# Homework 2

## Problem 1

Suppose $x_1, \ldots, x_n$ is a random sample from an exponential distribution with mean $1/\theta$.

1. Derive the Jeffrey's prior for $\theta$
2. Derive the posterior distribution of $\theta$ using the Jeffrey's prior.
3. Derive the predictive distribution of a future observation $z$.
4. Derive a 95% credible interval for $z$.

## Problems 2 and 3

Below two problems are exercises from Hoff's book.

3.10 Change of variables: Let $\psi = g(\theta)$, where $g$ is a monotone function of $\theta$, and let $h$ be the inverse of $g$ so that $\theta = h(\psi)$. If $p_\theta(\theta)$ is the probability density of $\theta$, then the probability density of $\psi$ induced by $p_\theta$ is given by $p_\psi(\psi) = p_\theta(h(\psi)) \times |\frac{dh}{d\psi}|$.

a) Let $\theta \sim \text{beta}(a, b)$ and let $\psi = \log[\theta/(1-\theta)]$. Obtain the form of $p_\psi$ and plot it for the case that $a = b = 1$.

b) Let $\theta \sim \text{gamma}(a, b)$ and let $\psi = \log\theta$. Obtain the form of $p_\psi$ and plot it for the case that $a = b = 1$.

3.12 Jeffreys' prior: Jeffreys (1961) suggested a default rule for generating a prior distribution of a parameter $\theta$ in a sampling model $p(y|\theta)$. Jeffreys' prior is given by $p_J(\theta) \propto \sqrt{I(\theta)}$ , where $I(\theta) = -\mathrm{E}[\partial^2 \log p(Y|\theta)/\partial\theta^2|\theta]$ is the *Fisher information*.

a) Let $Y \sim \text{binomial}(n, \theta)$. Obtain Jeffreys' prior distribution $p_J(\theta)$ for this model.

b) Reparameterize the binomial sampling model with $\psi = \log\theta/(1-\theta)$, so that $p(y|\psi) = \binom{n}{y} e^{\psi y}(1+e^\psi)^{-n}$. Obtain Jeffreys' prior distribution $p_J(\psi)$ for this model.

c) Take the prior distribution from a) and apply the change of variables formula from Exercise 3.10 to obtain the induced prior density on $\psi$. This density should be the same as the one derived in part b) of this exercise. This consistency under reparameterization is the defining characteristic of Jeffrey's' prior.

## Problem 4

Marginal distributions: Given observations $Y_1, \ldots, Y_n \sim$ i.i.d. normal$(\theta, \sigma^2)$ and using the conjugate prior distribution for $\theta$ and $\sigma^2$, derive the formula for $p(\theta|y_1, \ldots, y_n)$, the marginal posterior distribution of $\theta$, conditional on the data but marginal over $\sigma^2$. Check your work by comparing your formula to a Monte Carlo estimate of the marginal distribution, using some values of $Y_1, \ldots, Y_n$, $\mu_0$, $\sigma_0^2$, $\nu_0$ and $\kappa_0$ that you choose. Also derive $p(\tilde{\sigma}^2|y_1, \ldots, y_n)$, where $\tilde{\sigma}^2 = 1/\sigma^2$ is the precision.

## Problem 5

Studying: The files `school1.dat`, `school2.dat` and `school3.dat` contain data on the amount of time students from three high schools spent on studying or homework during an exam period. Analyze data from each of these schools separately, using the normal model with a conjugate prior distribution, in which $\{\mu_0 = 5, \sigma_0^2 = 4, \kappa_0 = 1, \nu_0 = 2\}$ and compute or approximate the following:

a) posterior means and 95% confidence intervals for the mean $\theta$ and standard deviation $\sigma$ from each school;

b) the posterior probability that $\theta_i < \theta_j < \theta_k$ for all six permutations $\{i, j, k\}$ of $\{1, 2, 3\}$;

c) the posterior probability that $\tilde{Y}_i < \tilde{Y}_j < \tilde{Y}_k$ for all six permutations $\{i, j, k\}$ of $\{1, 2, 3\}$, where $\tilde{Y}_i$ is a sample from the posterior predictive distribution of school $i$.

d) Compute the posterior probability that $\theta_1$ is bigger than both $\theta_2$ and $\theta_3$, and the posterior probability that $\tilde{Y}_1$ is bigger than both $\tilde{Y}_2$ and $\tilde{Y}_3$.

## Problem 6

Binomial and multinomial models.

1. Suppose data $(y_1, \ldots, y_J)$ follow a multinomial distribution with parameters $(\theta_1, \ldots, \theta_J)$. Also, suppose that $\theta = (\theta_1, \ldots, \theta_J)$ has a Dirichlet prior distribution. Let $\alpha = \frac{\theta_1}{\theta_1 + \theta_2}$. Write the marginal posterior distribution for $\alpha$ and show that this distribution is identical to the posterior distribution for $\alpha$ obtained by treating $y_1$ as an observation from the Binomial distribution with probability $\alpha$ and sample size $y_1 + y_2$ ignoring the data $y_3, \ldots, y_J$. This result justifies the use of the binomial distribution in multinomial problems where the interest is in two of the categories, such as in problem 2 below.

2. Comparison of two multinomial experiments: on September 25, 1988, the evening of a Presidential campaign debate, ABC News conducted a survey of registered voters in the United States; 639 persons were polled before the debate, and 639 different persons were pulled after. The results are displayed

below. Assume the surveys were independent simple random samples from the population of registered voters. Model the data with two different multinomial distributions. For $j = 1, 2$ let $\alpha_j$ be the proportion of voters who preferred Bush, out of those who had a preference for either Bush or Dukakis at the time of survey $j$. Plot a histogram of the posterior density for $\alpha_2 - \alpha_1$. What is the posterior probability that there was a shift towards Bush?

| Survey | Bush | Dukakis | No opinion/other | Total |
|---|---|---|---|---|
| pre-debate | 294 | 307 | 38 | 639 |
| post-debate | 288 | 332 | 19 | 639 |