

**Assignment #12 — Due Tuesday, December 14, 2010, by 4:00 P.M.**

Turn in homework to your TA's mailbox using this sheet as the cover page.

Fill in your name and also circle the *lecture section in which you are registered* and circle the *discussion section you expect to attend* to pick up this assignment.

**Name:**

<b>Lecture 1 (Larget).</b>	<b>311:</b> Tu 1:00 - 2:15pm	<b>312:</b> Th 8:00 - 9:15am	<b>313:</b> We 1:00 - 2:15pm
<b>Lecture 2 (Hanlon).</b>	<b>321:</b> Tu 1:00 - 2:15pm	<b>322:</b> We 2:30 - 3:45pm	<b>323:</b> We 1:00 - 2:15pm

Please answer the following questions.

**ANOVA problems:**

1. Use the context from p. 536, Problem 12, and the data in the file `beetle.txt` to answer these questions.

- (a) Make an interaction plot. Describe the pattern you see in words. Is there graphical evidence of an interaction?

**R hint:** The function `xyplot()` can be used to display a quantitative variable against a categorical variables and using a third categorical variable to color groups as follows.

```
> beetle = read.table("beetle.txt", header=T)
> library(lattice)
> xyplot(develop ~ species, data = beetle, groups=family, auto.key=list(space="right"), type="a")
```

- The `groups=family` argument will cause R to use a different color for each family, while plotting the points in a single panel.
  - The `type="a"` will cause R to draw straight lines connecting the averages (means) of each group instead of the points. You can see both points and averages with `type=c("p", "a")`.
  - The argument `auto.key=list(space="right")` causes R to add a key to the right of the plot. The alternative `auto.key=T` would put the default key on top of the graph. The alternative `auto.key=list(columns=3)` would put the key on top of the graph, but in three columns.
  - You can also produce an interaction plot by hand.
- (b) Use R to fit two separate models; a two-way additive ANOVA model (without interaction) and a two-way ANOVA with an interaction term.

```
fit.1 = lm(develop ~ species + family, data=beetle)
fit.2 = lm(develop ~ species*family, data=beetle)
```

Use `anova()` to summarize each fitted model. Answer these questions:

- For each ANOVA table, the hypotheses for each tested hypothesis test, and list the corresponding degrees of freedom for the  $F$  test statistic.
- Use the information about the sample sizes to explain why the values in the df columns are correct.
- Explain why the numerical values in the rows for `species` and `family` are the same in some columns, but different in other columns for the two ANOVA tables.

- iv. Each ANOVA table has an associated estimate for the standard deviation  $\sigma$  in the two fitted models. What are these values, and are they equal or different?
  - v. Summarize the statistical evidence about a possible interaction between family and species in this data.
- (c) Use `summary()` to summarize estimated parameters in each fitted model. Answer these questions:
- i. For each model, what does the intercept represent?
  - ii. For family #50, find the fitted value for the mean number of days to reach development using both fitted models.
  - iii. Both models have a parameter associated with `speciesI.purpurea` and a parameter associated with `family#50`. Explain in the context of the problem, what each of the parameters represents.
  - iv. Using each model, find a 95% confidence interval the parameter associated with `family#50`, and interpret these confidence intervals in the context of the problem.
2. p. 535, Problem 8 in the textbook.

**Regression problems:**

3. p. 503, Problem 13 in the textbook.
  4. p. 503, Problem 14 in the textbook.
  5. p. 505, Problem 19 in the textbook.
-