We've been considering the settings where

$$\begin{tikzcd}
\theta_1 \arrow[d] & \theta_2 \arrow[d] & \cdots & \theta_J \arrow[d] \\
Y_1 & Y_2 & & Y_J
\end{tikzcd}
\qquad\qquad
\begin{tikzcd}
 & & \theta \arrow[dll] \arrow[dl] \arrow[dr] & \\
Y_1 & Y_2 & \cdots & Y_J
\end{tikzcd}$$

Is there an inbetween? Hierarchy

$$\begin{tikzcd}
 & & \phi \arrow[dll] \arrow[dl] \arrow[dr] & & \text{(Hyper prior)} \\
\theta_1 \arrow[d] & \theta_2 \arrow[d] & \cdots & \theta_J \arrow[d] & \text{(Prior)} \\
Y_1 & Y_2 & \cdots & Y_J & \text{(data)}
\end{tikzcd}$$

Two main advantages

1. Borrowing of strength $\theta_j \mid Y$ depends on full dataset

2. More flexibility in modeling uncertainty – inherited from ϕ

$$\mathbb{P}(\theta, \phi \mid Y) \propto \mathbb{P}(Y \mid \theta)\, \mathbb{P}(\theta \mid \phi)\, \mathbb{P}(\phi)$$

This representation makes it simpler to sample from posterior.

(i) Sample $\phi^{(s)} \sim \mathbb{P}(\phi \mid y)$

(ii) Sample $\theta^{(s)} \sim \mathbb{P}(\theta \mid \phi = \phi^s, Y)$

(iii) [For post. pred.] $\tilde{Y} \sim \mathbb{P}(Y \mid \theta = \theta^s)$

Ex: Y_{ij} = Grade of student i in school j.

$$Y_{ij} \mid \theta_j \overset{ind}{\sim} N(\theta_i, \underbrace{\sigma^2}_{\text{known}})$$

$$\bar{Y}_{\cdot j} \mid \theta_j \sim N\left(\theta_j,\ \sigma_j^2 = \sigma^2/n_j\right)$$

Frequentist: $\hat{\theta}_j^{(1)} = \bar{Y}_{\cdot j}$ $\quad \hat{\theta}_j^{(2)} = \frac{1}{J}\sum_{j=1}^{m} \bar{Y}_{\cdot j} = \dfrac{\sum \bar{Y}_{\cdot j}/\sigma_j^2}{\sum 1/\sigma_j^2}$

Common approach is to then test

$H_0: \theta_1 = \theta_2 = \cdots = \theta_J$
- Use $\hat{\theta}_j^{(1)}$ if reject
- Use $\hat{\theta}_j^{(2)}$ if fail to reject

However we could want weighted means

$$\hat{\theta}_j = \lambda_j \hat{\theta}_j^{(1)} + (1-\lambda_j)\, \hat{\theta}_j^{(2)}$$

Consider the hierachial model

$$\begin{tikzcd}[column sep=small]
 & & \phi = (\mu, \tau^2) \arrow[dll, no head] \arrow[dl, no head] \arrow[dr, no head] & \\
\theta_1 \arrow[d] & \theta_2 \arrow[d] & \cdots & \theta_J \arrow[d] \\
\bar{Y}_{\cdot 1} & \bar{Y}_{\cdot 2} & \cdots & \bar{Y}_{\cdot J}
\end{tikzcd}$$

$$\bar{Y}_{\cdot j} \mid \theta_j \overset{ind}{\sim} N(\theta_j, \sigma_j^2)$$

$$\theta_j \mid \mu, \tau^2 \overset{iid}{\sim} N(\mu, \tau^2)$$

From this the full posterior is then

$$\mathbb{P}(\theta,\phi \mid Y) \propto \mathbb{P}(Y\mid\theta)\,\mathbb{P}(\theta\mid\mu,\tau)\,\mathbb{P}(\mu,\tau^2)$$

$$\propto \prod_{j=1}^{J} N(\bar{Y}_{\cdot j}\mid\theta_j,\sigma_j^2)\,N(\theta_j\mid\mu,\tau^2)\,\mathbb{P}(\mu\mid\tau^2)\,\mathbb{P}(\tau^2)$$

For partial conjugacy let's set $\mu\mid\tau^2 \sim N(\mu_0,\tau^2/k_0)$

Still need to understand

(1) $\mathbb{P}(\theta_j\mid\mu,\tau^2,y)$

(2) $\mathbb{P}(\mu,\tau^2\mid y)$

(3) $\mathbb{P}(\tau^2\mid Y)$

<u>Rmk</u>: Useful fact: $X\mid\theta \sim N(\theta,\sigma^2)$ $\quad\theta\sim N(\theta_0,\tau^2)$

(i) $X \sim N(\theta_0,\sigma^2+\tau^2)$

(ii) $\theta\mid X \sim N\left(\dfrac{X/\sigma^2+\theta_0/\tau^2}{1/\sigma^2+1/\tau^2},\ \dfrac{1}{1/\sigma^2+1/\tau^2}\right)$

(1) $\mathbb{P}(\theta_j\mid\mu,\tau^2,Y) \propto N(\bar{Y}_{\cdot j}\mid\theta_j,\sigma_j^2)\,N(\theta_j\mid\mu,\tau^2)$

$$\propto N\left(\frac{\bar{Y}_{\cdot j}/\sigma_j^2+\mu/\tau^2}{1/\sigma_j^2+1/\tau^2},\ \frac{1}{1/\sigma_j^2+1/\tau^2}\right)$$

$$\equiv N(\hat{\theta}_j, v_j)$$

Rmk: $\tau^2 \to 0 : \hat{\theta}_j \to \mu$ (pooled across all)

$\tau^2 \to \infty : \hat{\theta}_j \to \bar{Y}_{\cdot j}$ (no pooling)

(2) $$\mathbb{P}(\mu \mid \tau^2, y) = \int \mathbb{P}(\theta, \mu \mid \tau^2, y)\, d\theta$$

$$\propto \prod_{j=1} \underbrace{\int N(\bar{Y}_{\cdot j} \mid \theta_j, \sigma^2)\, \mathbb{P}(\theta_j \mid \mu, \tau^2)\, d\theta_j}_{\bar{Y}_{\cdot j} \mid \mu, \tau^2 \sim N(\mu, \sigma_j^2 + \tau^2)} \mathbb{P}(\mu \mid \tau^2)$$

$$\propto \prod_{j=1} N(\bar{Y}_{\cdot j} \mid \mu, \sigma_j^2 + \tau^2)\, N(\mu \mid \mu_0, \tau^2/k_0)$$

note

- $\bar{Y}_{\cdot 1} \mid \mu, \tau^2 \sim N(\mu, \sigma_j^2 + \tau^2)$

- $\mu \mid \tau^2 \sim N(\mu_0, \tau^2/k_0)$

$$\mu \mid \bar{Y}_{\cdot 1}, \tau^2 \sim N\left(\frac{\bar{Y}_{\cdot 1}/\sigma_j^2 + \tau^2 + \frac{\mu_0}{\tau^2/k_0}}{\frac{1}{\sigma_j^2 + \tau^2} + \frac{k_0}{\tau^2}}, \frac{1}{\frac{1}{\sigma_j^2 + \tau^2} + \frac{k_0}{\tau^2}} \right)$$

Iteratively updating

$$\mu \mid \tau^2, Y \sim N\left(\hat{\mu} = \frac{\sum_j \left(\bar{y}_{\cdot j} / \sigma_j^2 + \tau^2\right) + \frac{\mu_0 k_0}{\tau^2}}{\sum_j \frac{1}{\sigma_j^2 + \tau^2} + \frac{k_0}{\tau^2}}, \hat{r} = \frac{1}{\sum_j \frac{1}{\sigma_j^2 + \tau^2} + \frac{k_0}{\tau^2}}\right)$$

(3) $$P(\tau^2 \mid y) = \int P(\mu, \tau^2 \mid y)\, d\mu$$

$$\propto \int \prod_j N(\bar{Y}_{\cdot j} \mid \mu, \sigma_j^2 + \tau^2)\, N(\mu \mid \mu_0, \tau^2/k_0)\, P(\tau^2)\, d\mu$$

$$\propto \hat{r}^{1/2} \prod_{j=1}^{J} (\sigma_j^2 + \tau^2)^{-1/2} \exp\left\{ -\frac{1}{2} \left[\frac{\sum_j (\bar{Y}_{\cdot j} - \hat{\mu})^2}{\sigma_j^2 + \tau^2} + \frac{(\hat{\mu} - \mu_0)^2}{\tau^2/k_0} \right] \right\} P(\tau^2$$

$\longrightarrow$ Need numerical methods here.