Stat/For/Hort 572 **April 22, 2004**

Take-Home Midterm Examination

- This exam is due at the beginning of lecture **11:00 am on Thursday April 29**. For any change in due date (to be granted only under compelling circumstances), arrangements must be made in advance.
- Your solutions must be written on only **one side** of a page. You may use pen or pencil or you may use a word-processor (or use a combination); however, your solutions must be neat and legible. You should concentrate on being thorough in your thinking but concise in the write-up. It is important to provide your reasoning for the work you present. All computer output included must be integrated into the discussion; unnecessary computer output should be avoided. For each problem, you should begin with a brief summary of your findings. **The maximum number of pages for each problem is 7, including computer output.**
- **Hand in the two problems separately. For each problem please use a cover page with your name. Do not put your name on the exam otherwise.** (This cover page does not count toward the maximum number of pages described above.)
- You may use your notes and any books you wish. **However, the work is to be entirely your own. You may not collaborate with others or discuss this exam with others.** All questions concerning the exam must be addressed to the instructor — extra office hours during the exam are listed below. You may not consult with the TAs or computer assistants. There will be no discussion sections or TA office hours during the week of the exam (April 26 – 30).

Office hours for exam-related questions:	Thursday, April 23	3:30pm – 5:00pm
	Monday, April 26	1:00pm – 2:00pm
	Tuesday, April 27	2:30pm – 3:30pm
	Wednesday, April 28	3:00pm – 4:00pm

- You should strive to provide answers that are scientifically meaningful. You should attempt to analyze each problem as if it were your research problem. It is far better to provide an analysis that is simple, sensible, and easily interpreted than to demonstrate your knowledge of "fancy" statistical techniques. Use the simplest analysis that answers the question.
- The total score for the exam is 100. Each problem has equal weight.
- *Technical Note:* The following may or may not be useful to you. In R, to transform a variable in, say `y` you can type things like: `ty = log(y)` or `ty = log10(y)` or `ty = sqrt(y)` or

```
ty = y^2
```

. In SAS, suppose the data are in the variable `y`. If we use the notation `ty` to represent a transformed version of `y`, then we could say: `ty = log(y);` or `ty = log10(y);` or `ty = sqrt(y);` or `ty = y**2;` You place any of these SAS statements after the `input` and `infile` statements, but before any `proc` statements.

Problem 1: Abundance of Fire Ants

A researcher was interested in understanding possible influence of various environmental factors and habitat types on the abundance of fire ants. Four habitat types were considered according to 2 soil types (clay versus sandy soil) and whether the habitat was disburbed or not. The habitat types are coded as:

1 = clay undisburbed;

2 = clay disturbed;

3 = sandy undisburbed;

4 = sandy disturbed.

For each habitat type, 5 equal-sized plots were selected and hence there were a total of 20 plots. Within each plot, the number of fire ant mounds was counted and recorded. Also recorded for each plot were two environmental factor variables, namely, soil moisture levels and soil temperature.

The data appear on both the CALSHP and on the course web site in the file `abund.dat` . The 7 columns correspond to the variables:

`abund`: the number of fire ant mounds in a plot measuring abundance;

`moist`: soil moisture (in %);

`temp`: soil temperature (in ^{o}C);

`type`: habitat type (1-4);

`w1, w2, w3`: alternative coding of the habitat types using 0-1 variables.

You are to analyze the data with an ultimate goal of understanding how the fire ant abundance is related to the environmental factors and habitat types. You should address the following specific issues:

(a) Ignoring the environmental factor variables (i.e., soil moisture and soil temperature), what can you say about the relationship between the abundance and the habitat types?

(b) Ignoring the habitat types, what can you say about the relationship between the abundance and the environmental factor variables (i.e., soil moisture and soil temperature)?

(c) Based on all the data, how does the abundance relate to the habitat types and the environmental factor variables?

Problem 2: Nitrogen Sources and Levels

A field experiment was conducted to evaluate the effect of nitrogen on plant yield. Of particular interests were two nitrogen sources, namely fresh poultry and composted cow manure, and three levels of nitrogen levels, namely 67, 134, and 202 kg/ha. The experiment was performed at two sites. At each site, two large fields were selected. Hence there were a total of 4 fields. Each field was divided into 7 plots, where each plot was randomly assigned to one of the following 7 treatments:

1 = no nitrogen was applied (control);

2 = fresh poultry at 67 kg/ha;

3 = fresh poultry at 134 kg/ha;

4 = fresh poultry at 202 kg/ha;

5 = composted cow at 67 kg/ha;

6 = composted cow at 134 kg/ha;

7 = composted cow at 202 kg/ha;

The data consisted of plant yield in mg/ha in each plot and are tabulated as follows.

		Treatments						
Site	Field	1	2	3	4	5	6	7
1	1	10.75	18.66	21.15	20.76	13.97	19.84	17.24
1	2	12.37	16.36	19.39	17.57	15.89	19.56	15.89
2	3	6.33	14.97	19.02	20.25	11.64	14.15	11.69
2	4	11.19	15.48	16.78	17.91	11.19	10.02	9.06

Analyze these data with the goal of understanding the effect of nitrogen sources and nitrogen levels on plant yield and address the following issues.

(a) Assess an overall effect of nitrogen application compared to control (i.e., no nitrogen application).

(b) Which combination(s) of nitrogen source and nitrogen level would give the most plant yield?

(c) Focus on the fresh poultry manure, is the relationship between plant yield and nitrogen levels linear, quadratic, or cubic?

(d) Construct a 95% confidence interval for the among-plot variability.

(e) Evaluate the effect due to the differences among the fields at the two sites.