667 Homework 7

Due Friday (11/22), 4 pm

Problem 13.1. The data file is in the code and data folder.

Add 13.1.0: Present descriptive statistics. [Do not simply present copies of computer output. Do not present more than 3 significant digits. For the percentages in this problem 2 digits are enough. No p-values]

In 13.1.1 and 13.1.2, force the constraint that the treatment groups have the same mean at time 0 ("occasion" j=0).

In 13.1.3 and 13.1.4, fit the model that includes in its linear predictor all factors (clinic, time, treatment) and all two-factor interactions. Force the constraint that, within each clinic, the treatment groups have the same mean at time 0 ("occasion" j=0).

13.1 In a clinical trial of patients with respiratory illness, 111 patients from two different clinics were randomized to receive either placebo or an active treatment.

Patients were examined at baseline and at four visits during treatment. At each examination, respiratory status (categorized as 1 = good, 0 = poor) was determined.

These data are from Koch et al. (1990), and are reported in Davis (1991) and Stokes et al. (1995). The main objective of the analyses is to understand the joint effects of treatment and time on the probability that respiratory status is classified as good. It is also of interest to determine whether the effect of treatment is the same for patients from the two clinics.
The raw data are stored in an external file: respir .dat
Each row of the data set contains the following eight variables:
ID Clinic Treatment Y0 Y1 Y2 Y3 Y 4

Note: The respiratory status response variable Yj is coded 1 = good, and 0 = poor, at the jth occasion. The categorical (character) variable Treatment is coded A = Active drug, P = Placebo. The categorical variable Clinic is coded 1 = clinic 1, 2 = clinic 2.

13.1.1 Ignoring the clinic variable, consider a model for the log odds that respiratory status is classified as good, including the main effects of treatment and time (where time is regarded as a categorical variable with five levels), and their interaction.

Use generalized estimating equations (GEE), assuming separate pairwise log odds ratios (or separate pairwise correlations, if available software does not permit the within-subject association to be parameterized in terms of log odds ratios) among the five binary responses. Construct a test of the null hypothesis of no effect of treatment on changes in the log odds that respiratory status is classified as good based on the empirical standard errors.

13.1.2 What conclusions do you draw about the effect of treatment on changes in the log odds? Provide results that support your conclusions.

13.1.3 Patients in this trial were drawn from two separate clinics. Repeat the analysis for Problem 13.1.1, allowing the effects of treatment (and, possibly, time) to depend on clinic.

(a) Is the effect of treatment the same in the two clinics? Present results to support your conclusion.

(b) Find a parsimonious model that describes the effects of clinic, treatment, and time, on the log odds that respiratory status is classified as good. For the model selected, give a clear interpretation of the estimated regression parameters for the final model selected.

13.1.4 For the final model selected in Problem 13.1.3, construct a table of the estimated probabilities that respiratory status is classified as good as a function of both time and treatment group (and, possibly, clinic). What do you conclude from this table?