Quantitative Methods in Plant Biology

PBIO-415 (#05422) and PBIO-515 (#05428)
Winter 2007

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: Jyh-min (James) Chiang (jc293302@ohio.edu); Office (Porter Hall 419)

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

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	(25%)	Exam-I	(25%)
Exam-II	(25%)	Exam-II	(25%)
Final Exam	(30%)	Final Exam	(25%)
Problem Sets*	(20%)	R Problem Sets	(25%)

Statement on Grading:

EXAMS: At least half of quantitative exam material will come *directly* from the textbook. The rationale for this is to: (1) encourage you to work through the problems in the text (both exercises and examples), (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 (i.e., with stats functions). (Note that I reserve the right to make minor alterations to the data.) Exams for 415 and 515 will be identical and taken at the same time via pencil, paper, & calculator.

PROBLEM SETS (415): There will be ca. 10 problem sets (1-2 per week) worth 20% of your grade. These will involve 2-3 problems from your textbook. Problem sets must be solved using using computer software (e.g., Excel, NCSS, etc.) and not worked by pencil and calculator. Each assignment is worth 2 points; the grading rubric is simple: 0 = assignment not handed in by due date (NOTE: late assignments are not accepted without prior permission), 0.5 = less than 50% correct, 1 = less than 75% correct, 2 = greater than or equal to 75% correct.

PROBLEM SETS (515): Problem sets for 515 are expected to be done using the R programming language. These assignments will involve a detailed tutorial at the beginning (to teach you the necessary skills) and then an assigned problem. You must hand in your input, output, and any associated graphics. This is best done by copying all materials into a single DOC file. There will be 6-7 of these assignments, all of equal value, prorated to 25% of your grade. R webpage for class: http://www.plantbio.ohiou.edu/epb/instruct/quantmet/Rindex.htm

***SPECIAL NOTE (FOR 415):** Any student in 415 who wishes to learn R is encouraged to do so. This would be especially helpful to those thinking of graduate school in the biological sciences. As an incentive to do so, you will receive 25% possible points (as per 515), thus making your total possible score 105% (this has the potential to help you by a half letter grade at end of course). However, recognize up front that this option will require an additional commitment of time/work on your part.

Academic: *Attendance*, while not formally taken, is expected for all lectures and laboratories. Based on historical data, your final grade is highly correlated to your attendance rate. 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*! Ask for help if you need it.

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).

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	1/4	1	Introduction, types of data, frequency distributions, measures of central tendency	1,2
2	1/9	2	Measures of dispersion, Normal distribution, departures from normality	3,4
	1/11	3	Exploratory data analysis: box plots, stem-and-leaf plots, re-expression, outliers	4
3	1/16	4	Inference, sampling, statistics & parameters, estimation, confidence	6
	1/18	5	Hypothesis testing, 1&2-tailed tests, types of error, significance, z-test, t-distribution (independent samples)	7
4	1/23	***	**Exam-I** (descriptive statistics & hypothesis testing)	
	1/25	6	T-tests: independence, normality, equality of variance	8
5	1/30	7	Mann-Whitney U-test, Kolmogorov-Smirnov 2-sample test	8
	2/1	8	2-sample (non-independent) tests: paired t-test, Wilcoxon signed rank test	9
6	2/6	9	Enumeration data: χ^2 distribution, goodness-of-fit, contingency tables, median test	22,23
	2/8	10	Multi-sample tests, Intro. to analysis of variance (ANOVA), F-statistic, One-way ANOVA, Kruskal-Wallis test	10
7	2/13	11	Estimation of variance, ANOVA assumptions, tests for homogeneity of variance, normality, transformations	13
	2/15	12	Planned and unplanned comparison tests of means	11
8	2/20	***	**Exam-II** (2-sample tests, chi-square, ANOVA)	
	2/22	13	Regression analysis: liner regression, residuals analysis, regression diagnostics, prediction & inference	17,18
9	2/27	14	Correlation analysis, parametric & nonparametric, inferences	19
	3/1	15	The General Linear Model, Analysis of Covariance (ANCOVA)	
10	3/6	16	Completely randomized factorial designs	12
	3/8	17	Random- and mixed-effects ANOVA, ANOVA models with block effects, repeated measures	14,15
	3/16	***	**Final Exam** (08:00, Porter-417)	