CS7DS3 Assignment 1

February 25, 2019

To be handed into the SCSS Office by 12 noon on Monday 11th March, 2019. Please remember to print your name and student number on the front of your script. Answer ALL questions.

Question 1

I really like watching films directed by the Coen brothers. Say that they make up about about 40% of all the movies I watch. However, my wife Eve isn't so enthusiastic. Suppose that we only watch Coen brothers movies about 20% of the time if we're watching a film together. However, if I'm watching a film without Eve, then there's about a 70% chance that the film will be by the Coen brothers. In other words,

$$\begin{aligned}
\mathbb{P}(\text{Coen brothers}) &= 0.4 \\
\mathbb{P}(\text{Coen brothers} \mid \text{Eve}) &= 0.2 \\
\mathbb{P}(\text{Coen brothers} \mid \text{Not Eve}) &= 0.7.
\end{aligned}$$

a) Show that about 60% of the time, I watch movies with Eve, i.e., $\mathbb{P}(\text{Eve}) = 0.6$. Clearly show all workings. (Hint: Use the fact that $\mathbb{P}(A) = \mathbb{P}(A, B) + \mathbb{P}(A, \text{ not } B)$.)

b) Suppose I am watching a Coen brothers movie. What is the probability that Eve is watching with me, $\mathbb{P}(\text{Eve} \mid \text{Coen brothers})$?

Question 2

The amount of time a user spends viewing a webpage may be modelled by an exponential distribution with rate parameter θ: $T \sim \text{Exp}(\theta)$; with pdf $f(t) = \theta \exp\{-\theta t\}$ and moments $\mathbb{E}[T] = 1/\theta$ and $\mathbb{V}\text{ar}[T] = 1/\theta^2$.

Suppose we observe data $y = y_{1,}\ldots, y_n$. Let y follow a exponential distribution so that $y_i \sim \text{Exp}(\theta)$, for each i. Assume that all observations are independent.

a) Show that $\hat{\theta}$, the m.l.e for θ, is given by $\hat{\theta} = \bar{y}$.

b) Show that $\theta \sim \text{Gamma}(a_{0,}b_0)$ is a conjugate prior distribution for θ. In other words, show that if the prior $p(\theta|a_{0,}b_0)$ is a Gamma distribution, then the posterior distribution $p(\theta|y, a_0, b_0)$ will also be a Gamma. Derive the posterior parameters a_n and b_n. Compare the posterior parameters a_n and b_n to the hyperparameters a_0 and b_0, and explain their connection, if any, to the mle $\hat{\theta}$.

A company wants to know if a new layout encourages increased user engagement on their website. A/B testing is performed to determine how long users view a webpage with the existing layout (Layout A) compared to the new one (Layout B). Layout A is viewed 30 times, with an average viewing time $\bar{y_A} = 4.71$ minutes; Layout B is viewed 20 times with average viewing time $\bar{y_B} = 4.87$. The data can be downloaded from etc..

Let θ_A and θ_B denote the rate parameters of the respective webpage layouts, and assume a common gamma prior for both parameters.

c) Discuss how to choose a sensible prior distribution in this case, i.e., how might you decide to choose a_0 and b_0? For context, an industry veteran tells you that the worst layouts usually have mean engagements times of 45 seconds or so, (i.e., 0.75 of a minute), while the most optimistic projections for a new layout are never more than 10 minutes on average. Is it possible to specify a "non-informative" prior in this case? Explain why/why not, and if possible, specify the form of a non-informative prior in this case.

c) Using Monte Carlo methods or otherwise, estimate the probability that the new Layout B has increased user engagement compared to the existing Layout A, i.e., $\mathbb{P}(\theta_A > \theta_B)$. Provide either: i) a 95% interval for the difference in engagment rates; or ii) visualise and interpret a density plot or histogram of this difference. Interpret these results: do you think that Layout B is clearly superior to Layout A, or is more information required?

c) The company estimates that the value of user engagment can be quantified as follows: if y denotes the time a user spends on a webpage, and g denotes the value (in euro) of this time to the company, then $g(y) = 2(\log y + 5)$. For example, a user spending 5 minutes on a webpage is worth E3.41 to the company. Based on the m.l.e. estimate $\hat{\theta_A}$, and using Monte Carlo methods or otherwise, estimate the expected value to the company of website visits using Layout A. (N.B., $\mathbb{E}[g(\theta)] \neq g(E[\theta)]$! Do you seen any shortcomings with this projection? Outline an alternative approach that you might take.

Question 3

This question is based on a recent Cross Validated post. A colleague is trying to implement a Metropolis-Hastings (M-H) algorithm to infer the parameters of a complicated model. In the proposal step, they use a proposal

$$\theta^* = \theta^{(s)} + u^*,$$

where $u^* \sim \mathcal{U}(0, 1)$, and $\mathcal{U}$ denotes a uniform distribution. Your colleague is perplexed as to why the algorithm is not obtaining satisfactory results.

a) Sketch a trace plot of how such a M-H algorithm might behave. Compare this to a trace-plot of an optimally perfing M-H algorithm.

b) Explain why the current proposal fails to satisfy detailed balance, i.e., reversibility. (An intuitive explanation is fine.)

c) Suggest an alternative proposal distribution.