BIO 226: Homework Assignment 4

Due: 1:30pm (in class): Thursday March 26, 2015

Purpose: To provide an introduction to the joint parametric modeling of the mean response and covariance for longitudinal data.

Instructions:

1. For each question requiring data analysis, support your conclusions by including only the relevant SAS output in your answer.

2. Include your SAS program (but not your SAS output) as an appendix to your solutions.

Late homework will not be graded unless you make arrangements with the Instructor prior to the due date/time.

Joint Modeling of Mean and Covariance in the a Study on the Effect of Ozone on Pulmonary Function

In a study designed to examine the acute responses of normal subjects to ozone exposure, Follinsbee and colleagues randomized 20 subjects to exposure to either room air or 0.12 ppb ozone. The exposure lasted six hours. A baseline and 6 other measures of FEV1 (a measure of pulmonary function) made at hourly intervals were recorded for all study participants in the study. Subjects cycled on an exercise bike before a measure of FEV1 was obtained. The investigators were interested in determining whether changes in pulmonary function during the 6 hours of exposure were different in the ozone and room air exposed groups.

In the analyses of the data from this study, the response variable of interest was FEV1 (ml)/100. From here on out when we refer to FEV1 we mean this rescaled response. The measurement times were coded 0-6, with time=0 for the baseline measurement, time=1 for the measurement at hour 1,…,time=6 for the measurement at hour 6. For this homework, we will consider the data from hours 0, 2, 4, 6. The two exposure groups were coded 0 and 1, with group=1 denoting exposure to ozone and group=0 denoting exposure to room air.

The data are in the file "ozone0246.txt" on the course web page. Each row of the data set contains the following four variables: subject ID, hour, group and the y = FEV1/100 measurements, respectively.

1. Descriptive Analyses:

   a. Describe key aspects of the longitudinal design and completeness of data.

   b. Plot the FEV1 response against hour for each subject exposed to room air (overlaid on the same plot) and for each subject exposed to ozone (overlaid on a second plot). Comment on any patterns in the data or other notable aspects of the data.

   c. Obtain the mean FEV1 value at each hour of measurement for ozone and room air subjects separately. Plot the means against hour. Comment on the pattern of change in mean FEV1 with hour for ozone and room air subjects.

2. Fitting a "Maximal" Model and Evaluating Variance-Covariance Structure:

   a. Define a reasonable "maximal" mean model for this study. Fit this model using an unstructured variance-covariance matrix. Comment on the variance structure and on the correlation structure. What simplified variance-covariance structure(s) might be reasonable? Justify your answer.

   b. Keeping the same maximal mean model, evaluate whether your suggestion(s) for the variance-covariance structure from question 2a as well as the following models for the variance-covariance structure provide an adequate fit to the data compared with an unstructured variance-covariance:
      i. compound symmetry
      ii. heterogeneous compound symmetry
      iii. 1^{st}-order autoregressive
      iv. heterogeneous 1^{st}-order autoregressive.

      Using likelihood ratio tests and the AIC criterion as appropriate, identify a model for the variance-covariance structure that provides a good fit to the data. Provide estimates for the parameters used in defining this variance-covariance model. Also provide estimates of the variance-covariance and correlation matrices.

3. Analysis of Response Profiles:

   Fit the usual model for the analysis of mean profiles using room air exposure as the reference level for group and baseline as the reference group for time. Use the variance-covariance structure identified in your answer to question 2b. Based on this model:

      a. Test the null hypothesis that the pattern of means over hours is identical (coincides) for the two exposure groups. What do you conclude?

      b. Test the null hypothesis that the mean response profiles of the two groups are parallel. What do you conclude?

4. Fitting a Linear Model in Time:

Fit a model that includes hour as a continuous variable, group and their interaction. Use the model for the variance-covariance structure that you identified in question 2b.

a. What is the estimated rate of change in mean response for the room air group?

b. What is the estimated rate of change in mean response for the ozone group?

c. Test the hypothesis that the rates of change in mean response are identical in the two groups. What do you conclude? Give a possible reason for any difference in conclusion you make from this test and the analogous test based on the mean profiles analysis from 3b.

d. What is the estimated difference in rate of mean change between the two groups? By calculating a 95% confidence interval for this difference, identify what are plausible values for the underlying true difference.

5. Evaluating the Fit of a Linear Model in Time:

Does a model with a linear trend in hour for each exposure group adequately describe the pattern of change in the two groups? Justify your answer with appropriate statistical analysis.

6. Summarizing the Key Results and Conclusions:

Write a brief structured abstract (maximum 200 words) summarizing the objective, methods, results and conclusions that might be drawn concerning exposure differences in patterns of pulmonary function over time.