8.1 Components of variance: Consider the hierarchical model where

$$\theta_1, \ldots, \theta_m | \mu, \tau^2 \sim \text{i.i.d. normal}(\mu, \tau^2)$$
$$y_{1,j}, \ldots, y_{n_j,j} | \theta_j, \sigma^2 \sim \text{i.i.d. normal}(\theta_j, \sigma^2)\,.$$

For this problem, we will eventually compute the following:
$\text{Var}[y_{i,j}|\theta_i, \sigma^2]$, $\text{Var}[\bar{y}_{\cdot,j}|\theta_i, \sigma^2]$, $\text{Cov}[y_{i_1,j}, y_{i_2,j}|\theta_j, \sigma^2]$
$\text{Var}[y_{i,j}|\mu, \tau^2]$, $\text{Var}[\bar{y}_{\cdot,j}|\mu, \tau^2]$, $\text{Cov}[y_{i_1,j}, y_{i_2,j}|\mu, \tau^2]$
First, lets use our intuition to guess at the answers:

a) Which do you think is bigger, $\text{Var}[y_{i,j}|\theta_i, \sigma^2]$ or $\text{Var}[y_{i,j}|\mu, \tau^2]$? To guide your intuition, you can interpret the first as the variability of the Y's when sampling from a fixed group, and the second as the variability in first sampling a group, then sampling a unit from within the group.
b) Do you think $\text{Cov}[y_{i_1,j}, y_{i_2,j}|\theta_j, \sigma^2]$ is negative, positive, or zero? Answer the same for $\text{Cov}[y_{i_1,j}, y_{i_2,j}|\mu, \tau]$. You may want to think about what $y_{i_2,j}$ tells you about $y_{i_1,j}$ if θ_j is known, and what it tells you when θ_j is unknown.

8.1

(a) $\text{Var}(Y_{ij} | \mu, \tau^2)$ 은 within 과 between Variability를 모두 포함하고 있고, $\text{Var}(Y_{ij} | \theta_i, \sigma^2)$은 within variability 만 있다. 그러므로 $\text{Var}(Y_{ij} | \mu, \tau^2)$이 더 클것 같다.

(b) $\text{Cov}(Y_{i_1 j}, Y_{i_2 j} | \theta_j, \sigma^2)$: $Y_{1j}, Y_{2j} \cdots Y_{n_j j}$ 는 θ_j, σ^2 이 given일때, conditionally i.i.d를 만족한다. 그러므로 $\text{Cov}(Y_{i_1 j}, Y_{i_2 j} | \theta_j, \sigma^2) = 0$ 일것이다.
$\text{Cov}(Y_{i_1 j}, Y_{i_2 j} | \mu, \tau) > 0$ positive 일것이다

θ_j가 known 일때는 $Y_{i_1 j}, Y_{i_2 j}$ 의 cov가 0 이므로, $Y_{i_2 j}$가 $Y_{i_1 j}$ 에 대해 정보를 많이 주지 못할것이다. 하지만 θ_j가 unknown일때는 $Y_{i_2 j}$는 $Y_{i_1 j}$ 에 대해 정보를 줄것이다.

when θ_j is unknown.

c) Now compute each of the six quantities above and compare to your answers in a) and b).

d) Now assume we have a prior $p(\mu)$ for μ. Using Bayes' rule, show that

$$p(\mu|\theta_1,\ldots,\theta_m,\sigma^2,\tau^2,\boldsymbol{y}_1,\ldots,\boldsymbol{y}_m) = p(\mu|\theta_1,\ldots,\theta_m,\tau^2).$$

c) $\mathrm{Var}[y_{i,j}|\theta_i,\sigma^2]$, $\mathrm{Var}[\bar{y}_{\cdot,j}|\theta_i,\sigma^2]$, $\mathrm{Cov}[y_{i_1,j},y_{i_2,j}|\theta_j,\sigma^2]$
$\mathrm{Var}[y_{i,j}|\mu,\tau^2]$, $\mathrm{Var}[\bar{y}_{\cdot,j}|\mu,\tau^2]$, $\mathrm{Cov}[y_{i_1,j},y_{i_2,j}|\mu,\tau^2]$

① $\mathrm{Var}(Y_{IJ}|\theta_I,\sigma^2) = \sigma^2$

$Y_{IJ}|\theta_I,\sigma^2 \sim N(\mu,\tau^2)$

② $\mathrm{Var}(\bar{Y}_{\cdot J}|\theta_I,\sigma^2)$

$= \mathrm{Var}\left(\frac{\sum_{I}^{n_J} Y_{IJ}}{n_J}\,\Big|\,\theta_I,\sigma^2\right) = \frac{\sigma^2}{n}$

③ $\mathrm{Cov}(Y_{I_1J},Y_{I_2J}|\theta_J,\sigma^2)$

$= E[Y_{I_1J}Y_{I_2J}|\theta_J,\sigma^2) - E(Y_{I_1J}|\theta_J,\sigma^2)E(Y_{I_2J}|\theta_J,\sigma^2)$ ($\because$ conditionally indep.)

$=0$

④ $\mathrm{Var}(Y_{IJ}|\mu,\tau^2) = \mathrm{Var}(E(Y_{IJ}|\theta_J,\sigma^2)) + E(\mathrm{Var}(Y_{IJ}|\theta_J,\sigma^2)|\mu,\tau^2)$

$= \mathrm{Var}(\theta_J|\mu,\tau^2) + E(\sigma^2|\mu,\tau^2) = \tau^2+\sigma^2$

⑤ $\mathrm{Var}(\bar{Y}_{\cdot J}|\mu,\tau^2)$

$= \mathrm{Var}[E(\bar{Y}_{\cdot J}|\theta_J,\sigma^2)|\mu,\tau^2) + E(\mathrm{Var}(\bar{Y}_{\cdot J}|\theta_J,\sigma^2)|\mu,\tau^2)$

$= \mathrm{Var}(\theta_J|\mu,\tau^2) + E\left(\frac{\sigma^2}{n_J}\,\Big|\,\mu,\tau^2\right)$

$= \tau^2 + \frac{\sigma^2}{n_J}$

⑥ $\mathrm{Cov}(Y_{I_1J},Y_{I_2J}|\mu,\tau^2) = E[\mathrm{cov}(Y_{I_1J},Y_{I_2J}|\theta_J,\sigma^2)|\mu,\tau^2)$

$+ \mathrm{Cov}(E(Y_{I_1J}|\theta_J,\sigma^2), E(Y_{I_2J}|\theta_J,\sigma^2)]$

$= E[0|\mu,\tau^2) + \mathrm{Cov}($ 〃 $)$

$= \mathrm{Cov}(\theta_J,\theta_J) = \tau^2$

d) $$p(\mu|\underbrace{\theta_1,\dots,\theta_m}_{=\theta},\sigma^2,\tau^2,\underbrace{\boldsymbol{y}_1,\dots,\boldsymbol{y}_m}_{=Y}) = p(\mu|\theta_1,\dots,\theta_m,\tau^2).$$

$$= p(\mu|Y,\theta,\sigma^2,\tau^2) = \frac{p(\mu,Y,\theta,\sigma^2,\tau^2)}{\int p(\mu,Y,\theta,\sigma^2,\tau^2)\,d\mu}$$

$$= \frac{p(\mu)\,p(\tau^2)\,p(\sigma^2)\,p(Y|\theta,\sigma^2)\,p(\theta|\mu,\tau^2)}{\int \underline{p(\mu)}\,p(\tau^2)\,p(\sigma^2)\,p(Y|\theta,\sigma^2)\,\underline{p(\theta|\mu,\tau^2)}\,d\mu}$$

$$= \frac{p(\mu)\,\cancel{p(\tau^2)}\,\cancel{p(\sigma^2)}\,\cancel{p(Y|\theta,\sigma^2)}\,p(\theta|\mu,\tau^2)}{\cancel{p(\tau^2)}\,\cancel{p(\sigma^2)}\,\cancel{p(Y|\theta,\sigma^2)}\int p(\mu)\,p(\theta|\mu,\tau^2)\,d\mu}$$

$$= \frac{p(\mu)\,p(\theta|\mu,\tau^2)}{\int p(\mu)\,p(\theta|\mu,\tau^2)\,d\mu} = p(\mu|\theta,\tau^2)$$

μ는 Y, σ^2 에 dependent 하지 않음을 알 수 있다.