

## BIostatistics II (E10.2996, U10.2196, D60.7014) Spring 2008

Professors: Robert Norman, Ph.D. TA: Rashmi Shrestha  
Mariya Shiyko

Phone: Norman: 998-9507  
E-Mail: Norman: [robert.norman@nyu.edu](mailto:robert.norman@nyu.edu) Shiyko: [mps8@nyu.edu](mailto:mps8@nyu.edu)  
Shrestha: [rs2675@nyu.edu](mailto:rs2675@nyu.edu)  
Office: Norman: Rm 238, 421 1<sup>st</sup> Ave – By Appointment

Meeting Times: Class meets Tuesdays 4:55-7:35pm in the Nagle Auditorium (6<sup>th</sup> floor) located at the Dental College, 345 E. 24 St (at 1<sup>st</sup> Avenue). This time is divided into two parts: a lecture (usually led by Dr. Norman) that ends at 6:35pm followed by a recitation (usually led by M. Shiyko). The recitation session is optional, though has proven to be of great benefit in the past. Optional lab sessions are on Mondays 2:55-3:55 and Wednesdays 6:25pm-7:25pm and 7:40pm- 8:40 PM at 194 Mercer St, Room 304.

Required Texts: Introductory Biostatistics, Chap T. Le  
[British Medical Journal Statistics at Square One](#) (online book)  
NOTE: The same texts were used for Biostatistics I.

Optional Texts: Data Analysis for the Behavioral Sciences using SPSS, Weinberg&Abramowitz.  
SPSS for Windows Step-by-Step: A simple guide and reference, George&Mallery.

This course sequence is intended for graduate students in the epidemiology, public health and clinical research fields and is the second in the Biostatistics sequence. It will concentrate on more advanced methods of statistical analysis and research design that are typical of biological and medical applications of statistics. It is assumed that the student will be familiar with statistical techniques as presented in E10.2995 – U10.2996 – D60.7040. We will again be using SPSS to perform statistical processing. It is assumed that the student has basic skills in the use of SPSS for entering data sets and performing basic analyses and graphics.

**Objectives:** At the conclusion of this course the student will be able to:

- 1) Apply critical evaluation skills to the reading of scientific research presentations as they relate to: Regression, ANOVA, logistic regression, survival analysis and non-parametric analysis.
- 2) Combine their knowledge of statistical analysis and use of SPSS to perform an analysis of real data to answer scientific hypotheses.
- 3) Understand basic issues of experimental design and sample size determination.
- 4) Undertake further study of statistics based on the Advanced Quantitative Methods sequence.

### Course requirements:

**Homework:** Homework sets will be assigned periodically during the course, which concentrate on the interpretation of statistical results. In addition, there will be assignments that require the use of SPSS to perform statistical analyses. Homework will be collected and graded. Homework is due at the beginning of class on the date specified. Late homework will be penalized.

**Project:** The project consists of the analysis of a medically-oriented data set using the techniques learned in this course. The student will be given a series of questions and asked to choose and perform a statistical analysis that will provide insight into their answer. The student will then write up these analyses in the form of the results section of a scientific paper. The project is to be typed (word processed) and should be professional in appearance.

**Exams:** There will be both a midterm and final exam. Each exam will contain both general statistical knowledge questions and a series of statistical analyses with questions relating to the theory used, assumptions made and interpretation of the statistical data presented.

**Grades**

|                     |     |
|---------------------|-----|
| Homework            | 20% |
| Class participation | 10% |
| Project             | 10% |
| Midterm             | 25% |
| Final               | 35% |

**CLASS SCHEDULE**

- 01/22/2008 Regression (Le, Chapter 8) (Norman)
- 1) Relationship of correlation and regression
  - 2) Simple univariate regression
  - 3) Interpretation of coefficients.
  - 4) Hypothesis tests of coefficients.
  - 5) Regression in SPSS.
- 01/29/2008 Regression con't (Norman)
- 1) Assumptions in regression
  - 2) Regression diagnostics
  - 3) Residual analysis
  - 4) Influential points
  - 5) Examples
- 2/05/2008 Regression con't (Norman)
- 1) Multiple regression
  - 2) Interaction
  - 3) Dummy Coding and categorical predictors
  - 4) Examples
- 02/12/2008 Experimental design (Shiyko)
- 1) Longitudinal designs and cohort studies
  - 2) Case control studies
  - 3) Examples of experiments.
  - 4) Statistical control
  - 5) Examples of experimental designs
- 02/19/2008 Analysis of variance (Le, Chap 7) (Shiyko)
- 1) Introduction and rationale
  - 2) Hypotheses
  - 3) Mathematical derivation.
  - 4) 1 way analysis of variance.
- 02/26/2008 Analysis of variance and post hoc analysis (Shiyko)
- 1) 2 way analysis of variance. Additive effects vs interaction.
  - 2) Profile plots.
  - 3) Post hoc analysis.

- 03/04/2008 Categorical data analysis – (Norman)  
1) 2 x 2 x 2 tables.  
2) Confounding – Mantel-Haenszel Test  
3) Simpsons Paradox
- 03/11/2008 Midterm Exam
- 03/18/2008 Spring Recess
- 03/25/2008 Non-parametric tests. Permutation and randomization tests. (Norman)  
1) Descriptive statistics.  
2) Mann, Whitney, Wilcoxon and Wilcoxon matched pairs test.  
3) Freidman rank anova.  
4) Permutation test concept.  
5) Permutation test equivalent of 2 sample T test.
- 04/01/2008 Logistic Regression (Le, Chap 9) (Shiyko)  
1) Introduction  
2) Visualization  
3) Rationale (i.e., why not just use linear regression?)  
4) Components of a logistic regression model
- 04/08/2008 Logistic Regression (completion) (Shiyko)  
1) Predictions  
2) Interpretation  
3) Goodness of fit  
4) Examples
- 04/15/2008 Survival analysis (Le, Chapter 11) (Norman)  
1) Graphic techniques.  
2) Kaplan-Meier curves  
3) Hypothesis tests.
- 04/22/2008 Survival con't (Norman)  
1) Cox proportional hazards regression
- 04/29/2008 Power Analysis (Shiyko)  
1) Why do we need to plan sample size  
2) Error rates and sample size  
3) Power analysis  
4) Software
- 05/06/2008 READING DAY (NO CLASSES)
- 05/13/2008 Final Exam (NOTE: Exam time is from 6:00 PM to 7:50 PM)