Lecture 5

Random Effects Models

9/19/2018

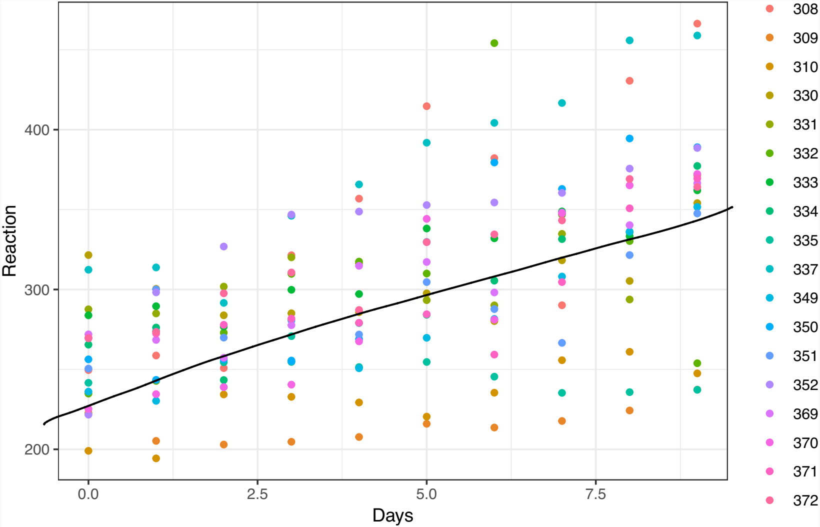
Random Effects Models

Sleep Study Data

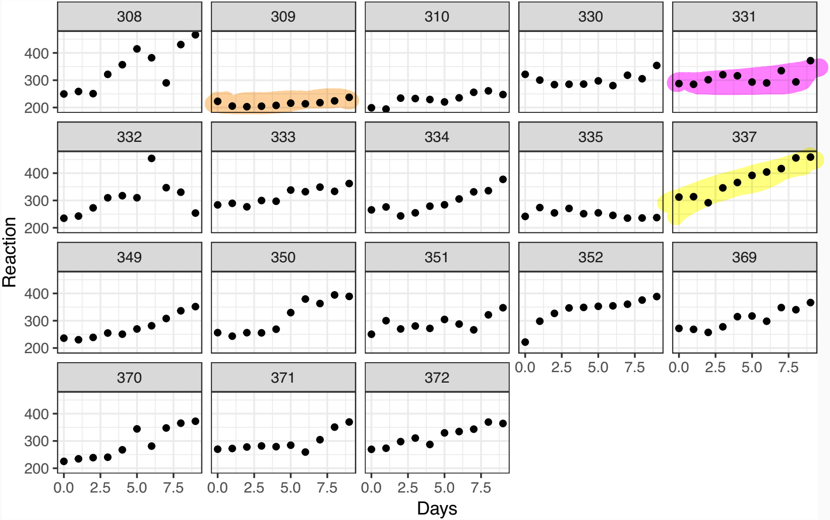
The average reaction time per day for subjects in a sleep deprivation study. On day 0 the subjects had their normal amount of sleep . Starting that night they were restricted to 3 hours of sleep per night. The observations represent the average reaction time on a series of tests given each day to each subject.

```
sleep = lme4::sleepstudy %>% tbl_df()

sleep
## # A tibble: 180 x 3
##    Reaction  Days Subject
##       <dbl> <dbl> <fct>
##  1     250.     0 308
##  2     259.     1 308
##  3     251.     2 308
##  4     321.     3 308
##  5     357.     4 308
##  6     415.     5 308
##  7     382.     6 308
##  8     290.     7 308
##  9     431.     8 308
## 10     466.     9 308
## # ... with 170 more rows
```

EDA


Reaction
400
300
200
0.0
2.5
5.0
7.5
Days
308
309
310
330
331
332
333
334
335
337
349
350
351
352
369
370
371
372

EDA (small multiples)



Bayesian Linear Model

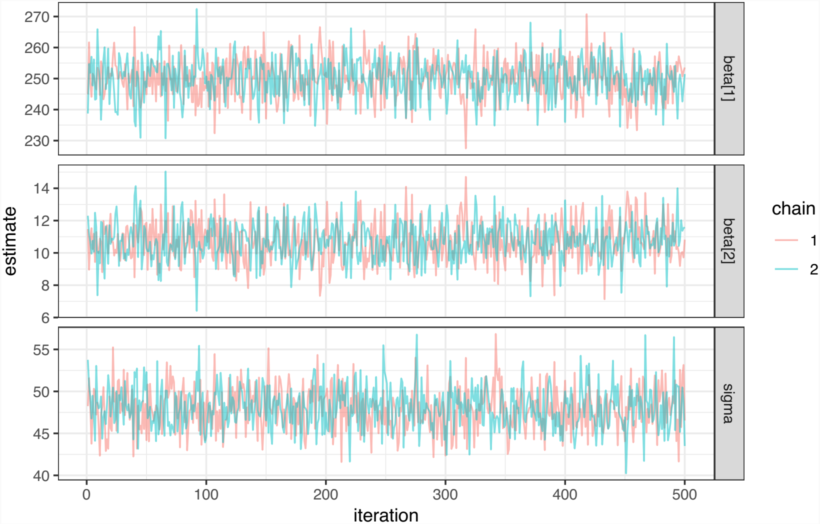
```
sleep_lm = ”model{
  # Likelihood
  for(i in 1:length(y)){
    y[i] ~ dnorm(mu[i], tau)
    mu[i] = beta[1] + beta[2]*x[i]

    y_pred[i] ~ dnorm(mu[i],tau)
  }

  # Prior for beta
  beta[1] ~ dnorm(0,1/10000)
  beta[2] ~ dnorm(0,1/10000)

  # Prior for sigma / tau
  sigma ~ dunif(0, 100)
  tau = 1/(sigma*sigma)
}”
```

$$y_i = \beta_0 + \beta_1 x_i + \varepsilon_i$$

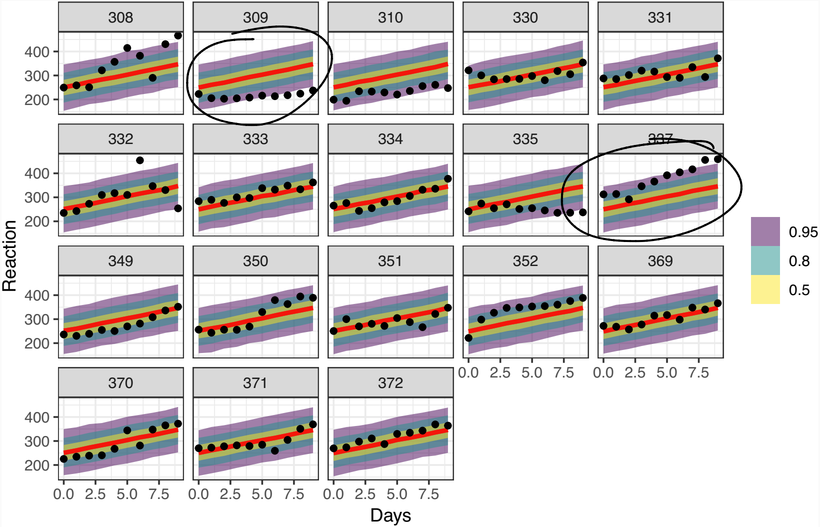
y_i → rxn time

x_i → days

MCMC Diagnostics

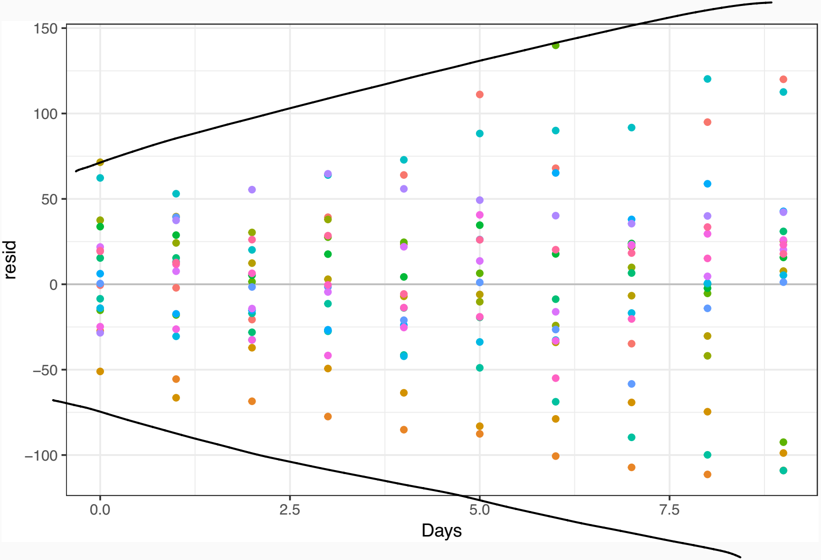


Model fit

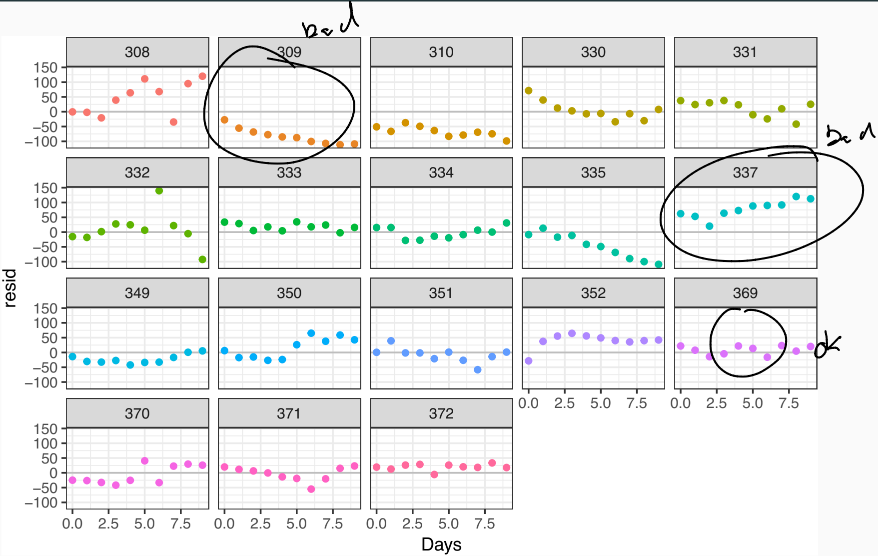


308
309
310
330
331
332
333
334
335
337
349
350
351
352
369
370
371
372
Reaction
400
300
200
Days
0.0
2.5
5.0
7.5
0.95
0.8
0.5

Residuals



Residuals by subject



Random Intercept Model

Coding

```
sleep = sleep %>%
  mutate(Subject_index = as.integer(Subject))

sleep[c(1:2,11:12,21:22,31:32),]
## # A tibble: 8 x 4
##   Reaction  Days Subject Subject_index
##      <dbl> <dbl> <fct>           <int>
## 1     250.     0 308                 1
## 2     259.     1 308                 1
## 3     223.     0 309                 2
## 4     205.     1 309                 2
## 5     199.     0 310                 3
## 6     194.     1 310                 3
## 7     322.     0 330                 4
## 8     300.     1 330                 4
```

Random Intercept Model

Let i represent each observation and $j(i)$ be subject in oberservation i then

$$y_i = \alpha_{j(i)} + \beta \times \text{Day}_i + \epsilon_i$$

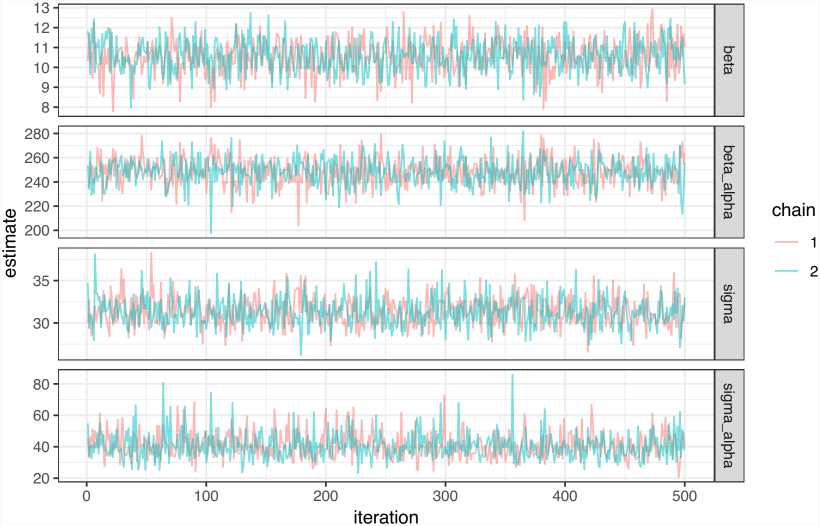
$$\alpha_j \sim \mathcal{N}(\beta_\alpha,\ \sigma_\alpha^2)$$
$$\epsilon_i \sim \mathcal{N}(0,\ \sigma^2)$$

$$\beta_\alpha \sim \mathcal{N}(0, 10^4)$$
$$\beta \sim \mathcal{N}(0, 10^4)$$
$$\sigma, \sigma_\alpha \sim \text{Unif}(0, 10^2)$$

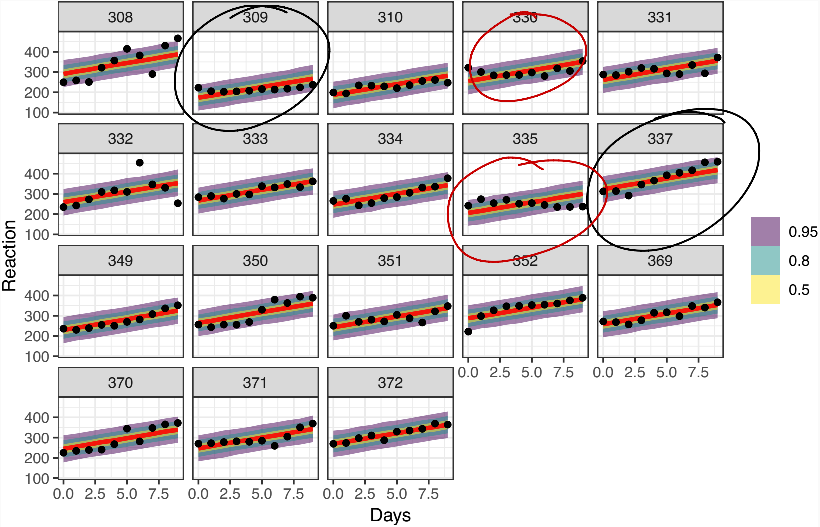
Random Intercept Model - JAGS

```
sleep_ri = "model{
  for(i in 1:length(Reaction)) {
    Reaction[i] ~ dnorm(mu[i],tau)
    mu[i] = alpha[Subject_index[i]] + beta*Days[i]

    y_pred[i] ~ dnorm(mu[i],tau)
  }

  for(j in 1:18) {
    alpha[j] ~ dnorm(beta_alpha, tau_alpha)
  }

  beta_alpha ~ dnorm(0,1/10000)
  sigma_alpha ~ dunif(0, 100)
  tau_alpha = 1/(sigma_alpha*sigma_alpha)


  beta ~ dnorm(0,1/10000)
  sigma ~ dunif(0, 100)
  tau = 1/(sigma*sigma)
}"
```

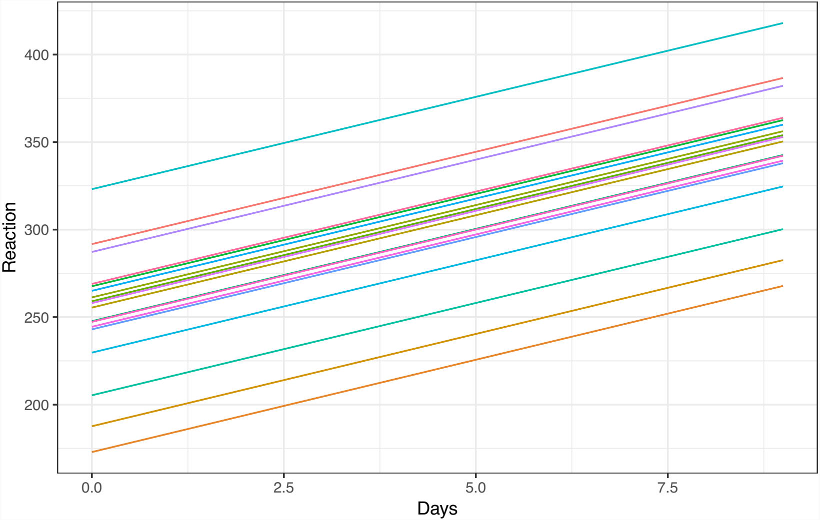
MCMC Diagnostics



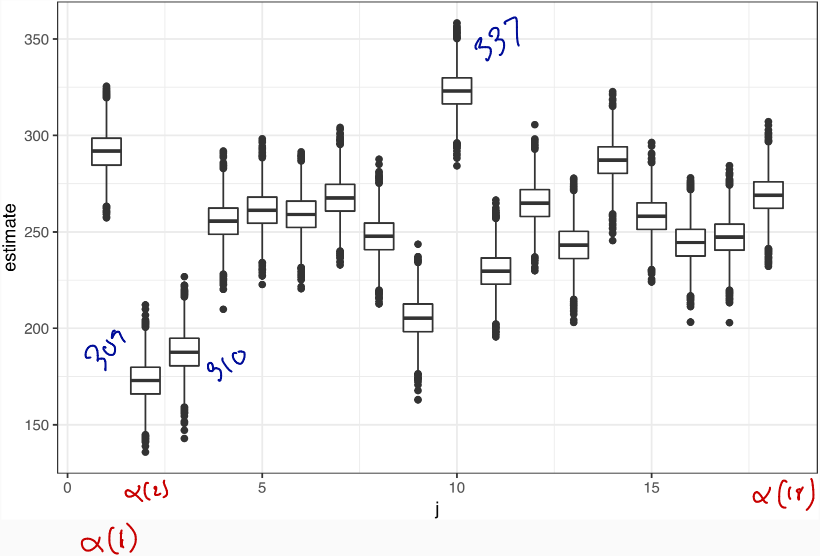
Model fit


308
309
310
330
331
332
333
334
335
337
349
350
351
352
369
370
371
372
Reaction
Days
0.95
0.8
0.5

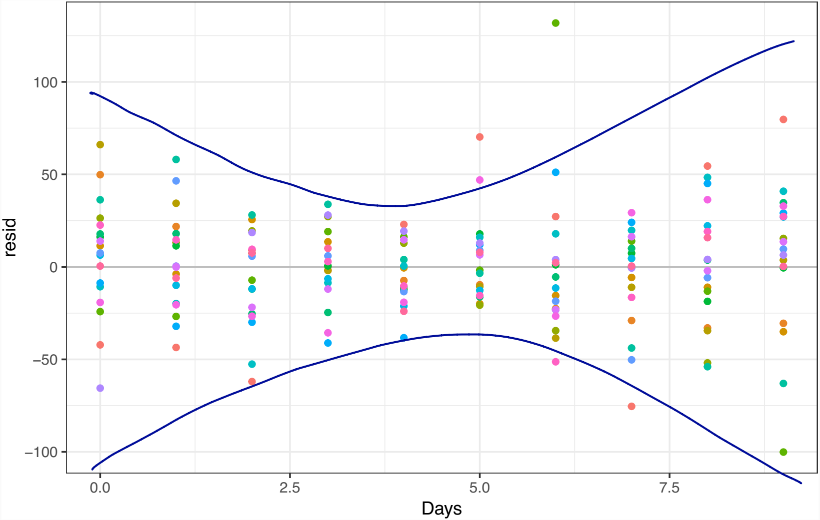
Model fit - lines



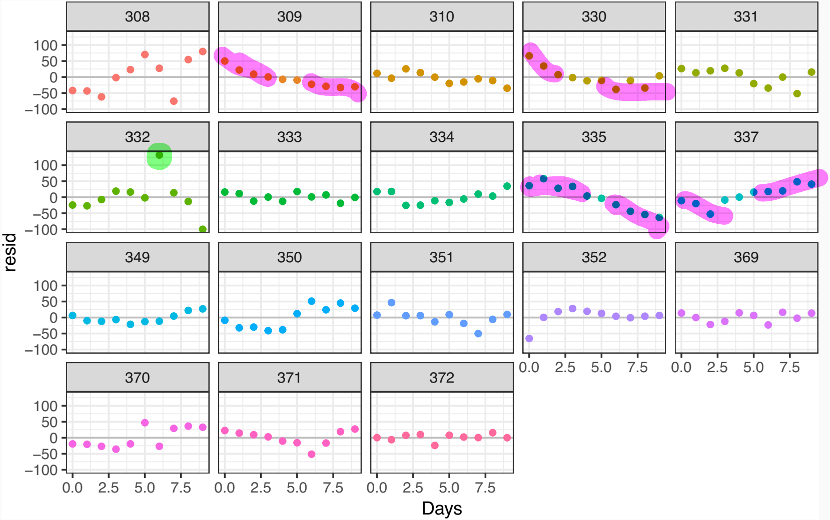
Random Effects



Residuals



Residuals by subject



Why not a fixed effect for Subject?

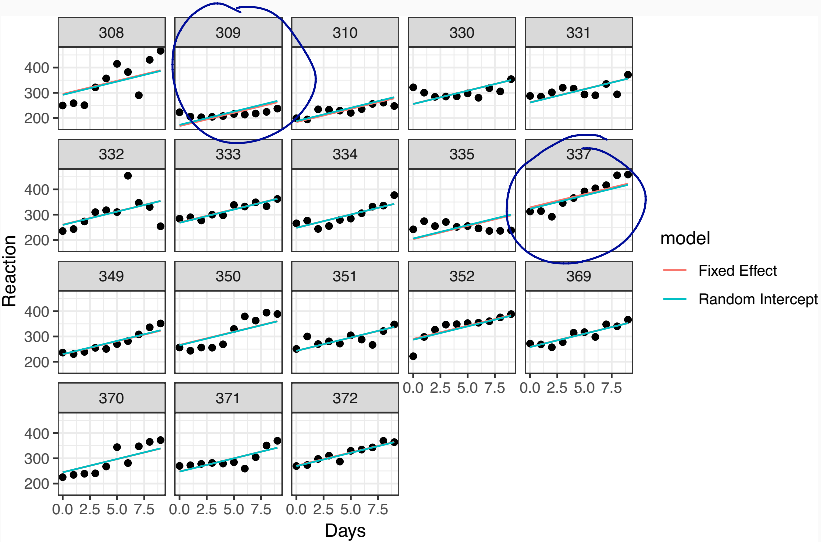
We are not going to bother with the Bayesian model here to avoid the headache of dummy coding and the resulting βs.

Why not a fixed effect for Subject?

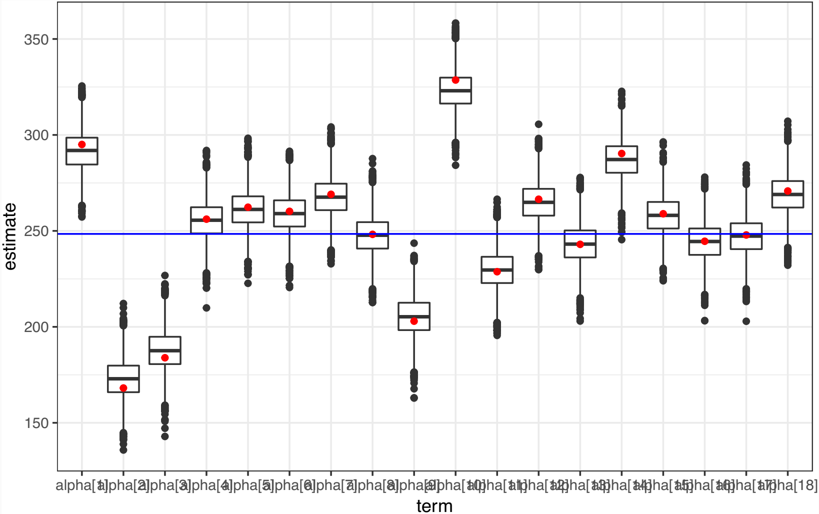
We are not going to bother with the Bayesian model here to avoid the headache of dummy coding and the resulting βs.

```
l = lm(Reaction ~ Days + Subject - 1, data=sleep)
summary(l)
##
## Call:
## lm(formula = Reaction ~ Days + Subject - 1, data = sleep)
##
## Residuals:
##      Min       1Q   Median       3Q      Max
## -100.540  -16.389   -0.341   15.215  131.159
##
## Coefficients:
##            Estimate Std. Error t value Pr(>|t|)
## Days        10.4673     0.8042   13.02   <2e-16 ***
## Subject308 295.0310    10.4471   28.24   <2e-16 ***
## Subject309 168.1302    10.4471   16.09   <2e-16 ***
## Subject310 183.8985    10.4471   17.60   <2e-16 ***
## Subject330 256.1186    10.4471   24.52   <2e-16 ***
## Subject331 262.3333    10.4471   25.11   <2e-16 ***
## Subject332 260.1993    10.4471   24.91   <2e-16 ***
## Subject333 269.0555    10.4471   25.75   <2e-16 ***
## Subject334 248.1993    10.4471   23.76   <2e-16 ***
## Subject335 202.9673    10.4471   19.43   <2e-16 ***
## Subject337 328.6182    10.4471   31.45   <2e-16 ***
## Subject349 228.7317    10.4471   21.89   <2e-16 ***
## Subject350 266.4999    10.4471   25.51   <2e-16 ***
## Subject351 242.9950    10.4471   23.26   <2e-16 ***
## Subject352 290.3188    10.4471   27.79   <2e-16 ***
## Subject369 258.9319    10.4471   24.79   <2e-16 ***
## Subject370 244.5990    10.4471   23.41   <2e-16 ***
## Subject371 247.8813    10.4471   23.73   <2e-16 ***
## Subject372 270.7833    10.4471   25.92   <2e-16 ***
## ---
## Signif. codes:  0 '***' 0.001 '**' 0.01 '*' 0.05 '.' 0.1 ' ' 1
```

Comparing Model fit



Random effects vs fixed effects

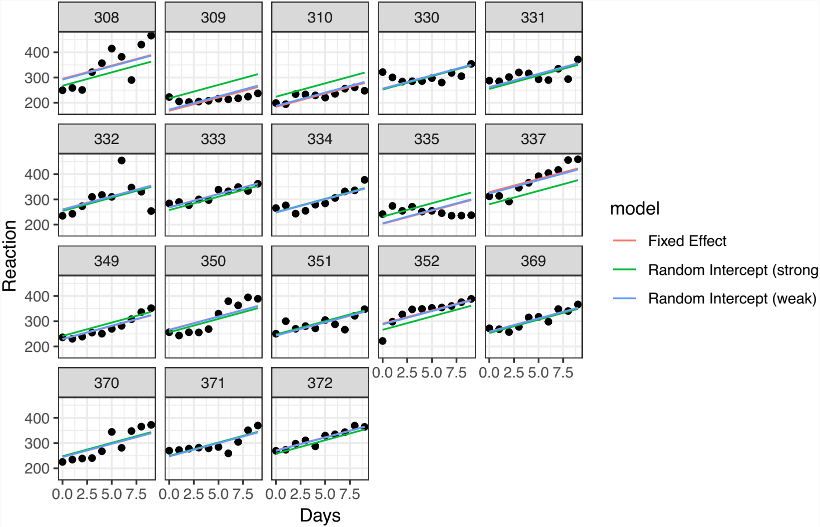


Random Intercept Model (strong prior for σ_α)

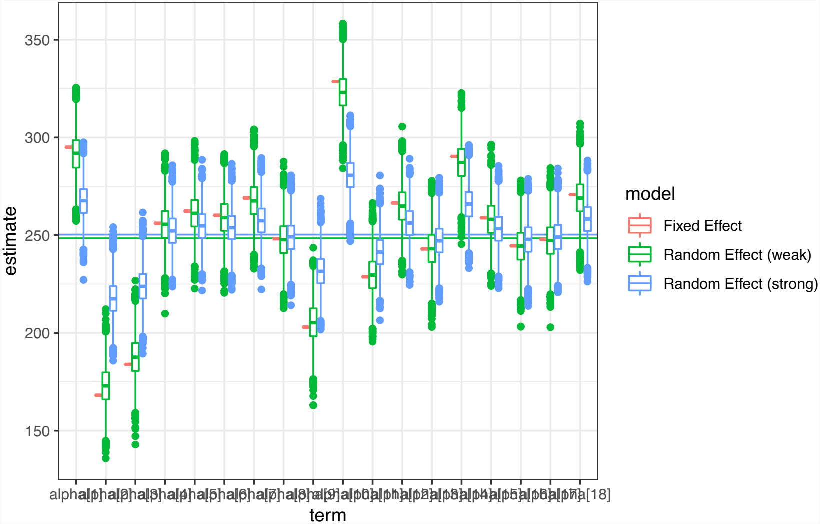
```
sleep_ri2 = "model{
  for(i in 1:length(Reaction)) {
    Reaction[i] ~ dnorm(mu[i],tau)
    mu[i] = alpha[Subject_index[i]] + beta*Days[i]

    y_pred[i] ~ dnorm(mu[i],tau)
  }

  for(j in 1:18) {
    alpha[j] ~ dnorm(beta_alpha, tau_alpha)
  }

  beta_alpha ~ dnorm(0,1/10000)
  sigma_alpha ~ dunif(0, 10)
  tau_alpha = 1/(sigma_alpha*sigma_alpha)


  beta ~ dnorm(0,1/10000)
  sigma ~ dunif(0, 100)
  tau = 1/(sigma*sigma)
}"
```

vs dunif(0, 100)

Comparing Model fit



Prior Effect on α



$$Y_i = \alpha_{j(i)} + \beta_1 X_i + \varepsilon_i$$
$$\alpha_j \sim N(\beta_\alpha, \sigma_\alpha^2)$$
$$\varepsilon_i \sim N(0, \sigma^2)$$

$$\underline{Y} \sim N(\underline{\mu}, \underline{\Sigma})$$
$$\underline{\mu} = ?$$
$$\underline{\Sigma} = ?$$

$$E(Y_i \mid \underline{\theta}) = E(\alpha_{j(i)} + \beta_1 X_i + \varepsilon_i \mid \underline{\theta})$$
$$= E(\alpha_{j(i)} \mid \underline{\theta}) + E(\beta_1 X_i \mid \underline{\theta}) + E(\varepsilon_i \mid \underline{\theta})$$
$$= \boxed{\beta_\alpha + \beta_1 X_i} + 0$$

$$\underline{\mu} = \beta_\alpha \underset{n\times 1}{\underline{\mathbb{1}}} + \beta_1 \underline{X}$$

$$\underline{\theta} = \{\beta_1, \beta_\alpha, \sigma_\alpha^2, \sigma^2\}$$

Some Distribution Theory

$$\{\underline{\Sigma}\}_{ik} = \text{Cov}(Y_i, Y_k \mid \underline{\theta})$$

$$= E\left((Y_i - E(Y_i \mid \underline{\theta}))(Y_k - E(Y_k \mid \underline{\theta})) \mid \underline{\theta} \right)$$

$$= E\left(\left(\alpha_{j(i)} + \beta_1 X_i + \varepsilon_i - (\beta_\alpha + \beta_1 X_i) \right) \left(\alpha_{j(k)} + \beta_1 X_k + \varepsilon_k - (\beta_\alpha + \beta_1 X_k) \right) \right)$$

$$= E\left((\alpha_{j(i)} + \varepsilon_i - \beta_\alpha)(\alpha_{j(k)} + \varepsilon_k - \beta_\alpha) \mid \underline{\theta} \right)$$

$$= E\left(\alpha_{j(i)} \alpha_{j(k)} + \varepsilon_i \varepsilon_k - \alpha_{j(i)} \beta_\alpha - \alpha_{j(k)} \beta_\alpha + \beta_\alpha^2 \mid \underline{\theta} \right)$$

$$= E\left(\alpha_{j(i)} \alpha_{j(k)} \mid \underline{\theta} \right) + E\left(\varepsilon_i \varepsilon_k \mid \underline{\theta} \right) - \beta_\alpha^2$$

$$= E\left(\alpha_{j(i)}\alpha_{j(k)} \mid \underline{\theta}\right) + E\left(\varepsilon_i \varepsilon_k \mid \underline{\theta}\right) - \beta_\alpha^2$$

$i = k$

$$\text{Cov}(\cdot) = E\left(\alpha_{j(i)}^2 \mid \underline{\theta}\right) + E\left(\varepsilon_i^2 \mid \underline{\theta}\right) - \beta_\alpha^2$$

$$= \text{Var}\left(\alpha_{j(i)}^2 \mid \theta\right) + E\left(\alpha_{j(i)}\right)^2$$

$$+ \text{Var}(\varepsilon_i) + E(\varepsilon_i)^2$$

$$- \beta_\alpha$$

β_α^2, 0

$$= \sigma_\alpha^2 + \sigma^2$$

Some Distribution Theory

$$= E\left(\alpha_{j(i)}\alpha_{j(k)} \mid \underline{\theta}\right) + E\left(\varepsilon_i \varepsilon_k \mid \underline{\theta}\right) - \beta_\alpha^2$$

$\alpha_j(i) = \alpha_j(k)$
$i \neq k$

Diff obs Same subj

$$\text{Cov}(\cdot) = E\left(\alpha_j(i)^2\right) + E(\varepsilon_i \varepsilon_k) - \beta_\alpha^2 \quad (E(\varepsilon_i\varepsilon_k) \to 0)$$

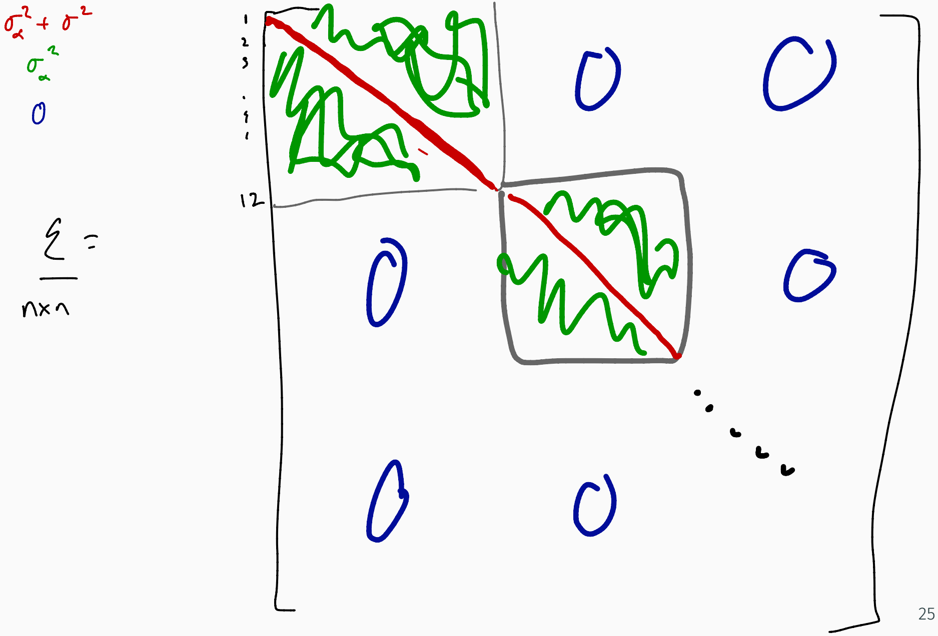
$$= \text{Var}(\alpha_j(i)) + E(\alpha_j(i))^2 - \beta_\alpha^2 \quad (E(\alpha_j(i))^2 = \beta_\alpha^2)$$

$$= \sigma_\alpha^2$$

$i \neq k \quad \alpha_j(i) \neq \alpha_j(k)$

$$\text{Cov}(\cdot) = E(\alpha_j(i))\,E(\alpha_k(k)) + E(\varepsilon_i)\,E(\varepsilon_k) - \beta_\alpha^2$$

$$= \beta_\alpha \beta_\alpha + 0 \cdot 0 - \beta_\alpha^2 = 0$$


$\sigma_\alpha^2 + \sigma^2$
σ_α^2
0
$\underset{n\times n}{\Sigma} =$
1
2
3
12
0
0
0
0
0
0

Random intercept and slope model

Model

Let i represent each observation and $j(i)$ be the subject in oberservation i then

$$Y_i = \alpha_{j(i)} + \beta_{j(i)} \times \text{Days} + \epsilon_i$$

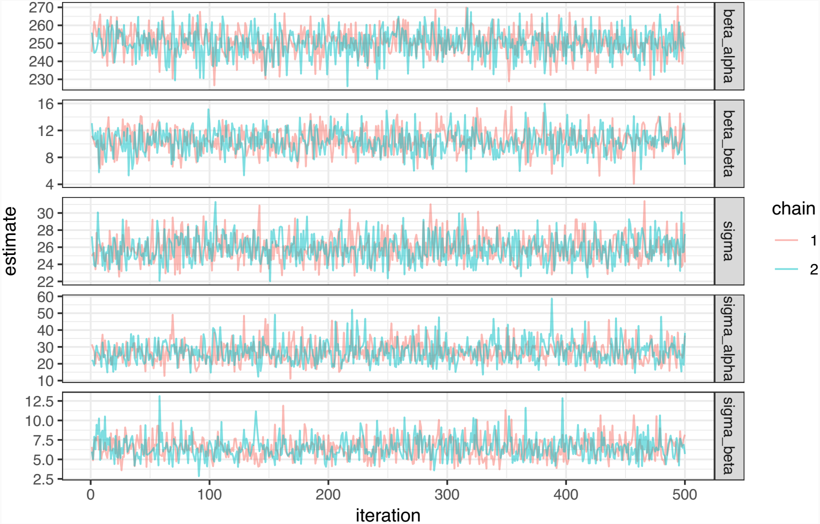
$$\alpha_j \sim \mathcal{N}(\beta_0,\ \sigma^2_\alpha)$$
$$\beta_j \sim \mathcal{N}(\beta_1,\ \sigma^2_\beta)$$
$$\epsilon_i \sim \mathcal{N}(0,\ \sigma^2)$$

$$\beta_\alpha, \beta_\beta \sim \mathcal{N}(0, 10000)$$
$$\sigma, \sigma_\alpha, \sigma_\beta \sim \text{Unif}(0, 100)$$

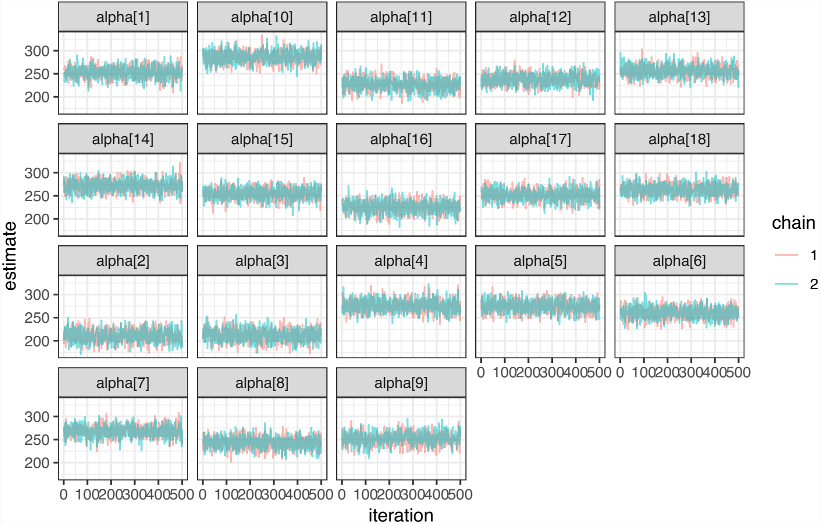
Model - JAGS

```
sleep_ris = "model{
  for(i in 1:length(Reaction)) {
    Reaction[i] ~ dnorm(mu[i],tau)
    mu[i] = alpha[Subject_index[i]] + beta[Subject_index[i]] * Days[i]
    y_pred[i] ~ dnorm(mu[i], tau)
  }

  sigma ~ dunif(0, 100)
  tau = 1/(sigma*sigma)

  for(j in 1:18) {
    alpha[j] ~ dnorm(beta_alpha, tau_alpha)
    beta[j]  ~ dnorm(beta_beta, tau_beta)
  }

  beta_alpha ~ dnorm(0,1/10000)
  beta_beta ~ dnorm(0,1/10000)

  sigma_alpha ~ dunif(0, 100)
  tau_alpha = 1/(sigma_alpha*sigma_alpha)

  sigma_beta ~ dunif(0, 100)
  tau_beta = 1/(sigma_beta*sigma_beta)
}"
```

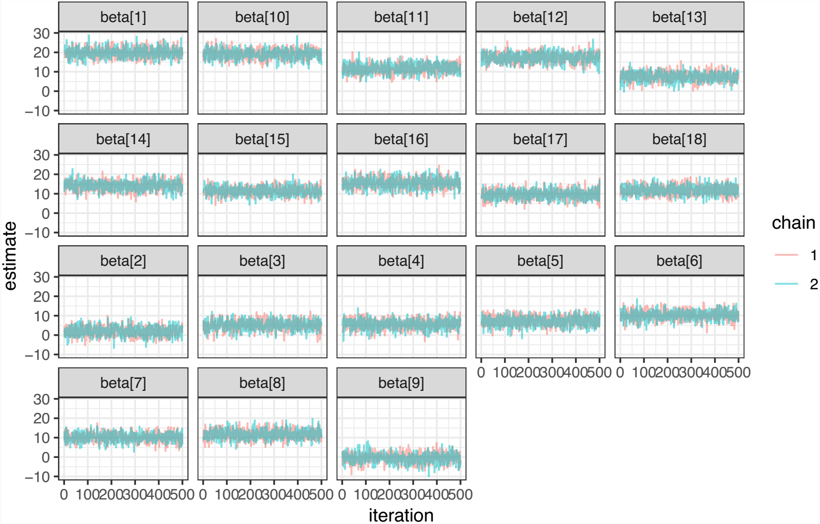
MCMC Diagnostics



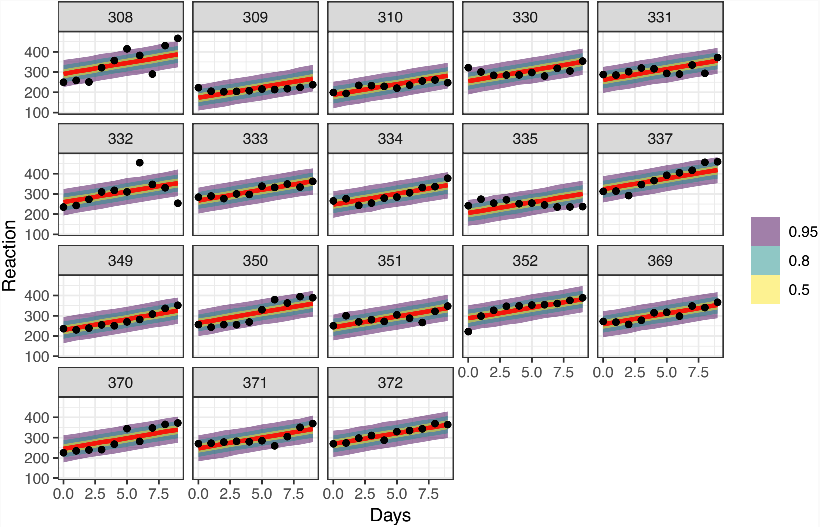
MCMC Diagnostics - α



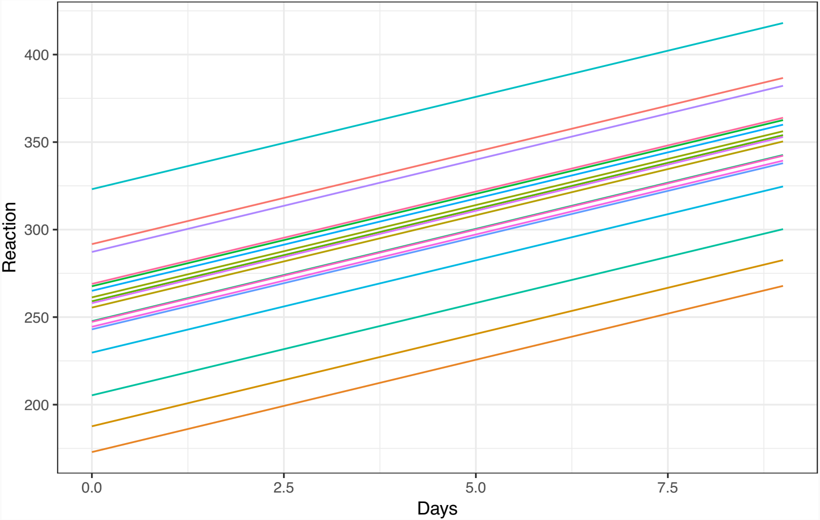
MCMC Diagnostics - β



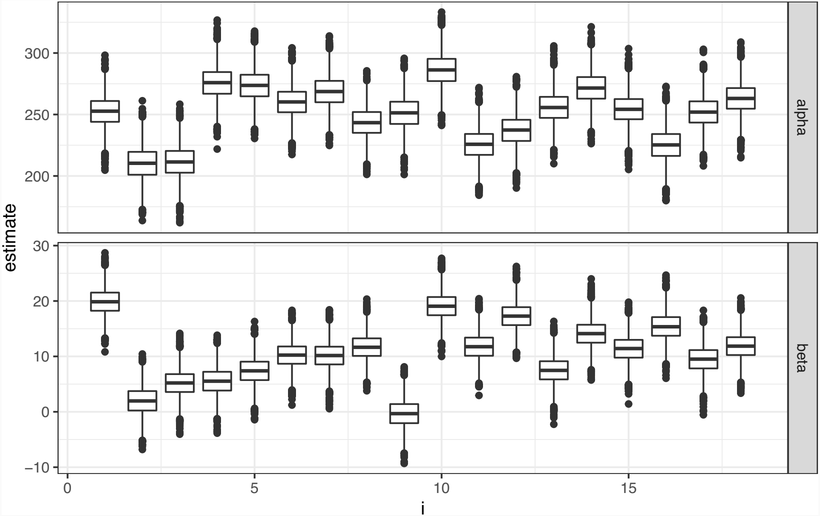
Model fit


308
309
310
330
331
332
333
334
335
337
349
350
351
352
369
370
371
372
Reaction
100
200
300
400
Days
0.0
2.5
5.0
7.5
0.95
0.8
0.5

Model fit - lines



Random Effects



Residuals by subject