Exam 4
Statistics 426
Due May 15, 2015

This exam will involve analysis of the `toenail` dataset from the `faraway` package, to answer some specific questions that are listed below. Guidelines are given on how to prepare your document, and you should imagine this is a report done for some professional purpose. A substantial part of the scoring will be due to how well your procedures and results are communicated. **Keep in mind, this is an exam and there can be absolutely no collaboration or discussion of your analysis or report with other students, or anyone else.**

The aim is to compare two oral treatments for toenail infection. Patients were evaluated for the degree of separation of the nail. Patients were randomized into two treatments and were followed over 7 visits: four in the first year and yearly thereafter. The patients were not treated prior to the first visit, so this should be regarded as baseline.

1. Fit regression models to uncover treatment effects on the outcome of the toenail infection, that consider possible effects due to the time of the test as well as possible treatment effect interactions with time.

2. Perform this analysis using each of three methods: generalized linear mixed models, generalized estimating equations, and transition models. Compare and contrast the three approaches, and state how they compare with regard to estimated treatment effects. Note that for transition models, some work might need to be done to get the dataset in shape to run with the glm() function.

3. Also, compare the three methods above with a naive approach that ignores that observations are taken within the same person, and regards each line in the dataset as independent of one another.

Outline for reports
1.Introduction Paraphrase the main objectives of your analysis, and perhaps summarize the de Backer and the Lesaffre papers that are referenced in the R helpfile for the toenail data.
2.The Data Provide some summaries of your data, search for outliers and consider what to do with them in your analyses.
3.Methods Discuss your methods for developing your models, such as variable coding, utilization of interactions (if needed), selection of covariance structure (in gee case), and so forth.
4.Results Show the results and details of the fitted regression models in appropriate tables and possibly any relevant graphs. Provide an interpretation of the results and a comparison of the four techniques for modeling (gee, glmm,transition, glm with independent assumption).
5.Summary Give a very brief summary of what you have found in a non-technical manner
6.Appendix Provide the R-code that you used to conduct your analysis.