BIOST 2011: Principles of Statistical Reasoning
Spring 2017

Course Instructor:	Jenna C. Carlson Teaching Fellow E-mail: jnc35@pitt.edu
Course Mentor:	Ada O. Youk, PhD Associate Professor A412 Crabtree Hall Phone: 412-624-5451 E-mail: ayouk@pitt.edu
Instructor Office Hours:	TBD
Teaching Assistants:	TBD
TA Office Hours:	TBD
Class Time:	Monday 4:00 – 5:25 pm, G23 Parran Hall Wednesday 4:00 – 5:25 pm, G23 Parran Hall
Recitation:	**Register for one, either:** Monday 5:30 – 6:25 pm, G23 Parran Hall OR Thursday 1:00 – 1:55 pm, A522 Crabtree Hall
Prerequisites:	College Algebra with grade of C or better Working knowledge of PC or Mac computer ***Scientific calculator***: If you do not already own a scientific calculator, you may need to purchase one for this course. Consider a calculator with at least one memory and a variety of mathematical functions. ***Computer Software***: The software for BIOST 2011 is STATA and is available on all university supported computers in the recitations and libraries. Personal copies of STATA can be downloaded for free (or $5 if you want it on CD) through the University of Pittsburgh Software Licensing Services.
Required text:	Introductory Applied Biostatistics by Ralph D'Agostino, Sr., Lisa Sullivan and Alexa Beiser; ISBN # 978-0-534-42399-5
Recommended Text:	Using Stata for Quantitative Analysis by Kyle C. (Clayton) Longest; ISBN # 978-1483356631

Catalogue Description

This course acquaints students with the concepts of statistical reasoning as applied to the study of Public Health problems. Students learn the general principles of statistical analysis and acquire the ability to utilize a statistical software package (STATA) as a tool to facilitate the processing, editing, storing, displaying, analysis and interpretation of health research related data.

Course rationale

This is the Biostatistics core course for Graduate School of Public Health for non-Biostatistics majors and provides a basic introduction to the concepts of statistical reasoning as applied to the study of public health problems. This course is designed for public health students that expect to primarily to be able to read and understand statistical procedures in the form of books, journal articles, reports, grants, etc. The course will also give students the ability to perform some basic analyses. Students who intend to be professional research workers in public health areas requiring the daily application of quantitative procedures and statistics should consider taking BIOST 2041 and 2042 (Introduction to Statistical Methods I and II).

Course Objectives

By the end of this course, each student will be able to:

- Identify the appropriate statistical procedures to be applied in different public health situations, especially in social science research.
- Identify advantages and challenges of working with different types of data from real public health examples, including ordinal, scale, and quantitative variables. Perform exploration of data, including descriptive tables and plots, basic transformations (e.g. categorizing a continuous variable), and inter-rater reliability, and be able to interpret the findings.
- Use statistical software to conduct hypothesis tests, including t-tests, ANOVAs, linear regression, analyses of proportions as well as the non-parametric versions of these tests.
- Perform (using statistical software) and interpret analyses such as computing confidence intervals, sensitivity and specificity analyses, simple linear regression models and basic survival analyses.

MPH Competencies

This course will help students to meet the Biostatistics competencies developed by the Association of Schools of Public Health (ASPH):

- Define commonly used statistical terminology.
- Demonstrate the ability to correctly select the most appropriate statistical procedures for given research hypotheses and types of data.
- Demonstrate the ability to interpret the results of statistical analysis given the results of a statistical analysis.

Ground Rules for Class

- Be respectful of your fellow students, teaching assistants, and instructors. This includes being on time and refraining from using your cell phone during class.
- Students may work together on homework assignments, but each student must submit their own homework. Copying of another student's assignment will NOT be tolerated.

How to Succeed in this Class

BIOST 2011 can be a challenging course. Here are some important things to remember to keep on track:

- Learning statistics is similar to learning a new language – it is done over time and with lots of practice! Come to class and attend your recitation. Just reviewing the course handouts via Courseweb will not be sufficient.
- One of the most important things to do for success is to not get behind on readings and assignments. The course content is cumulative so if you get behind, it is very difficult to catch up.
- Although recitation attendance is optional, it is strongly encouraged. This is your opportunity to get feedback and help from the TA and solidify the concepts presented in lecture.
- Utilize the instructor and TA office hours.
- Do all of the homework.
- Communicate!

Course Website

All readings and course material will be found on the Blackboard site for this class. The website for Blackboard is http://courseweb.pitt.edu. Your login ID and password are the same as for your Pitt account.

Course Requirements

- Lecture attendance.
- Satisfactory completion and submission of all required assignments.
- Satisfactory performance on three exams.

Student Performance Evaluation

Course grades are based on performance on:

- Exams (60%)

  You will take three in-class exams, each worth 20% of the overall grade. Exams are closed book and closed note. You should bring a calculator. The use of computers, cell phones or other internet-attached devices will **NOT** be permitted during exams. The exams may consist of true/false, multiple choice, and short answer questions. You will be allowed one sheet of 8.5"x11" (both sides) of handwritten notes. This notesheet will be turned in with the exam. Tentative dates for the exams are:

  Exam #1 – Monday, February 6

Exam #2 – Monday, March 20
Exam #3 – Wednesday, April 26

These exams cannot be taken early or late without a compelling reason and supporting documentation. Students who cannot be present on the day of an exam will be required to take the exam on the first earlier available date. Notification must be given to the instructor in advance.

- Homework (30%)

  There will be eight homework assignments, graded for accuracy and completion. Homework assignments are due at the start of class on the due date, unless otherwise stated. No credit will be given for late assignments. Note: only the top six scores will count toward your grade, but I may consider how many assignments you completed when deciding between "borderline" grades (e.g. having turned in all assignments on time may make the difference between a B+ and an A-).

- Discussions (10%)

  There will be seven discussion posts on blackboard, graded for completion. Discussion posts will be assigned and due via the course website by 11:59pm on the due date, unless otherwise stated.

Grade scale

98-100%	A+
90-97%	A
89%	A-
85-88%	B+
80-84%	B
79%	B-
70-78%	C
<70%	F

Academic Integrity

All students are expected to adhere to the school's standards of academic honesty. Any work submitted by a student for evaluation must represent his/her own intellectual contribution and efforts. The GSPH policy on academic integrity, approved by EPCC on 10/14/08, which is based on the University policy, is available online at http://www.publichealth.pitt.edu/Portals/0/Main/Prospective%20Students/Academics/Pitt%20Public%20Health%20Academic%20Handbook-Part%20IV%20G_AcademicIntegrity_AY2014-15.pdf

These guidelines are based on the University policy found here: http://www.provost.pitt.edu/info/acguidelinespdf.pdf

The policy includes obligations for faculty and students, procedures for adjudicating violations, and other critical information. Please take the time to read this policy.

Students committing acts of academic dishonesty, including plagiarism, unauthorized collaboration on assignments, cheating on exams, misrepresentation of data, and facilitating dishonesty by others, will receive sanctions appropriate to the violation(s) committed. Sanctions include, but are not limited to, reduction of a grade for an assignment or a course, failure of a course, and dismissal from GSPH.

All student violations of academic integrity must be documented by the appropriate faculty member; this documentation will be kept in a confidential student file maintained by the GSPH Office of Student Affairs. If a sanction for a violation is agreed upon by the student and instructor, the record of this agreement will be expunged from the student file upon the student's graduation. If the case is referred to the GSPH Academic Integrity Hearing Board, a record will remain in the student's permanent file.

Accommodation for Students with Disabilities

If you have any disability for which you may require accommodation, you are encouraged to notify both your instructor and the Office of Disability Resources and Services, 140 William Pitt Union (412-648-7890) during the first two weeks of the term (http://www.studentaffairs.pitt.edu/drswelcome).

Video / Audio recording of class lectures

Audio recording of the class is permissible provided you first receive approval from the course instructor. You also agree that the recording is for your own personal use and will not be redistributed in any form.
Video recording of the class, in full or in part, is NOT permitted.

Copyright of course material

Unless otherwise stated all course material is protected by copyright. United States copyright law, 17 USC section 101, et seq., in addition to University policy and procedures, prohibit unauthorized duplication or retransmission of course materials. See Library of Congress Copyright Office and the University Copyright Policy. As such the material is to be used for academic purposes only. Redistribution of this material to web sites and repositories (e.g., Course Hero) is strictly prohibited.

Course Calendar

Date	Topic	Corresponding Reading (D'Agostino)	Lecture - Mon	Lecture - Wed
Week 1 Jan 2 - Jan 6	Intro and Motivation	1.1-1.4		
Week 2 Jan 9 - Jan 13	Summarizing Data, Probability	2.1-2.3, 3.1-3.3		**HW #1 due**
Week 3 Jan 16 - Jan 20	Probability	3.4-3.6	MLK holiday	**DP #1 due**
Week 4 Jan 23 - Jan 27	Sampling Distributions, Intro to Hypothesis Testing	4.1-4.3		**HW #2 due**
Week 5 Jan 30 - Feb 3	Statistical Inference: Procedures for μ	5.1-5.3	**DP #2 due**	**HW #3 due**
Week 6 Feb 6 - Feb 10	Statistical Inference: Procedures for $\mu_1 - \mu_2$	6.1-6.3	**EXAM #1***	
Week 7 Feb 13 - Feb 17	Categorical Data	7.1-7.3	**DP #3 due**	**HW #4 due**
Week 8 Feb 20 - Feb 24	Categorical Data	7.4-7.7		
Week 9 Feb 27 - Mar 3	Comparing Risks in Two Populations	8.1-8.4	**DP #4 due**	**HW #5 due**
Week 10 Mar 6 - Mar 10	spring break			
Week 11 Mar 13 - Mar 17	Analysis of Variance	9.1-9.7		
Week 12 Mar 20 - Mar 24	Correlation and Regression - 1	10.1-10.2	**EXAM #2****	**HW #6 due**
Week 13 Mar 27 - Mar 31	Correlation and Regression - 2	10.3-10.5	**DP #5 due**	
Week 14 Apr 3 - Apr 7	Logistic Regression Analysis	11.1-11.5		**HW #7 due**
Week 15 Apr 10 - Apr 14	Nonparametric Tests	12.1-12.6	**DP #6 due**	
Week 16 Apr 17 - Apr 21	Introduction to Survival Analysis	13.1-13.3		**HW #8 due**
Week 17 Apr 24 - Apr 28	Review		**DP #7 due**	**EXAM #3*****

Please note: recitations are not included on this schedule but will be necessary to solidify concepts introduced in lecture and to help with homework.

*** Exam #1 covers material from weeks 1-5**
**** Exam #2 covers material from weeks 6-11**
***** Exam #3 covers material from weeks 12-16**

Course Objective Outline

Introduction and Motivation

- Understand the purpose and goals of this course
- Be able to explain why statistics are used and what types of conclusions they provide
- Be able to explain the difference between population and sample parameters
- Be able to identify population, sample, and hypothesis of a research study

---

Summarizing Data

- Learn to distinguish between continuous and discrete variables
- Be able to describe the shape, center, and spread of data
- Be able to compute (using Stata) and interpret numerical summaries of data
- Be able to generate (using Stata) and interpret graphical displays to summarize data

---

Probability

- Understand the concept of probability
- Be able to compute simple probabilities
- Understand the rules of probability
- Understand the difference between combinations and permutations
- Be able to identify important distributions

Sampling Distributions

- Know the definition of a sampling distribution
- Understand the Central Limit Theorem
- Be able to compute statistics for certain sampling distributions

---

Introduction to Hypothesis Testing

- Identify the key steps of a hypothesis testing procedure
- Be able to set up statistical hypotheses based on a research question

---

Statistical Inference for μ

- Learn how to estimate a population mean μ
- Learn how to compute and interpret a confidence interval for μ
- Understand the concept of precision
- Be able to perform hypothesis tests for μ

Statistical Inference for $\mu_1 - \mu_2$

- Learn how to compare means from two different populations under different scenarios
- Distinguish between independent and dependent observations

Categorical Data

- Learn how to estimate a population proportion and the difference in two population proportions
- Learn how to compute and interpret a confidence interval for p and for p_1-p_2
- Learn how to perform hypothesis testing for the comparison of two populations proportions

- Be able to create cross-tabulation tables for discrete variables
- Be able to compute and understand sensitivity and specificity measures
- Learn when to use a chi-square test
- Be able to distinguish between goodness of fit tests and tests of independence

---

Comparing Risks in Two Populations

- Be able to define and interpret an effect measure
- Be able to compute and interpret a confidence interval for an effect measure
- Learn how to use Chi-square tests of homogeneity and Fisher's exact tests

---

Analysis of Variance

- Learn how to compare more than two means using ANOVA
- Be able to distinguish between a fixed and random effect
- Understand the concept of multiple comparisons and when they apply

---

Correlation and Regression

- Learn how to estimate a population correlation coefficient ρ
- Be able to compute and interpret a sample correlation coefficient r
- Be able to perform hypothesis tests for ρ
- Learn the assumptions that must be met for linear regression
- Learn how to test these assumptions
- Be able to fit a simple linear regression model and interpret the coefficients
- Be able to fit a multiple linear regression model and interpret the coefficients

---

Logistic Regression Analysis

- Know when it is appropriate to use logistic regression
- Be able to fit a simple logistic regression model and interpret the coefficients
- Be able to generate and interpret an ROC curve

Nonparametric Tests

- Learn to use the nonparametric counterpart tests of the tests we've discussed in earlier lectures

---

Introduction to Survival Analysis

- Be able to generate and interpret a Kaplan-Meier curve and compute median survival time

---