Practice questions

1. In question 1 of homework 1, Romeo figured out that girls are Exponential(Θ) hours late to a date, where Θ 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 Θ after the first and second dates, respectively.

2. 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?

   (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?

   (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?

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 θ 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?

   (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]*

4. Let X be a uniform Uniform$(0, 2^\Theta)$ random variable with unknown Θ. Your prior on Θ is Geometric$(1/2)$.

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

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

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

[1] This is the Erdős-Rényi random graph model.

Additional ESTR 2020 questions

5. In this question you will investigate the Shannon entropy $H(X)$ of a Geometric(p) random variable X.
   (a) Derive a formula for $H(X)$.
   (b) Describe a prefix-free encoding of expected length $H(X)$ when $p = 1/2$.
   (c) In general, the expected length of the best prefix-free encoding may not be $H(X)$ but is never larger than $H(X) + 1$. What is the best prefix-free encoding of X that you can come up with for other values of p?

6. You can use Bayesian statistics to empirically estimate probabilities that may be hard to calculate. Suppose for example that you want to know the probability Θ that three random points in the unit square form an acute triangle.
   (a) Write a computer program that samples 500 triangles and tracks the number X of acute ones in the sample. Assuming a Uniform$(0, 1)$ prior on Θ, what is your MAP estimate of Θ given X? What is the probability that $|MAP - \Theta| > 0.05$ (given X)?
   (b) Now suppose that, before doing the experiment, your friend was "quite sure" that $\Theta = 1/2$, so her prior on Θ is Beta(n, n) for some large n. How would that change the conclusions in part (a)?