BIOSTATISTICS I (E10.2995)

Instructor: Robert Norman, PhD
Phone: Leave message at 995-4700 (Mailbox 90851)
E-Mail: robert.norman@med.nyu.edu weekdays (9-5).
        rnorman@magiccarpet.com evenings and weekends.
Office Hours: By Appointment
Office: Kimball Hall, 3rd Floor or 3R (no specific room)

Required Texts: Introductory Biostatistics, Chap T. Le
        British Medical Journal Statistics at Square One (online book)
Optional Text:  Data Analysis for the Behavioral Sciences using SPSS, Weinberg&Abramowitz
        The Research Methods Knowledge Base, William M. K. Trochim
        NOTE: The latter text is required for Biostatistics II.

This course sequence is intended for consumers of statistics in the biological and medical fields. 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. There will be assignments that require the use of SPSS. We will go over the necessary parts of SPSS in class as needed but you will be expected to work on these assignments outside of class. SPSS is installed on the computers in the Tisch Hall computer lab. You will have a class account that gets you access to the lab during prime access hours.

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 from 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 Advanced Quantitative Methods sequence.

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. Five homework sets will be assigned over the course, which concentrate on the interpretation of statistical results. In addition, there will be three assignments that require the use of SPSS to perform simple statistical calculations. Homework will be collected and graded.

Projects: The project consists of the analysis of a medically oriented data set. The student will be given a series of hypotheses and asked to perform a statistical analysis that will either support or refute the hypotheses. The student will then write up these analyses in the form of a results section of a scientific paper.
Exams: There will be both a midterm and final exam. Each will take the form of a series of statistical analyses with questions relating to the theory used, assumptions made and interpretation of the statistical data presented.

Grades

<table>
<thead>
<tr>
<th>Grade</th>
<th>Percentage</th>
</tr>
</thead>
<tbody>
<tr>
<td>Homework</td>
<td>30%</td>
</tr>
<tr>
<td>Project</td>
<td>20%</td>
</tr>
<tr>
<td>Midterm</td>
<td>20%</td>
</tr>
<tr>
<td>Final</td>
<td>30%</td>
</tr>
</tbody>
</table>

Class 1  Introduction, measurement, frequency tables and graphics. (Le, Chapter 1)
1) What is statistics? Definitions and types of problems
2) Example of a statistical problem.
4) Measured values vs. underlying characteristics.
5) Concept of distribution.
6) Descriptive methods for categorical data - numeric.
7) Descriptive methods for categorical data - graphic.
8) Use of SPSS.

Class 2  Ranks and Percentiles. Measures of central tendency, graphic methods (Le, 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
6) Calculation of statistics to date with SPSS

Class 3  Measures of dispersion, skew and graphic methods.
1) Range and interquartile range.
2) Variance, standard deviation.
3) The standard error of a statistic.
4) Pearson and spearman skew statistics
5) Calculation of statistics with SPSS. Interactive graphics.
6) Linear transforms
7) Non-linear transforms

Class 4  Introduction to probability (Le, Chapter 3)
1) Definitions of probability.
2) Mathematical definitions
3) Practical examples – coins, urns etc.
4) Transforms in SPSS. SPSS syntax.

Class 5  Discrete probability theory and problems.
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.
Class 6  Continuous probability theory and introduction to hypothesis testing
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.
5) Maximum likelihood and testing
6) The purpose of hypothesis testing

Class 7  Hypothesis tests  (Le, Chapters 5, 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.
6) T tests in SPSS

Class 8  Confidence intervals
1) Examples of use of T tests
2) Confidence interval for the mean.
3) Confidence interval for a standard deviation.
4) Other types of confidence intervals.
5) Confidence intervals in SPSS.

Class 9  Correlation and scatterplots (Le, Chapter 8)
1) Introduction to bivariate analysis.
2) The scatterplot.
3) Pearson correlation. Definition, use examples
4) Spearman correlation
5) Point biserial correlation
6) Phi
7) Correlation in SPSS.

Class 10  Introduction to regression analysis
1) Relationship of correlation and regression
2) Simple univariate regression
3) Interpretation of coefficients.
4) The standardized coefficient.
5) Hypothesis tests of coefficients.
6) Regression in SPSS.

Class 11  Regression continued
1) $R^2$ calculation and interpretation.
2) ANOVA and regression.
3) Multiple regression
4) Examples of use of regression analysis.

Class 12  Simple research design (Trochim, Chapter 1, 7 [pp 191-202], 8 [pp215-222]
1) Introduction to experimentation based on the solution of a research problem worked through.
<table>
<thead>
<tr>
<th>Class 13</th>
<th>Introduction to experimental design</th>
</tr>
</thead>
<tbody>
<tr>
<td>1)</td>
<td>Observation, quasi-experimental vs. experimental</td>
</tr>
<tr>
<td>2)</td>
<td>Cross sectional designs</td>
</tr>
<tr>
<td>3)</td>
<td>The control</td>
</tr>
<tr>
<td>4)</td>
<td>Randomized clinical trial</td>
</tr>
<tr>
<td>5)</td>
<td>Double blind placebo controlled crossover study</td>
</tr>
<tr>
<td>6)</td>
<td>Examples of experimental designs</td>
</tr>
</tbody>
</table>

| Class 14 | Miscellaneous topics and review |