March 6 Lecture Notes

1. ESS F-test: What is the contribution of a subset of explanatory variables in explaining variability in the response variable? Compare the difference in residuals for the larger and smaller model normalized by the estimated model error for the larger model.

2. See handouts on damage to red spruce in App. Mtns. (week of 2/18) An example of comparing separate lines model to equal lines model. Hypotheses in terms of regression coefficients? In terms of the problem?

   - "For the red spruce example, location (N,S) and the interaction between location and elevation explain a significant proportion of the variability in percent of damaged trees given elevation (ESS F-test for $\beta_2 = \beta_3 = 0$, $F_{...,...}$, p-value = ...)"
   - "There is strong evidence that the regression model in which the linear relationship between percent of damaged trees and elevation depends on region represents a statistically significant improvement over the simple linear regression model in terms of explaining the variability in percentage of damaged forest (summary of ESS test here)."

3. Bat echolocation data

   - See Splus commands handout from week of 3/4 for the ANOVA tables for the 3 bat models. See also Display 10.12.

	Model SS, df	RSS, df
Sep. Lines	29.47,5	0.5,14
Parallel Lines	29.42,3	0.55,6

   - Test $H_o : \beta_4 = \beta_5 = 0$ vs. $H_a : \beta_4, \beta_5$ free to vary. That is, whether the linear relationship (slopes) between median energy and mass differs according to species.
   - ESS = 0.05/2 = 0.025; $\sigma^2_{full} = MSE_{full} = 0.036$
   - F=0.025/0.036=0.67 compare to F on 2, 14 df.
   - $P(F_{2,14} > 0.67 | H_0$: parallel lines model holds)=p-value
   - $F = \frac{\text{avg. amt. by which the unexpl. var decreases by adding beta4 and beta5 to parallel lines model}}{\text{avg amt of unexplained var. in full model}}$
   - Calculate the p-value and write in terms of the problem. See also page 285 for an explanation in terms of the problem. Note that it is still appropriate and informative to form a confidence interval here – see Sleuth.

4. Discussion of Pimm extinction data

   - In the parallel lines log-log model, $\hat{\beta}_2$ is the estimated amount by which the log of time to extinction given log of mating pairs for the large group exceeds that of the small group. In class we stated that the statement above is also valid for the logy~x model. *However*, for the two models the calculated median time to extinction for an individual species with pairs taking a value, say 10, will be different for the log-log and the logy~x models.
   - Refer back to Chapter 8 for interpretations of the slope in the different types of log models. Also, review p 273-4 of Sleuth in the case of $\beta_3 = 0$ analysis.

5. Topics yet to cover in Ch. 10: Polynomial regression; Adjusted R2 versus R2; dealing with functions of coefficients in situations that are too complicated to use the centering trick. We'll combine this with the Delta method discussion in Chapter 11.