**BIOS6643 HW4 Due Friday, October 14th**

1. Consider a basic science experiment conducted where cell counts are measured at 4 time points for samples taken from individual subjects or animals. A linear mixed model will be fit for the data (perhaps after log transformation), and fixed effects will be included for time, and possibly treatment group as well as their interaction. (To answer this question we do not need to know the specific form of **Xβ**.) Determine the structure for **V***i* if a random intercept for subjects will be included, plus an AR(1) structure for the error covariance matrix (**R***i*). What does the combination of non-simple **R** and **G** allow you to do in modeling covariances that using only one cannot do? Discuss in a few sentences.
2. For the Dog Data, complete the following.
   1. First, use group and time as class variables, plus group\*time, and determine the best Kronecker Product structure to use for the error covariance structure. (Note: there are 3 options in SAS, and you may be limited by what will work.) Highlight results.
   2. Add the R and RCORR options in the REPEATED statement. (Note that this is equivalent to the fitted V matrix since there are no random effects.) In 3-4 sentences, interpret the correlations and variances in the data.
   3. For the model in part a, write a contrast to test the null hypothesis that the means for the two treatment groups are equivalent for all time points and summarize the results.
   4. Is there a time-as-continuous model that can ‘beat’ the time-as-class model in terms of AIC? (Remember that the highest degree of polynomial is one less than the number of time points.) Summarize your results.
3. In a paragraph, describe your plan for your project. Include the data that you plan to use, and a couple of research questions of interest. (You can change these later if you need to.) If you plan to do an alternative project, describe that (e.g., special research topic; advanced data analysis with R).

Not to turn in but review before class on Wednesday (we will complete in class):

* + 1. One model we used for the Mt. Kilimanjaro data included random effects for subject, up to the quadratic term (plus covariances between random effects), along with a simple **R** structure. (We did find at least one model with a better AIC, but let’s focus on this one for now.) We talked about how including multiple random effects can induce a covariance structure that is time sensitive (or in this case, altitude sensitive). Show this by considering a simple data set and model. In particular, let times be *t*=0, 1, 2, and consider a model that includes a random intercept and slope for time by subject, plus covariance between them (i.e., UN structure in **G**). Show that it is possible to obtain Cov(*Yi*1,*Yi*2) > Cov(*Yi*1,*Yi*3) < Cov(*Yi*2,*Yi*3), i.e., decaying covariance as distance between time points is increased. For what covariance parameter values will these hold.
    2. Consider a study where children are sampled from schools, and then measured over time. We will include a random intercept for schools and for subjects within schools (but simple **R**). Determine **V***h*, the covariance matrix for school *h*, if there are 3 children sampled from this school, where the first two kids have 3 measures and the last has 2. You might find it helpful start by writing the model for outcome *Yhij* and determining the design matrix for the random effects. (You can just write something generic for the fixed-effect part of the model.) For thought, not to turn in: how would **V***h* change if we had more measures for subjects and employed the AR(1) structure for **R***i(h)* (the error covariance structure for subject *i* within school *h*)?