Quantitative Methods in Plant Biology

PBIO-415 (#05275) and PBIO-515 (#05281)
Winter 2006

Instructor: Brian C. McCarthy
Office: Porter Hall, Rm. 416
Telephone: 740-593-1615
Email: mccarthy@ohio.edu
WWW: http://www.plantbio.ohiou.edu/epb/instruct/quantmet/quantmet.htm
Office Hours: By appointment (email or call)

TA: Rochelle Jacques (rj262604); Office (Porter Hall 416A)

Scheduling:
Lecture: Tuesday & Thursday 9:10-11:00 (Porter 417)
Laboratory: Tuesday & Thursday 11:10-12:00 (Porter 417/407)

Required Text: Zar, J. H. 1999. Biostatistical Analysis, 4th ed. Prentice Hall, Upper Saddle River, NJ. 663 p. + appendices.

Optional Text: Sokal, R.R. and F.J. Rohlf. 1995. Biometry: The Principles and Practice of Statistics in Biological Research, 3rd ed. W.H. Freeman, San Francisco, CA. (Strongly Recommended for PBIO-515.)

Grading:

PBIO-415		**PBIO-515**	
Exam-I	(20%)	Exam-I	(20%)
Exam-II	(25%)	Exam-II	(20%)
Final Exam	(25%)	Final Exam	(20%)
Problem Sets (4)	(30%)	Problem Sets (4)	(30%)
		SAS Worksets	(10%)

Statement on Grading: At least half of quantitative exam material will come *directly* from the textbook. (Note that I reserve the right to make minor alterations to the data.) The rationale for this is to: (1) encourage you to work through the problems in the text, (2) confirm that you have reached a correct solution, and (3) to limit the scope of material you are likely to encounter.

Most exam material will be quantitative in nature (problem solving) and will require a good quality calculator [sensu stricto]. (Note that calculators may be programmable, but all PDAs and other electronic devises are *not* permitted.) Problem sets and Work sets will be graded on content and correctness (see "Problem Sets" and "Work Sets" below).

Tutorials: The TA will minimally provide two evening tutorials on software used in the laboratory. This will include introductions to: Excel (only needed if you have no or little experience with spreadsheets) and SAS (required for grad students). The instructor will review other software applications in class, as appropriate.

Academic: *Attendance*, while not formally taken, is expected for all lectures and laboratories. The *Code of Student Conduct* should be strictly adhered to. Academic misconduct (cheating, plagiarism, etc.) will result in a grade of "F" for the course with subsequent referral to the Office of University Judiciaries. The university policies regarding the student code of conduct and academic integrity are clearly and completely summarized at: http://www.cats.ohiou.edu/judiciaries/index.htm. If you are unclear as to what these policies are, please consult the website.

Goals of Course: The goal of this course is to introduce upper-level students in the biological sciences to the quantitative skills and technological tools necessary to evaluate the literature and be able to carry out original research in the discipline. *Lectures* will focus on the quantitative skills associated with proper experimental design, statistical analysis, data interpretation, and presentation of results. Numerous worked examples and problem sets will be used to demonstrate and enhance your understanding of important biostatistical procedures. *Laboratories* will emphasize micro-computer technology and software applications likely to be encountered in the biological sciences. The goal is to provide an overview of common software applications and technological solutions to standard research problems in biology.

Course Progress: Please note that the development of this course is deceptively slow at first. The purpose of this is to provide a strong grounding and review of the basics (introductory level basic statistics). Much of this material is essential to latter parts of the course. By virtue of the subject material, the knowledge is cumulative (although exams are not, per se). The second half of the course will have a greater momentum with increasingly technical material (intermediate to advanced biostatistics). *Be careful not to fall behind*.

Feedback: An anonymous feedback form has been set up on the university server to provide feedback or input to the instructor or TA. We welcome any comments about the course as it proceeds (e.g., "slow lectures down", "provide more examples", "assign more homework", "we could use a lab tutorial on blank", etc.). Point your browser to: http://teach.citl.ohiou.edu/pages/anonymousFeedback/PBIO%20415-515mccarthyA01.html. This system is 100% anonymous and provides an avenue for you to provide input or feedback to me without providing your identity (if you feel uncomfortable doing so).

Problem Sets: There will be four (4) problem sets during the quarter. Typically, you will be provided with a data set and asked to implement an analysis, test for assumptions, summarize data, and provide an interpretation. Each assignment should be typed, double-spaced, and concise (< 4 pages + computer output). Assignments are due in class on the day specified. Late papers will be assessed a penalty of 10% loss per day up to 5 days. You must hand in all 4 problem sets in order to receive a final grade for the course.

Each paper should include the following:

1. Cover page with assignment title, your name, etc.
2. Body with 4 component sections (ca. 2-3 pages):
    a. *Problem Overview* (background, statement of hypothesis, experimental design, number and type of variables, appropriate method of analysis, statement of assumptions)
    b. *Results* (verification of assumptions, analysis, results summarized in tables and/or figures, etc.)
    c. *Discussion* (interpretation & analysis of results)
3. Appendix (computer output)

PBIO-515:

Graduate students are expected to complete additional assignments in the form of weekly SAS computer Work Sets. The purpose of these assignments is to give you exposure to the SAS programming language which will ultimately permit you to perform complex statistical analysis on research-grade data sets. Typically, this will involve you solving a simple textbook problem by writing a SAS program. Both the source code and the output should be handed in each week. There will be six of these assignments.

The TA will be your primary contact in assisting you with these Work Sets. Because of the cost, we have only several designated SAS workstations in the Porter Hall computer lab. Please write your code elsewhere and use this workstation only to run the program (i.e., be considerate of others). Other workstations may be available in indivudual research or computer laboratories.

There are many fine resources available for using SAS (it is widely regarded as the "industry standard"), in part because it has a steep learning curve. Printed manuals provided by SAS are the most comprehensive (see www.sas.com). Because our university has a site license, we also have the complete documentation for SAS v. 9.0 online: http://v9doc.sas.com/sasdoc/. Third-party manuals also provide excellent [and thankfully laconic] introductions. I *strongly* recommend Cody and Smith (1997; Applied Statistics and the SAS Programming Language, 4th ed., Prentice Hall; available at www.amazon.com (and elsewhere online).

Optional:

For those wishing additional information and exposure to recent problems in biostatistics, I have posted a reading list (with direct library links to online material) composed of 17 carefully selected papers from the primary literature. The core of these papers represent some of the most important and contentious statistical issues in modern experimental ecology, evolution, and organismal biology. The goal of these readings is to make you aware of current problems regularly encountered in research and possible solutions.

I have no expectations regarding this material (i.e., you will not be tested over it), but I would hope that grad students in particular would take the time to read selected papers that may pertain to their research area. The readings are posted at: http://www.plantbio.ohiou.edu/epb/instruct/quantmet/gradread.htm.

Lecture Topics & Reading Assignments

Wk	Date	Lec	Topic	Read (Zar)
1	1/3	1	Introduction, types of data, frequency distributions, measures of central tendency	1,2
	1/5	2	Measures of dispersion, Normal distribution, departures from normality	3,4
2	1/10	3	Exploratory data analysis: box plots, stem-and-leaf plots, re-expression, outliers	4
	1/12	4	Inference, sampling, statistics & parameters, estimation, confidence	6
3	1/17	5	Hypothesis testing, 1&2-tailed tests, types of error, significance, z-test, t-distribution (independent samples)	7
	1/19	***	**Exam-I** (descriptive statistics & hypothesis testing)	
4	1/24	6	T-tests: independence, normality, equality of variance	8
	1/26	7	Mann-Whitney U-test, Kolmogorov-Smirnov 2-sample test	8
5	1/31	8	2-sample (non-independent) tests: paired t-test, Wilcoxon signed rank test	9
	2/2	9	Enumeration data: χ^2 distribution, goodness-of-fit, contingency tables, median test	22,23
6	2/7	10	Multi-sample tests, Intro. to analysis of variance (ANOVA), F-statistic, One-way ANOVA, Kruskal-Wallis test	10
	2/9	11	Estimation of variance, ANOVA assumptions, tests for homogeneity of variance, normality, transformations	13
7	2/14	12	Planned and unplanned comparison tests of means	11
	2/16	***	**Exam-II** (2-sample tests, chi-square, ANOVA)	
8	2/21	13	Regression analysis: liner regression, residuals analysis, regression diagnostics, prediction & inference	17,18
	2/23	14	Correlation analysis, parametric & nonparametric, inferences	19
9	2/28	15	The General Linear Model, Analysis of Covariance (ANCOVA)	
	3/2	16	Completely randomized factorial designs	12
10	3/7	17	Random- and mixed-effects ANOVA, ANOVA models with block effects	14,15
	3/9	18	Repeated measures designs	14
	3/14	***	**Final Exam** (10:10, Porter-417)	