Homework # 8—Due Mon 17 Apr

This problem comes directly from examples in Chapters 18-20 of Milliken and Johnson. The SAS runs are available in the public data directory `/p/stat/Data/MJ` as files of the form `wheat.*` and `random.*`. There are a number of errors in the MJ book. For instance, the E(MS) below Table 18.1 are not all correct. Keep this in mind as you compare your findings with the text. Note also that their notation changes quite a bit in this region of the text.

1. Consider the one-way random experiment concerning wheat varieties.
(a) Justify in words the one-way random model for this experiment. Why is it appropriate.
(b) Write down the model with all assumptions. Indicate in a small table the sample sizes.
(c) Test whether there is any variation among wheat varieties in terms of insect damage.
(d) Estimate the (two) variance components using one of the four methods discussed in class. Briefly justify your choice of method–are you selecting it because it is easy to explain to scientist or because it has some (specified) "optimal" properties?
(e) Construct a 95% confidence interval for each variance component using the estimate from (d). Justify your approximation, whether it be based on chi-square or normal distribution (see ch. 20 of MJ).
(f) Estimate the grand mean insect damage for wheat.
(g) Estimate the variance of the grand mean. Careful–the design is not balanced and you may need to make some adjustments. See details in my chapter.
(h) Construct a 95% confidence interval for the grand mean.

2. Consider the two-way random experiment labelled simply as row and column treatments.
(a) Describe in words an experiment resulting in this data. That is, identify realistic treatments instead of "row" and "column". Justify your choice as a random rather than fixed model.
(b) Write down the model with all assumptions. Indicate in a small table the sample sizes.
(c) Test whether there is any variation among row and column treatments, and any evidence of row*column interaction.
(d) Why are the Type I estimates of main effects variance components negative? Is this consistent with the results of tests from part (c)?
(d) Estimate the interaction and error variance components using ML or REML. Briefly justify your choice of method–are you selecting it because it is easy to explain to scientist or because it has some (specified) "optimal" properties?
(e) Construct a 95% confidence interval for the two variance components estimated in (d). Justify your approximation, whether it be based on chi-square or normal distribution (see ch. 20 of MJ).
(f) Estimate the grand mean.
(g) Determine the variance of the mean if all variance components were known for this experiment. Careful–the design is not balanced and you may need to make some adjustments. Thus my notes provide a start, but you must adjust for unequal sample sizes. [Hint: use the model and determine what the sample grand mean ($\bar{y}_{...}$) is in terms of model components. Then find its variance.]
(h) Estimate the variance of the grand mean using the variance component estimates. [Do not try to construct a confidence interval–you would need to estimate `df` which is rather complicated!]