Solutions for Module 3 and 4

1. Assume a priori that $p \sim Be(\alpha, \beta)$. Then we need to solve

$$E[p] = \frac{\alpha}{\alpha+\beta} = \frac{1}{3} \quad \text{and} \quad Var[p] = \frac{\alpha\beta}{(\alpha+\beta)^2(\alpha+\beta+1)} = \frac{1}{32}.$$

From the first equation we obtain $\beta = 2\alpha$. Inserting this in the second equation and isolating α gives $\alpha = \frac{55}{27}$, which in turn implies that $\beta = \frac{110}{27}$. Observing $x = 8$ success in $n = 20$ trials, it follows from Section 2.1 that $p|x \sim Be(x+\alpha, n-x+\beta) = Be(8+\frac{55}{27}, 12+\frac{110}{27})$.

2. Observing x_1 successes in n_1 trials gives posterior $p|x_1 \sim Be(\alpha_1, \beta_1)$, where $\alpha_1 = x_1+\alpha$ and $\beta_1 = n_1 - x_1 + \beta$. Now use this as our prior, and assume we observe a further x_2 successes in the next n_2 trials. The posterior is then $p|x_1, x_2 \sim Be(\alpha_2, \beta_2)$ where $\alpha_2 = x_2 + \alpha_1 = x_1 + x_2 + \alpha$ and $\beta_2 = n_2 - x_2 + \beta_1 = n_1 + n_2 - x_1 - x_2 + \beta$. Notice that $\alpha_1 + \beta_1 = n_1 + \alpha + \beta$, $\alpha_2 + \beta_2 = n_1 + n_2 + \alpha + \beta$ and so on if we repeat everything. Therefore, in some sense, we can interpret $\alpha + \beta$ as representing the number of experiments that our prior knowledge corresponds to.

3. (a) A priori we assume $\lambda \sim \text{Gamma}(\alpha, \beta)$, i.e.

$$\pi(\lambda) = \frac{\lambda^{\alpha-1}e^{-\lambda/\beta}}{\Gamma(\alpha)\beta^\alpha}.$$

The posterior density is then

$$\begin{aligned}\pi(\lambda|x) &\propto \pi(x|\lambda)\pi(\lambda)\\ &= \frac{e^{-\lambda}\lambda^x}{x!}\frac{\lambda^{\alpha-1}e^{-\lambda/\beta}}{\Gamma(\alpha)\beta^\alpha}\\ &\propto \lambda^{x+\alpha-1}e^{-\lambda(1+1/\beta)}\end{aligned}$$

and so $\lambda|x \sim \text{Gamma}(x+\alpha, \beta/(1+\beta))$.

Remark: A more common situation is when $x \sim \text{Pois}(\lambda t)$, which corresponds to x being the random number of events in a Poisson process with rate λ on an interval of length t. In this case the posterior is $\lambda|x \sim \text{Gamma}(x+\alpha, \beta/(1+t\beta))$. Here the posterior mean is

$$E[\lambda|x] = \frac{(x+\alpha)\beta}{1+t\beta} = \frac{x\beta}{1+t\beta} + \frac{\alpha\beta}{1+t\beta}.$$

Now, as t increases, $E[\lambda|x]$ will tend towards x/t which is the usual frequentist estimator.

(b) If we let $x_1, \ldots, x_6$ be the six observations and set $x = x_1 + \ldots + x_6$, we obtain the likelihood

$$\pi(x_1, \ldots, x_6|\lambda) = \prod_{i=1}^{6}\pi(x_i|\lambda) = \prod_{i=1}^{6}\exp(-\lambda)\frac{\lambda^{x_i}}{x_i!} \propto \lambda^{\sum_i x_i}\exp(-6\lambda) = \lambda^x\exp(-6\lambda).$$

Then along similar lines as in (a) it is seen that $\lambda|x \sim \text{Gamma}(x+\alpha, \beta/(1+6\beta))$ (in agreement with the remark above!). Finally, a priori we want $\alpha\beta = 3$ and $\alpha\beta^2 = 4$, i.e. $\beta = 4/3$ and $\alpha = 9/4$.

4. Let $\underline{x} = (x_1, x_2, \ldots, x_n)$ denote the vector of observations. The posterior density for

the mean is

$$\begin{aligned}
\pi(\mu|\underline{x}) &\propto \pi(\underline{x}|\mu)\pi(\mu) \\
&= \Big(\frac{\tau}{2\pi}\Big)^{\frac{n}{2}} \exp\Big(-\frac{1}{2}\tau\sum_{i=1}^{n}(x_i-\mu)^2\Big)\sqrt{\frac{\tau_0}{2\pi}}\exp\Big(-\frac{1}{2}\tau_0(\mu-\mu_0)^2\Big) \\
&\propto \exp\Big(-\frac{1}{2}\tau\sum_{i=1}^{n}\mu^2+\tau\mu\sum_{i=1}^{n}x_i-\frac{1}{2}\tau_0\mu^2+\tau_0\mu_0\mu\Big) \\
&= \exp\Big(-\frac{1}{2}(\tau_0+n\tau)\mu^2+(\tau\sum_{i=1}^{n}x_i+\tau_0\mu_0)\mu\Big) \\
&= \exp(-\frac{1}{2}\tau_1\mu^2+\tau_1\mu_1\mu).
\end{aligned}$$

Comparing this to equation (2) we see that $\mu|\underline{x} \sim N(\mu_1, \tau_1)$, where

$$\tau_1 = \tau_0 + n\tau \quad \text{and} \quad \mu_1 = \frac{\tau_1\mu_1}{\tau_1} = \frac{\tau\sum_{i=1}^{n} x_i + \tau_0\mu_0}{\tau_0+n\tau} = \frac{\tau n\bar{x}+\tau_0\mu_0}{\tau_0+n\tau}.$$

The posterior density for the precision is

$$\begin{aligned}
\pi(\tau|\underline{x}) &\propto \pi(\underline{x}|\tau)\pi(\tau) \\
&= \Big(\frac{\tau}{2\pi}\Big)^{\frac{n}{2}} \exp\Big(-\frac{1}{2}\tau\sum_{i=1}^{n}(x_i-\mu)^2\Big)\frac{\tau^{\alpha-1}e^{-\tau/\beta}}{\Gamma(\alpha)\beta^{\alpha}} \\
&\propto \tau^{\frac{n}{2}+\alpha-1}\exp\Big(-\tau\Big(\frac{1}{2}\sum_{i=1}^{n}(x_i-\mu)^2+\frac{1}{\beta}\Big)\Big)
\end{aligned}$$

Comparing this to the density of a gamma distributed random variable we see that $\tau|\underline{x} \sim \text{Gamma}(\alpha_1, \beta_1)$, where β_1 denotes the scale parameter and

$$\alpha_1 = \frac{n}{2} + \alpha \quad \text{and} \quad \beta_1 = \frac{1}{\frac{1}{2}\sum_{i=1}^{n}(x_i-\mu)^2+\frac{1}{\beta}}.$$

5. (a) A priori μ is drawn from $\mathcal{N}(0,1)$ with probability 1/3, and else it is drawn from $\mathcal{N}(1,1)$.

   (b) Follows by a straightforward calculation using (2).

   (c) A posteriori μ is drawn from $\mathcal{N}(\frac{\tau x}{1+\tau}, 1+\tau)$ with probability

$$\frac{\exp\Big(\frac{1}{2}\frac{(\tau x)^2}{1+\tau}\Big)}{\exp\Big(\frac{1}{2}\frac{(\tau x)^2}{1+\tau}\Big)+2\exp\Big(\frac{1}{2}\frac{(1+\tau x)^2}{1+\tau}-\frac{1}{2}\Big)},$$

and else it is drawn from $\mathcal{N}(\frac{1+\tau x}{1+\tau}, 1+\tau)$. So the prior and posterior distributions are conjugate within the family of mixture distributions given by two normal distributions.