Due: Monday, March 13, by 5 pm.
You can turn in the exam in class, during office hours, to my mailbox in 3010 Agronomy (available until 4pm), or under my office door (3405 Agronomy).

There are three questions, each with multiple parts.

I grade anonymously. Please do not put your name on the front of your answers. Instead, put your name on the back of the last page of your written/typed answers and also the last page of your SAS output (if any). Make sure all pages are securely attached.

Reminders: You are not allowed to ask classmates or friends for help. Please ask me, not a classmate, for help with SAS (or other computer program), even if you are working next to that classmate in the same computer room. Please sign the appropriate line on the last page when you are finished working on the exam.

You are allowed to use your book(s) and your notes. You are always welcome to e-mail me with questions or requests for clarification. I will not provide SAS code, but I am very happy to answer questions about syntax, output, or errors. If you do not understand the context for each problem, ask for clarification.

Some questions should be easy, some are harder, and a few require you to integrate ideas, which many folks find hard. You don’t have to get a perfect score to do well on the exam. I adjust my letter grades to account for the difficulty of the exams. Your answers may be written or typed. You may submit just your answers. But, it is easier to determine partial credit if you include your SAS code and output.

Arrange your answers in the following order:

- Answers and SAS code to question 1, ditto to question 2, ditto to question 3,
- your name on the back of the last page of question 3 answers,
- signature page
- SAS output for Q 1, ditto for Q 2, ditto for Q 3 (if included)
- your name again on the back of the last page of SAS output

You DO NOT need to copy your answers. All I’m asking is that you put the SAS output after the main part of your answers.

Reminder: e-mail, call (4-6828 or 4-2142), or come to office hours for help or clarification. Don’t spend more than 30 minutes trying to fix a SAS issue. If you haven’t found the problem in 30 minutes, get help from me.

Office hours during the exam:
Weds, 4/8: 2-4 pm, 3405 Agronomy
Thursday, 4/9: 4-5 pm, 3405 Agronomy
Friday, 4/10, 9-10, 1436 Wilson
Friday, 4/10, 4-5, 3405 Agronomy
over the weekend: e-mail to pdixon@iastate.edu
Monday, 4/13, 12-1, 3405 Agronomy
1. 40 pts. The following problem is based on an study of two species of tupelo, \textit{Nyssa}. Both species are swamp trees native to the Southeastern part of the US. The study considered the effect of 3 factors on the height growth of 1 year-old seedlings: LIGHT (3 levels: 35\%, 60\% or 100\% of ambient light), WATER depth (3 levels: 0=surface, 1=5cm deep, 2=10cm deep), and SPECIES (2 levels: A=\textit{N. aquatica}, S=\textit{N. sylvatica}). All 18 combinations of the three factors were used as treatments in this study. The tupelo seedlings were grown individually in large pots placed in wading pools filled with water. Water depth was manipulated individually for each pot using bricks under some pots to raise them up (and so decrease the water depth). There were 6 pots in each pool. Hence, all combinations of species and water levels were randomly assigned to pots within each pool. Light levels were manipulated with shade cloth placed over the entire pool. Hence, light was randomly assigned to the pool. Pools were arranged in blocks of 3 pools, with each light level present in each block. There were a total of 5 blocks. The response is the height growth in cm.

The data are in \textit{nyssa.txt}. The variables are the block number, the pool number, the light level, the species code, the water depth code, and the height growth.

(a) 2 pts. What is (or are) the experimental unit (or units, if more than one) in this study?
(b) 3 pts. Name the treatment design and the experimental design(s) used in this study.
(c) 10 pts. Write out an appropriate skeleton ANOVA table for this study. This ANOVA table should indicate sources of variability, degrees of freedom and whether an effect should be considered fixed or random.
(d) 5 pts. One of the major questions in the study is whether the two species have the same patterns of response to the environmental variables (light and water depth either separately or together). Analyze the data and answer the investigator's question. Include an appropriate test statistic (or test statistics) and p-value(s) with your answers.
(e) 5 pts. The investigators are also interested in the effect of light level, averaged over all species and all water depths. Test the null hypothesis of no differences between light levels. Report an appropriate test statistic and p-value.
(f) 5 pts. Estimate the difference in mean height growth between 100\% light - the mean height growth for 35\% light, averaged over species and water levels. Report the difference and its standard error.
(g) 5 pts. Estimate the difference in mean height growth between 0 cm water depth and 10 cm water depth, averaged over species and light levels. Report the difference and its standard error.
(h) 5 pts. Are the s.e.'s in part 1f and 1g the same? Explain why they should be (or should not be) the same.
2. 28 pts. The following data come from an experiment on the effectiveness of a fungicide to control potato scab virus. This fungicide has never before been tried on this disease. It is not clear that it has any effect. One of 4 AMOUNTs of fungicide (0, 300, 600 or 1200 lbs/acre) was applied either in the fall or spring SEASON. If it helps you interpret the results, the fall application was after the field was harvested; the fungicide was in the soil through the winter. The spring application was just before the field was planted with potatoes.

A field was divided into 32 plots, arranged in 4 BLOCKs, each with 8 plots. Each plot was randomly assigned to one of the 8 combinations of fungicide amount and time of application in a randomized complete block design. The response is the amount of potato scab DISEASE in that plot.

The data are in scab.txt on the class web site. This file includes BLOCK, two treatment codes, AMOUNT, SEASON, and the DISEASE response. trtcode1 numbers the treatments from 1 to 8. trtcode2 numbers the treatments from 1 to 7; 0/Fall and 0/Spring are both numbered 1.

The researchers want to know:
A) is there a difference between season of application (fall and spring) when the fungicide is applied. How large is that difference? How precisely is it estimated?
B) whether the amount of fungicide added has any effect on the amount of disease
C) especially, does adding fungicide (in any amount, in any season) reduce scab? How large is the difference between the control (0 fungicide) and the average of all non-zero treatments?

(a) 8 pts. Construct an appropriate skeleton ANOVA table for this study, indicating sources of variation, d.f., and whether an effect should be considered fixed or random.

(b) 5 pts. Is there a difference between season of application (fall and spring) when the fungicide is applied. How large is that difference? How precisely is it estimated?

(c) 5 pts. The investigators want to know whether the amount of fungicide (0, 300, 600, or 1200 lbs/ac) has any effect on the amount of disease. Is it more appropriate to answer this question separately for each application time (fall and spring) or to average over application times? Explain your choice.

(d) 5 pts. Answer the investigator’s question in part 2c. Report your test statistic(s) and p-value(s).

(e) 5 pts. Does adding fungicide (in any amount, in any season) reduce scab? How large is the difference between the control (0 fungicide) and the average of all non-zero fungicide treatments? How precisely is that difference estimated?
3. 32 pts. The following problem is motivated by a recent microarray experiment. Fungal infections of crop plants are a common problem in the midwest where the summers are warm and humid. There is often some genetic control of which fungi infect which crop plants. This is a study of 2 genetic isolates of one fungus species (A and B) and 3 genetic isolates of barley (1, 2, and 3). Barley genotypes 1 and 3 are suspected to be sensitive to fungus B and resistant to fungus A while Barley genotype 2 is sensitive to fungus A and resistant to fungus B.

This experiment considered all 6 combinations of fungus genotype and barley genotype. A flat (a large plastic tray) was randomly assigned to one of the six treatments. The appropriate barley genotype was planted in each flat then innoculated with the appropriate fungus. These trays were then grown in a growth chamber. The response is the expression level of a particular gene thought to be involved in resistance. There is one response for each flat of plants (The real study was much more complicated than this, but one response per flat will simplify the problem). Because fungi spread easily, you can’t have two different fungal genotypes in the same growth chamber. Here we consider four possible designs and their analysis. Each treatment is replicated 3 times in each design, but the characteristics of a replicate are not the same in each of the designs. Each design has a total of 18 observations.

For each of the following ways of conducting the experiment, write out the skeleton ANOVA (source of variation and d.f.) and indicate the appropriate error term for the F test of each effect (barley, fungus and the interaction).

(a) 8 pts. If you had 18 growth chambers, you could randomly assign each of the six treatments to growth chambers (3 chambers per treatment).

(b) 8 pts. If you had 6 growth chambers at one time, you could randomly assign each of the six treatments to a growth chamber (1 chamber per treatment). There is one flat per growth chamber. After you collect the data, you repeat the study with a new randomization of the 6 treatments to 6 growth chambers. You repeat once again, for a total of 3 repetitions. Because growth chambers are used by many different people, each repetition of the study is in a different set of 6 growth chambers.

(c) 8 pts. Imagine you only had access to 2 growth chambers at a time. Three flats (one of each barley genotype) are grown in chamber 1 and three more are grown in chamber 2. The entire study is repeated a total of 3 times, as in the previous design. As before, different growth chambers are used each repetition.

(d) 8 pts. One disadvantage of the design in part 3c is the time it requires. Your professor proposes that you make 18 flats of plants (6 of each barley genotype). One growth chamber is randomly assigned to fungus A; the other chamber is assigned to fungus B. 9 flats (3 of each barley genotype) are placed in growth chamber 1 (Fungus A); the other 9 flats are placed in growth chamber 2.

That’s all!
Please check the appropriate box, sign, and include at the end of your answers.

_____ I completed this exam without assistance from friends or classmates.

_____ I received the following assistance from friends or classmates (please describe below).

Signed: ____________________