Linear Random Effects Model

Yingbo Li

Clemson University

MATH 9810

Mercury in Bass

High levels of mercury in fish known to cause health problems, especially in children. Researchers studied mercury concentrations in largemouth bass from the Wacamaw and Lumber rivers in NC.

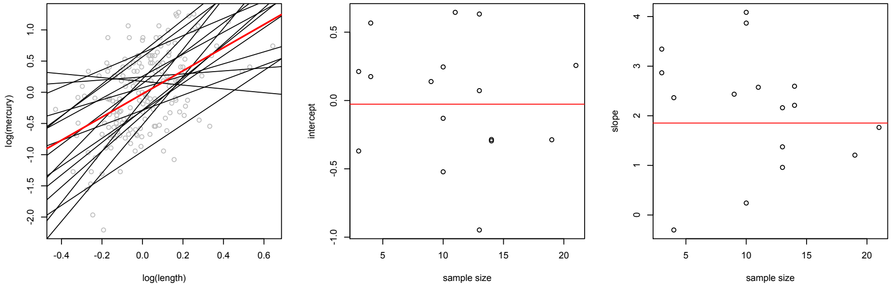
- Mercury cannot be excreted and accumulates in tissue over the lifetime of a fish
- Direct measurement of mercury requires sacrificing fish
- Build model for predicting mercury concentration given easy to measure variables (weight and/or length)
- Fish were caught, weighed and measured at 16 stations
- Filets sent to the lab for mercury measurements

Question: Can we predict mercury levels for fish caught in these rivers?

Initial Analysis

Let's do some exploratory data analysis to learn the association between:

- Response $y_{i,j}$: log(mercury concentration).
- Predictor $x_{i,j}$: log(length).
- We center the $x_{i,j}$'s within each station.
- To control for where fish caught: run a regression for each station.



Borrow Information Across Stations

- From exploratory data analysis and OLS regressions, clear differences across stations
- But not many fish in some stations, so makes sense to borrow strength across stations
- Bayesian hierarchical regression, also known as random effects models, allow us to do that

Random Effects Models

- Hierarchical models can be applied to regression contexts where observations are grouped
- We discuss examples for linear regression, but ideas apply to Generalized Linear Models (GLMs) such as logistic regression, etc.
- Recall linear model with one predictors can be written $y_i = \beta_0 + \beta_1 x_i + \epsilon_i$, where $\epsilon_i \sim N(0, \sigma^2)$
- Suppose that observations fall in J groups, indexed by j

Random Intecepts Models

Let j index groups and i index observations. The random intercepts model is

- Centered parameterization:

$$\begin{aligned} y_{ij} &= \beta_{0j} + \beta_1 x_{ij} + \epsilon_{ij}, \quad \epsilon_{ij} \sim N(0, \sigma_j^2) \\ \beta_{0j} &\sim \mathsf{N}(\beta_0, \tau^2) \end{aligned}$$

- This is also written sometimes in the uncentered parameterization:

$$\begin{aligned} y_{ij} &= \beta_0 + \gamma_j + \beta_1 x_{ij} + \epsilon_{ij}, \quad \epsilon_{ij} \sim N(0, \sigma_j^2) \\ \gamma_j &\sim \mathsf{N}(0, \tau^2) \end{aligned}$$

Random Intecepts Models

- Allows separate intercepts for each group, but shrinks estimates towards common value
- Useful for repeated measurements, when the "groups" are individuals (e.g., we take a subject's blood pressure three times and include all three measurements in the model)
- Also useful when some groups have small sample sizes, so that estimation of intercept is highly variable
- Model implies same slope of x for each group

Random Slopes (and Intercepts) Models

- Centered parameterization:

$$
\begin{aligned}
y_{ij} &= \beta_{0j} + \beta_{1j} x_{ij} + \epsilon_{ij}, \quad \epsilon_{ij} \sim N(0, \sigma_j^2) \\
(\beta_{0j}, \beta_{1j}) &\sim \mathsf{N}_2(\beta, \Sigma)
\end{aligned}
$$

- Uncentered parameterization:

$$
\begin{aligned}
y_{ij} &= \beta_0 + \gamma_{0j} + \beta_1 x_{ij} + \gamma_{1j} x_{ij} + \epsilon_{ij}, \quad \epsilon_{ij} \sim N(0, \sigma_j^2) \\
(\gamma_{0j}, \gamma_{1j}) &\sim \mathsf{N}_2(0, \Sigma)
\end{aligned}
$$

Random Slopes Models

- Allows separate slopes (and intercepts) for each group, but shrinks estimates towards common value
- Useful when some groups have small sample sizes, so that estimation of slopes (and intercepts) is highly variable
- (β_0, β_1) are called <u>fixed effects</u>
- $(\gamma_{0j}, \gamma_{1j})$ are called <u>random effects</u>
- Models with fixed and random effects are called <u>mixed effects</u> models

Mercury in NC River Example

Consider the following models for y_{ij} log(mercury) as a function of x_{ij} log(length):

1. $y_{ij} = \beta_0 + \beta_1 x_{ij}$ (common line for all stations)
2. $y_{ij} = \beta_{0j} + \beta_1 x_{ij}$ (parallel regression lines)
3. $y_{ij} = \beta_{0j} + \beta_{1j} x_{ij}$ (separate lines for each station)

EDA suggests that separate slopes and intercepts model is most appropriate. Mixed effects model:

$$\begin{aligned} y_{ij} &\sim \mathsf{N}(\beta_{0j} + \beta_{1j} x_{ij}, \sigma^2) \\ (\beta_{0j}, \beta_{1j}) &\sim \mathsf{N}_2((\beta_0, \beta_1), \Sigma) \end{aligned}$$

Noninformative Priors for Bass Model

We will use independent priors for (β_{0j}, β_{1j}), so that Σ is a diagonal matrix with elements τ_0^2 and τ_1^2 on first and second diagonal positions, respectively

$$
\begin{aligned}
p(\beta_0) &\propto 1 \\
p(\beta_1) &\propto 1 \\
p(1/\sigma^2) &\propto (1/\sigma^2)^{-1} \\
p(1/\tau_0^2) &\propto (1/\tau_0^2)^{-\frac{3}{2}} \\
p(1/\tau_1^2) &\propto (1/\tau_1^2)^{-\frac{3}{2}}
\end{aligned}
$$

Note that priors are uniform in τ, not τ^2. R code is available on Blackboard.

Full Conditionals: Centered Model

Denote $\beta_j = (\beta_{0j}, \beta_{1j})', \beta = (\beta_0, \beta_1)', \Sigma = \mathrm{diag}(\tau_0^2, \tau_1^2)$, $Y_j = (y_{1,j}, \ldots, y_{n_j,j})'$ and $X_j = [(1, x_{1,j}), \ldots, (1, x_{n_j,j})]'$, then

$$\beta_j \mid \beta, \sigma^2, \tau_0^2, \tau_1^2, Y \sim \mathsf{N}\left(\tilde{\beta}_j, (\Sigma^{-1} + X_j'X_j/\sigma^2)^{-1}\right)$$

$$\tilde{\beta}_j = (\Sigma^{-1} + X_j'X_j/\sigma^2)^{-1}(\Sigma^{-1}\beta + X_j'Y_j/\sigma^2)$$

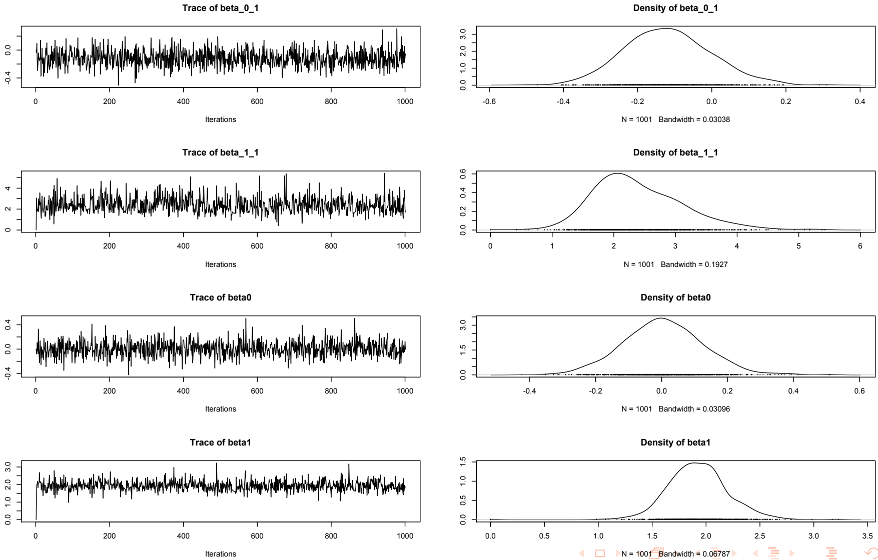
$$\beta \mid \beta_j, \tau_0^2, \tau_1^2, Y \sim \mathsf{N}\left(\frac{\beta_1 + \cdots + \beta_J}{J}, \frac{\Sigma}{J}\right)$$

$$\tau_0^2 \mid \beta_j, \beta, Y \sim \mathsf{IG}\left(\frac{J-1}{2}, \frac{\sum_j (\beta_{0j} - \beta_0)^2}{2}\right)$$

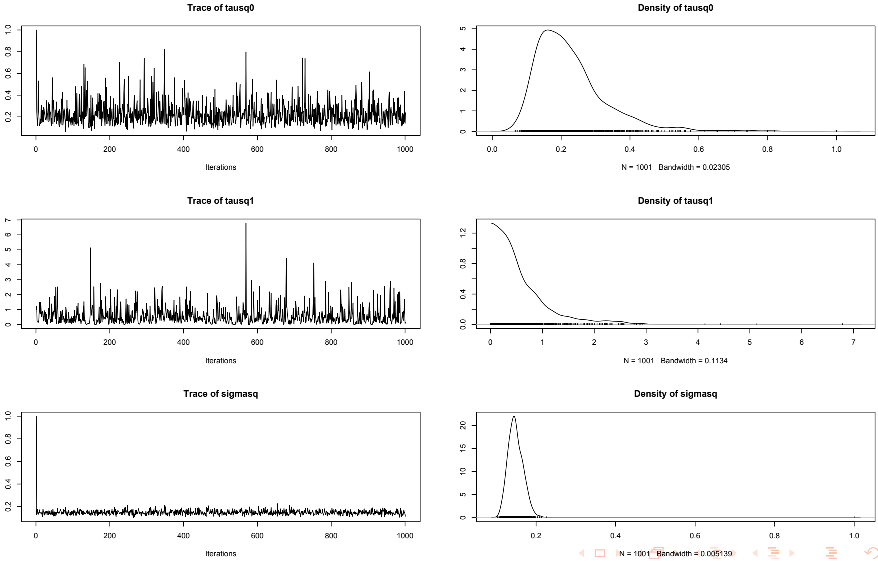
$$\tau_1^2 \mid \beta_j, \beta, Y \sim \mathsf{IG}\left(\frac{J-1}{2}, \frac{\sum_j (\beta_{1j} - \beta_1)^2}{2}\right)$$

$$\sigma^2 \mid \beta_j, \beta, Y \sim \mathsf{IG}\left(\frac{\sum_j n_j}{2}, \frac{\sum_j (Y_j - X_j\beta_j)'(Y_j - X_j\beta_j)}{2}\right)$$

Centered Model

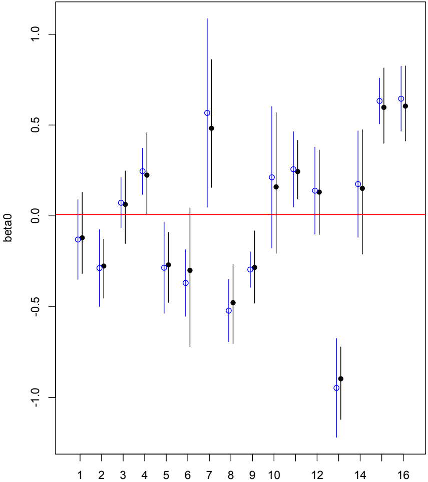


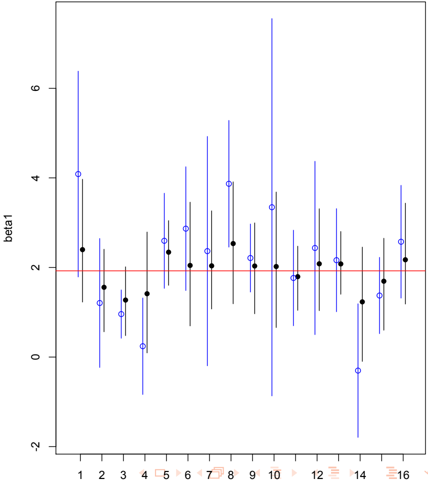
Centered Model



Centered Model

Blue: MLE for separate regressions; Black: mixed effects model





Full Conditionals: Uncentered Model

$$\gamma_j \mid \beta, \sigma^2, \tau_0^2, \tau_1^2, Y \sim \mathsf{N}\left(\tilde{\gamma}_j, (\Sigma^{-1} + X_j' X_j/\sigma^2)^{-1}\right)$$

$$\tilde{\gamma}_j = (\Sigma^{-1} + X_j' X_j/\sigma^2)^{-1} X_j' (Y_j - X_j \beta)/\sigma^2$$

$$\beta \mid \gamma_j, Y \sim \mathsf{N}\left(\tilde{\beta}, (\sum_j X_j' X_j/\sigma^2)^{-1}\right)$$

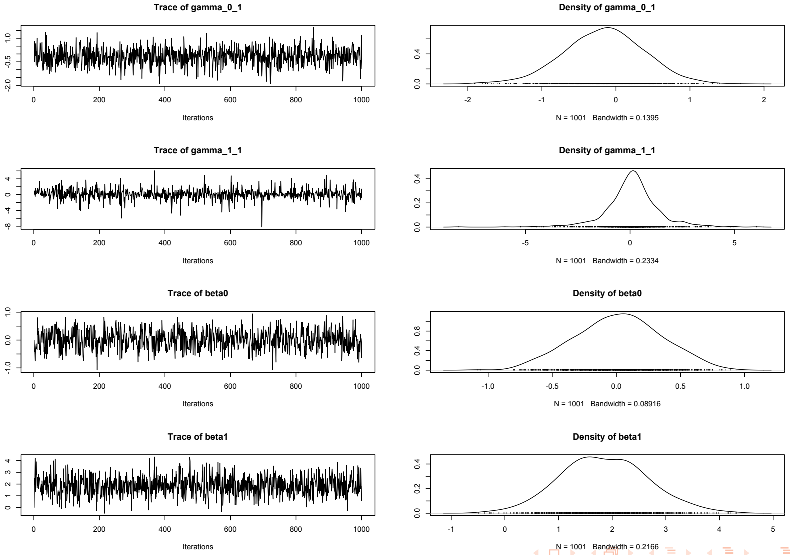
$$\tilde{\beta} = (\sum_j X_j' X_j)^{-1} \sum_j X_j' (Y_j - X_j \gamma_j)$$

$$\tau_0^2 \mid \gamma_j, Y \sim \mathsf{IG}\left(\frac{J-1}{2}, \frac{\sum_j \gamma_{0j}^2}{2}\right)$$

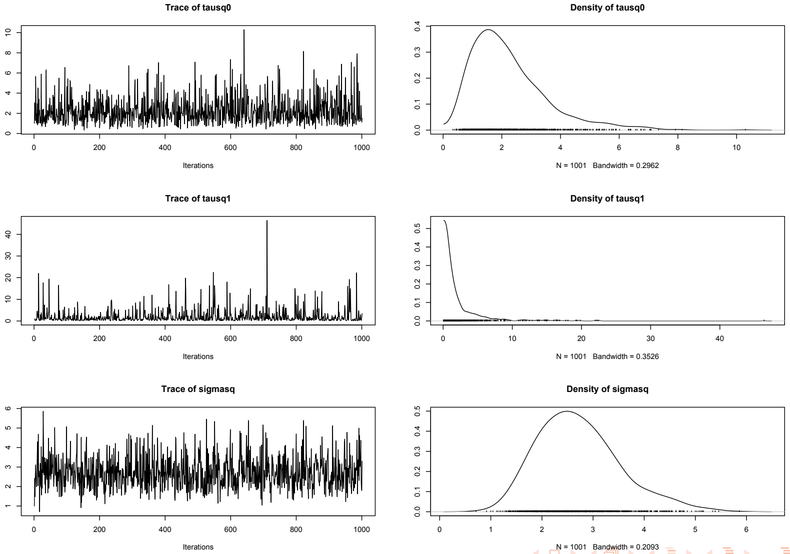
$$\tau_1^2 \mid \gamma_j, Y \sim \mathsf{IG}\left(\frac{J-1}{2}, \frac{\sum_j \gamma_{1j}^2}{2}\right)$$

$$\sigma^2 \mid \gamma_j, \beta, Y \sim \mathsf{IG}\left(\frac{\sum_j n_j}{2}, \frac{\sum_j (Y_j - X_j(\gamma_j + \beta))^2}{2}\right)$$

Uncentered Model

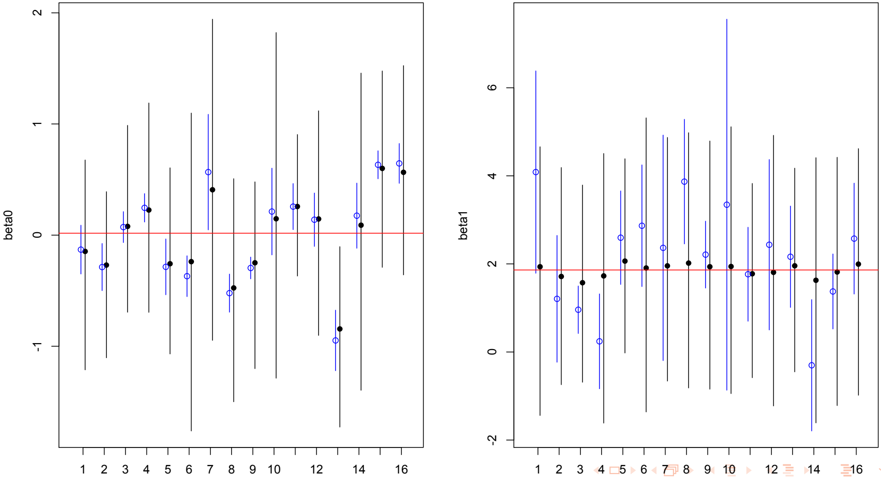


Uncentered Model



Uncentered Model

Blue: MLE for separate regressions; Black: mixed effects model



Mercury Prediction

- Observed stations: can make posterior predictions for fish of log length x_{ij} caught at any of 16 stations by
  1. Take sampled values of $(\beta_{0j}, \beta_{1j}, \sigma^2)$ from their posterior distributions
  2. For each value, generate a new y_{ij} for given x_{ij} from normal distribution with appropriate mean and variance
- Unobserved stations: can make posterior predictions for fish of log length x_j caught in some station other than 16 observed ones, say station k, by
  1. Take sampled values of (β, Σ) from their posterior distributions
  2. For each, sample new values of (β_{0k}, β_{1k}) from $N_2(\beta, \Sigma)$
  3. For each value of $(\beta_{0k}, \beta_{1k}, \sigma^2)$, generate a new y_i for given x_i from normal distribution with appropriate mean and variance