Bayesian Inference PSTAT 215A
Homework 1 (of 4) RBG/Fall 2009

Assigned 6 Oct 2009; due 15 Oct 2009. *You may study the questions/answers with others, but what you hand in for credit must be your own work.*

1. Let X, Y, Z be RVs with joint density $p(x, y, z) \propto f(x, z)g(y, z)h(z)$ for functions f, g, and h. Show that
   (a) $p(x|y, z) \propto f(x, z)$, i.e., $p(x|y, z)$ is a function of x and z
   (b) $p(y|x, z) \propto g(y, z)$, i.e., $p(y|x, z)$ is a function of y and z
   (c) X and Y are conditionally independent given Z

2. The US Senate contains a proportion p of Republicans, who are incapable of changing their minds about anything, and a proportion $1 - p$ of Democrats, who change their minds completely at random, with probability r, between successive votes on the same issue. A randomly chosen senator is observed to have voted twice in succession in the same way. What is the probability that this senator will vote in the same way next time?

3. A cancer laboratory is estimating the rate of tumorigenesis in two strains of mice, A, and B. They have tumor count data for 10 mice in strain A and 13 mice in strain B. Type A mice have been well studied, and information from other laboratories suggests that each type A mouse has a tumor count that is approximately Poisson–distributed with a mean of 12. Tumor count rates for type B mice are unknown, but type B mice are related to type A mice. The observed tumor counts for the two populations are

$$\mathbf{y}_A = (12, 9, 12, 14, 13, 13, 15, 8, 15, 6)$$
$$\mathbf{y}_B = (11, 11, 10, 9, 9, 8, 7, 10, 6, 8, 8, 9, 7)$$

   (a) Find (and plot where appropriate) the posterior distributions, means, variances and 95% quantile–based credible intervals for θ_A and θ_B, assuming a Poisson sampling model for each group and the following prior distribution:

$$\theta_A \sim G(120, 10) \qquad \theta_B \sim G(12, 1) \qquad p(\theta_A, \theta_B) = p(\theta_A)p(\theta_B)$$

   (b) Compute and plot the posterior expectation of θ_B under the prior distribution $\theta_B \sim G(12n_0, n_0)$ for each value of $n_0 \in \{1, 2, \dots, 50\}$. Describe what sort of prior beliefs would be necessary in order for the posterior expectation of θ_B to be close to that of θ_A.
   (c) Should knowledge about population A tell us anything about the population B? Discuss whether or not it makes sense to have $p(\theta_A, \theta_B) = p(\theta_A)p(\theta_B)$.

4. The exponential distribution is commonly used to model "waiting times" and other continuous, positive, real–valued RVs, usually measured on a time scale. The sampling distribution of an outcome y, given parameter θ, is

$$p(y|\theta) = \theta e^{-y\theta}, \qquad \text{for } \; y > 0,$$

and $\theta = 1/\mathbb{E}\{Y|\theta\}$ is called the "rate". The exponential sampling model is notated as $Y \sim \text{Exp}(\theta)$. Mathematically, $\text{Exp}(\theta) \equiv G(1, \theta)$, i.e., it is a special case of the gamma distribution.

Suppose that we were to use the exponential sampling model for data

$$Y_1, \ldots, Y_n \stackrel{\text{iid}}{\sim} \text{Exp}(\theta).$$

(a) Derive the conjugate prior and form of the corresponding posterior distribution for this sampling model.

(b) Derive the Jeffreys' prior for this sampling model. Show that this prior is improper, but that it leads to a proper posterior under a condition which you should describe.

(c) Argue that, mathematically speaking, the Jeffreys' prior is a member of the conjugate family you found in part (a).

(d) Calculate the MLE $\hat{\theta}$ for θ. Under what setting(s) of the conjugate prior parameter(s) [from in part (a)] does the posterior expectation coincide with the MLE? Argue that this prior is improper but the resulting posterior is proper under a condition which you should describe.

(e) The length of life of a light bulb Y manufactured by a certain process is exponentially distributed with unknown failure rate θ. Suppose we take the prior $\theta \sim G(a, b)$ with parameters a and b set so that the *coefficient of variation* (standard deviation divided by the mean) is 0.5. A random sample of light bulbs is to be tested and the lifetime of each $y_1, \ldots, y_n$ obtained. If the coefficient of variation is to be reduced to 0.1, how many light bulbs (n) need to be tested?

5. Let Y denote the number of heads in n flips of a coin, whose probability of heads is θ.

(a) Suppose you assign a $\text{Beta}(a, b)$ prior distribution for θ, and then you observe y heads out of n flips. Show algebraically that your posterior mean of θ always lies between your prior mean and the observed relative frequency of heads.

(b) Show that, if the prior distribution on θ is uniform, the posterior variance of θ is always less than the prior variance.

(c) Give an example of a $\text{Beta}(a, b)$ prior distribution and data (y, n), in which the posterior variance of θ is higher than the prior variance.