Homework 5

Casey Gibson

Problem 1

Let

$$x_i = \text{ entry score of ith student}$$
$$p_i = \text{parents education ith student}$$
$$q_j = \text{school quality of the jth student}$$
$$y_i \sim N(\alpha_{j[i]} + \beta_{1,j[i]} x_i + \beta_2 p_i, \sigma_y^2)$$
$$\alpha_j = \gamma_0^\alpha + \gamma_1^\alpha q_i + \eta_j^\alpha$$
$$\beta_{1,j} = \gamma_0^{\beta_1} + \gamma_1^{\beta_1} q_i + \eta_j^{\beta_1}$$

where

$$\begin{pmatrix} \eta_j^\alpha \\ \eta_j^{\beta_1} \end{pmatrix} \sim N(\begin{pmatrix} 0 \\ 0 \end{pmatrix}, \begin{pmatrix} \sigma_\alpha^2 & \rho\sigma_\alpha\sigma_{\beta_1} \\ \rho\sigma_\alpha\sigma_{\beta_1} & \sigma_{\beta_1}^2 \end{pmatrix}$$

$$y_i \sim N(\alpha_{j[i]} + \beta_{1,j[i]} x_i + \beta_2 p_i, \sigma_y^2)$$

b)

$$y_i \sim N(\alpha_{j[i]} + \beta_{1,j[i]} x_i + \beta_{2,j[i]} p_i, \sigma_y^2)$$

$$\alpha_j = \gamma_0^\alpha + \gamma_1^\alpha q_i + \eta_j^\alpha$$
$$\beta_{1,j} = \gamma_0^{\beta_1} + \gamma_1^{\beta_1} q_i + \eta_j^{\beta_1}$$

$$\beta_{2,j} = \gamma_0^{\beta_2} + \gamma_1^{\beta_2} q_i + \eta_j^{\beta_2}$$

where

$$\begin{pmatrix} \eta_j^\alpha \\ \eta_j^{\beta_1} \\ \eta_j^{\beta_2} \end{pmatrix} \sim N(\begin{pmatrix} 0 \\ 0 \\ 0 \end{pmatrix}, \begin{pmatrix} \sigma_\alpha^2 & \rho_{\alpha,\beta_1}\sigma_\alpha\sigma_{\beta_1} & \rho_{\alpha,\beta_2}\sigma_\alpha\sigma_{\beta_2} \\ \rho_{\beta_1,\alpha}\sigma_{\beta_1}\sigma_\alpha & \sigma_{\beta_1}^2 & \rho_{\beta_1,\beta_2}\sigma_{\beta_1}\sigma_{\beta_2} \\ \rho_{\beta_2,\alpha}\sigma_{\beta_2}\sigma_\alpha & \rho_{\beta_2,\beta_1}\sigma_{\beta_2}\sigma_{\beta_1} & \sigma_{\beta_2}^2 \end{pmatrix}$$

where $\rho_{a,b} = \rho_{b,a}$

Problem 2

a)

```
## Compiling model graph
##    Resolving undeclared variables
##    Allocating nodes
## Graph information:
##    Observed stochastic nodes: 2400
##    Unobserved stochastic nodes: 43
##    Total graph size: 4856
##
## Initializing model

## [1] "Mean"

## [1] 20.1

## [1] "95 PI"

##  2.5% 97.5%
##  17.9  22.3
```

b)

```
## [1] "Mean"

## [1] 20.1

## [1] "95 PI"

##  2.5% 97.5%
##  15.6  24.5
```

In **a)** we sampled from the posterior predictive distribution of α_j, that is we generated samples according to

$$\alpha_{j[k]}^{new}|\mu_\alpha^{(s)}, \sigma_\alpha^{(s)} \sim N(\mu_\alpha^{(s)}, (\sigma_\alpha^2)^{(s)})$$

and similarly for **b)** we generated samples from

$$y_k^{new}|\alpha_{j[k]}^{(s)}, \sigma_y^{(s)} \sim N(\alpha_{j[k]}^{(s)}, (\sigma_y^2)^{(s)})$$

It makes sense that we see a larger PI for the observation level, rather than the cluster level since we are incorporating two sources of uncertainty.