

Statistical Reasoning in Public Health II (Biostatistics 140.612) Term 2, 2009

Lecture: Tuesday, Thursday 10:30-11:50
W1020 (Becton Dickinson Auditorium) Wolfe Street Building

Instructor: John McGready
Office Hours: After class Tuesday and Thursday, and by appointment

Description: This introductory course is intended to provide students with a broad overview of biostatistical methods and concepts used in the public health sciences. The emphasis is on interpretation and concepts rather than calculations or mathematical details.

Textbooks: There is no required textbook: the course e-reserves will contain selections from some of the recommended course texts plus journal articles and other reading of relevance. The URL for the course ereserves will be posted on the course web page. This site requires a password that will be given out in class.

Recommended For Reference:

- *Practical Statistics for Medical Research*, Douglas Altman. (1991)
- *Introduction to the Practice of Statistics*, David Moore, George McCabe. (2003)
- *Statistics*, David Freeman, Robert Pisani, Roger Purves. (1998)
- *Intuitive Biostatistics*, Harvey Motulsky. (1995)
- *Fundamentals of Biostatistics*, Bernard Rosner. (2000)
- *Principles of Biostatistics*, Marcello Pagano, Kimberlee Gauvreau. (2000)

All above books are on reserve at the Lilienfeld Library (Hampton House, 9th floor).

Web Page: <http://www.biostat.jhsph.edu/~jmcgread/bio612/> : This is the address for the main page, from which you will also be able to link to pages containing the handouts in .pdf format, solutions to homeworks and exams (posted of course after the due date/exam date), live recordings of the on campus classes, and narrated lecture slides as used in the distance education version of this course.

Courseplus There is a Courseplus page for this course: this will be used for homework submissions via the Dropbox feature: otherwise, all pertinent course information with II on the web page described in the preceding paragraph.

Lecture Recordings All lectures will be digitally recored, and the the mp3 will be posted on the course website within rougly two business days. Additionally, narrated slide versions of the lectures are available through the course website

Teaching Assistants: Russel (Taki) Shinohara
Vidya Venugopal

Office Hours John: Tuesday, Thursday 12:15-1:15 pm, Room W4007, and by appointment (schedule permitting)
Taki and Widua: TBA

Office Hours will begin on Thursday, October 29.

Student Evaluation: There will be one midterm exam, a final exam, and 3 homework assignments.

<u>Grading</u>	Midterm Quiz -	20%
	Final Exam -	30%
	Homework -	50%

Homeworks – There will be 3 homework assignments. This must be submitted by the stated deadline.. The three assignments will count for 50% of the final grade. You may discuss ways to approach homework with other students , the TA's, and John McGready. However, each homework submission must be your own independent work.

Class Notes: distributed by instructor

Statistical Reasoning in Public Health Biostatistics 612 Learning Objectives

140.612 Statistical Reasoning in Public Health II, 2009

- Recognize different study designs, and understand the pros and cons of each.
- Learn methods for randomly assigning subjects to two groups
- Understand the concepts of confounding and statistical interaction; know how to recognize each.
- Explain the relationship between power and sample size; use Stata to perform sample size calculations.
- Create a scatterplot to visually assess the nature of an association between two continuous variables
- Interpret the calculated values of the correlation coefficient and the coefficient of determination, and understand the relationship between these two measures of association
- Perform a simple linear regression using Stata and use the results to assess the magnitude and significance of the relationship between a continuous outcome variable and a continuous predictor variable, and for predicting values of the outcome variable
- Understand why multiple regression techniques allow for the analysis of the relationship between an outcome and a predictor in the presence of confounding variables
- Perform a multiple linear regression using Stata and use the results to assess the magnitude and significance of the relationship between a continuous outcome variable and multiple continuous and categorical predictor variables, and for predicting values of the outcome variable
- Perform a multiple logistic regression using Stata and use the results to assess the magnitude and significance of the relationship between a dichotomous outcome variable and multiple continuous and categorical predictor variables
- Interpret the results from a proportional hazards regression model

Statistical Reasoning in Public Health Biostatistics 612 Topics

1. Study Design Issues

- Randomized, Observational, and Prospective study designs
- Case Control Study Design
- Assessing Confounding and Interaction
- The Relationship of Power, Detectable Difference, and Sample Size
- Necessary Sample Size of Comparing Two Means
- Necessary Sample Size for Comparing Two Proportions
- Sample Size and Power Computations with Stata

2. Linear Regression

- Making the Case for Linear Regression
- Simple Linear Regression and Correlation
- The t-test as a Simple Linear Regression
- Multiple Linear Regression
- ANOVA as a Multiple Linear Regression
- Confidence Intervals and Inference in Linear Regression
- Assessing Confounding and Interaction with Multiple Linear Regression
- Prediction with Linear Regression
- Linear Regression with Stata

3. Logistic Regression

- Making the Case for Logistic Regression
- Probability, Odds and log-Odds
- Simple Logistic Regression
- The Two Sample Problem as a Simple Logistic Regression
- Multiple Logistic Regression
- Confidence Intervals and Inference in Logistic Regression
- Assessing Confounding and Interaction with Multiple Logistic Regression
- Prediction with Logistic Regression
- Logistic Regression with Stata

4. Cox Regression

Making the Case for Cox Regression
The Proportional Hazards Assumption
Hazard Ratios
Kaplan-Meier and Log-Rank vs. Cox Regression
Multiple Cox Regression
Confidence Intervals and Inference in Cox Regression
Assessing Confounding and Interaction with Cox Regression
Cox Regression with Stata