

Biostatistics

Math4397/6397

Fall 2009

Class: TuTh 2:30pm-3:50pm, AH 301

Instructor: Bernhard Bodmann, bgb@math.uh.edu

Office: PGH 604; Wed 2:00-2:50pm, Th 11:00-11:50am

Objectives: This course covers applications of statistics in biology and medicine, motivated by typical case studies. The students will learn a variety of uses, and abuses, of statistical methods. The material will be interspersed with simple programming projects, which allows the students to become familiar with R, the open-source software package used in this course.

Contents: The first part of the course is a rapid review of essentials in probability and statistics. The main part of the material focuses on typical estimation problems and hypothesis testing applied to data from medicine as well as population, molecular and physiological biology.

<i>General topic</i>	<i>Approximate Time</i>
Probability and statistics essentials	2 weeks
Inferences for one sample	3 weeks
Summarizing and describing data	1 week
The two sample problem	2 weeks
Contingency tables	2 weeks
Case-control and cross-sectional studies	2 weeks
Introduction to non-parametric methods	1 weeks
Large datasets	1 week

Detailed topics include: Independence, Bayes rule, sensitivity and specificity of a test, likelihood ratio; normal and chi-squared distribution, confidence intervals; students t-distribution; empirical quantiles, boxplot, qauntile-quantile plot; kernel density estimates, stem and leaf plots, histograms; bootstrap principle; binomial confidence intervals; group comparisons; Pearsons chi-squared test; retrospective case/control studies; multiplicity: Bonferroni adjustment for family-wise error, false-discovery rate; stratified tables; matched pairs; Poisson processes and rate estimate.

Prerequisites: MATH 1432 and MATH 2311, or equivalent.

Text: Bernard Rosner, Fundamentals of Biostatistics, 6th edition, Thomson Brooks/Cole, 2006.

- Website:** All course-related materials such as homework assignments, solutions, and practice exams will be posted on www.math.uh.edu/~bgb/biostats
- Software:** We will use R, an open-source variant of the common statistics language and environment S. R runs on many platforms including Microsoft Windows, Linux or Apple OS X and is freely available at the website www.cran.r-project.org
- Midterm Exam:** Tuesday, October 20, 2009, in class.
- Assignments:** You will be asked to hand in approximately ten assignments, which will be due on Thursdays in the lecture.
- Grad students:** To obtain full credit for this course at the graduate level, students will need to complete 4 additional projects on biological datasets. These projects will be included with part of the assignments.
- Final Grade:** Final exam contributes 40%, midterm 30%, assignments 30%. All grades are summed and divided by the total number of points you can collect in the course. A percentage of 46% or more is D- , 54% or more is D, 62% or more is C, 70% is B-, 77% is B, 85% or more is A- , of 90% or more is A.