# Statistical Methods in Public Health III

Biostatistics 140.623

**January 19 - March 10, 2016**

**Department of Biostatistics**
**Johns Hopkins University**
**Bloomberg School of Public Health**

**Instructors:**

Marie Diener-West,PhD
John Mc Gready,PhD

# COURSE INFORMATION
# STATISTICAL METHODS IN PUBLIC HEALTH III (140.623)
# THIRD TERM
**January 19 - March 10, 2016**

**Instructors:** **Marie Diener-West, Ph.D. (Section 140.623.01)**
**Office E3622, 410-502-6894, mdiener@jhu.edu**

**John McGready, Ph.D. (Section 140.623.02)**
**Office E3543, 410-614-9405, jmcgrea1@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 140.623.01**
**Sheldon Lecture Hall (W1214)- Section 140.623.02**

**Lab 140.923 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 for questions: 3:00 p.m – 3:30 p.m. Monday through Friday**

**COURSE INFORMATION**
**STATISTICAL METHODS IN PUBLIC HEALTH III (140.623)**
**THIRD TERM**
**January 19 - March 10, 2016**
**(continued)**

**Lab Instructors:**

**Marie Diener-West, PhD**
**John McGready, PhD**
**Aaron Fisher (Lead TA)**
**Jean-Philippe Fortin (Lead TA)**
**David Lenis(Lead TA)**
**Elizabeth Sweeney (Lead TA)**

**Teaching Assistants:**

**Yibing Chen**
**Yu Du**
**Weixiang Fang**
**Emily Huang**
**Jordan Johns**
**Youjin Lee**
**Lu Li**
**Shuiqing Liu**
**Haidong Lu**
**Leslie Myint**
**Huitong Qiu**
**Claire Ruberman**
**Genevieve Stein-O'Brien**
**Yuting Xu**
**Chao Yang**

**Teaching Assistant Office Hours (starting Wednesday, January 20, optional):**

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

**Stata Office Hours in Computer Lab (starting Wednesday, January 21, optional):**

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

**COURSE INFORMATION**
**STATISTICAL METHODS IN PUBLIC HEALTH III (140.623)**
**THIRD TERM**
**January 19 - March 10, 2016**
**(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/bio623/ 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 III (140.623)**
**THIRD TERM**
**January 19 - March 10, 2016**
**(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 (via Quiz Generator in CoursePlus)
  5% quiz 2 (via Quiz Generator in CoursePlus)
  35% midterm examination (in class)
  35% final examination (in class)
- Contact your 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 III (140.623)**
**THIRD TERM**
**January 19 - March 10, 2016**
**(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) **Critique a proposed public health hypothesis to determine its suitability for testing using regression methods and the available data.**
   (b) **Formulate and correctly interpret a multivariable linear, logistic or survival regression model to estimate a health effect while minimizing confounding and identifying possible effect modification.**
   (c) **Evaluate the limitations of observational data as evidence for a health effect**
   (d) **Appreciate the importance of relying upon many regression models to capture the relationships among a response and predictor in observational studies.**

2. Conduct statistical computations and construct graphical and tabular displays for regression analysis :

   (a) Use the statistical analysis package Stata to perform multivariable regression models.

   (b) Document and archive the steps of your statistical analysis by creating a Stata do-file.

   (c) Create and interpret scatter-plots and adjusted variable plots that display the relationships among an outcome and multiple risk factors.

   (d) Create and interpret tables of regression results including unadjusted and adjusted estimates of coefficients with confidence intervals from many models.

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

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

   (b) Recognize the key assumptions underlying a multivariable regression model and judge whether departures in a particular application warrant consultation with a statistical expert.

**COURSE INFORMATION**
**STATISTICAL METHODS IN PUBLIC HEALTH III (140.623)**
**THIRD TERM**
**January 19 – March 10, 2016**
**(continued)**

4. **Use statistical methods for inference in multiple regression to draw valid public health inferences from data:**
   (a) **Conduct a simple linear, logistic or survival regression and correctly interpret the regression coefficients and their confidence interval.**
   (b) **Conduct a multiple linear, logistic or survival regression and correctly interpret the coefficients and their confidence intervals.**
   (c) **Examine residuals and adjusted variable plots for inconsistencies between the regression model and patterns in the data and for outliers and high leverage observations.**
   (d) **Fit and compare different models to explore the association between outcome and predictor variables in an observational study.**

**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**
**STATISTICAL METHODS IN PUBLIC HEALTH III (140.623)**
**THIRD TERM**

**January 19 – March 10, 2016**

| Class | Date | Topic | Optional Reading* |
|---|---|---|---|
| 1 | Jan 19 | Multiple linear regression;<br>Adjusted variable plots<br>Linear spline models<br>Analysis of Variance revisited | 516-522 |
| 2 | Jan 21 | Analysis of Covariance revisited | 553-555 |
| 3 | Jan 26 | MLR: model selection and checking | 468-483 |
| 4 | Jan 28 | MLR example in detail | |
| 5 | Feb 2 | Logistic regression (LR) revisited<br>PROBLEM SET 1 DUE | 377-380<br>628-643 |
| 6 | Feb 4 | LR example in detail<br>QUIZ 1 | |
| 7 | Feb 9 | Propensity scores to control for potential confounding | |
| 8 | Feb 11 | Review Lecture<br>PROBLEM SET 2 DUE | |
| 9 | Feb 16 | MIDTERM EXAMINATION | |
| 10 | Feb 18 | Poisson regression for event counts and exposure;<br>person-years | |
| 11 | Feb 23 | Survival analysis | 758-767 |
| 12 | Feb 25 | Survivor analysis (continued) | |
| 13 | Mar 1 | Log-rank statistic for comparing survival curves<br>Proportional hazards regression (PHR)<br>PROBLEM SET 3 DUE | 767-773<br>774-783 |
| 14 | Mar 3 | PHR example in detail<br>QUIZ 2 | |
| 15 | Mar 8 | Review Lecture<br>PROBLEM SET 4 DUE | |
| 16 | Mar 10 | FINAL EXAMINATION | |

* Fundamentals of Biostatistics by Rosner 2012