www.stat.duke.edu/courses/Fall02/sta240/quiz4/quiz4data.html

Quiz 4: Lab Exercise, 11/11/02	Name:
I will follow the NSEES Honor Code.	Signature:

1. [3 points] Circle the terms that describe the meadowfoam study:

- (a) completely randomized design
- (b) a factorial treatment arrangement
- (c) an incomplete design
- (d) an unbalanced design
- (e) a study that allows conclusions about causality to be made.
- 2. **[5 points]** Use the meadowfoam data to fit the fully saturated model with flowers as the response. (You can use the formula "flowers~time*intensity" or "flowers~time+intensity+time:intensity") Evaluate whether the interaction term is significant by giving the hypotheses in terms of the problem, reporting the test statistic, the distribution of the test statistic, the p-value and conclusion in terms of the problem.

- 3. [5 points] Consider an experiment with 2 factors, A and B, and a response, Y. Circle all that apply:
 - (a) The interaction term is included when the change in the true average response Y when factor A changes is the same for each level of B.
 - (b) The interaction term is used when the main effects provide an incomplete description of the data.
 - (c) The inclusion of the interaction term will change our estimate of mean squared error (and s_{pooled}).
 - (d) The interaction term is used when the effect of A on the true average response depends on what level of factor B is considered.
 - (e) If we create an interaction plot (for example, in meadowfoam data below) and the lines are not exactly parallel, we can conclude that there is sufficient statistical evidence for an interaction effect.



4. **[3 points]** Conceptual exercise 3(b) of the *Sleuth* states that the assumption of normality is assisted by considering averages of groups of 10 plants in the analysis. However the use of averages as the responses in an analysis may have an impact on the estimate of variability in plant responses.

True or False. *Circle one.* The effect of using averages as responses in the meadowfoam study could decrease variability in the response (number of flowers), thereby decreasing the width of associated confidence intervals for the study and increasing Type I error.

- 5. Fit the model "flowers~intensity+timing". Under "Results", check the box for "means". Under "Plot", click the button for "Residuals vs. Fit". Copy the ANOVA table to a Word file so you can save it for future questions. We will refer to this model as "Model A".
 - (a) **[5 points]** Is there a light intensity effect? Give hypotheses in terms of the problem, test statistic, distribution of test statistic, and p-value. Give a 1-sentence conclusion in terms of the problem.

(b) **[5 points]** Give the assumptions of the 2-way ANOVA that can be verified by the Residuals vs. Fit plot.

(c) [10 points] Give a 95% confidence interval for the difference in the number of flowers at 600 and at 300 μ mol/m³ / sec based on the additive model.

- (d) [2 points] True or False. *Circle one*. The p-value you gave in 5(a) can be viewed as inaccurate because it does not reduce variation by explicitly accounting for the effect of timing.
- 6. Now we will perform a one-way analysis of variance on the same dataset. Go to the quiz web page to find the re-formatted data you will need. We will consider each combination of timing and intensity as a separate group. Thus we have 16 groups in our ANOVA, with 2 measurements per group.

Fit the ANOVA model, making sure to check "means" under "Results" and "Residuals Normal QQ" under "Plot". Copy the ANOVA table to a Word file so you can save it for future questions. We will refer to this model as "Model B".

(a) **[10 points]** Use the 1-way ANOVA to calculate a 95% confidence interval for the difference in the number of flowers at 600 and at 300 μ mol/m³ / sec. Clearly write out the contrasts you use to construct the interval. Is the confidence interval here different from the one in 5(c)? Why or why not?

- (b) **[2 points]** Use the QQ plot of the residuals to compare the shape of the distribution of the residuals to that of the normal distribution. *Select one.*
 - (i) The distribution of the residuals has longer tails than the normal distribution.
 - (ii) The distribution of the residuals has shorter tails than the normal distribution.
 - (iii) The distribution of the residuals is skewed right.
 - (iv) The distribution of the residuals is skewed left.

PAGE 1:_____ PAGE 2:_____ PAGE 3:_____

TOTAL:_____