

BIOS 6612 HW 6: Generalized Estimating Equations

The data for this assignment comes from a clinical trial of 59 epileptics. For each patient, the number of epileptic seizures was recorded during a baseline period of 8 weeks. Number of seizures was then recorded in 4 consecutive 2-week intervals.

The data is referenced in Lecture 12.3 and given as part of `geepack` package.

- Data set is in wide format (one row per subject)

Each row of the data set contains the following variables:

- `y1`: the outcome (number of seizures) at the first time point
- `y2`: number of seizures at the second time point
- `y3`: number of seizures at the third time point
- `y4`: number of seizures at the fourth time point
- `base`: number of seizures at baseline
- `trt`: the treatment (0 = placebo, 1 = progabide)
- `age`: in years, at baseline

You can load the data into R by loading the `geepack` library and calling `data(seizure)`.

Exercise 1: Exploratory Data Analysis (40 points)

- 1a. (10 pts)** Load the data and convert it from wide to long format. Is this data well-balanced? Why or why not?
- 1b. (15 pts)** Construct plot(s) to explore the mean structure of your outcome over time, and relationships between the outcome and covariates. Interpret trends that you see.
- 1c. (15 pts)** Construct plot(s) to explore the correlation structure of your data. Interpret trends that you see. What might be a good choice of working correlation structure?

Exercise 2: GEE Model for Count Data (60 points)

- 2a. (15 pts)** Your clinical collaborators are interested in a Poisson GEE model with treatment, age, and time as predictors. They also want to control for number of seizures at baseline. Write out the form of the GEE model. Assume an exchangeable correlation structure.
- 2b. (15 pts)** Fit this model in R using the `geepack` package. Provide a table of estimates, standard errors, and P-values. What is the estimated correlation coefficient for your model?
- 2c. (5 pts)** Is this an appropriate model for your data? Should you consider other correlation structures?
- 2d. (15 pts)** Provide exponentiated coefficients for your model. Interpret the coefficients of your model on the exponentiated scale.