Math Tests

a.) $y_{ij} \sim N(\theta_i, \sigma^2)$

$$f(y;\theta_1,\dots,\theta_p) = \left(\prod_{i=1}^{p} \left[\prod_{j=1}^{N_i} \frac{1}{\sqrt{2\pi\sigma^2}} \exp\left\{\frac{-1}{2\sigma^2}(y_{ij}-\theta_i)^2\right\}\right.\right.$$

$$\propto \exp\left\{\frac{-1}{2\sigma^2}\sum_{i=1}^{p}\sum_{j=1}^{N_i}(y_{ij}-\theta_i)^2\right\}$$

So the log likelihood is

$$\ell(\theta_1,\dots,\theta_p) = \frac{-1}{2\sigma^2}\sum_{i=1}^{p}\sum_{j=1}^{N_i}(y_{ij}-\theta_i)^2$$

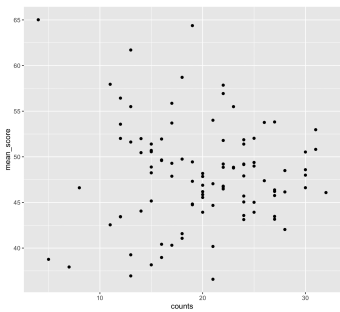
Now to maximize this we take the derivative wrt each θ_i:

$$\frac{d}{d\theta_i}\ell(\theta) = \frac{-1}{2\sigma^2}\sum_{j=1}^{N_i}2(y_{ij}-\theta_i)$$

then set it to 0:

$$\frac{-1}{\sigma^2}\sum_{j=1}^{N_i} y_{ij}-\hat{\theta}_i^{MLE}=0 \Rightarrow \boxed{\hat{\theta}_i^{MLE} = \bar{y}_i}$$

b.) The reason extreme $\bar{y}_i$'s occurs when N_i is low is due to the sample variance being bigger when the sample size is small.



c) $p(y_{ij}|\theta_i,\sigma^2) \sim N(\theta_i,\sigma^2)$

$p(\theta_i|\tau^2,\sigma^2) \sim N(\mu, \tau^2\sigma^2)$

$p(\mu) \propto 1$

$p(\sigma^2) \propto \frac{1}{\sigma^2}$

$p(\tau^2) \sim IG(\frac{1}{2},\frac{1}{2})$

First we find the form of the posterior:

MVN$(\mu\vec{1}, \tau^2\sigma^2 I_p)$

$$p(\theta_i,\mu,\sigma^2,\tau^2|y) \propto p(\mu)\,p(\sigma^2)\,p(\tau^2)\,p(\theta|\tau^2,\sigma^2,\mu)\,p(y|\theta,\sigma^2)$$

$$\propto (1)\left(\frac{1}{\sigma^2}\right)(\tau^2)^{-(\frac{1}{2}+1)}\exp\left\{\frac{-1}{2\tau^2}\right\}\left(\underbrace{\det(2\pi\tau^2\sigma^2 I_p)^{-1/2}}_{(2\pi\tau^2\sigma^2)^{-p/2}}\exp\left\{\frac{-1}{2}(\theta-\mu\vec{1})'(\tau^2\sigma^2 I_p)^{-1}(\theta-\mu\vec{1})\right\}\right)\left(\prod_{i=1}^{p}\prod_{j=1}^{N_i}\frac{1}{\sqrt{2\pi\sigma^2}}\exp\left\{\frac{-1}{2\sigma^2}(y_{ij}-\theta_i)^2\right\}\right)$$

$$\propto (\sigma^2)^{-1}(\tau^2)^{-3/2}\exp\left\{\frac{-1}{2\tau^2}\right\}(\tau^2)^{-p/2}(\sigma^2)^{-p/2}\exp\left\{\frac{-1}{2\tau^2\sigma^2}(\theta-\mu\vec{1})'(\theta-\mu\vec{1})\right\}(\sigma^2)^{-N/2}\exp\left\{\frac{-1}{2\sigma^2}\sum_{i}^{p}\sum_{j}^{N_i}(y_{ij}-\theta_i)^2\right\}$$

$$\propto (\sigma^2)^{-(\frac{N+p}{2})-1}(\tau^2)^{-(\frac{p+3}{2})}\exp\left\{\frac{-1}{2\tau^2}\right\}\exp\left\{\frac{-1}{2\tau^2\sigma^2}(\theta-\mu\vec{1})'(\theta-\mu\vec{1})\right\}\exp\left\{\frac{-1}{2\sigma^2}\sum_{i}^{p}\sum_{j}^{N_i}(y_{ij}-\theta_i)^2\right\}$$

From this we obtain the conditional distributions:

$$p(\sigma^2) \propto (\sigma^2)^{-(\frac{N+p}{2})-1}\exp\left\{-\frac{1}{\sigma^2}\left[\frac{1}{2\tau^2}(\theta-\mu\vec{1})'(\theta-\mu\vec{1})+\frac{1}{2}\sum_{i=1}^{p}\sum_{j}^{N_i}(y_{ij}-\theta_i)^2\right]\right\} \rightarrow IG\left(\frac{N+p}{2}, \frac{1}{2}\left[\frac{1}{\tau^2}(\theta-\mu\vec{1})'(\theta-\mu\vec{1})+\sum_{i=1}^{p}\sum_{j=1}^{N_i}(y_{ij}-\theta_i)^2\right]\right)$$

$$p(\tau^2) \propto (\tau^2)^{-(\frac{p+3}{2})}\exp\left\{\frac{-1}{\tau^2}\left(\frac{1}{2}+\frac{1}{2\sigma^2}(\theta-\mu\vec{1})'(\theta-\mu\vec{1})\right)\right\} \rightarrow IG\left(\frac{p+1}{2}, \frac{1}{2}+\frac{1}{2\sigma^2}(\theta-\mu\vec{1})'(\theta-\mu\vec{1})\right)$$

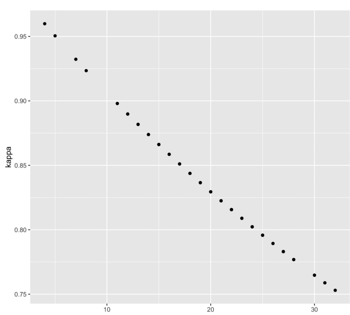
$$p(\theta_i) \propto \exp\left\{\frac{-1}{2\tau^2\sigma^2}(\theta_i-\mu)^2 - \frac{1}{2\sigma^2}\sum_{j=1}^{N_i}(\theta_i-y_{ij})^2\right\} \rightarrow \text{Normal, mean} = \frac{\frac{1}{\tau^2\sigma^2}\mu + \frac{N_i}{\sigma^2}\bar{y}_i}{\frac{1}{\tau^2\sigma^2}+\frac{N_i}{\sigma^2}} \text{ and variance} = \frac{N_i}{\sigma^2}+\frac{1}{\sigma^2\tau^2}$$

$$p(\mu) \propto \exp\left\{\frac{-1}{2\tau^2\sigma^2}(\theta-\mu\vec{1})'(\theta-\mu\vec{1})\right\} = \exp\left\{\frac{-1}{2\tau^2\sigma^2}\sum_{i=1}^{P}(\mu-\theta_i)^2\right\} = \exp\left\{\frac{-1}{2\tau^2\sigma^2}\left(P\mu^2-2\mu\sum_{i=1}^{P}\theta_i+\sum_{i=1}^{P}\theta_i^2\right)\right\} \rightarrow N\left(\bar{\theta}, \frac{\tau^2\sigma^2}{P}\right)$$

Implement Gibbs sampling...

d.) $$E(\theta_i \mid y, \tau^2, \sigma^2, \mu) = \frac{\frac{1}{\tau^2\sigma^2}\mu + \frac{N_i}{\sigma^2}\bar{y}_i}{\frac{1}{\tau^2\sigma^2}+\frac{N_i}{\sigma^2}} = \underbrace{\left(\frac{\frac{1}{\tau^2\sigma^2}}{\frac{1}{\tau^2\sigma^2}+\frac{N_i}{\sigma^2}}\right)}_{k_i}\mu + \left(1-\frac{\frac{1}{\tau^2\sigma^2}}{\frac{1}{\tau^2\sigma^2}+\frac{N_i}{\sigma^2}}\right)\bar{y}$$

$$k_i = \frac{\frac{1}{\tau^2\sigma^2}}{\frac{1+\tau^2 N_i}{\tau^2\sigma^2}} = \boxed{\frac{1}{1+\tau^2 N_i}}$$



e.) $y_{ij} = \mu + \delta_i + e_{ij} \qquad \delta_i \sim N(0, \tau^2\sigma^2) \qquad e_{ij} \sim N(0,\sigma^2)$

$$\text{cov}(y_{ij}, y_{ik}) = \text{Cov}(\mu+\delta_i+e_{ij}, \mu+\delta_i+e_{ik}) = E[(\mu+\delta_i+e_{ij}-\mu-E(\delta_i))(\mu+\delta_i+e_{ik}-\mu-E(\delta_i))]$$
$$= E[(\delta_i-E(\delta_i)+e_{ij})(\delta_i-E(\delta_i)+e_{ik})] = \text{var}(\delta_i) = \boxed{\tau^2\sigma^2}$$

$$\text{cov}(y_{ij}, y_{i'k}) = \text{cov}(\mu+\delta_i+e_{ij}, \mu+\delta_{i'}+e_{i'k}) = \boxed{0}$$

f.) No — the data has higher variance when N_i is low

Blood Pressure