Second Midterm Exam

Instructions:

• The examination is due at 9:30 am at the beginning of lecture on Thursday, April 13.

• For your solution for each problem, you may write neatly or use a word processor. Begin the second solution on a new sheet of paper. Each separate solution should contain a maximum of five pages, including graphs and computer output. Many excellent solutions will use fewer than five pages. You will likely create many graphs in your analysis. Only include graphs in your solution that are important to illustrate points. If using a word processor, use standard margins and font. Do not decrease the margins or the font size. Rather work on a concise explanation of your reasoning and results. Do not exceed the page limit by appending pages of graphs and computer output.

• Begin each solution with a summary. This should be a short paragraph that summarizes the main results and refers to the statistical evidence in support of your conclusions. It should be a well written, self-contained description of your results, and should make sense to a reader who has read the problem. This summary paragraph should not include computer output, tables, or too many details of the analysis. You may think of it as a paragraph of a publication in a biology journal.

• Follow the summary paragraph with a concise description of the data analysis and your thinking and reasoning to justify the analysis. Any computer output should be incorporated into your discussion. Only include computer output that is directly relevant to support your main conclusions.

• You may use any books and notes that you desire, but the work that you turn in must be completely your own. In particular, you may not discuss the examination with anyone other than the instructor. Any questions you have can only be addressed to me. I will copy the question and my answer to all in the class. You may not post questions on the forum this week.

• Strive to provide a solution that addresses the main scientific questions of interest, and use the simplest and most straightforward statistical methods that are sufficient to this task. It is much more preferable to provide a simple sensible statistical analysis than to attempt to demonstrate the full range of your mastery of the most complex statistical methods.

Grading criteria:

• Presentation: Are the results neatly presented? Is the presentation well organized and in the page limit? Is supporting computer output provided in edited form?

• Statistical model: Is one “final model” presented? Is this model justified by important hypothesis tests? If there are two or more models that seem to fit equally well, you may describe them briefly but decide on one final model based on the experimental design and the goals of the experimenter. Remember there is no one right answer. Is the model appropriate for the design and the questions of interest? Does the model fit the data and are appropriate diagnostics (including plots) given? Are correct interpretations given for all the parameters in the model?

• Conclusions: Are they correctly drawn from the model and do they answer the question of interest? Are the conclusions illustrated with plots? Are the appropriate hypothesis tests supporting the conclusions described and documented?
1. (60 points) Plant defense theory predicts that short-lived plants should invest their resources for fast growth mostly, while long-lived, slower growing plants should invest more of their resources to fight against herbivores. To test this hypothesis, an experiment was run in a large field, where aphids were the most important plant herbivores and were naturally present. Plants of 14 closely related species were planted under different levels of soil fertilization and under either a high density or low density of aphid predators (these predators were mostly lady bugs, and were controlled with a savvy bagging system). 5 plants were planted for each species and each experimental condition. Assignment of all experimental treatments to positions within the field was random. The data are available in the file named `plantgrowth.txt`, with the following variables:

- **species**: identifier for the plant species.
- **predator**: whether aphid predators were reduced in number or not.
- **fertilizer**: amount of fertilizer added to the soil. Either 0 (none) or 5g of slow-release fertilizer.
- **plant**: plant replicate number, from 1 to 5 for each treatment and species.
- **biomass**: dry above-ground biomass of the plant at the end of the experiment, in grams.

Aphid numbers were also counted, but not presented here for this exam.

(a) Present plots to describe the variation in biomass, depending on the predictors that appear to be most important.

(b) Determine which experimental factors influence plant biomass, based on the evidence from these data and on an appropriate statistical analysis of these data.

(c) For species number 3 and in the absence of fertilizer addition, what is the effect of reducing aphid predators on plant biomass? Answer this question quantitatively and in an easily interpretable way. For the same species and in the presence of aphid predators, what is the effect of adding fertilizer on plant biomass, under conditions similar to those in the experiment?

(d) Based on plant defense theory, one would predict that the plants that most benefit from aphid predators are the fast-growing plants, or those plants that are using most of their resources for growth rather than for resisting herbivores. To test this prediction, calculate for each species two measures of plant growth strategy:

   i. plant biomass, averaged over all treatments,
   
   ii. responsiveness to soil fertility: estimated mean effect of fertilizer addition on plant biomass, averaged across all levels of predator exposition,
   
and calculate one measure of resistance to herbivores:

   iii. mean predator effect on plant biomass, averaged across all amounts of fertilizer addition. For aphid-resistant plants, aphid predator presence would have little effect. For aphid-sensitive plants however, aphid predator presence would have a beneficial effect on the plant.

Present your three measures either in a table or preferably in one or 2 graphs. For each measure of plant growth strategy (i and ii), determine if the data provide evidence that fast growing species benefit more from predators than slow growing species (as the theory predicts). Note: if you cannot calculate the 3 variables above as specified, then find another sensible way to calculate plant biomass, soil fertility effect and predator effect, and explain what your variables represent. The goal is to obtain one measure of each of these 3 variables for each species, and then to use them to test the pattern predicted by the theory.
2. (40 points) An experiment was conducted to better understand how castes are determined in ants: is caste determined genetically, or by the diet the ant larva receives? The data and questions in this problem are focused on one particular ant colony of one particular species. Newly emerged worker female ants and reproductive female ants were randomly selected from this colony. Their caste (worker vs. reproductive female) was determined based on their morphology. The genotype of each ant was determined. Based on the genotype data, it was determined which ants were full sisters (from the same queen and same male) and which ants were half-sisters (from the same queen but different males). The ants were then crushed and analyzed, in an effort to determine what diet there were given as larvae. The following variables were measured: percent Carbon (C), percent Nitrogen (N), C-to-N ratio, $\delta^{13}C$ and $\delta^{15}N$. A lower C-to-N ratio or a higher percent N, $\delta^{15}N$ or $\delta^{13}C$ indicate nutrition at a higher trophic level as when the larva is fed with insect preys rather than seeds. Therefore, reproductive females are expected to show lower C-to-N ratio, or higher percent N, $\delta^{15}N$ or $\delta^{13}C$ than workers. The file named antCastes.txt contains the data collected on these ants, with the following variables:

- father: identification number for the father shared by full sister ants
- percentN: percent Nitrogen (in %)
- CtoN: Carbon to Nitrogen ratio
- d15N: $\delta^{15}N$ (per mil)
- d13C: $\delta^{13}C$ (per mil)
- caste.isqueen: 0 is the ant is a worker, 1 if the ant is a reproductive female
- caste: “worker” or “reproductive”.

(a) Present plots to visualize how the proportion of reproductive females depends on nutrition variables or on genetic background (father-line). Choose to present the most informative relationships.

(b) Based on an appropriate analysis of these data, determine if caste depends on nutrition, or on the ant’s genetic background, or on both.

   i. If caste is determined by the larva’s diet, which nutrition variable(s) does caste depend on?
   
   ii. If caste depends on the ant’s genetic background, determines the father-line that appears to be most biased toward producing worker females and the father-line that appears to be most biased toward producing reproductive females, given equal nutrition variables. Quantify the difference between these two father lines quantitatively and in an easily interpretable way.