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?

   (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?

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.

   (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?

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

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?

   (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?

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?

   (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]*

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

Additional ESTR 2020 questions

5. In this exercise you will explain the accuracy of the MAP estimate given sufficiently many samples. For concreteness assume there are only two possible models $\Theta = 0$ and $\Theta = 1$ with equal prior probabilities, and that the values are discrete. Assume also that the true value of Θ is zero, so that the PMF of the samples is $f(x) = \mathrm{P}(X = x|\Theta = 0)$. Let $\rho(x)$ denote the *likelihood ratio* $\mathrm{P}(X = x|\Theta = 0)/\,\mathrm{P}(X = x|\Theta = 1)$.

   (a) Show that the *expected log-likelihood ratio* $\mathrm{E}[\ln \rho(X)]$, where X has PMF f, is always nonnegative. (**Hint:** Study the function $d(q_1, \ldots, q_n) = p_1 \log q_1 + \cdots + p_n \log q_n$ for fixed probabilities $p_1, \ldots, p_n$, $p_1 + \cdots + p_n = q_1 + \cdots + q_n = 1$.)

   (b) Show that the log-likelihood ratio is strictly positive unless X is independent of Θ.

   (c) Show that unless X is independent of Θ,

   $$\lim_{n\to\infty} \mathrm{P}\big(\mathrm{P}(\Theta = 0|X_1, \ldots, X_n) > \mathrm{P}(\Theta = 1|X_1, \ldots, X_n)\big) = 1,$$

   where $X_1, \ldots, X_n$ are independent with PMF f, assuming $\mathrm{E}[(\ln \rho(X))^2]$ is finite. (**Hint:** Apply the law of large numbers to a suitable random variable.)

   (d) Argue that the conclusion in part (c) holds even if you change the prior probabilities, as long as both are nonzero.

   (e) Investigate when $\mathrm{E}[(\ln \rho(X))^2]$ is finite for some common likelihood functions (e.g. Binomial, Geometric, Poisson).

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)?