STT 814 Advanced Statistics for Biologists
Sections 1&2
Spring 2011

Instructor: Sasha Kravchenko
Email kravche1@msu.edu
Phone 355-0271 ext.1241
Office A376 Plant & Soil Sci Bldg
Office Hours: Wednesday 10 am – noon

TA: TBA
Office Hours: TBA

Course webpage:
https://courses.css.msu.edu/
Username: your MSU Net ID
Password: the last 5-digits of your PID

Class meets:
Lecture Sections 1&2: MW 12:40-2:00 PM at 149 Plant and Soil Science Bldg.
Lab Section 1: Tuesday 12:40-2:30 PM at 152A Plant and Soil Science Bldg.
Lab Section 2: Tuesday 3:00 - 4:50 PM at 152A Plant and Soil Science Bldg.

Required:
3) Course notes on the web

Recommended additional reading:
Regression

Design of experiments:
**Grading scale:**
90% of the total points for the semester 4.0  
80% of the total points for the semester 3.5  
70% of the total points for the semester 3.0  
60% of the total points for the semester 2.5  
50% of the total points for the semester 2.0  
40% of the total points for the semester 1.5  
Less than 40% of the total points for the semester 1.0  

**Percentage distribution:**
Lab presentation/project 20%  
Homework assignments, paper critiques, and lab reports 20%  
   *All assignments - one day late 10 % off, not accepted after that*

Tests:
Midterm  
First: *February 2*\textsuperscript{nd} Two-sided page of *handwritten* notes 20%  
Second: *~ March 21*\textsuperscript{st} Two-sided page of *handwritten* notes 20%  
Final  
Monday *May 2*\textsuperscript{nd} 12:45-2:45 p.m.  
   *Two-sided page of *handwritten* notes* 20%  

**Course outline:**
Review of the simple linear regression: model, parameter estimation, hypothesis testing, prediction.

Multiple regression: introduction, parameter estimation and hypothesis testing, coefficient interpretations, predictions.

Multiple regression: checking model assumptions, graphical diagnostic, influence, leverage, outliers. Variable transformation. Weighted least squares and dealing with correlated errors.

Multiple regression: Multicollinearity and variable selection.


The data and the model: validity of the assumptions. Residuals are the basis of diagnostic tools. Statistical tests for homogeneous variances. Looking for outliers. Data transformations to stabilize variances.


Complete block designs. Randomized complete block designs with one blocking criterion. Latin square designs with two blocking criteria. Missing data in blocked designs.


Split-plot designs. Plots of different size in the same experiment. Analysis of split-plot designs. Relative efficiency of subplot and whole-plot comparisons. The split-block designs.


Analysis of covariance. Local control with a measured covariate. Analysis of covariance for completely randomized designs. Analysis of covariance for blocked experiment designs.
Lab presentation/project info:

1) Sign up for a 20-minute initial discussion of your experiment during the first week of classes. The available times are on sign-up sheets. Please consider bringing visual materials (diagrams, schemes, plot outlines, graphs etc.) that will facilitate description of your study.

2) Based on the initial discussion, your experiment will be placed in one of the class topics. You will be given a tentative date for your lab presentation. Final date may vary but will be within 1-2 weeks of the tentative date. The final date will be set at least 2 weeks prior to your presentation.

3) The presentation assignment will consist of the following components:

   a. Power Point (overhead transparencies are OK) presentation of your research topic and experiment design delivered to the class during the lab.
   The presentation should be no more than 15-20 minutes in length and should include the following components:
      i. Introduction
      ii. Research hypothesis and objectives
      iii. Description of the treatment design
      iv. Description of the experimental design
      v. Your assessment of why your experiment fits a particular class topic

   The presentation does not need to include the data analyses – we will discuss the data analysis in class after your presentation.

   The draft of your presentation should be prepared one week prior to the presentation date. During the lab (Tuesday) preceding the lab when your presentation is scheduled you need to informally rehearse the presentation for me and provide me with your data. If draft is presented late the same week (Wednesday-Friday) – 10% off. No presentation will be allowed without rehearsal of the draft.

   b. Data analysis report submitted prior to the presentation.
   The report should include the following components:
      i. A brief introduction of the study topic
      ii. Research hypotheses and objectives
      iii. Description of the treatment and experimental design
      iv. Detailed description of the statistical methods that you have used for the data analysis
      v. Presentation of the key results, including tables, graphs, and discussion of the findings and conclusions. The tables and figures must be formatted consistent with format of the main journals of your field.
      vi. SAS code used for data analyses

   Parts i-iii should not exceed 300 words. Parts i-v should not exceed 2 two-sided pages (either single or double spaced).

   No page limit on SAS code. However, for full credit, it must be the final code used to get the results presented in the report; it must be well annotated with comments briefly identifying the purpose and key outcome of each portion of the code.

   The data analysis report should be a result of your independent work using relevant class and reference material and only general recommendations from me or the TAs.

   The data analysis report is due one day prior to the presentation (Monday in class). 10% off if late. No report will be accepted after the lab presentation is delivered.