Name:

STAT 215 – Midterm 1

1. [10pts] Consider a model

$$y_i \overset{ind}{\sim} Poi(\lambda_i),\ i = 1, \dots, n,$$

with prior distribution

$$\lambda_i \overset{iid}{\sim} Ga(\alpha, \beta),\ i = 1, \dots, n,$$

and hyperprior

$$\beta \sim Gamma(c, d),$$

where α, c and d are known.

Find the marginal posterior $p(\beta | y_1, \dots, y_n)$.

2. When a coin is spun on its edge, instead of being thrown in the air, the proportion of *heads* is rarely close to 1/2, but rather 1/3 or 3/4 because of irregularities in the edge.

When spinning, n times, a given coin on its edge, we observe a number of heads

$$x \sim Bin(n, \theta).$$

Assume as prior distribution on the probability of *head* θ a mixture of two Beta distributions:

$$p(\theta) = \frac{1}{2}Be(10, 20) + \frac{1}{2}Be(20, 10),$$

where

$$Be(a, b) = \frac{\theta^{a-1}(1-\theta)^{b-1}}{B(a, b)}$$

denotes the Beta p.d.f. with parameters a and b. Here $B(a, b) = \frac{\Gamma(a)\Gamma(b)}{\Gamma(a+b)}$.

Hint: The solution for 2a and 2b will involve ratios of Beta functions. You can leave expressions $B(a, b)$ etc. in the solution, i.e., you do not need to simplify further.

2a. [5pts] Find the posterior distribution $p(\theta|x)$.

2b. [5pts] Find the posterior mean $E(\theta|x)$.

3. $x_i \overset{iid}{\sim} Exp(\theta)$, $i = 1, \ldots, n$.

3a. [2pts] Assuming θ is distributed as $Ga(\theta|\alpha, \beta)$, with α and β fixed, find the posterior distribution $p(\theta|x_1, \ldots, x_n)$.

3b. [4pts] Under the same assumptions, compute the predictive distribution $p(x_{n+1}|x_1, ..., x_n)$.

3c. [4pts] Consider now a slightly different situation. Imagine now that the exponential random variables represent waiting times for the recurrence of disease in n patients surgically treated at time 0. We are at time a, and all patients are still alive. Some of the n patients may have experienced recurrence. We record their recurrence time x_i[1]. Other patients may still be healthy, so we don't know exactly what the recurrence time is. We only know $x_i > a$. (This is called censored sampling). Repeat steps (a) and (b) for this case.

[1] Assume patients are not followed up after the recurrence, i.e. we have no further data