

StatR 502 Homework 8 (Last HW!)

Gregor Thomas

Due Thursday, March 10, 2016, 6:30 pm

Submission guidelines: please submit a knitted PDF or Word document, and optionally your `.Rmd` file. As always, ask in the discussion forum if you're having trouble!

Remember that this homework covers two weeks of class. Do try to get started early, but some topics won't be covered in the first lecture—the lab and next lecture will help.

All of these problems rely on the CD4 data available at <http://www.stat.columbia.edu/~gelman/arm/examples/cd4/allvar.csv>. As a response, you should use the square root of the CD4 percentage (`CD4PCT`). (For some context, CD4 cells are white blood cells that get infected by HIV. The count of CD4 cells in a blood sample is often used as a measure of progression to AIDS, with lower counts indicating a weaker immune system.) Each subject is uniquely identified by the `newpid` column. As the main time variable (and variable whose coefficient we will call “slope”), use the time since treatment began in years, that is, the difference `visage - baseage`.

The subjects are divided into two treatment groups, with `treatmnt = 1` as the control and `treatmnt = 2` as the experimental group. I'm not sure what the treatment being evaluated was, so we'll follow G&H's lead in ignoring it (though, optionally, you could add in a treatment covariate and assess whether or not you think the treatment is promising. If we assume the treatment was randomly assigned at the first visit, it *shouldn't* have any effect on the intercept. Thus, this is a case where we would only include the treatment interacted with the time variable, leaving its main effect out as that would fit adjust the intercept.)

Do:

- **Gregor 1** In your *final project data*, do you have categorical variables that might work as groupings for random effects (they have enough levels)? Is there nesting? Do you think mixed methods might improve your modeling plans?
- **G&H 11.4** Chapter 11, #4 (page 249)
- **G&H 12.2, 12.3, 12.4** Chapter 12, #2, #3, #4 (page 277)
- **G&H 13.5 (a)** Chapter 13, #5 part (a) (page 299)
- **Gregor Optional:** Use your favorite model of the CD4 models and add a treatment-time interaction term. Bootstrap confidence intervals to determine whether the treatment effects the slope.