**Lecture 6**

What we covered last time:

1. Review of previous lecture material (RM ANOVA, DDF in the LMM, Distributions for quantities in the RM ANOVA table, 2 types of predicted values in SAS).
2. Means versus effects models (considering data with group and time as class variables).
3. Less-than-full-rank versus full-rank models.
4. Using a sum-to-0 constraint in a LTFR model.
5. Writing **L** (a contrast) for the test/estimate involving **Lβ** for a 2-way effects model.

Plan for today:

1. Continue with the previous model *Yhij*=*µ+αh+τj+γhj+εhij*, where *h*=1,2, *j*=1,2,3 (2 groups and 3 times). [Can add random effects, but our main focus is on the fixed effects, since we’re developing specific tests/estimates of interest.
   1. Write a test to compare groups, averaged over time (main effect test for group).
   2. Look at dog data (similar model, but 3 groups and 5 times). Perform an interaction test to compare 2 drug groups. NOTE: if there is more than 1 equation in H0, must use F-test (and the CONTRAST statement in SAS).
2. Talk about estimability: what functions of parameters have unique estimators? I.e., which functions of parameters do not depend on how the linear dependency is dealt with?
3. Walk through HW questions.
4. For Wednesday: write contrast/estimate statements for models that include continuous predictors.

**Estimability (in a nutshell)**

1. Estimability is an important concept in less-than-full-rank models since fixed-effect parameters in these models do not have unique unbiased estimators.
2. For example, say you have a class variable with 3 levels. You can set the effect associated with highest level to 0, or the first level… Alternatively, you could place a sum-to-0 constraint on the effects. So clearly the estimates for the effects will depend on which approach you take.
3. An alternative is to leave the linear dependency in the **X** matrix and solve quantities of interest using generalized inverses. There are multiple g-inverses you can use for the same non-singular **X**.
4. This begs the question: “What linear combinations of parameters do not depend on the approach one takes to deal with the linear dependencies?” Theory of estimability tells us precisely which ones…
5. First, we need to introduce a generalized (or ‘g’) inverse, which can handle a matrix with linear dependencies. A g-inverse satisfies **AA**–**A** = **A**, where **A**– is the g-inverse of matrix **A**. See the course notes for one specific approach to creating a g-inverse.
6. This allows us to solve quantities such as MLE estimator of **β** in the LMM:
7. The projection (or hat) matrix in the GLM is

which is unique, i.e., does not depend on the g-inverse used for

(**X***t***X**)–.

1. Also, a linear combination of parameters, **Lβ**, is uniquely estimable if **L**=**LH**, where **H**=(**X***t***X**)–(**X***t***X**). Keeping **L** generic [e.g., **L**=(*L*1 *L*2 … *Lp*)] allows you to determine the general form of estimable functions. See the course notes for an example of how to use this in practice and justification for the quantity.
2. In SAS, you can quickly determine the form of the estimable functions by including the ‘e’ option in the MODEL statement.
3. For purposes of this course, what do I expect you to know? Know the basic idea; I will not quiz you over specifics (i.e., the math), but much more detail is given in the course notes if you are interested in learning more.
4. Fun fact: SAS’s g-inverse looks exactly like a set-to-0 option, setting the effects associated with highest levels of class variables to 0. This is because it looks for linear dependencies in the X matrix moving from left to right; it is generally not found until you reach the highest level of each factor… Please see an example of the calculation of a g-inverse (using SAS’s approach) in the course notes.