STAT 8630 — Mixed-Effect Models and Longitudinal Data Analysis
Homework 4 – Due Friday, April 5, 2013

Homework Guidelines:

- Homework is due by 4:30 on the due date specified above. You may turn it in to me during class, slip it under my door or send it to me via e-mail. I will post homework solutions shortly after all homeworks have been collected. **No late homeworks will be accepted without permission granted prior to the due date.**
- Use only standard (8.5 × 11 inch) paper and use only one side of each sheet.
- Homework should show enough detail so that the reader can clearly understand the procedures of the solutions.
- Problems should appear in the order that they were assigned.

Assignment:

1. Do problem 10.1 from our text. This problem is at the end of the chapter on Residual Analysis and Diagnostics. In addition to parts 10.1.1–10.1.7 do the following:

   10.1.8 The model described in this problem was fit in our in-class example, Fevexample1.sas. In that example, we called it Model 1, and I'll continue to refer to it with that name. To investigate the adequacy of the mean assumed in Model 1, refit it after dropping the random intercept and slope from the model. From this purely fixed effect model, produce a plot of the cumulative sum of residuals versus age along with 500 simulated realizations from the null distribution of the cumulative residual profile versus age. What does this plot say about the appropriateness of Model 1?

   10.1.9 Now refer to the in-class example FEVexample2.sas. Fit the model called Model 3 in that example. That is, fit a model of the form

   $$\text{Model 3:} \quad \begin{aligned} y_{ij} &= (\beta_1 + b_{1i}) + (\beta_2 + b_{2i})x_{1ij} + (\beta_3 + b_{3i})(x_{1ij} - \kappa)_+ \\ &\quad + \beta_4 x_{1i1} + \beta_5 x_{2ij} + \beta_6 x_{2i1} + \varepsilon_{ij} \end{aligned}$$

   where $\kappa = 14$; x_{1ij}, x_{2ij} are age and log height, respectively, at the jth measurement occasion ($j = 1$ corresponds to baseline) for subject i; and where $\mathbf{b}_i = (b_{1i}, b_{2i}, b_{3i})^T \sim N(\mathbf{0}, \mathbf{D})$, $\mathbf{D}$ a symmetric positive definite matrix. Assume the $\boldsymbol{\varepsilon}_i$s are independent with $\boldsymbol{\varepsilon}_i \sim N(\mathbf{0}, \mathbf{R}_i)$ where, for now, $\mathbf{R}_i = \sigma^2\mathbf{I}$.
   As in part 10.1.8, produce a plot of the cumulative sum of residuals versus age along with 500 simulated realizations from the null distribution of the

cumulative residual profile versus age. What does this plot say about the appropriateness of Model 3? Does the mean structure of the model still need improvement? If so, try to develop a better model and justify your final model choice.

10.1.10 Based on the final model you chose in part 10.1.9, produce a plot of the empirical semivariogram of the normalized residuals provided by R (the residuals given by *residuals(model,type="n")*). What does this plot tell you about the appropriateness of the model. Explain.

2. Do problem 13.1 from our text.

3. Do problem 13.3 from our text.