Statistical Methods in Public Health I

Biostatistics 140.621

September 1 - October 22, 2015

Department of Biostatistics
Johns Hopkins University
Bloomberg School of Public Health

Instructors:

Marie Diener-West, PhD
Karen Bandeen-Roche, PhD

COURSE INFORMATION

STATISTICAL METHODS IN PUBLIC HEALTH I (140.621)
FIRST TERM

September 1 - October 22, 2015

Faculty Lecturers: Marie Diener-West, PhD (Section 140.621.01)
Office E-3622, 410-502-6894, mdiener@jhu.edu

Karen Bandeen-Roche, PhD (Section 140.621.02)
Office E-3527, 410-955-3067, kbandee1@jhu.edu

Department of Biostatistics
Johns Hopkins University
Bloomberg School of Public Health

Lectures: 10:30 a.m. - 12:00 p.m. – Tuesday, Thursday
Sommer Hall (E2014)- Section 140.621.01
Sheldon Lecture Hall (W1214)- Section 140.621.02
Overflow Rooms: W4030 and W5030

Lab 140.921 for review of material through a structured exercise and time for questions:

Lab 01 -	1:30 p.m. - 3:00 p.m. – Monday	W3008
Lab 02 -	1:30 p.m. - 3:00 p.m. – Tuesday	W3008
Lab 03 -	1:30 p.m. - 3:00 p.m. – Wednesday	W3008
Lab 04 -	1:30 p.m. - 3:00 p.m. – Thursday	W3008
Lab 05 -	1:30 p.m. - 3:00 p.m. – Friday	W3008
Lab 06 -	3:30 p.m. - 5:00 p.m. – Monday	W3008
Lab 07 -	3:30 p.m. - 5:00 p.m. – Tuesday	W3008
Lab 08 -	3:30 p.m. - 5:00 p.m. – Wednesday	W3008
Lab 09 -	3:30 p.m. - 5:00 p.m. – Thursday	W3008

Open time in lab with two lab instructors for questions: 3:00 p.m – 3:30 p.m. on Monday through Friday

COURSE INFORMATION

STATISTICAL METHODS IN PUBLIC HEALTH I (140.621)
FIRST TERM
September 1 - October 22, 2015
(continued)

Lab Instructors:
Karen Bandeen-Roche, PhD
Marie Diener-West, PhD
Jean-Philippe Fortin (Lead TA)
Emily Huang (Lead TA)
Prasad Patil (Lead TA)

Teaching Assistants:
Yibing (Oliver) Chen
Yu Du
Youssef Farag
Jordan Johns
Shuiqing Liu
Yi-Chen Liu
Nick Lu
Gina Norato
Yuting Xu
Chao Yang
Claire Ruberman

Teaching Assistant Office Hours (starting Thursday, September 3, optional):

Monday through Friday	12:15 p.m. - 1:15 p.m.	W2009

Stata Office Hours in Computer Lab (starting Thursday, September 3, optional):

Monday through Friday	2:30 p.m. - 3:20 p.m.	W3017

COURSE INFORMATION

STATISTICAL METHODS IN PUBLIC HEALTH I (140.621)
FIRST TERM
September 1 - October 22, 2015

(continued)

Lecture Notes: Copies of the course materials are distributed during class. Purchase of these materials is included in registration. Copies of most materials are available for downloading in the "Classes"section of the course web site.

Web Site:

Available through CoursePlus or http://www.biostat.jhsph.edu/courses/bio621/
Contains course schedule, office hours, lecture notes, self-evaluation problems, Stata lecture notes, problem set solutions, and quiz and exam solutions.

Audio files: An audio lecture is available and posted after each lecture on the course website in the "Classes"section.

Recommended Book: Rosner, B.
Fundamentals of Biostatistics
2011, Duxbury,Thomson Brooks/Cole, Belmont, California

Suggested Book: Lawrence C. Hamilton
Statistics with Stata 12
2012, Duxbury,Thomson Brooks/Cole, Belmont, California

Calculator: Basic functions (+, -, ×, ÷), logarithms and exponents, simple memory and recall, factorial key.

Statistical Computing Package: Stata 14 Intercooled, Stata Press, College Station, Texas
(Buy through http://www.stata.com/order/new/edu/gradplans/campus-gradplan/)

COURSE INFORMATION

STATISTICAL METHODS IN PUBLIC HEALTH I (140.621)
FIRST TERM

September 1 - October 22, 2015

(continued)

Course Policies:

- Attendance is required for quizzes and exams and expected for lectures and labs.
- Laptops and iPads may be used during lecture for class-related purposes. Common courtesy should be followed.
- Please email your faculty lecturer regarding extenuating circumstances or conflicts regarding course deadlines.
- Availability for course questions: after lecture, during labs, TA office hours, and Stata office hours.
- Course Grade based on:
  20% completion of 4 problem sets (points deducted if turned in late)
  5% quiz 1 (through CoursePlus)
  5% quiz 2 (through CoursePlus)
  35% midterm examination (in class)
  35% final examination (in class)
- Quizzes and examinations are individual work for which a student must work by himself or herself.
- Problem sets may be worked on together and discussed. However, each student must write up the problem set individually using his or her own words. Copying work is not allowed.
- Disability Support Services
  If you are a student with a documented disability who requires an academic accommodation, please contact Betty H. Addison in the Office of Career Services and Disability Support: dss@jhsph.edu, 410-955-3034, or Room E-1140.

COURSE INFORMATION

STATISTICAL METHODS IN PUBLIC HEALTH I (140.621)
FIRST TERM

September 1 - October 22, 2015

(continued)

- Academic Ethics Code
  The code, discussed in the Policy and Procedure Memorandum for Students, March 31, 2002, will be adhered to in this class (https://my.jhsph.edu/Resources/PoliciesProcedures/ppm/PolicyProcedureMemoranda/Students_01_Academic_Ethics.pdf)

- Students enrolled in the Bloomberg School of Public Health of The Johns Hopkins University assume an obligation to conduct themselves in a manner appropriate to the University's mission as an institution of higher education. A student is obligated to refrain from acts which he or she knows, or under the circumstances has reason to know, impair the academic integrity of the University.

Course Objectives:
Students who successfully master this course will be able to:

1. Use statistical reasoning to formulate public health questions in quantitative terms:
   (a) Understand the role of statistical reasoning within the scientific model.
   (b) Understand and apply the counterfactual definition of cause in public health research.
   (c) Distinguish between continuous, categorical, binary and time-to-event data.
   (d) Understand that evidence for establishing an association between a risk factor and health outcome is generated by comparing the distribution of the outcome in otherwise similar populations with different levels of the risk factor.
   (e) Use stratification in design and analysis to minimize confounding and identify risk modification

2. Design and interpret graphical and tabular displays of statistical information:
   (a) Create by hand and interpret stem and leaf plots, boxplots, Q-Q plots and frequency tables.
   (b) Use the statistical analysis package Stata to make basic statistical computations and graphical displays.
   (c) Graphically compare two groups of observations of otherwise similar units (e.g. people or treatments) and interpret the display.
   (d) Characterize the distribution of a variable using the concepts of typical value, variability, and shape.
   (e) Use a variable transformation, such as the logarithm, to study a right skewed distribution such as hospital costs.
   (f) Analyze survey information and identify sources of error in variables.
   (g) Explore study results for associations among multiple variables and interpret the findings.
3. Use probability models to describe trends and random variation in public health data
   (a) Use the statistical analysis package Stata to make basic statistical computations and graphical displays.
   (b) Use the concepts of probability to describe the effect of a treatment on a health outcome in a randomized trial.
   (c) Use the binomial distribution and the Poisson approximation to the binomial to calculate probabilities of events.
   (d) Use the Gaussian or normal probability model to approximate the distribution of a continuous public health measure and to assess the quality of this approximation;
   (e) Use a quantile-quantile (Q-Q) plot to compare the shape of an empirical with a theoretical distribution.

4. Use statistical methods for inference, including tests and confidence intervals, to draw public health inferences from data:

   (a) Generate random numbers and appreciate variation among multiple observations of a random process.

   (b) Explain the implications of the Central Limit Theorem in determining the sampling distribution of the mean of n observations.

   (c) Use bootstrapping to determine confidence intervals.

   (d) Use sampling distribution theory for the mean and for differences between two means to create confidence intervals and hypothesis tests.

   (e) Use stratification to eliminate the influence of a possible confounding variable in a study of the association of a risk factor and outcome.

   (f) Use the appropriate two-sample t-test and confidence interval to assess whether average outcome is different between two groups and draw inferences.

   (g) Use the paired-sample t-test and confidence intervals to assess if the average change is different from zero.

   (h) Examine the consequence of using an inappropriate unpaired (two-sample) analysis when a paired analysis is appropriate.

   (i) Define and apply the term effect modification or equivalently interaction in a randomized trial.

The course is designed to enable students to develop their data analysis skills. Four important datasets will be analyzed by the students using the statistical package Stata throughout the 621-624 course sequence.

Class	Date	Topic	Suggested Reading*
1	Sept 1	Statistical Reasoning in Public Health	Chapter 1
2	Sept 3	Exploring and Organizing Data to Address Public Health Questions:Continuous and Discrete Data	Chapter 2.1 - 2.5 Chapter 2.8-2.9
3	Sept 8	Exploring and Organizing Data (continued) Probability Concepts and Distributions	Chapter 3
4	Sept 10	Binomial and Poisson Distributions **PROBLEM SET 1 DUE**	Chapter 4
	Sept 11	**QUIZ 1**	
5	Sept 15	Binomial and Poisson Distribution (cont'd)	Chapter 5
6	Sept 16	Summary and Review **PROBLEM SET 2 DUE**	
7	Sept 22	**MIDTERM EXAMINATION**	
8	Sept 24	Normal Distribution; Populations and Samples Introduction to Statistical Inference: Sampling Distributions - Single Sample Mean	Chapter 6.1-6.4 Chapter 6.5
9	Sept 29	- Difference between Two Sample Means - Single Sample Proportion - Difference between Two Sample Proportions	
9	Oct 1	Introduction to Statistical Inference: The Bootstrap	
11	Oct 6	Estimation Hypothesis Testing **PROBLEM SET 3 DUE**	
12	Oct 8	Confidence Intervals and Hypothesis Tests: Single Sample Mean **QUIZ 2**	 Chapter 7.1-7.4,7.7
13	Oct 13	Confidence Intervals and Hypothesis Tests: Difference between Two Sample Means Pre-Post Designs and Other Paired Comparisons	 Chapter 8.4-8.7 Chapter 8.2
14	Oct 15	Confidence Intervals and Hypothesis Tests: Estimating a Proportion in a Single Population Comparing Proportions from Two Populations	 Chapter 7.10 Chapter 10.2
15	Oct 20	Summary and Review **PROBLEM SET 4 DUE**	
16	Oct 22	**FINAL EXAMINATION**	

* Fundamentals of Biostatistics by Rosner (2011)