## Homework # 11—Due Mon 8 May

This problem comes from an horticulturalist investigating the relative energy balance between vegetative and flowering parts of cranberry uprights in bogs of Northern Wisconsin. [The upright is like a plant stem. It is very difficult in a bog to determine which uprights belong to which plants, as they tend to overlap considerably. In fact, a large bog may be essentially one plant!]

She has one cranberry bog which has been divided into 6 `samples`. On seven `periods` through the summer she measured the soluble sugar and starches (`conc`) in these two parts of uprights. She is interested in assessing the change in sugars and starches over the course of the summer. That is, is one consistently higher than the other, or does this depend on `julian` date?

Your task is to address her question using the variety of repeated measures methods we discussed in class. Note that the `repeated` option does not work quite right for this problem in either `proc glm` or `proc mixed`! The former considers `sample` as a factor and assumes you are interested in `period*sample` interactions. In fact, if you include `sample` in the model, the following message in the log indicates that multivariate tests cannot be performed:

```
NOTE: No multivariate tests performed for JULIAN due to insufficient error
      degrees of freedom.
```

The second run of `proc glm` drops the `sample`, which is assumed to only affect the whole plot (as a blocking factor). The new `proc mixed`, on the other hand, seems to handle `sample` as a whole plot blocking factor appropriately. Notice that the final two runs of `proc mixed` seem to have the right `F` test values (check for yourself based on earlier printout), but have the wrong degrees of freedom and `p`-values. This is a problem in our version of this package, which I mentioned in class. YOU should count degrees of freedom so that YOU know what you are working with!

1. Analyze the data, date by date, for differences between flowering and vegetative parts. Discuss dates individually and give an overall assessment. You may find tables of means and/or interaction plots to be helpful. Include measures of variation as appropriate.

2. Conduct an analysis as if this were a split plot design. [Use whatever segment of the output you think is appropriate, but clearly indicate your choice in terms of model and assumptions.]

3. Adjust the degrees of freedom for the split plot using the Greenhouse-Geiser `epsilon`. Perform relevant tests. How does this change the results?

4. Adjust the degrees of freedom for split plot using Huynh-Feldt `epsilon` (but use 1 if it is greater than 1). Comment on why this is so different from 3 based on discussion in class (see also SAS/STAT manual discussion of Huynh-Feldt conditions).

5. Conduct the sphericity test using the `printe` option for the `repeated` statement. This is a variant of the test discussed on p. 359 of MJ (sec 27.1). Interpret the results.

6. Examine the orthogonal polynomials individually. As with separate analyses by dates, give some indication of where differences in `tissue` appear to be significant, and give an overall assessment. [Again, you can find the needed material either in the `proc glm` or `proc mixed` listings. Indicate which you use.

7. As stated above, the multivariate analyses (and the sphericity test) had to be performed with sample removed from the model. Briefly consider the multivariate tests and compare these results with the earlier findings.

8. Based on a "gestalt" of the analyses, what are your conclusions? How would you share this with the scientist? Be brief, presenting relevant information in a short paragraph, perhaps with adjoining table or figure.