BIOS 6643 Quiz 1, Fall 2013

Note: please show any additional work on the back side of pages. If you have any (general) questions on the SAS code, let me know.

1. Consider the Beta Carotene data that you analyzed on the last homework. Recall that y is the level of plasma carotene stored up after taking one type of vitamin over time. Drug codes are: 1: Solatene (30mg capsules), 2: Roche (60mg capsules), 3: BASF (30mg capsules), 4: BASF (60mg capsules). Times of measurement were 0, 6, 8, 10, 12 weeks. Here are two models that we considered in the analysis, based on the linear mixed model . Note that ‘V’ stands for *Var*(**Y***i*) and ‘R’ stands for *Var*(**ε***i*)=**R***i*.

Model I

**data** test; set long.beta\_carotene\_univar; t=time; run;

**proc** **mixed** data=test;

class id prepar time;

model y= prepar time prepar\*time;

random intercept t / subject=id type=un v vcorr; **run**;

Model II

**proc** **mixed** data=test;

class id prepar time;

model y= prepar time prepar\*time;

repeated / subject=id type=un r rcorr; **run**;

1. Write out the statistical model for Model I. Write enough detail so that all parameters (fixed effect and covariance) are mentioned. Also, include indices on parameters and variables, using *h* for group (i.e., ‘prepar’), *i* for subject and *j* for time. You can either write a full-rank or less-than-full-rank model.
2. Determine the number of covariance parameters for Model I and Model II.
   * 1. Model I:\_\_\_\_\_\_
     2. Model II:\_\_\_\_\_\_
3. Identify the number of columns in  and  for each model. For , consider both the full-rank (FR) and less-than-full-rank (LTFR) models.

Columns in  Columns in 

LTFR FR

* + 1. Model I: \_\_\_\_\_\_ \_\_\_\_\_\_ \_\_\_\_\_\_
    2. Model II: (same as above) \_\_\_\_\_\_

1. Say we want to test whether the mean change from 0 to 12 weeks differs between at least 2 groups (the null would be that the mean change is the same across all groups). Write a statement that could be added to the SAS code above to accomplish this.
2. Say that two additional covariates (predictors) are added to Model I, but otherwise the code is the same as above (for Model I). The AIC drops. Does this suggest that the new model is better? Justify your response.
3. For Model II, there is no ‘V’ or ‘VCORR’ option in the REPEATED statement. Explain why this is the case (and why you really don’t need it).
4. Note that the UN structure for **G** in Model I allows for modeling of a covariance parameter between the subject (random) intercept and slope. The estimate of this covariance is positive. Explain what this means, in terms of the data, and write your statement meaningful to the application.
5. Mention another potential **R** structure that you might consider to model these data other than that for Model II. Briefly justify your response. Your credit will depend on your justification.
6. **Answer 3 of the next 4 questions.** PLEASE CROSS OUT THE LETTER OF THE PART THAT YOU WISH TO ELIMINATE. I will be somewhat picky about giving full credit on these, but will be more lenient about partial credit…
   1. We have talked about the fact that a random intercept model (i.e., random intercept but ‘independent’ structure for ) is equivalent to having no random effects, but specifying the CS structure for . Explain why ML or REML estimates of parameters in the models will be the same for the two approaches (if we let  for the 2nd approach).
   2. Name one advantage and one disadvantage of REML estimation compared with ML estimation.
   3. How does inference (hypothesis tests, estimation) differ for mixed model methods relative to RM ANOVA?
   4. What are drawbacks of RM ANOVA compared with mixed model methods?