

Biostatistics 560
Statistical Methods for Epidemiology
Fall 2008
Course Information & Syllabus

INSTRUCTOR: Dr. Alex Tsodikov
Department of Biostatistics
Office: M4142 SPH II
Tel: (734) 615-6416
Email: tsodikov@umich.edu
Office Hours: T/Th 9:30 – 10:30 am, or by appointment

GSI: Elizabeth Love
Department of Biostatistics
Office: MG 519
Tel: 763-9054 or 763-9055
Email: ellove@umich.edu
Office Hours: Tue, Thu 3-5pm

CLASS TIME: T/TH 1:00 – 3:00 PM

LOCATION: 1690 SPH1A

COURSE DESCRIPTION: Biostat 560 is a 4-credit course on statistical methods commonly used in epidemiology. The emphasis is on applying appropriate statistical methods and subsequent interpretation of epidemiologic data. Upon entering this course, you are expected to have completed course in introductory statistics or biostatistics and multiple linear regression. You should understand the basic statistical concepts of sampling variation, parameter estimation and statistical hypothesis tests. You should know how to fit multiple linear regression models, how to interpret regression coefficients in these models and how to perform hypothesis tests about regression coefficients. You should be familiar with case-control, cross-sectional and cohort study designs.

PREREQUISITES: EPI 503 or 601, BIOST 523; or permission of the instructor.

TEXTS:

Highly Recommended (on reserve in SPH library):

1. N.E. Breslow and N.E. Day: *Statistical Methods in Cancer Research*, Volume 1. 1980. International Agency for Research on Cancer.
2. N.E. Breslow and N.E. Day: *Statistical Methods in Cancer Research*, Volume 2. 1987. International Agency for Research on Cancer.
3. D.W. Hosmer and S. Lemeshow: *Applied Logistic Regression (Second Edition)*. 2000. John Wiley & Sons.

4. D.W. Hosmer and S. Lemeshow: *Applied Survival Analysis*. 1999. John Wiley & Sons.
5. M. Woodward: *Epidemiology Study Design and Data Analysis (Second Edition)*. 2005. Chapman & Hall/CRC.

Recommended:

1. P.D. Allison. *Logistic Regression Using the SAS System: Theory and Application*. 1999. Cary, NC: SAS Institute.
2. P.D. Allison. *Survival Analysis Using the SAS System: A Practical Guide*. 1995. Cary, NC: SAS Institute.
3. P. Dalgaard, *Introductory Statistics with R*, Springer-Verlag, NY. 2004

READINGS: Given the breadth of topics covered in this course, it is impossible to find a single textbook that does justice to all the topics covered at the level of this course. Materials for the course will be derived from a variety of sources, such as the recommended texts, other books, peer-reviewed articles, etc. in order to provide you with lecture notes. Using lecture notes and recommended reading materials will be needed to complete the homework assignments and exams.

COMPUTING: The computing requirement for the course is quite heavy. Statistical software examples used in the course will be in SAS and R for Windows. Upon entering the course, students are expected to have basic skills in one of statistical programming languages (e.g. reading in data, creating new variables, merging data sets, case selection, subsetting, sorting, stratification, etc.). The choice of statistical software for course assignments is yours.

CLASS WEBPAGE: Available on CTools: <https://ctools.umich.edu/> All lecture notes, handouts, required reading materials, homework assignments, their solutions, and various other help materials for the course will be posted on the course website on a regular basis throughout the semester. Please bring your own downloaded copy of the lecture notes to each day's lecture.

HOMEWORKS: There will be weekly homework assignments, typically on Thursdays, covering theoretical material, applications, data analysis using statistical software, and interpretation of results. Homeworks will be due the following Thursdays, with exceptions for holidays. You are encouraged to submit your homeworks electronically using CTools drop box or email of the instructor. Alternatively, you could submit hardcopies on standard size paper, with the sheets stapled together provided there is more than one page. Computer outputs should be attached where necessary. An un-annotated computer output will not constitute an acceptable answer unless it is entirely self-explanatory, which is seldom the case. It is advisable to present a report-style integration of written answers and computer printouts by a “cut-and-paste” job if necessary.

You may work on homework problem sets in groups; however the write-up and interpretations must be your own work. Any form of copying from another person's work will be construed as violation of academic conduct and will be subject to appropriate disciplinary actions as per university rules.

Late assignments will be assessed a penalty of 20% for each day late, so you are encouraged to submit them on time. No late homeworks will be accepted after the solutions have been posted.

COURSE EVALUATION:

Homeworks: 20%

Exam 1 (Thursday, October 9th, 1:00 – 3:00 pm, in class): 25%

Exam 2 (Thursday, November 6th, 1:00 – 3:00 pm, in class): 25%

Final (Thursday, December 11th, TBA, in class): 30%

Your course grade will be based on the weighting scheme presented above. There is no fixed grading scale for this course; conversion from your cumulative score to letter grades will be carried out at the end of the course.

ACADEMIC INTEGRITY: The faculty of the School of Public Health believes that the conduct of a student registered or taking courses in the School should be consistent with that of a professional person. Courtesy, honesty and respect should be shown by students toward faculty members, guest lecturers, administrative support staff and fellow students. Similarly, students should expect faculty to treat them fairly, showing respect for their ideas and opinions and striving to help them achieve maximum benefits from their experience in the School. Student academic misconduct refers to behavior that may include plagiarism, cheating, fabrication, falsification of records or official documents, intentional misuse of equipment or materials (including library materials), and aiding and abetting the perpetration of such acts. The preparation of reports, papers, and examinations, assigned on an individual basis, must represent each student's own effort. Reference sources should be indicated clearly. The use of assistance from other students or aids of any kind during a written examination, except when the use of aids such as electronic devices, books or notes has been approved by an instructor, is a violation of the standard of academic conduct.

COURSE OUTLINE:

Upon entering the course students will be responsible for knowledge of introductory materials on the following topics: probability (conditional probability, Bayes' theorem), discrete and continuous distributions (binomial, poisson, normal, chi-square), statistical estimation (point estimates, confidence intervals) and hypothesis testing (significance level, type I and II errors, power, p-value), and multiple linear regression (model and assumptions, categories and dummy variables, inference about parameters, interactions, polynomial, transformations, and model diagnostics). Some brief notes and handouts on these topics will be posted on the course website as review materials, and will be discussed in the lectures, time permitting.

Topics

1. Background
 - Measures of disease occurrence and association
2. Contingency Tables
 - 2 by 2 Tables
 - 2 by K Tables
3. Logistic Regression
4. Analyses of Matched Data
 - Tests
 - Conditional logistic regression
5. Polytomous logistic regression
6. Ordinal logistic regression
7. Survival Analysis
 - Univariate methods, tests
 - Proportional hazards regression
8. Poisson Regression
9. Analysis of rates, Standardized incidence/mortality ratio, Age-Period-Cohort models
10. Longitudinal Data

Time permitting other topics will be included

Learning Objectives
Biostatistics 560
Statistical Methods for Epidemiology
Fall 2008

After completing the course, the student can ordinarily expect to be able to:

1. Use simple statistical techniques for analyzing matched and unmatched data in contingency tables
2. Fit appropriate logistic regression models to data from epidemiologic cohort studies and case-control studies, interpret regression coefficients in these models, test hypotheses about them, perform model diagnostics and recognize problems.
3. Understand the use of ordinary vs. conditional inference.
4. Understand the use of ordinal vs. polytomous models.
5. Use appropriate univariate methods, tests, and multivariate regression models, interpret regression coefficients from these models, test hypotheses about them and evaluate the fit of these regression models.
6. Understand the concept of rates such as incidence and mortality, age-adjustment, and be able to use Poisson regression models to analyze rates data.
7. Fit appropriate regression models to longitudinal data from epidemiological cohort studies, interpret regression coefficients from these models and test hypotheses about them.
8. Understand the steps in analyzing data from an epidemiological case-control study or cohort study and be able to implement them.
9. Present results of analyses using advanced regression models to readers who are familiar with these models as well as to readers who are not.