P8157 - Analysis of Longitudinal Data, Fall 2019
Homework - 2
DO NOT SUBMIT

October 22, 2019

Marginal Models. For this homework you will need the R package *gee*.

Part A

Read the Analysis requirement in part B. For the both analysis parts (1) and (2) in part B, clearly set up the model i.e. clearly state

1. Distribution assumption for the outcome
2. The systematic component of the model
3. The canonical link function
4. The variance function
5. The correlation structure

Part B [1]

In a clinical trial of patients suffering from epileptic seizures, patients were randomized to receive either a placebo or the drug progabide, in addition to standard therapy. A baseline count of the number of epileptic seizures in a 8-week period prior to randomization was obtained. In addition, counts of the number of epileptic seizures in each of the four successive 2-week (post baseline) treatment periods were obtained. The goal of the analysis is to make a comparison between the two treatment groups in terms of changes in the rates of epileptic seizures throughout the duration of the study. The data epilepsy.dat is available on Courseworks. The variables in the data set are

```
Patient ID
Treatment (0=Placebo, 1=Progabide)
Age at baseline
Baseline 8 week seizure count
First 2 week seizure count
Second 2 week seizure count
Third 2 week seizure count
Fourth 2 week seizure count
```

[1] Always explore the options available in the commands in your software

1. (a) Consider a model for the seizure rate that includes the main effects of treatment and time (time is regarded as a ordinal variable with 5 levels) and their interaction. (*library gee, command gee()*).

   (b) What conclusions do you draw about the effect of treatment on changes in the seizure rate.

   (c) Construct a variable *Ptime* such that $Ptime = 0$ if baseline and $Ptime = 1$ if post baseline. Repeat the above analysis using Ptime instead of the 5 categorized time variables. Compare the two models.

   (d) Are there any potential outliers in the dataset. If so do they have any effect on the analysis.

   (e) Summarize your results and provide relevant interpretation.

2. Construct a new outcome variable NS at each time point such that $NS = 1$ if number of seizures is greater than 20 and $NS = 0$ If the number of seizures is less than or equal to 20.

   (a) Repeat the analysis in (1) above to model this new outcome.