CS7DS3 Assignment 1

February 12, 2018

To be handed into the SCSS Office by 12 noon on Friday 23rd February, 2018. Please remember to print your name and student number on the front of your script.

Question 1

Let $y_1, \ldots, y_n$ follow a Poisson distribution so that $y_i \sim \mathcal{P}(\theta)$, for each i. Assume that all observations are independent.

a) Show that $\theta \sim$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.

b) Represent the Poisson distribution in exponential family form. Explicitly identify the sufficient statistic $s(y)$, natural parameter $\phi(\theta)$, and link function $g(\theta)$. Hence or otherwise represent the prior and posterior distributions for θas members of an exponential family. Specify ν and η in both cases.

Question 2

Let $y_1, \ldots, y_n$ follow a Bernouilli distribution. We assume a Beta(a_0, b_0) prior for θ.

a) A colleague proposes you use a *Jeffreys* prior for θ. This means setting $a_0 = b_0 = 0.5$. Visualise this distribution, e.g., using `dbeta` in R. Comment on the properties of the distribution. Compare this choice of prior to setting $a_0 = b_0 = 1$.

b) A different colleague recommends that you use a logit transformation of the data, such that $g(\theta) = \log\{\theta/(1-\theta)\}$, as that is the standard way to report this information in the

area of application. Using Monte Carlo methods or otherwise, comment on the effect of choosing of hyperparameters in this setting. Does the prior distribution $p(g(\theta))$ have a uniform distribution when $a_0 = b_0 = 1$? Compare to the case when $a_0 = b_0 = 0.5$.

c) A company conducts A/B testing to determine which advert placement is more effective. Placement A receives 13 clickthroughs from 200 page views; placement B receives 7 clickthroughs from 75 page views. Denote by θ_A and θ_B the probability of clickthrough for each placement strategy, and assume a common beta prior for both parameters.

Using Monte Carlo methods or otherwise, estimate the probability that placement B has a higher clickthough rate than placement A, i.e., $\mathbb{P}(\theta_A > \theta_B)$. Provide either: i) a 95% interval for the difference in clickthrough rates; or ii) visualise and interpret a density plot of this difference. Briefly discuss your choice of hyperparameters a_0 and b_0.

Question 3

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

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

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

a) Suggest an alternative proposal that would lead to improved sampler performance.

b) Explain why the current proposal is inadequate. In your answer, explicitly identify which of the required conditions of the transition kernel $q(\theta^{(s)}|\theta^{(s-1)})$ have not been met.