STAT 638: Solution for Homework #3

3.8 a) It is convenient to parameterize each beta in terms of $\mu = a/(a+b)$ and $n_0 = a+b$, in which case $a = \mu n_0$ and $b = n_0(1-\mu)$. Based on the results of Diaconis and Ylvisaker it makes sense to use a prior that is a mixture of beta$(n_0/3, 2n_0/3)$, beta$(n_0/2, n_0/2)$ and beta$(2n_0/3, n_0/3)$, where the mixing weights are 0.4, 0.2 and 0.4, respectively.

b) Using a 2014 nickel, I got 32 heads in 50 spins.

c) Assume that y heads are obtained in the 50 spins. The posterior is proportional to

$$\theta^y(1-\theta)^{50-y}\left[0.4A_1\theta^{n_0/3-1}(1-\theta)^{2n_0/3-1} + 0.2A_2\theta^{n_0/2-1}(1-\theta)^{n_0/2-1} + 0.4A_3\theta^{2n_0/3-1}(1-\theta)^{n_0/3-1}\right],$$

where

$$A_1 = \frac{\Gamma(n_0)}{\Gamma(n_0/3)\Gamma(2n_0/3)}, \quad A_2 = \frac{\Gamma(n_0)}{\Gamma(n_0/2)\Gamma(n_0/2)} \quad \text{and} \quad A_3 = A_1.$$

It follows that the posterior is a mixture of beta$(y+n_0/3, 50-y+2n_0/3)$, beta$(y+n_0/2, 50-y+n_0/2)$ and beta$(y+2n_0/3, 50-y+n_0/3)$. The respective mixing weights are $w_i = a_i/(a_1+a_2+a_3)$, $i = 1, 2, 3$, where

$$a_1 = 0.4\frac{\Gamma(y+n_0/3)\Gamma(50-y+2n_0/3)}{\Gamma(n_0/3)\Gamma(2n_0/3)}, \quad a_2 = 0.2\frac{\Gamma(y+n_0/2)\Gamma(50-y+n_0/2)}{\Gamma(n_0/2)^2}$$

and

$$a_3 = 0.4\frac{\Gamma(y+2n_0/3)\Gamma(50-y+n_0/3)}{\Gamma(2n_0/3)\Gamma(n_0/3)}.$$

d) I used a 1996 quarter and got 22 heads in 50 spins. I don't see why information from the 2014 nickel is relevant for the quarter. Therefore I'll use the same prior as before. This leads to the following posterior.

3.9 a) By inspection, a prior proportional to $\theta^{2(c-1/2)}e^{-d\theta^2}$ would be a conjugate prior. For $c > 0$, $d > 0$, it is easy to show that the following function is a density:

$$p(\theta|c,d) = \frac{2d^c}{\Gamma(c)}\theta^{2(c-1/2)}e^{-d\theta^2}I_{(0,\infty)}(\theta).$$

The plot below shows a few members of this family. In each case $d = 1$, and from left to right $c = 1/2$, 1, 2, 4. It's worth noting that when $c < 1/2$ the density is monontone decreasing on $(0, \infty)$ and unbounded at 0.

b) The posterior $p(\theta|\boldsymbol{y})$ is such that

$$p(\theta|\boldsymbol{y}) \propto p(\theta|c,d)\theta^{2an}\exp(-\theta^2\sum_{i=1}^n y_i^2) \propto \theta^{2(c+an-1/2)}\exp[-\theta^2(d+\sum_{i=1}^n y_i^2)].$$

The prior is therefore of the form in a) with first parameter $c + an$ and second parameter $d + \sum_{i=1}^n y_i^2$.

c) We have

$$\frac{p(\theta_1|\boldsymbol{y})}{p(\theta_2|\boldsymbol{y})} = \left(\frac{\theta_1}{\theta_2}\right)^{2(c+an-1/2)} \exp\left[-\left(d + \sum_{i=1}^n y_i^2\right)\left(\theta_1^2 - \theta_2^2\right)\right].$$

A sufficient statistic is $\sum_{i=1}^n Y_i^2$.

d) The posterior mean is

$$\begin{aligned}
E(\theta|\boldsymbol{y}) &= \frac{2\left(d + \sum_{i=1}^n y_i^2\right)^{c+an}}{\Gamma(c+an)} \int_0^\infty \theta^{2(c+an+1/2-1/2)} \exp[-\theta^2(d + \sum_{i=1}^n y_i^2)]\, d\theta \\
&= \frac{2\left(d + \sum_{i=1}^n y_i^2\right)^{c+an}}{\Gamma(c+an)} \cdot \frac{\Gamma(c+an+1/2)}{2\left(d + \sum_{i=1}^n y_i^2\right)^{c+an+1/2}} \\
&= \frac{\Gamma(c+an+1/2)}{\Gamma(c+an)} \left(d + \sum_{i=1}^n y_i^2\right)^{-1/2}.
\end{aligned}$$

e) We'll assume that $\tilde{Y}$ and $Y_1, \ldots, Y_n$ are independent given θ. Then

$$\begin{aligned}
p(\tilde{y}|\boldsymbol{y}) &= \frac{p(\tilde{y}, \boldsymbol{y})}{p(\boldsymbol{y})} \\
&= \frac{1}{p(\boldsymbol{y})} \int_0^\infty p(\tilde{y}|\theta) p(\boldsymbol{y}|\theta) p(\theta)\, d\theta \\
&= \int_0^\infty p(\tilde{y}|\theta) p(\theta|\boldsymbol{y})\, d\theta \\
&= \frac{2\left(d + \sum_{i=1}^n y_i^2\right)^{c+an}}{\Gamma(c+an)} \cdot \frac{2\tilde{y}^{2a-1}}{\Gamma(a)} \int_0^\infty \theta^{2(c+a(n+1)-1/2)} \times \\
&\qquad \exp\left[-\theta^2\left(\tilde{y}^2 + d + \sum_{i=1}^n y_i^2\right)\right] d\theta \\
&= \frac{2\left(d + \sum_{i=1}^n y_i^2\right)^{c+an}}{\Gamma(c+an)} \cdot \frac{2\tilde{y}^{2a-1}}{\Gamma(a)} \cdot \frac{\Gamma(c+a(n+1))}{2\left(\tilde{y}^2 + d + \sum_{i=1}^n y_i^2\right)^{c+a(n+1)}} \\
&= \frac{2\Gamma(c+a(n+1))}{\Gamma(a)\Gamma(c+an)} \cdot \frac{\tilde{y}^{2a-1}\left(d + \sum_{i=1}^n y_i^2\right)^{c+an}}{\left(\tilde{y}^2 + d + \sum_{i=1}^n y_i^2\right)^{c+a(n+1)}}.
\end{aligned}$$

3.14 a) The likelihood function is $L(\theta) = \theta^y(1-\theta)^{n-y}$, where $y = \sum_{i=1}^n y_i$. This is maximized at the same θ as is the log-likelihood function: $\ell(\theta) = y \log\theta + (n-y)\log(1-\theta)$. We have

$$\ell'(\theta) = \frac{y}{\theta} - \frac{n-y}{1-\theta},$$

and $\ell'(\theta) = 0$ if and only if $\theta = y/n = \hat{\theta}$. To find the observed Fisher information, we need

$$\frac{\partial^2 \ell(\theta)}{\partial\theta^2} = -\frac{y}{\theta^2} - \frac{n-y}{(1-\theta)^2}.$$

It follows that

$$\begin{aligned} J(\hat{\theta}) &= \frac{n\hat{\theta}}{\hat{\theta}^2} + \frac{n(1-\hat{\theta})}{(1-\hat{\theta})^2} \\ &= n\left(\frac{1}{\hat{\theta}} + \frac{1}{(1-\hat{\theta})}\right) \\ &= \frac{n}{\hat{\theta}(1-\hat{\theta})}. \end{aligned}$$

b) Since $\ell(\theta)/n = \hat{\theta}\log\theta + (1-\hat{\theta})\log(1-\theta)$, the prior is such that

$$p_U(\theta) \propto \theta^{\hat{\theta}}(1-\theta)^{1-\hat{\theta}},$$

and hence the prior is beta$(\hat{\theta}+1, 2-\hat{\theta})$. The same calculations as in a) show that

$$-\frac{\partial^2 \log p_U(\theta)}{\partial\theta^2} = \frac{\hat{\theta}}{\theta^2} + \frac{(1-\hat{\theta})}{(1-\theta)^2}.$$

c) The posterior is proportional to $\theta^{y+\hat{\theta}}(1-\theta)^{n+1-y-\hat{\theta}}$, and so would be a beta$(y+\hat{\theta}+1, n+2-y-\hat{\theta})$ density. The only reason this might not be considered a posterior for θ is that the prior depends on the data, which strictly speaking does not conform with Bayesian principles. However, choosing a prior in this way is commonly done in practice.

d) The maximum likelihood estimate is $\hat{\theta} = \bar{y} = \sum_{i=1}^n y_i/n$, and $J(\hat{\theta})/n = 1/\bar{y}$. The unit information prior is gamma$(\bar{y}+1, 1)$, and

$$-\frac{\partial^2 \log p_U(\theta)}{\partial\theta^2} = \frac{\bar{y}}{\theta^2}.$$

The posterior is gamma$((n+1)\bar{y}+1, n+1)$.