BIOS6643 HW3 Due Friday, October 4, 2019, 5pm

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. 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?
3. Fit the Beta Carotene data using a continuous model for time, including group and group\*time in the model. (For a description of the data, see the file in the HW folder.) Determine the degree of polynomials for time that is important and sufficient for the model. For covariance structure, define the UN structure for **R**.
   1. Write your final model, fit it and compare it to the model that used group, time and group\*time as class variables. Which would you go with in a final report? Explain. NOTE: in comparing model AIC’s use method=ML for a more apples-to-apples comparison, particularly when changes are being made to the fixed effects.
   2. Write an estimate or contrast statement for your continuous model based on what you think is interesting. The custom estimate and/or test could involve a subset of the data (e.g., comparing 2 specific groups), or the whole data.
4. 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*)?