Biol 206/306 – Advanced Biostatistics
Course Syllabus
Fall 2015

Instructors

Instructor: Dr. Philip J. Bergmann
Office: Lasry 331
Phone: 508-793-7553
E-mail: pbergmann@clarku.edu
Office hours: Tuesday 10:15-11:00 am

Lecture and Lab

Lecture:	Tuesday & Thursday	Time: 9:00-10:15 am	Place: Lasry 124
Laboratory:	Thursday	Time: 1:15-4:10 pm	Place: Lasry 124

Textbooks

None; readings will be from the primary literature (journal articles and book chapters). Students should use their own laptops for lab.

Prerequisites

- Biol106 – Introduction to Biostatistics, or Permission

LEEP Learning Objectives

During Advanced Biostatistics, students will build on the basic statistical skills they gained in introductory statistics courses and learn to apply more advanced techniques to real data. In the context of the five LEEP Learning Outcomes, students will:
1. Knowledge of the Natural World and Human Culture. Learn a statistical approach to thinking about and understanding the natural world. Statistically-informed thought alters how one thinks about what they observe in nature and society.
2. Intellectual and Practical Skills. Intellectually, students will learn to approach problems in the world around them and learn how to read, critically evaluate, and discuss primary biostatistical literature. At the end of the course, they should have the ability to learn new techniques on their own. Practically, students will learn a range of techniques that will allow them to analyze real quantitative data. Students will also learn how to design rigorous experiments. Finally, students will learn to do statistics in the R Computing Environment, which has become the standard in the sciences for conducting complex statistical analysis.
3. Personal and Social Responsibility. Students will learn about the use and misuse of statistics. They will learn to properly and ethically use statistics to help dispel the often voiced view that "you can show anything with statistics".
4. Ability to Integrate Knowledge and Skills. Students will integrate knowledge gained in reading the literature, lab activities devoted to specific techniques, and activities on experimental design

into a term project that will analyze a complex, real dataset, and present their finding in a report and an oral presentation.
5. Capacities for Effective Practice. Ample opportunities for discussion will allow undergraduate, M.S., and Ph.D. students to interact with one another and with the professor to gain a deeper understanding of the course material. The techniques learned are those frequently used by Biology Faculty at Clark University, allowing students to proficiently work with data they generate in research laboratories.

Course Website

All course material will be posted on Moodle, including any announcements, lab and lecture schedules, and PDFs of course handouts, readings, and Powerpoint presentations. Check Moodle regularly – it contains information and materials you will need for this course.

Grading, Components and Expectations

Biol 206/306 will be graded based on the following components:

Paper Discussion		
Participation in discussion	25%	Ongoing
Laboratory Exercises		
Worksheets and analyses	30%	Ongoing
R coding assignment	5%	October 15
R function notebook	5%	November 10
Term Project		
Choose Dataset	0%	September 10
Hypothesis activity	5%	October 1
Experimental design	8%	October 22
Term Project report	15%	November 24
Presentation	7%	December 1 & 3
Total	**100%**	

Biol 206 versus 306
Students working on their B.A. should register in Biol 206, while M.S. and Ph.D. students should register in Biol 306. Discussion participation, lab worksheets, term project reports, and presentations will be graded more rigorously for students in 306. The R coding assignment will also be more involved for students in 306.

Paper Discussion | 25% of the course | Ongoing
Each week, the class will read a paper on the technique covered during that week. Depending on the paper, this will happen either on Tuesday or Thursday morning of each week. All students are expected to participate in the discussion, as this represents a considerable amount of the grade.

Lab Exercises | 30% | Ongoing
During the lab, students will have the opportunity to analyze real data using the technique that is learned each week. The class will be given biological datasets collected and previously published by Clark faculty or other researchers, and each student will analyze the data in R. For each lab, students will either be provided with a worksheet to complete for a grade.

R Coding Assignment | 5% | October 15
Students will be asked to synthesize the knowledge they have gained so far to implement a statistical test that will be identified in class. This assignment will be a brief introduction to how powerful R can be to implementing any analysis that can be conceived. Biol 206 students will write a verbal algorithm (in plain English) to logically reason through the steps needed to implement the test. This should be a bulleted or numbered list of steps needed for implementation. Biol 306 students should also do the verbal algorithm, but should then implement the test in R by writing an R function. This will help to consolidate their knowledge of R and increase their proficiency with the language. This exercise will also help students design an algorithm and logically reason through the steps needed to implement it. Further guidance will be provided in class.

R Function Notebook | 5% | November 10
Students should maintain a notebook as they proceed through the course and learn how to do analyses in R. This assignment involves a notebook that lists the syntax and provides a brief explanation of each function that students learn during the course. The notebook will be handed in near the end of the semester and handed back to the students after grading. Such a notebook can be an extremely useful resource both during and after the course to quickly refresh one's memory on the details of a particular function.

Term Project | 35% over multiple assignments | Multiple Deadlines
The term project has been broken down into multiple components due throughout the semester to provide further guidance for students. On September 10, students will select a dataset to use for their term project. Students can either provide their own dataset (if they have collected data during their research), or will be given the opportunity to analyze a complex dataset supplied by faculty as their term project.

For the Hypothesis Activity (October 1), students will take a close look at their dataset and formulate biological hypotheses that they would like to test statistically. The assignment will be handing in these hypotheses.

For the Experimental Design assignment (October 22), students will outline which analyses they will use to test their biological hypotheses and provide the specific explicit statistical hypotheses that they will test.

The Term Project Report (November 24) will be written after students complete their analyses. The report for Biol 206 students will include a Statistical Methods and a Results section, complete with tables and figures. Methods should include sufficient detail to redo the analyses. The results should include everything necessary for interpretation of their analyses and data, but

not superfluous material. Biol 306 students and any students analyzing their own datasets should also include an Introduction and a Discussion. These sections should be no more than two pages, double spaced, each, but should cite multiple literature sources. Term Project Reports for all students should include a title page with a title, student name, course number and name, and assignment name. The text of the report should be double spaced, with indented paragraphs, 1" margins, 12pt Times New Roman Font, and page numbers. Tables should be single spaced with headings above each table. Figures should have captions below each figure. Figures and tables can be embedded in the text or provided at the end of the document. Literature cited should follow the format for the journal *Evolution*. Assignments that do not follow these formatting instructions will be returned to the student for correction prior to grading.

Finally, students will give a short, in-class presentation about their study, analyses and findings. Presentations will be in PowerPoint and will be done during the last week of classes. Further details will be provided as the semester progresses.

Course Time Commitment and Rules for Late Assignments

- All components of the course are mandatory to receive a passing grade. A student may choose not to do a component, but must E-mail Dr. Bergmann, stating that they are not doing the assignment (in this case a zero will be assessed for that component only).
- Late assignments will only be accepted if accompanied by documentation of a valid excuse (*e.g.* doctor's note, funeral certificate, etc.). Every effort will be made to accommodate these situations.
- The time commitment expected of students in this course each week is as follows:
    - 2.5 hours of lecture, plus 4 hours of preparation and studying
    - 3 hours of scheduled lab, plus 0.5 hours of preparation for the lab, and 2 hour for finishing lab worksheets
    - 1.5 hours for working on various components of the term project & coding assignment
    - Over a 14 week semester, this amounts to approximately 189 hours

Academic Integrity, Honesty, and Plagiarism

Academic dishonesty includes any effort to circumvent the evaluation procedures of the course to improve a grade for yourself or other students ("cheating"). Academic dishonesty includes but is not limited to unauthorized examination of written materials (i.e., notes, neighbor's paper during an exam), misrepresentation of the cause of an absence during an exam or laboratory, submitting the work of another (partially or entirely) as one's own, alteration of an exam answer to be submitted for regrading, and alteration of data. You are encouraged to report academic dishonesty. Anonymity will be protected if requested. If I believe that academic dishonesty has occurred and I have supporting evidence, I will report the case to the College Board immediately after informing the student that I am doing so, and why. I will recommend that a grade of F be given for the course. All students are expected to adhere to Clark University's rules of Academic Integrity, available at: http://www.clarku.edu/offices/aac/integrity.cfm.

Students with Disabilities

Clark University is committed to providing students with documented disabilities equal access to all university programs and facilities. If you have or think you have a disability and require academic accommodations, you must register with Student Accessibility Services (SAS), which is located in room 430 on the fourth floor of the Goddard Library. If you have questions about the process, please contact The Director of Accessibility Services, Emily Tarconish, at etarconish@clarku.edu or (508)798-4368. If you are registered with SAS, and qualify for accommodations that you would like to utilize in this course, please request those accommodations through SAS in a timely manner.

Course Schedule

Date		Information Session	Practical Session
Aug	25	Intro, Hypothesis Testing Review	
	27	Introduction to R	R Installation and Introduction
Sep	1	Discussion of Hurlbert 1984	
	3	Experimental Design and Sampling	Expt. Design, Multiple Comparisons
	8	ANOVA Variations & Models	
	10	Discussion of Bennington & Thayne 1994	Analysis of data with ANOVA **Choose dataset for Term Project**
	15	Bivariate Regression Choices	
	17	Discussion of McArdle 1988	Different Regressions in R
	22	Multiple Regression & ANCOVA	
	24	Discussion of Oliveira et al. 2012	Mult. Reg. and Dummy Variables
	29	Resampling Techniques	
Oct	1	Discussion of Schnell et al. 1985	Mantel Test & Randomiz. ANOVA **Hypothesis Assignment Due**
	6	Discussion of Betz 1987	
	8	Intro: Multivariate Stats, MANOVA	Using MANOVA & DFA
	13	**Fall Break - No Class**	
	15	Principal Component Analysis	PCA Re-analysis of DFA Data **R Coding Assignment Due**
	20	Discussion of Anderson et al. 2000	
	22	Model Selection & Likelihood	Comparing & Averaging Models **Expt. Design Assignment Due**
	27	Phylogenetic Regression	
	29	Discussion Freckleton 2009	Analysis of Trait Evolution
Nov	3	Extensions of Phylogenetic Statistics	
	5	Discussion of Butler & King 2004	Fitting models of Trait Evolution
	10	Discussion of Murray & Conner 2009	
	12	TBA	Work on Term Projects
	17	Discussion of Ellison 2004	
	19	Intro to Bayesian Statistics	BayesTraits Analysis
	24	Discussion of Stinchcomb et al. 2012	**Term Project Report Due**
	26	**Thanksgiving - No Lecture**	**No Lab**
Dec	1	**Term Project Presentations**	
	3	**Term Project Presentations**	

Note: This schedule may change as the course progresses. Please check Moodle regularly.