Biostatistical Methods I for Public Health and Medical Sciences
Introduction to Statistics: Statistical Summaries and Inference

Professor: Ron Schrader
Office: Humanities 435
Phone: 277-7423
e-mail: schrader@math.unm.edu

James Degnan
Office: Humanities 330
Phone: 277-6163
e-mail: james@math.unm.edu

Class: Tuesday: 9:30 - 11:30 FPCT 340 (lecture)
Thursday: 9:30 - 11:00 School of Med. Bldg. 2 -- Electronic Classroom (lab)

Office Hours:
- Tues 11:30 - 12:00 Schrader (after class)
- Tues 1:00 - 1:30 Schrader (in HUM 435)
- Thur 11:30 - 1:30 Schrader (in HUM 435)
- TBA Degnan

Texts:

Required (available at UNM Bookstore – Main Campus):


Optional:

Calculator (one that does square roots, logarithms and statistical functions)

Course Notes:

Most of the course notes will be available in PDF format (readable by Adobe Acrobat) at
http://www.stat.unm.edu/~schrader/biostat/bio1.html. Our schedule is ambitious; material
may be excised as necessary.
**Homework:**
Assigned approximately every week. Due at next lecture, unless otherwise stated. Late homework normally will not be accepted.

**Grade:**
<table>
<thead>
<tr>
<th>Component</th>
<th>Percentage</th>
</tr>
</thead>
<tbody>
<tr>
<td>Homework</td>
<td>90%</td>
</tr>
<tr>
<td>Class participation/Computer lab</td>
<td>10%</td>
</tr>
<tr>
<td><strong>Total</strong></td>
<td><strong>100%</strong></td>
</tr>
</tbody>
</table>

**Prerequisites:**
Math 121 (College algebra) or permission of the instructor.

**OBJECTIVES**

1. Understand basic statistical and probability concepts.
2. Be able to interpret and prepare graphical and numerical summaries of data.
3. Understand the basics of statistical inference with respect to estimation and hypothesis testing.
4. Be able to determine appropriate statistical methods to use and implement them in simple analyses involving inferences for the population mean from one-sample, population means from two samples, simple discrete data analysis, and simple linear regression models.
5. Be able to use computer software to conduct simple statistical analyses.
6. Understanding basic research designs used in Public Health
7. Determining appropriate use of data and statistical methods
8. Finding sources of relevant data and information, and how to access these sources

**COURSE DESCRIPTION**

This course covers basic statistical methods used in the medical sciences. Types of data will be discussed. Methods of summarizing data through graphical displays and numerical summaries (measures of central tendency, percentiles, and variability) will be studied. Probability concepts will be covered to form the basis of statistical inference. Confidence intervals and hypothesis testing will be studied. Methods for statistical inference will focus on population means for one-sample, paired samples and two independent samples. Both normal-theory and nonparametric approaches will be studied. Methods of summarizing and analyzing discrete data will include proportions and tests of association and independence for two-way tables. The course will conclude with an introduction to simple linear regression. Emphasis will be placed on conducting statistical analyses on the computer.
COURSE OUTLINE

1. INTRODUCTION (Aug. 24)
   
   Reading: Samuels and Witmer (hereafter SW), Chapter 1
   
   i. Examples of the use of statistics
   ii. Lab - Introduction to computing

   DESCRIPTIVE STATISTICS: NUMERICAL
   
   Reading: SW, Chapter 2, Sections 1-7
   
   i. Types of variables
   ii. Frequency distributions
   iii. Numerical summaries of location: mean, median, mode, geometric mean, percentiles
   iv. Numerical summaries of spread: standard deviation, variance, range, interquartile range

2. DESCRIPTIVE STATISTICS: GRAPHICAL (Aug. 31)

   Reading: SW, Chapter 2, Sections 3 and 5
   
   i. Graphical displays of data: histograms, stem-and-leaf plots, box plots

3. PROBABILITY - BASIC IDEAS, DISCRETE AND CONTINUOUS DISTRIBUTIONS
   (Sept. 7 and lab)

   Reading: SW, Chapter 2, Section 8; Chapter 3, Sections 1-6
   
   i. Populations and samples
   ii. Definition of probability
   iii. Rules for obtaining probabilities: Multiplication and addition rules
   iii. Trees
   iv. Binomial distribution
   v. Normal distribution
   vi. Standard normal distribution
4. PROBABILITY – SAMPLING DISTRIBUTIONS AND THE CENTRAL LIMIT
THEOREM (Sept. 14 and lab)

Reading: SW, Chapter 3, Section 7-8; Chapter 4, Sections 1-4, Chapter 5, Sections 1-3

i. Sampling distributions
ii. Central Limit Theorem

5. STATISTICAL INFERENCE: ESTIMATION IN THE ONE-SAMPLE SITUATION
(Sept. 21)

Reading: SW, Chapter 6, Sections 1 – 7

i. t-distribution
ii. Standard errors and sampling distributions
iii. Confidence intervals for the population mean
iv. Confidence intervals for a population proportion

6. STATISTICAL INFERENCE: HYPOTHESIS TESTING IN THE ONE-SAMPLE
SITUATION (Sept. 28)

Reading: Instructor’s notes

i. Hypothesis testing for the population mean: significance levels and p-values
ii. Relationship between confidence intervals and hypothesis testing
iii. Hypothesis tests for a population proportion

7. TWO SAMPLE PROBLEMS (Oct. 5)

Reading: SW Chapter 7, Sections 1 – 7, 9 and 11; Chapter 9, Sections 1 - 3

i. Independent Samples: CI for difference in population means
ii. Independent Samples: Hypothesis tests
iii. Tests and CI for dependent samples: paired data
8. **ONE-WAY ANALYSIS OF VARIANCE (Oct. 12)**

Reading: SW, Chapter 11, Sections 1-4

i. ANOVA table and F-tests
ii. Multiple comparisons

9. **DISCRETE DATA (Oct. 19 & 26)**

Reading: SW, Chapter 10, Sections 1-10

i. Comparison of two proportions: large sample tests and confidence intervals
ii. Two-by-two contingency tables
iii. Fisher’s exact test
iv. Relative risk and odds ratio
iv. Tests of association and independence in two-way tables

10. **SIMPLE LINEAR REGRESSION (Nov. 2)**

Reading: SW, Chapter 12, Sections 1 - 7

i. Scatter plots
ii. Modeling: deterministic and stochastic
iii. Simple linear regression model: interpretation
iv. Inference concerning the coefficients
v. Correlation coefficients: Pearson and Spearman

11. **MULTIPLE REGRESSION (Nov. 9)**

Reading: Notes

i. Model
ii. Interpretation of regression parameters
iii. Diagnostics and model selection
12. LOGISTIC REGRESSION (Nov. 16 & Nov. 23)

Reading: Notes

i. Model and data
ii. Interpretation of regression parameters and odds ratios
iii. Diagnostics

13. SURVIVAL ANALYSIS (Nov. 30)

Reading: Notes

i. Basic ideas: survival curves, censored data
ii. Empirical and Kaplan-Meier estimate of survival
iii. Proportion hazards regression model