

# BIostatISTICS I (E10.2995, U10.2995, D60.7040, N41.3501) FALL 2008

- Professors: Robert Norman, Ph.D.  
Mariya Shiyko, Ph.D. (cand.)
- TA: Joseph Palamar, MPH, Ph.D. (cand.)
- Phone: Norman: 998-9507
- E-Mail: Norman: [robert.norman@nyu.edu](mailto:robert.norman@nyu.edu) Shiyko: [mps8@nyu.edu](mailto:mps8@nyu.edu)  
Palamar: [jjp250@nyu.edu](mailto:jjp250@nyu.edu)
- Office: Norman: Rm 238, 421 1<sup>st</sup> Ave
- Office Hours: 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.
- Required Texts: Fundamentals of Biostatistics, Bernard Rosner.  
[British Medical Journal Statistics at Square One](#) (online book)  
NOTE: The same texts are used for Biostatistics II.
- 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 nursing, epidemiology, public health and clinical research fields. This course provides both the foundations necessary for Biostatistics II and serves as a stand alone introductory statistics course. It will concentrate on the interpretation and comprehension of graphical and statistical techniques that are important components of scientific literature. Mathematical ability at the level of high school algebra is required. We will also be using the statistical program SPSS to perform statistical processing and there will be assignments that require the use of this program. We will go over the necessary parts of SPSS in class sessions as needed, but you will be expected to work on these assignments on your own. SPSS is installed on the computers in the computer labs at NYU. There are three attended lab sessions per week for those seeking additional help. They are on Mondays from 3:20-4:20 PM and Fridays from 7:30-9:35 PM and are held at 194 Mercer St, room 304. If these times are not used we will lose them, so please take advantage of this opportunity.

**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: sample statistics, simple hypothesis tests, correlation and regression.
- 2) Understand the principle of the use of probability theory as a basis for making statistical decisions.
- 3) Perform statistical analyses covered in class using the SPSS statistics program.
- 4) Combine their knowledge of statistical analysis and use of SPSS to perform an analysis of real data to answer scientific hypotheses.
- 5) Undertake further study of statistics based on the Biostatistics II course (N41.3502, E10.2996, U10.2996, D60.7014).

## Course requirements:

**Homework:** It is essential that the student practice the skills covered in class sessions. In addition, some material not covered in class will be presented in homework problems. Homework sets will be assigned during the course, which concentrate on the use of SPSS and the interpretation of statistical results, with only minimal manual calculation. Homework will be collected and graded. Homework is due on the date specified in the syllabus unless changed in class. Late homework will not be accepted unless arranged in advance. While you may work together on homework assignments, you **MUST** write up the homework **individually**. Homework papers with substantially identical sections will receive a grade of 0.

**Project:** The project consists of the analysis of a biologically-oriented data set. The student will be given a series of hypotheses and asked to choose and perform a statistical analysis that will either support or refute the hypotheses. 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. Hand written projects will not be graded. The project is to be worked on and written up **independently**.

**Exams:** There will be both a midterm and final exam. Each 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. They do NOT emphasize calculations and formulas. Material from both the lectures and from the text may appear on the exam and all exams are cumulative.

## Grades

Homework	30%
Project	10%
Midterm	20%
Cumulative Final	40%

## COURSE SCHEDULE

- 09/02/2008 Introduction, measurement, frequency tables and graphics. (*Chapter 1*)
- 1) What is statistics? Definitions and types of problems
  - 2) Example of a statistical problem.
  - 3) Types of measures. Nominal, ordinal, interval, ratio. Continuous, discrete.
  - 4) Unit of analysis
  - 5) Measured values vs. underlying characteristics.
  - 6) Concept of a distribution.
  - 7) Descriptive methods for categorical data - numeric.
  - 8) Descriptive methods for categorical data - graphic.
- 09/09/2008 Ranks/Percentiles. Measures of central tendency, graphic methods. (*Chapter 2*)
- 1) Percentile rank vs percentile. Definition and calculation.
  - 2) Quartiles, quintiles, deciles.
  - 3) Numerical summary of a distribution. The concept of a summary statistic.
  - 4) Descriptive methods for continuous data – graphics
  - 5) Descriptive methods for continuous data – central tendency

- 09/16/2008 Measures of dispersion, skew and graphic methods. (*Chapter 2*)  
**SB 1 due**  
1) Range and interquartile range.  
2) Variance, standard deviation.  
3) The standard error of a statistic.  
4) Pearson and spearman skew statistics  
5) Density plots
- 09/23/2008 Introduction to probability. (*Chapter 3*)  
**HO 1 due**  
1) Definitions of probability.  
2) Mathematical definitions  
3) Practical examples – coins, urns etc.
- 09/30/2008 Discrete probability theory and problems. (*Chapter 4*)  
**SB 2 due**  
1) Counting theory. Permutations and combinations.  
2) Binomial distribution. Definitions, use and examples.  
3) Poisson distribution. Definitions, use and examples.  
4) Hypergeometric distribution. Definitions, use and examples.  
5) Populations and samples
- 10/07/2008 Continuous probability theory. Introduction to hypothesis testing. (*Chapter 5*)  
Populations and sampling  
**HO 2 due**  
1) Normal distribution. Definitions, use and examples.  
2) T Distribution. Definitions, use and examples.  
3) Chi-square distribution, Definitions, use and examples.  
4) F distribution. Definitions, use and examples.
- 10/14/2008 **NO CLASS – COLUMBUS DAY**
- 10/21/2008 Hypothesis tests. (*Chapters 6,7*)  
1) Null hypothesis vs. practical hypothesis  
2) Power, Type I and Type II errors.  
3) Introduction to the T test using one sample T test  
4) Independent sample T test  
5) Paired sample T test.
- 10/28/2008 **MIDTERM EXAM**

- 11/04/2008 Hypothesis Tests Continued and confidence intervals. (*Chapters 6,7,8*)  
**SB 3 due**  
 1) Confidence interval for the mean.  
 2) Other types of confidence intervals.
- 11/11/2008 Hypothesis Tests Continued and confidence intervals. (*Chapters 7,8*)  
**HO 3 due**  
 1) Confidence interval for the mean.  
 2) Other types of confidence intervals.
- 11/18/2008 Measures of association : correlation and scatterplots. (*Chapter 11 skip 511-539*)  
**SB 4 due**  
 1) Introduction to bivariate analysis.  
 2) The scatterplot.  
 3) Pearson correlation. Definition, use examples  
 4) Spearman correlation  
 5) Point biserial correlation  
 6) Phi
- 11/25/2008 Regression (*Chapter 11 skip 511-539*)  
**SB 5 due**  
 1) Relationship of correlation and regression  
 2) Simple univariate regression  
 3) Interpretation of coefficients.  
 4) Hypothesis tests of coefficients.  
 5) Regression in SPSS.
- 12/02/2008 Regression con't (*Chapter 11 skip 511-539*)  
**SB 6 due**  
 1) Assumptions in regression  
 2) Regression diagnostics  
 3) Residual analysis  
 4) Influential points
- 12/09/2008 Experimental design (*Chapter 6:174-177*)  
**HO 4 due**  
 1) Longitudinal designs and cohort studies  
 2) Case control studies  
 3) Examples of experiments.  
 4) Statistical control
- 12/16/2008 **FINAL EXAM (TIME 6:00 PM to 7:50 PM)**

SB refers to SPSS/Book based homework  
 HO refers to handout based homework.