is not the same).

(c) Repeat parts (a) and (b) with an Exponential(Θ) likelihood.

Solution: For part (a) we would get that $\mathrm{P}(X \geq 1/6|\Theta = \theta) = e^{-\theta/6}$ for $\theta \geq 0$. As the prior remains the same, $\mathrm{P}(X \geq 1/6) = \int_0^1 e^{-\theta/6}d\theta = 6(1 - e^{-1/6}) \approx 0.921$. In Homework 2 the posterior PDF for Θ given $X = 1/6$ was calculated to be $f_{\Theta|X}(\theta|1/6) \approx 2.233\theta e^{-\theta/6}$ for $0 \leq \theta \leq 1$. Therefore the answer in part (b) changes to

$$\mathrm{P}(X' \geq 1/6|X = 1/6) \approx \int_0^1 e^{-\theta/6} \cdot 2.233\theta e^{-\theta/6}d\theta \approx 0.897.$$

The probability decreased by a little bit.

3. In a group of ten people, including Alice and Bob, each pair is friends with probability P, independently of the other pairs.[1]

(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:

$$\mathrm{P}_{A|P}(A = a|P = p) = \binom{10-1}{a}p^a(1-p)^{10-1-a}$$

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

[1] 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 $6/17$.

4. The ENGG2780A TAs want to estimate the hardness of the problem that they prepared for an upcoming homework. They know from experience that the time X (in minutes) it takes them to solve the problem is Exponential$(1/15)$ if the problem is easy and Exponential$(1/25)$ if the problem is hard (like this one). Based on previous quizzes, the prior probability that the problem is hard is 0.3.

(a) Fanbin solves the problem in 20 minutes. According to the MAP rule, is the problem easy or hard? What is the probability of error?

Solution: Let Θ be the difficulty of the problem. The problem is difficult if $\Theta = 1$, otherwise $\Theta = 0$.

$$\mathrm{P}(\Theta = 1|X = 20) \propto f_{X|\Theta}(X = 20|\Theta = 1)\,\mathrm{P}(\Theta = 1) = \frac{1}{25}e^{-20/25} \cdot 0.3 \approx 0.00539$$
$$\mathrm{P}(\Theta = 0|X = 20) \propto f_{X|\Theta}(X = 20|\Theta = 0)\,\mathrm{P}(\Theta = 0) \propto \frac{1}{15}e^{-20/15} \cdot 0.7 \approx 0.01230$$

The MAP rule predicts that the problem is simple ($MAP = 0$). The probability of error is

$$\mathrm{P}(MAP \neq \Theta|X = 20) = \mathrm{P}(\Theta = 1|X = 20) \approx \frac{0.00539}{0.00539 + 0.01230} \approx 0.305.$$

(b) The other five TAs try it out and their recorded solution times are 10, 10, 15, 25, and 35 minutes, respectively. How do the answers in part (a) change? *[Adapted from Bertsekas-Tsitsiklis problem 8.2.6]*

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

$$\begin{aligned}\mathrm{P}(\Theta = 1|X_1 = 20, X_2 = 10, \cdots, X_6 = 35) &\propto f_{X_1|\Theta}(20|1) \cdots P_{X_6|\Theta}(35|1)\,\mathrm{P}(\Theta = 1)\\ &\propto \frac{1}{25^6}e^{-(20+10+10+15+25+35)/25} \cdot 0.3\\ &\approx 1.235 \cdot 10^{-11}\end{aligned}$$

Similarly,

$$\begin{aligned}\mathrm{P}(\Theta = 0|X_1 = 20, X_2 = 10, \cdots, X_6 = 35) &\propto f_{X_1|\Theta}(20|0) \cdots P_{X_6|\Theta}(35|0)\,\mathrm{P}(\Theta = 0)\\ &\propto \frac{1}{15^6}e^{-(20+10+10+15+25+35)/15} \cdot 0.7\\ &\approx 2.870 \cdot 10^{-11}\end{aligned}$$

Again the MAP rule predicts the problem is easy, and the probability of error is about $1.235/(1.235 + 2.870) \approx 0.300$.