**BIOS6643 Fall 2022 HW3 Due Wednesday, Oct. 5, 5pm (upload to Canvas)**

Consider the Fitness data that includes 3 groups (low intensity=a, high intensity=b, control=c). Group and time will be modeled as class variables, a group\*time interaction will be included, and the error covariance matrix will be fit using a first-order autoregressive [AR(1)] structure. There are no random effects in this model. Note that you can either use PROC MIXED in SAS or the gls function in R (nlme package) to fit the model. NOTE: use Satterthwaite DDF, mainly so we’re all on the same page. To do this in SAS, use ddfm=satterth as an option in the MODEL statement. Please include your SAS and/or R code in an Appendix.

* 1. Write the less-than-full-rank (LTFR) statistical model, i.e., include all levels of class variables, letting α denote group, τ denote time and γ denote interaction; let *h* indicate group, *i* indicate subject and *j* indicate time. Include distributions, assumptions, etc. for the model.
  2. How many (a) fixed-effect parameters and (b) covariance parameters are in the LTRF model?
  3. Fit the model using SAS or R. How does this model that uses an AR(1) structure compare to the random intercept model (with the same fixed effects), based on AIC goodness of fit? (What I call the ‘random intercept model’ has a random intercept but simple error covariance structure – i.e., you do not need to specify the latter.)
  4. Write out the form of **V***i*=*Var*(**Y***i*) for the 2 models in Q3 (should be 5×5 matrices). [Note for SAS: you can also determine what these matrices look like numerically in SAS by including the *v* option (after the slash) if there is a RANDOM statement, and an *r* option in the REPEATED statement if there is no random statement. Similarly, the *vcorr* and *rcorr* options put these in terms of correlation matrices.]
  5. Based on the model in #1-3 (using AR(1) structure on errors), determine coefficients for the following customized tests/estimates, and run then with SAS or R to obtain results.
     + 1. Write an estimate (or contrast) to determine the change from baseline to 4 weeks (4-0wk) for group a.
       2. Write an estimate (or contrast) to determine the change from baseline to 4 weeks (4-0wk) for group b.
       3. Write an estimate (or contrast) to compare change over time (4-0wk) for group a versus b.
       4. Write an estimate (or contrast) to compare the average of programs b vs. c, but only for the last 2 weeks.
       5. Write a contrast for the group\*time interaction for group b (high intensity) versus c (control).
  6. Using the model determined from Q3 to be the best in terms of AIC, refit a model considering time as a continuous variable. What order of polynomial is sufficient? How do you decide? Include group\*time interactions, up to the specified order. E.g., if quadratic is sufficient, include group\*time and group\*time2 as predictors in addition to group and time. Based on all the models you have fit so far, which one would be your ‘final’ model, e.g., if you were to submit a paper on it? For practice: write out some meaningful contrast/estimate statements that involve group and/or time as continuous, similar to those discussed during lectures (do not need to turn in).
  7. Write a brief (half page) summary of your findings overall, including the statistical lingo that you would ordinarily include in a research article. (Talk to me if you have questions about this.) In your write up, feel free to include other tests if you’d like. For example, in 6e you did an interaction test for program b versus c; you could compare a and c or a and b. This is not necessary for full credit, but might help in your learning and to make a more complete write up..