## **Boston University School of Public Health**

## **BS701 B1 Elementary Biostatistics (3 credits)**

#### Fall 2010

Instructor:	Lisa Sullivan, Ph.D. Office: Crosstown Center (801 Mass Ave) Tel: (617) 638-5047 Fax: (617) 638-6484 E-mail: <u>lsull@bu.edu</u>	, Room 324				
Class Meeting Times:	BS701 B1 Tuesdays 2:30-4:30 PM	L-110 Auditorium				
Teaching Assistant:	Carly Milliren <u>cmillire@bu.edu</u>					
<b>Review Sessions</b> :	Review sessions will be held weekly on Tuesday afternoons from 1:30-2:20 PM in room L-110 - No appointment is necessary. BS701 A1 will also hold weekly review sessions – dates and times to be determined.					
Textbook:	<i>Essentials of Biostatistics for Public Health</i> , by LM Sullivan. Jones and Bartlett Publishers, Inc. 2008. ISBN-13:978-0-7637-3737-5					
Workbook:	Essentials of Biostatistics Workbook: Statistical Computing Using Excel, by LM Sullivan. Jones and Bartlett Publishers, Inc. 2008. ISBN-13:978-0-7637-5477-8					
Course website:	BS 701 B1 ELEMENTARY BIOSTATISTICS (FALL 2010) (10FALLSPHBS701 B1)					
Calculator:	Students must have a calculator with square root, exponentiation and log Capabilities and access to Microsoft Excel.					

## **Overall Course Objectives:**

- To develop an understanding of probabilistic and statistical terminology.
- To learn how to appropriately apply statistical procedures in the public health environment.
- To learn how to interpret output from statistical computing packages to draw appropriate inferences, and to report results effectively.

## Course Requirements

*Problems Sets.* Problems sets will be assigned weekly and are contained in the course text (See course schedule). Problems are to be solved by hand and using Microsoft Excel. Selected problems will be graded. Late homework will not be accepted.

*Examinations.* There will be a midterm examination and a final examination. The dates are listed in the course outline. Examinations are closed-book. However, students may prepare a one-page note sheet which can be used during the examinations.

All electronic devices, including cell phones, PDAs, iPods, and lap top computers must be turned off (not just silenced). If a student taking an exam must be available to receive emergency calls, the proctor will take control of the telephone for the purpose of receiving such calls and notifying the student of the call.

Non-programmable calculators are allowed for examinations. Students may not use the calculator on their PDA, cell phone, or other prohibited device.

*Grades*. Course grades will be determined based on the following: Problem Sets 20%, Midterm Exam 40%, Final Exam 40%.

*Policy on Minimum Grades:* All MPH degree candidates must achieve a minimum grade of a B- in each core course. Other degree programs may have different requirements for courses core to their curriculum; please check your program guide. A degree student who fails to achieve the minimum grade in a core course must retake the course within one year and is allowed to do so at no cost, one time only, to remediate the failing grade. Ideally the course should be remediated the next semester in which it is offered so the student can make degree progress. Students who wish to retake a class must contact the Registrar's Office at least three months before class starts to have a seat reserved. Students SHOULD NOT register for the class again.

The full grade change and minimum grade requirement policies are available at: <u>http://www.bu.edu/bulletins/sph/item09.html#anchor02</u>

## Boston University School of Public Health - Information regarding Academic Honesty:

Academic honesty is essential for students to attain the competencies the University and School expect of graduates, and any action by a student that subverts these goals seriously undermines the integrity of the educational programs at the School. Students at the Boston University School of Public Health are expected to adhere to the highest standards of academic honesty.

Academic misconduct is any intentional act or omission by a student which misrepresents his or her academic achievements, or attempts to misrepresent these achievements. While not an exhaustive list, the following acts constitute academic misconduct:

- Cheating on examinations. The use or attempted use of any unauthorized books, notes or other materials in order to enhance the student's performance in the examination, copying or attempting to copy from another student's examination, permitting another student to copy from an examination or otherwise assisting another student during an examination, or any other violation of the examination's stated or commonly understood ground rules.
- Plagiarism: Any representation of the work of another as his or her own constitutes plagiarism. This includes copying or substantially restating the work of another person without the use of quotation marks or other indication that the words of another have been copied, the use of any written or oral work from which the student has obtained ideas or data without citing the source, or collaborating with another person in an academic endeavor without acknowledging that person's contribution.
- Submitting the same work in more than one course without the consent of all the instructors
- Misrepresentation or falsification of data
- Allowing another student to represent your work as his or her own
- Violating the rules of an examination or assignment

Charges of academic misconduct will be brought to the attention of the Associate Dean for Education, who will review all such cases and decide upon the appropriate action. A student who is found guilty of academic misconduct may be subject to disciplinary action, up to and including dismissal from the School.

The full academic misconduct policy is available at: <u>www.bu.edu/bulletins/sph/item09.html</u>

#### **Course Outline/Learning Objectives**

# Class 1. Introduction & Study Designs Chapters 1 & 2; Excel Workbook Chapter 1: Basics & 2: Formulas, Functions and the Data Analysis Toolpak

Introduction to Biostatistics. Students learn the difference between a population and a sample, the issues involved in designing a statistical study, collecting and analyzing data and in generating appropriate inferences. Current examples are selected from the medical and public health literature to serve as discussion points. The introduction is focused on the bigger picture but sets the stage for specific issues addressed in the course.

Study Designs. Students learn the difference between observational and randomized study designs, between prospective and retrospective study designs. The advantages and disadvantages of each type of study are discussed. Students learn the key features of case series, case-control studies, retrospective and prospective cohort designs, the randomized controlled trial and the crossover trial. Students learn different methods of sampling including simple random sampling, stratified sampling, cluster sampling and multi-stage sampling, and will be able to distinguish the different sampling schemes.

- □ Define biostatistical applications and their objectives
- □ Explain the limitations of biostatistical analysis
- □ Compare and contrast a population and a sample
- □ Explain the importance of random sampling
- Develop research questions and select appropriate outcome variables to address important public health problems
- Identify the general principles and explain the role and importance of biostatistical analysis in medical,
  public health and biological research
- □ List and define the components of a good study design
- □ Compare and contrast observational and experimental study designs
- □ Summarize the advantages and disadvantages of alternative study designs
- Describe the key features of a randomized controlled trial
- □ Identify the study designs used in public health and medical studies

#### Class 2. Summarizing Data

#### Chapter 4; Excel Workbook Chapter 4: Summarizing Continuous Variables in a Sample

Summarizing Data Collected in the Sample. Students learn to distinguish dichotomous, ordinal, categorical and continuous outcome variables. They learn to compute relative frequencies, means, medians, standard deviations, quartiles and ranges. Students also learn to construct and interpret graphical displays including relative frequency histograms, bar charts and box-whisker plots.

Learning objectives:

- Distinguish between dichotomous, ordinal, categorical and continuous variables
- □ Identify appropriate numerical and graphical summaries for each variable type
- Compute a mean, median, standard deviation, quartiles, and range for a continuous variable
- □ Construct a frequency distribution table for a discrete variable
- $\Box$  Provide an example of when the mean is a better measure of location than the median
- □ Interpret the standard deviation of a continuous outcome
- Generate and interpret a box plot for a continuous variable
- □ Produce and interpret side-by-side box plots
- Differentiate between a histogram and a bar chart

# Class 3. Probability Chapter 5 (5.1-5.5); Excel Workbook Chapter 5: Working with Probability Functions

Probability. Students learn basic principles of probability as they apply to public health and medical research. Students also learn to compute and interpret conditional probabilities. They learn to assess independence and implement Bayes' Theorem. Students will learn to compute probabilities for the Binomial and Normal probability distributions using probability tables, and to compute and interpret percentiles of the normal distribution. Students also learn the Central Limit Theorem and its application for statistical inference.

- □ Define the terms "equally likely" and "at random"
- □ Compute and interpret unconditional and conditional probabilities
- □ Evaluate and interpret independence of events

#### Class 4. Probability

#### Chapter 5 (5.6); Excel Workbook Chapter 5: Working with Probability Functions

Learning objectives:

- **□** Explain the key features of the binomial distribution model
- □ Calculate probabilities using the binomial formula
- □ Explain the key features of the normal distribution model
- □ Calculate probabilities using the standard normal distribution table
- □ Compute and interpret percentiles of the normal distribution
- □ Define and interpret the standard error
- □ Explain sampling variability
- □ Apply and interpret the results of the Central Limit Theorem

## Class 5. Confidence Interval Estimates (One Sample Procedures) Chapter 6 (6.1-6.3); Excel Workbook Chapter 6: Confidence Interval Estimates (6.1-6.2)

Confidence Interval Estimates. Students learn vocabulary associated with estimation techniques including point estimates, standard errors and confidence interval estimates. Students learn to generate confidence interval estimates for the mean and proportion of a population, for the difference in means in independent and matched populations and for the difference in proportions in independent populations in small and large sample situations. Students learn the assumptions and appropriate interpretation of confidence interval estimates using examples from published studies.

- Define point estimate, standard error, confidence level and margin of error
- □ Compare and contrast standard error and margin of error
- □ Compute and interpret confidence intervals for means and proportions

## Class 6. Confidence Interval Estimates (Two Sample Procedures) Chapter 6 (6.4-6.6); Excel Workbook Chapter 6: Confidence Interval Estimates (6.3-6.5)

Learning objectives:

- Differentiate independent and matched or paired samples
- Compute confidence intervals for the difference in means and proportions in independent samples and for the mean difference in paired samples
- Identify the appropriate confidence interval formula based on type of outcome variable and number of samples

## Class 7. MIDTERM EXAM

# Class 8. Hypothesis Testing Procedures (One and Two Samples Procedures) Chapter 7 (7.1-7.3, 7.5-7.7); Excel Workbook Chapter 7: Hypothesis Testing Procedures (7.1,7.2,7.4-7.6)

Hypothesis Testing Procedures. Students learn vocabulary associated with hypothesis testing techniques including the null and alternative hypothesis, the test statistic, the decision rule, Type I and II error rates, and exact significance levels or p-values. Students learn assumptions and appropriate interpretation of tests of hypothesis, as well as issues related to the use of historical or external comparators. Students learn to conduct tests for the mean and proportion of a population, for the difference in means in independent and matched populations and for the difference in proportions in independent populations in small and large sample situations. Students learn to conduct chi-square tests and perform analysis of variance.

- Define null and research hypothesis, test statistic, level of significance and decision rule
- Distinguish between Type I and Type II errors and discuss the implications of each
- □ Explain the difference between one and two sided tests of hypothesis
- □ Estimate and interpret p-values
- □ Explain the relationship between confidence interval estimates and p-values in drawing inferences
- Identify the appropriate hypothesis testing procedure based on type of outcome variable and number of samples

#### Class 9. Hypothesis Testing Procedures (Analysis of Variance)

```
Chapter 7 (7.8); Excel Workbook Chapter 7: Hypothesis Testing Procedures (7.7)
```

Learning objectives:

- □ Perform analysis of variance by hand and using Excel
- □ Appropriately interpret results of analysis of variance tests
- Distinguish between one and two factor analysis of variance tests
- Identify the appropriate hypothesis testing procedure based on type of outcome variable and number of samples

# Class 10. Hypothesis Testing Procedures (Chi-square tests) Chapter 7 (7.4, 7.9); Excel Workbook Chapter 7: Hypothesis Testing Procedures (7.3, 7.8)

Learning objectives:

- □ Perform chi-square tests by hand and using Excel
- □ Appropriately interpret results of chi-square tests
- Identify the appropriate hypothesis testing procedure based on type of outcome variable and number of samples

# Class 11.Power and Sample SizeChapter 8; Excel Workbook Chapter 8: Power and Sample Size Determination

Power and Sample Size Determination. Students learn to determine the sample size required to generate a precise estimate of a population parameter from a sample and to test a hypothesis with adequate statistical power.

- Provide examples demonstrating how the margin of error, effect size and variability of the outcome affect sample size computations
- □ Compute the sample size required to estimate population parameters with precision
- □ Interpret statistical power in tests of hypothesis
- □ Compute the sample size required to ensure high power in tests of hypothesis

# Class 12. Multivariable Methods Chapter 9; Excel Workbook Chapter 9: Regression Analysis

Multivariable Methods. Students learn to distinguish confounding and effect modification. They learn the assumptions and logic of analysis of relationships among variables. Students learn the logic of regression analysis and will use real data to estimate simple linear regression coefficients and to test for statistical significance. Students are introduced to multiple linear regression applications using examples from published studies.

Learning objectives:

- Define and provide examples of dependent and independent variables in a study of a public health problem
- □ Explain the principle of statistical adjustment to a lay audience
- □ Organize data for regression analysis
- □ Define and provide an example of confounding
- □ Define and provide an example of effect modification
- □ Interpret coefficients in multiple linear regression analysis

# Class 13. Introduction to Logistic Regression and Survival Analysis Methods Chapter 9; Excel Workbook Chapter 9: Regression Analysis Chapter 11: Survival Analysis (Draft Chapter)

Introduction to Logistic Regression and Survival Analysis. Students learn to distinguish linear regression, logistic regression and survival analysis methods. They learn the assumptions and logic of logistic regression analysis and survival analysis. Students are introduced to multiple logistic and survival analysis methods using examples from published studies with focus on interpretation.

- □ To distinguish between continuous, dichotomous and time to event outcomes
- □ To compute and interpret unadjusted and adjusted odds ratios
- □ To interpret coefficients in multiple logistic regression analysis

- □ To construct a life table using the actuarial approach
- □ To construct a life table using the Kaplan-Meier approach
- $\Box$  To perform and interpret the log rank test
- □ To compute and interpret a hazard ratio
- □ To interpret coefficients in Cox proportional hazards regression analysis

Class 14. Review

Class 15. FINAL EXAM

## **Course Schedule and Problem Sets**

Class	Date	Торіс	Chapters (Sections) in	Chapters: Title (Sections) in	HW Set	Problems
			Textbook	Excel Workbook	Due Date	
1	9/7	Introduction	1	1: Basics	1	Ch. 2 #1,3,5,8,10
		Study Designs	2	2: Formulas, Functions and the Data Analysis Toolpack	9/14	
2 9/14	9/14	Summarizing Data	4	3: Creating Tables and Graphs (3.1-3.3)	2	Ch. 4 #1,2,4,9,10
					9/21	EXCEL Ch 3 #1,2,3
				4: Summarizing Continuous Variables in a Sample		and Ch.4 #1,2,3
				variables in a bample		
3	9/21	Probability	5 (5.1-5.5)	5: Probability Functions	3	Ch. 5 #2,4,7,13
					9/28	
4	9/28	Probability	5 (5.6)	5: Probability Functions	4	Ch. 5 #1,3,10,11,15
					10/5	EXCEL Ch.5 #1,5
5	10/5	Confidence Interval Estimates	6 (6.1-6.3)	6: Confidence Interval Estimates	5	Ch. 6 #1,4,10
		(One Sample Procedures)		(6.1-6.2)	10/19	EXCEL Ch.6 #1
6	10/19	Confidence Interval Estimates	6 (6.4-6.6)	6: Confidence Interval Estimates	6	Ch. 6 #2,3,7,8
		(Two Sample Procedures)		(6.3-6.5)	10/26	EXCEL Ch. 6 # 2,4,5
7	10/26	Midterm Exam				
8 11/2	11/2	Hypothesis Testing Procedures	7 (7.1-7.3,7.5-7.7)	7: Hypothesis Testing Procedures (7.1,7.2,7.4-7.6)	7	Ch. 7 #4,6,7,8
		(One and Two Samples Procedures)			11/9	EXCEL Ch. 7 # 4,6
9 11/9	11/9	Hypothesis Testing Procedures	7 (7.8)	7: Hypothesis Testing Procedures	8	Ch. 7 #5,14,19
		(Analysis of Variance)		(7.7)	11/16	EXCEL Ch. 7 #3
10 11/16	//16      Hypothesis Testing Procedures	7 (7.4,7.9)	7: Hypothesis Testing Procedures	9	Ch. 7 #1,2,11	
		(Chi-square Tests)		(7.3,7.8)	11/23	EXCEL Ch. 7 # 1,2
11	11/23	Power and Sample Size	8	8: Power and Sample Size	10	Ch. 8 # 2,3,4,7
				Determination	11/30	Excel Ch.8 #2,3,4,7
12	11/30	Multivariable Methods	9	9: Regression Analysis	11	Ch. 9 NEW
					12/7	Handout
13	12/7	Logistic Regression and Survival Analysis	9			Ch. 9 NEW and
			11 (Draft)			Ch. 11 NEW
						Handout
14	12/14	Review				
15	12/21	Final Exam	-			