Statistical Methods in Public Health II

Biostatistics 140.622

October 27 - December 17, 2015

Department of Biostatistics
Johns Hopkins University
Bloomberg School of Public Health

Instructors:

Marie Diener-West, Ph.D.
Karen Bandeen-Roche, Ph.D.

COURSE INFORMATION

STATISTICAL METHODS IN PUBLIC HEALTH II (140.622)
SECOND TERM

October 27 - December 17, 2015

Faculty: **Marie Diener-West, PhD (Section 140.622.01)**
Office E-3622, 410-502-6894, mdiener@jhu.edu

Karen Bandeen-Roche, PhD (Section 140.622.02)
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 140.622.01
Sheldon Lecture Hall (W1214)- Section 140.622.02
Overflow Rooms - W3030 and W4030

Lab 140.922 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 II (140.622)
SECOND TERM
October 27 - December 17, 2015
(continued)

Lab Instructors:
Karen Bandeen-Roche, PhD
Marie Diener-West, PhD
Leonardo Collado Torres (Lead TA)
Aaron Fisher (Lead TA)
David Lenis (Lead TA)
Prasad Patil(Lead TA)

Teaching Assistants:
Yibing (Oliver) Chen
Yu Du
Youssef Farag
Emily Huang
Jordan Johns
Shuiqing Liu
Yi-Chen Liu
Haidong Lu
Gina Norato
Claire Ruberman
Genevieve Stein-O'Brien
Yuting Xu
Chao Yang

Teaching Assistant Office Hours (starting Tuesday, October 27 optional):

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

Stata Office Hours in Computer Lab (starting Tuesday, October 27 optional):

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

COURSE INFORMATION

STATISTICAL METHODS IN PUBLIC HEALTH II (140.622)
SECOND TERM

October 27 - December 17, 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/bio622/ 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
2013, Duxbury,Thomson Brooks/Cole, Belmont, California

Calculator: Basic functions (+, -, $\times$, $\div$), 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 II (140.622)
SECOND TERM

October 27- December 17, 2015

(continued)

Course Policies:

- **Attendance is required for 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 section 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)
- **Contact your faculty section lecturer if you have a conflict, illness, or other issue.**
- **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 II (140.622)
SECOND TERM

October 27- December 17, 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) Use the statistical analysis package Stata to construct statistical tables and graphs of journal quality.

3. Use probability models to describe trends and random variation in public health data:

   (a) Distinguish among the underlying probability distributions for modeling continuous, categorical, binary and time-to-event data.

   (b) Calculate the sample size necessary for estimating either a continuous or binary outcome in a single group.

   (c) Estimate the sample size necessary for determining a statistically significant difference in either a continuous or binary outcome between two groups.

   (d) Recognize the assumptions required in performing statistical tests assessing relationships between an outcome and a risk factor.

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

   (a) Estimate two proportions and their difference, and confidence intervals for each. Interpret the interval estimates within a scientific context. Recognize the importance of other sources of uncertainty beyond those captured by the confidence interval

   (b) Estimate an odds ratio or relative and its associated confidence interval. Explain the difference between the two and when each is appropriate.

   (c) Perform and interpret one-way analysis of variance to test for differences in means among three or more populations. Evaluate whether underlying probability model assumptions are appropriate.

COURSE INFORMATION

STATISTICAL METHODS IN PUBLIC HEALTH II (140.622)
SECOND TERM

October 27- December 17, 2015

(continued)

5. Express the relationship between a response and an independent variable using a generalized linear model:

- **(a) Formulate and correctly interpret relationships in a linear or logistic model.**
- **(b) Estimate the regression coefficients and their associated confidence intervals and interpret them**
- **(c) Assess whether the relationship between a response variable and an independent variable varies by the level of a second independent variable using analysis of covariance.**

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.

COURSE OUTLINE AND READINGS

October 27- December 17, 2015

Class	Date	Topic	Suggested Readi
1	Oct 27	Statistical Comparisons - Regression Ideas	
2	Oct 29	Introduction to Generalized Linear Models: Linear, Logistic, Poisson, Survival Models	
3	Nov 3	Linear Regression for Comparing Multiple Groups: ANOVA	516-522
4	Nov 5	Partitioning Variation: Tests of Hypotheses PROBLEM SET 1 DUE	522-538
5	Nov 10	Simple Linear Regression for a Continuous Variable QUIZ 1	427-465
6	Nov 12	Multiple Linear Regression	468-474
7	Nov 17	Linear Regression for Comparing Groups: Dependence in a Covariate: ANCOVA	553-555
8	Nov 19	Review Lecture PROBLEM SET 2 DUE	
9	Nov 24	MIDTERM EXAMINATION	
10	Nov 26	No class- Thanksgiving	
11	Dec 1	Logistic Regression for Binary Responses: 2x2 Tables and More	391-394, 377-380
12	Dec 3	Logistic Regression: ANCOVA Revisited	
r 13	Dec 8	Sample Size PROBLEM SET 3 DUE	233-234
14	Dec 10	Sample Size (cont'd) QUIZ 2	301-304, 381-389
15	Dec 15	Review Lecture PROBLEM SET 4 DUE	
16	Dec 17	FINAL EXAMINATION	

* Fundamentals of Biostatistics by Rosner (2011)