STAT 638: Solution for Homework #7

8.1 a) Intuitively, $\text{Var}(Y_{i,j}|\mu,\tau^2)$ should be bigger because it includes both variance within and between groups.

b) The model says that, given θ_j and σ^2, $Y_{1,j},\ldots,Y_{n_j,j}$ are i.i.d., and so definitely $\text{Cov}(Y_{i_1,j},Y_{i_2,j}|\theta_j,\sigma^2)=0$. Intuitively, the covariance between $Y_{i_1,j}$ and $Y_{i_2,j}$ given μ and τ^2 should be positive since data values in the same group will tend to be on the same side of μ.

c) $\text{Var}(Y_{i,j}|\theta_j,\sigma^2)=\sigma^2$
$\text{Var}(\bar{Y}_{\cdot,j}|\theta_j,\sigma^2)=\sigma^2/n_j$
$\text{Cov}(Y_{i_1,j},Y_{i_2,j}|\theta_j,\sigma^2)=0$

$$\begin{aligned}\text{Var}(Y_{i,j}|\mu,\tau^2) &= E\left(\text{Var}(Y_{i,j}|\mu,\tau^2,\theta_j,\sigma^2)\right)+\text{Var}\left(E(Y_{i,j}|\mu,\tau^2,\theta_j,\sigma^2)\right)\\ &= E(\sigma^2)+\text{Var}(\theta_j)\\ &= \sigma^2+\tau^2\end{aligned}$$

$$\begin{aligned}\text{Var}(\bar{Y}_{\cdot,j}|\mu,\tau^2) &= E\left(\text{Var}(\bar{Y}_{\cdot,j}|\mu,\tau^2,\theta_j,\sigma^2)\right)+\text{Var}\left(E(\bar{Y}_{\cdot,j}|\mu,\tau^2,\theta_j,\sigma^2)\right)\\ &= E\left(\frac{\sigma^2}{n_j}\right)+\text{Var}(\theta_j)\\ &= \frac{\sigma^2}{n_j}+\tau^2\end{aligned}$$

$$\begin{aligned}\text{Cov}(Y_{i_1,j},Y_{i_2,j}|\mu,\tau^2) &= E\left(\text{Cov}(Y_{i_1,j},Y_{i_2,j}|\mu,\tau^2,\theta_j,\sigma^2)\right)\\ &\quad+\text{Cov}\left(E(Y_{i_1,j}|\mu,\tau^2,\theta_j,\sigma^2),E(Y_{i_2,j}|\mu,\tau^2,\theta_j,\sigma^2)\right)\\ &= 0+\text{Cov}(\theta_j,\theta_j)\\ &= \text{Var}(\theta_j)\\ &= \tau^2.\end{aligned}$$

d)

$$\begin{aligned}p(\mu|\theta_1,\ldots,\theta_m,\sigma^2,\tau^2,\boldsymbol{y}_1,\ldots,\boldsymbol{y}_m) &= \frac{p(\boldsymbol{y}_1,\ldots,\boldsymbol{y}_m,\mu,\theta_1,\ldots,\theta_m,\sigma^2,\tau^2)}{p(\boldsymbol{y}_1,\ldots,\boldsymbol{y}_m,\theta_1,\ldots,\theta_m,\sigma^2,\tau^2)}\\ &= \frac{p(\boldsymbol{y}_1,\ldots,\boldsymbol{y}_m|\mu,\theta_1,\ldots,\theta_m,\sigma^2,\tau^2)}{p(\boldsymbol{y}_1,\ldots,\boldsymbol{y}_m|\theta_1,\ldots,\theta_m,\sigma^2,\tau^2)}\cdot\frac{p(\mu,\theta_1,\ldots,\theta_m,\sigma^2,\tau^2)}{p(\theta_1,\ldots,\theta_m,\sigma^2,\tau^2)}\\ &= \frac{p(\boldsymbol{y}_1,\ldots,\boldsymbol{y}_m|\theta_1,\ldots,\theta_m,\sigma^2)}{p(\boldsymbol{y}_1,\ldots,\boldsymbol{y}_m|\theta_1,\ldots,\theta_m,\sigma^2)}\cdot\frac{p(\mu,\theta_1,\ldots,\theta_m,\sigma^2,\tau^2)}{p(\theta_1,\ldots,\theta_m,\sigma^2,\tau^2)}\\ &= \frac{p(\theta_1,\ldots,\theta_m|\mu,\sigma^2,\tau^2)}{p(\theta_1,\ldots,\theta_m|\sigma^2,\tau^2)}\cdot\frac{p(\mu,\sigma^2,\tau^2)}{p(\sigma^2,\tau^2)}\\ &= \frac{p(\theta_1,\ldots,\theta_m|\mu,\tau^2)}{p(\theta_1,\ldots,\theta_m|\tau^2)}\cdot\frac{p(\mu,\sigma^2,\tau^2)}{p(\sigma^2,\tau^2)}\end{aligned}$$

If we assume that, given τ^2, μ is independent of σ^2, then the last expression may be written as $p(\mu|\theta_1,\ldots,\theta_m,\tau^2)$. This implies that the full conditional of μ is free of the data and σ^2.

8.3 a) The mixing was extremely good. The autocorrelation function for each of σ^2, μ and τ^2 died out after two or three lags, and hence the effective sample size for each parameter is at least 1/3 the number of values generated.

b) The posterior means for σ^2, μ and τ^2 were 14.52, 7.58 and 5.66, respectively. Respective 95% credible regions were $(11.80, 17.86)$, $(5.96, 9.10)$ and $(1.96, 14.59)$.

c) The density indicates that the most likely values for R are between 0.2 and 0.6. So there is evidence of between schools variation, but the variation is not large in comparison to within schools variance.

d) $P(\theta_7 < \theta_6|\text{data}) \approx 0.52$ and $P(\theta_7 = \min(\theta_1, \ldots, \theta_8)|\text{data}) \approx 0.32$

e) The plot shows that the posterior means are closer to the overall sample mean than are the sample means themselves. This is the shrinkage phenomenon. The overall sample mean is 7.691 and the posterior mean of μ is 7.643.