Bayesian Theory Workshop 4: Solutions

1. (5 marks.) For a Normal random variable inference about μ, given known σ^2: posterior, HPDI, and posterior predictive distribution.

   (a) From Lecture Notes 3, equations 3.8 (setting $n = 8$, m=1, $\bar{y} = 0.2625$, $\mu_0 = 0$, and $\sigma_0^2 = 1$), we have,

   $$\text{Posterior for } \mu|\mathbf{y},\sigma^2: \quad \text{Normal}\left(\frac{m\mu_0 + n\bar{y}}{m+n}, \frac{\sigma^2}{m+n}\right)$$
   $$= \text{Normal}\left(\frac{1*0+8*0.2625}{1+8}, \frac{1}{1+8}\right) = \text{Normal}(0.2333, 0.111)$$

   (b) Because the posterior distribution is symmetric (and unimodal), the 95% HPDI is equal to the 95% symmetric credible interval and hence the lower and upper 2.5% quantiles. Thus, the 95% HPDI is $0.2333 \pm 1.96\sqrt{0.111} = (-0.420, 0.887)$. Similarly, the 90% HPDI is $0.2333 \pm 1.6449\sqrt{0.111} = (-0.315, 0.782)$.

   (c) From Lecture Notes 3, equation 3.13, with the posterior mean for μ denoted μ_1 and the posterior variance denoted σ_1^2:

   $$y^{new}|\mathbf{y}^{old} \sim \text{Normal}\left(\mu_1, \sigma_1^2 + \sigma^2\right) = \text{Normal}\left(0.2333, 1 + 0.111\right) = \text{Normal}\left(0.2333, 1.111\right)$$

2. (2 marks.) This exercise is showing how to calculate posterior model probabilities from Bayes Factors. By definition,

   $$BF_{01} = \frac{\Pr(H_0|\mathbf{y})/\Pr(H_1|\mathbf{y})}{p_0/p_1} = \frac{\Pr(H_0|\mathbf{y})/(1-\Pr(H_0|\mathbf{y}))}{p_0/p_1}$$
   $$\Rightarrow$$
   $$BF_{01}\frac{p_0}{p_1} = \frac{\Pr(H_0|\mathbf{y})}{(1-\Pr(H_0|\mathbf{y}))}$$
   $$\Rightarrow$$
   $$\Pr(H_0|\mathbf{y}) = \frac{p_0}{p_0 + p_1/BF_{01}}$$

3. (5 marks.) Calculating BF_{01} for two simple hypotheses about the mean, μ, of a Normal(μ, σ^2) random variable, where σ^2 is known.

   (a) As was shown in Lecture Notes 5, BF_{01} is the ratio of the likelihood under H_0 to the likelihood for H_1. However, we rederive the results here. For hypothesis i, i=0,1:

   $$\Pr(H_i|\mathbf{y}) = p(\mu_i|\mathbf{y}) \propto p_i f(\mathbf{y}|\mu_i) \propto p_i e^{-\frac{\sum_{j=1}^{n}(y_j-\mu_i)^2}{2\sigma^2}} \propto p_i e^{-\frac{n(\bar{y}-\mu_i)^2}{2\sigma^2}}$$

   where the last term is eq'n 3.2 in Lecture Notes 3. Then

   $$BF_{01} = \frac{\left[p_0 e^{-\frac{n(\bar{y}-\mu_0)^2}{2\sigma^2}}\right] \Big/ \left[p_1 e^{-\frac{n(\bar{y}-\mu_1)^2}{2\sigma^2}}\right]}{p_0/p_1}$$
   $$= e^{-\frac{n(\bar{y}-\mu_0)^2}{2\sigma^2} + \frac{n(\bar{y}-\mu_1)^2}{2\sigma^2}} = e^{-\frac{n}{2\sigma^2}(\bar{y}^2 - 2\bar{y}\mu_0 + \mu_0^2 - \bar{y}^2 + 2\bar{y}\mu_1 - \mu_1^2)}$$
   $$= e^{-\frac{n}{2\sigma^2}(\mu_0^2 - \mu_1^2 - 2\bar{y}(\mu_0 - \mu_1))} = e^{-\frac{n}{2\sigma^2}((\mu_0-\mu_1)(\mu_0+\mu_1-2\bar{y}))}$$

(b) For the given values,

$BF_{01} = e^{-\frac{9}{2*1}((0-1)(0+1-2*0.645))} = e^{-9*0.145} = 0.272. Without\ intermediate\ rounding, 0.27117.$

Examining the Bayes factor of H_1 against H_0 as well:

$$BF_{10} = \frac{1}{BF_{01}} = 1/0.272 = 3.69.$$

(c) Interpretation: Using the Kass and Raftery rule of thumb guide, there is no support for H_0 over H_1, but there is positive evidence to support H_1, that $\mu = 1$, over H_0, where $\mu = 0$.

(d) Effect of increasing n on BF_{01} and BF_{10}:

$$BF_{01} = e^{-n*0.145} \quad \text{vs} \quad BF_{10} = e^{+n*0.145}$$

As we increase n, BF_{01} decreases, (or equivalently, BF_{10} increases), increasing the support for H_1 and $\mu = 1$.

4. (8 marks.) Calculating a joint posterior distribution for both parameters of a Normal(μ, σ^2) distribution.

(a) Use the change of variable formula to find the "induced" prior for σ^2 given the prior for σ: $\sigma \sim$ Uniform(0,T), where T is large.
Define $g(\sigma) = \sigma^2$, then $g^{-1}(\sigma^2) = \sqrt{\sigma^2}$. Note given that σ is defined on (0,T), thusσ^2 is defined on (0,T^2). Using the change of variables formulat, we have the probability density function for σ^2,

$$\pi_{\sigma^2}(\sigma^2) = \pi_\sigma(g^{-1}(\sigma^2)) \left| \frac{\mathrm{d}g^{-1}(\sigma^2)}{\mathrm{d}\sigma^2} \right| I(\sigma^2 \leq T^2) = \frac{1}{T} \left| \frac{\mathrm{d}\sqrt{\sigma^2}}{\mathrm{d}\sigma^2} \right| I(\sigma^2 \leq T^2)$$
$$= \frac{1}{T} \left| \frac{1}{2}(\sigma^2)^{-1/2} \right| I(\sigma^2 \leq T^2) = \frac{1}{2T} \frac{1}{\sqrt{\sigma^2}} I(\sigma^2 \leq T^2).$$

In practice T is set very high with the assumption that Pr($\sigma > T$)=0.

(b) To calculate the posterior conditional distributions, we begin by calculating the joint posterior distribution. We then calculate the conditional distributions by considering each parameter in turn, and treating all other parameters as "fixed".
The joint posterior distribution for μ and σ^2, given the support $-\infty < \mu < \infty$ and $\sigma^2 \in (0, T^2]$:

$$p(\mu, \sigma^2|\mathbf{y}) \propto e^{-\frac{\mu^2}{2\sigma_0^2}} \cdot (\sigma^2)^{-\frac{1}{2}} \cdot (\sigma^2)^{-\frac{n}{2}} e^{-\frac{\sum_{i=1}^n (y_i-\mu)^2}{2\sigma^2}} I(0 \leq \sigma^2 \leq T^2)$$

To find the conditionals for μ and σ^2, just examine the terms in joint posterior containing each.

$$p(\mu|\sigma^2, \mathbf{y}) \propto e^{-\frac{\mu^2}{2\sigma_0^2}} \cdot e^{-\frac{\sum_{i=1}^n (y_i-\mu)^2}{2\sigma^2}}$$

which is the kernel for a Normal, and after some algebra yields Normal$\left(\frac{\sigma_0^2 n}{n\sigma_0^2+\sigma^2}\bar{y}, \frac{\sigma^2\sigma_0^2}{\sigma_0^2 n+\sigma^2}\right)$.
The posterior conditional distribution of σ^2 given μ is then,

$$p(\sigma^2|\mu, \mathbf{y}) \propto (\sigma^2)^{-(n/2+1/2)} \exp\left(-\frac{\sum_{i=1}^n (y_i - \mu)^2}{2\sigma^2}\right) I(0 \leq \sigma^2 \leq T^2)$$
$$\Rightarrow \quad \sigma^2|\mu, \mathbf{y} \sim \text{Inverse Gamma}\left(\frac{n}{2} - \frac{1}{2}, \frac{\sum_{i=1}^n (y_i - \mu)^2}{2}\right) \frac{1}{F_{\Gamma^{-1}}(T^2)} I(0 \leq \sigma^2 \leq T^2).$$

where $F_{\Gamma^{-1}}(T^2)$ is the cumulative distribution function of this inverse Gamma distribution, and $\sigma^2|\mu, \mathbf{y}$ is a truncated inverse Gamma distribution (truncated for $\sigma^2 \leq T^2$).