Assignment #8 — Due Friday, March 30, 2007, by 4:00 P.M.

Turn in homework in lecture, discussion, or your TA's mailbox. Please indicate the discussion section you expect to attend to pick up this assignment.

311: Tues. 1:00–2:15 **312:** Wed. 2:30–3:45 **313:** Wed. 1:00–2:15

1. Ten needles were randomly selected from a large branch of a loblolly pine tree. The stomata (microscopic breathing holes) on loblolly pine needles are arranged in rows. On each needle, 4 rows were randomly chosen and the number of stomata per centimeter were determined for each row. The resulting data are shown below and are in the file `needle.txt`.

Needle #	1	2	3	4	5	6	7	8	9	10
	149	136	143	121	148	129	127	134	117	129
	143	139	142	133	121	134	130	137	128	132
	138	129	124	126	124	127	123	119	117	131
	131	143	134	130	128	113	125	130	118	137

   (a) Write down the random effects model appropriate to this problem identifying all terms used. State the distributional assumptions. Why is a random effects model more appropriate than a fixed effects model?

   (b) Draw a nesting diagram for the model variables as in the notes.

   (c) Examine a plot of the stomata counts versus needle. Are the random effects model assumptions reasonable?

   (d) Estimate all relevant variance components defined in (a) using both `lmer` and from computations with sums of squares using an ANOVA table from an analysis using `lm` (or `aov`) based on the expected mean square error (EMS) expressions of $\sigma_e^2 + n\sigma_a^2$ for "treatment" and σ_e^2 for error. (See end of notes for *Random Effects in R*.) Are the estimates similar?

2. Suppose it is of interest to estimate the mean number (μ) of parasitic insects per unit leaf weight for a particular tree. Eight leaves were randomly selected from the tree. From each leaf, four small disks were cut. For each disk the number of insects per unit leaf weight was determined. The data presented below are also in the file `leaf.txt`.

Leaf #	I	II	III	IV	V	VI	VII	VIII
	11.4	20.2	14.3	23.6	8.4	18.3	21.6	12.8
	19.3	17.0	11.1	23.1	10.7	16.2	15.8	9.3
	16.2	15.8	12.8	19.9	12.3	23.0	17.2	11.5
	13.6	18.9	8.9	21.0	9.8	19.4	19.8	10.1

   (a) Examine a dotplot of the data that shows the insect data plotted against each leaf. Do the mean counts look similar for each leaf? Is the spread of counts similar for each leaf? (You do not need to include the plot in your solution, but you may if it makes you happy.)

   (b) Give a suitable model for describing these data, identifying all terms in the model and identifying any distribution assumptions. Draw a nesting diagram for the model variables as in the notes, indicating which variables should be modeled as random effects and which as fixed.

   (c) Find a 95% CI for μ assuming a t distribution with 7 degrees of freedom.

   (d) Use the `mcmcsamp` function with a sample size of 10,000 to find a 95% credible region for μ. How does this region differ from that found in (c)? Is its width much larger, much smaller, or about the same?

(e) Fit a model for insects using leaf as a fixed effect using `lm`. Based on this model find a 95% CI for μ. How many degrees of freedom are used here? How does this interval compare with the intervals from parts (c) and (d)? Which interval or inervals are most appropriate?

(f) In a balanced experiment with k leaves and s disks sampled from each leaf, the expression for the variance of the intercept treating leaf as a random effect is as follows.

$$\mathsf{V}(\hat{\mu}) = \frac{\sigma_a^2}{k} + \frac{\sigma_e^2}{ks}$$

Verify that this expression is consistent with the summary in R using `lmer`.

(g) Suppose that 16 leaves had been selected and two disks per leaf had been taken. Assume that the same estimates for σ_e^2 and σ_a^2 are obtained as in the actual experiment (using notation from lecture). What would the estimate of the variance of $\hat{\mu}$ be in this case? Compare this with the estimate for the variance of $\hat{\mu}$ in the actual experiment. Interpret the comparison. Is it better to sample more leaves or more disks per leaf given a fixed total sample size?

3. An experiment was conducted to compare the amount of calcium in loaves of bread made with three different recipes. Two loaves were made from each of the three recipes. Measurements of calcium were made on small sections obtained from the loaf interior; three small sections were obtained for each loaf and measured for calcium concentration. The data were recorded as follows. Note that the "loaf number" was used to identify the two loaves for a given recipe. The three observations in each group are the calcium measurements on the small sections; the units are mg of calcium per gm of bread. The data is also in the file `bread.txt`.

	Recipe		
	A	B	C
Loaf 1	0.18	0.19	0.07
	0.15	0.16	0.10
	0.16	0.18	0.08
Loaf 2	0.14	0.23	0.09
	0.12	0.20	0.12
	0.14	0.20	0.10

(a) Write down a random effects model appropriate for this experiment. Identify the terms of the model. Draw a nesting diagram for the model variables as in the notes.

(b) Plot the data versus loaf and versus recipe. Summarize your observations. Is the model reasonable?

(c) Is there much difference in the recipes effect on calcium content in bread? Summarize your findings.