

BIO 226 Homework Assignment 5
Due in class on Thursday April 2, 2015

Purpose:

To review concepts of linear mixed effects models

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 as an appendix to your solutions.
- 3) Late homework will not be graded unless you make prior arrangements with the Instructor.

Analysis of Data from Dental Growth Study

In a study of dental growth, measurements of the distance (mm) from the center of the pituitary gland to the pteryomaxillary fissure were obtained on 11 girls and 16 boys at ages 8, 10, 12 and 14 years (Potthoff and Roy, 1964).

The data are in the file “dental.txt” on the course web page. Each row of the data set contains the following six variables: subject ID, gender (“F” or “M”), and the measurements at ages 8, 10, 12 and 14 years, respectively.

Note: In HW2 you considered this study, but analyzed data only from the girls. In this homework please analyze ALL the data.

Problems:

1. Descriptive Analysis

- a. Describe key features of the study’s design and the completeness of data.
- b. Create two plots, one for each gender, showing the trajectories of each subject’s distance over time. Comment on the notable features in these trajectories.
- c. Create a single plot of the distance over time for each gender. Comment on the notable features in this plot.

2. Mixed Effects Analysis

Using PROC MIXED, fit a model for distance over (continuous) time which includes subject-specific intercepts and slopes as random effects and allows both the mean intercept and the mean slope (fixed effects) to differ by gender.

- a. State the model being fitted including any distributional assumptions. As much as possible, use the same notation as in the course notes.
- b. Compare the within-subject “error” variance to the between subject variance of the intercepts. What do you conclude?
- c. Obtain estimates for the differences in mean intercept and mean slope by gender and their associated standard errors. What do you conclude about these differences?
- d. Obtain and provide an interpretation for the correlation of the intercept and slope among subjects.
- e. From your model, conditional upon gender, obtain the standard deviation of the slopes among subjects. Provide a clinically meaningful interpretation of this standard deviation.
- f. Obtain the marginal correlation matrix for the vector of responses, $Y_i = (Y_{i0}, Y_{i1}, Y_{i2}, Y_{i3})$, for the subject with ID=1. Considering your findings in response to question 2b, why are the correlations relatively large?
- g. Are the random slopes necessary for these data? Answer this question by fitting the model with the same fixed effects but now with just random intercepts. Which model do you prefer for these data and why?
- h. The data for this study could also be analyzed using the repeated measures approach discussed extensively earlier in the course.
 - i) Define a repeated measures model with an unstructured variance-covariance matrix and the same form of fixed effects model as in question 2a which could be used.
 - ii) Fit this model and use the AIC statistic to assess the goodness of fit of the variance-covariance structure induced by the mixed effects model. What do you conclude?
 - iii) Now fit the model with a compound symmetry covariate pattern instead of an unstructured variance covariance matrix. Compare the results you get from this model to those obtained from the random intercept model you fit in 2g, and explain the reason behind any similarities or differences between these two sets of results.