

Homework #6

BSTA 519 Applied Longitudinal Data Analysis

Fall 2021

Due - Monday 11/22/2021 by midnight

Respiratory Data (Binary) and Marginal GEE model

A clinical trial was conducted to compare two treatments for a respiratory disorder. Patients in each of two centers were randomly assigned to groups receiving the active treatment or a placebo. During the trial, **respiratory status** (coded here as 0=poor, 1=good) is determined **at baseline** and **four follow up visits**, the outcome variable of this study. For independent variables, we have **center** (1 or 2), **treatment** (placebo or active treatment), **age** and **gender**; all are categorical variables with two levels except for age.

The research question here is to determine whether the respiratory statuses of the patients change over time and whether the treatment worked better than the placebo to improve respiratory status. In addition, we want to know whether there is any difference in the outcome by center.

The dataset for the homework is called **Respiratory data HWK.xls**. Import the file into the statistical software of your choice, and do the following:

- a) For each treatment group, calculate and plot the proportion of good respiratory status over time and describe the trend you see. Does the plot suggest a treatment effect?

For each treatment group, calculate the proportion of good respiratory status by center, and by gender separately. Comment on whether these proportions would give you some idea of whether the treatment effect differs by center, or by gender.

- b) The primary research question is to determine whether the treatment works better than the placebo to improve respiratory status. Using **GEE method** and an **unstructured working correlation matrix** (and pay attention to the ID variable), fit a marginal logistic regression model with **treatment group**, **time** as a discrete variable and **interaction term between group and time** (like a response profile analysis). In this analysis, treat baseline respiratory status as part of the repeated measures of outcome.

Write out your model and **calculate the predicted probability of good respiratory status at each visit** (including baseline) for each group and compare these probabilities with the results you get in a).

Provide an **estimate of treatment effect at each visit** (including baseline) and interpret your estimates. *(I did not provide the codes that explicitly show you how to get these estimates but I totally believe that you are all very capable to figure these out based on what you have learned so far...☺)*

Based on above estimates comment on whether the treatment effect change over time.

Homework #6

BSTA 519 Applied Longitudinal Data Analysis

Fall 2021

Due - Monday 11/22/2021 by midnight

Does the treatment work better than the placebo to improve respiratory status?

Answer the questions and interpret your results.

Fit a marginal logistic regression model with treatment group, time as a discrete variable and interaction term between group and time but using the exchangeable working correlation matrix. Compared to the unstructured working correlation matrix, does the exchangeable working correlation matrix provide an adequate fit?

- c) In addition to the treatment effect, it is also of interest to determine whether the treatment effects differ by center, or gender. Using GEE method and an unstructured working correlation matrix, fit an appropriate marginal logistic regression model to answer this question.
- d) Considering all independent variables (treatment group, time, center, gender and age), find a parsimonious model to describe the relationship between good respiratory status and these variables. That is, you can only consider the significant terms in the model, and provide appropriate interpretation for your results. For the purpose of building this model, consider: 1) whether treatment worked better than the placebo to improve respiratory status and whether the treatment effects changed over time; 2) for other variables including center, gender and age, only consider how they are associated with the outcome (as main effects). Write out your model and interpret your results.
- e) Instead of treating baseline respiratory status as part of the repeated measures of outcome, treat baseline respiratory status as a covariate and fit a similar response profile model as in b). That is, using the unstructured working correlation matrix, fit a marginal logistic regression model with treatment group, time as a discrete variable interaction term between group and time, and baseline respiratory status. You may also fit a simplified model by excluding the interaction term if the interaction term is not significant.

In this case, how do you answer the question whether the treatment works better than the placebo to improve respiratory status? Interpret your results.

For more information of whether treating baseline as a covariate or part of the outcome, see your textbook Chapter 5.7 ALTERNATIVE METHODS OF ADJUSTING FOR BASELINE RESPONSE.