PUBH 741 (2017) Section 001
Quantitative Methods for Health Care Professionals

Instructor: Kim Faurot, PA, MPH, PhD               email: faurot@med.unc.edu
Teaching Assistants:   Henry Zhang                     email: htzhang@live.unc.edu
                       Jessica Islam                    email: islamjy@email.unc.edu

Assigned Readings: Recorded copies of the class lectures and the accompanying pdf files are posted on Sakai under “Course Documents” and “Lectures”; you will need to listen to the lectures before you come to class so that you will be prepared for class exercises. Also read the chapter in the textbook prior to beginning the exercise. See syllabus below for assigned lectures and textbook chapters. Journal articles you will need to read for the exercises are listed on the syllabus and posted on Sakai.

Required Texts:
H. Bush: Biostatistics- an Applied Introduction for the Public Health Practitioner. Although this book is intended for a two-semester class, we will cover most of the material this semester.
This book can be purchased as an eTextbook for just $43—a really good deal for a statistics textbook.

Suggested Texts:
Stanton A. Glantz: Primer of Biostatistics
Juul & G Frydenberg: An Introduction to Stata for Health Researchers

Enhancements:
It is impossible to teach in depth the wide variety of statistics that you will need to understand the medical literature. Hence, to enhance your learning, especially for those of you who want to understand concepts more deeply, we have embedded links to material on the internet that we have vetted. Please feel free to make use of these materials.

If we find additional good examples online during the semester, we will add them as suggested enhancements.

Schedule:
   Time: Tues/Thurs 9:30-11:10
   Location: MacNider 321

August 9, 2017
Class structure: the class consists of recorded lectures, readings, readiness quizzes, demonstrations, and class exercises. The readiness quizzes enable students to ensure that they understand the lecture and textbook material sufficiently well to apply it to class exercises. Before each exercise, the instructor and/or TA will demonstrate how to perform a similar analysis in Stata. Additional demonstrations will supplement the recorded lecture material.

Individuals in the class will sort themselves into groups on the first day of class. We will ask you to choose group mates to maximize the variability in comfort and experience with biostatistics. You may confer with group members in completing your exercise, but we strongly suggest that you attempt each problem individually, and then compare your answers to check on your results! For each exercise, one member of your group will upload the group portion of the exercise into Sakai. You will also turn in individual reports based on the exercises. Each group has the opportunity present an analysis of the statistical section of an article. Quizzes are designed to help you identify what you need to learn more about; you will have immediate feedback and can change your answers to receive full credit (low stakes). The individual reports and final must be completed without help from your group. The final exam is designed to let you pull together everything you have learned in the semester (exercises, lectures, readings) and apply it to an analysis. Your instructor and TAs will be available to answer questions during class periods and by appointment.

Grades: Exercises (group): (5*4 points) =20 points
Reports (individual): (4*5 points) = 20 points
Presentations (group): 8 points (Critical review of an article, peer-reviewed)
Quizzes: (8*1.5 points) = 12 points (Immediate feedback, low stakes)
Final Exam:  40 points (Comprehensive Analysis Report)
You must earn 95 points or greater to achieve an H.  P:  80-94, L:  70-79, F: <70

Course Goals:
At the end of this course, you should be able to:
1) REMEMBER:
   a. The appropriate statistical analysis procedure for each research question;
   b. Basics of statistical inference (e.g., Type I and II error);
   c. How to use a statistical software package for elementary analyses.
2) UNDERSTAND:
   a. Statistical estimation and hypothesis testing concepts;
   b. Statistical concepts encountered in the medical and public health literature.
3) APPLY:
   a. Data management principles in an approach to data;
   b. Basic software techniques for data analysis.
4) ANALYZE and INTERPRET:
   a. Results of statistical software procedures;
   b. Common statistics encountered in the medical and public health literature;
5) COMMUNICATE:
   a. Results of statistical analyses to a clinical audience.
Software: Class examples and data sets will use Stata 14; Stata is available for use on the campus virtual computer lab and the Odum Institute (in Davis Library). Personal copies of Stata can be purchased at student rates. (See Sakai for how to purchase Stata; also will be explained in class).

Additional Stata Resources:
- [www.stata.com](http://www.stata.com) (main website for Stata)
- [http://www.youtube.com/user/statacorp](http://www.youtube.com/user/statacorp) (Brief tutorials on the use of Stata)
- [www.ats.ucla.edu/stat/stata/](http://www.ats.ucla.edu/stat/stata/) (UCLA site with helpful Stata information and tutorials)

Alan Neustadl has the best Stata videos for both basic and more complex analyses—I highly recommend his Introduction to Stata and Visualizing Data with Stata series as well as Cleaning Data in Stata and Using Stata dialogs to Learn Stata commands. (His basic stats videos are good, too!)

SAS Users: You may choose to do your Exercises using SAS. SAS users should congregate into groups. We cannot teach you SAS, but we will provide SAS code necessary to complete most of the exercises. For additional practice, please see the SAS exercises provided with your textbook.

Turning in assignments: Upload exercises and analysis reports into the assignment section on Sakai. You will turn the final in on paper with an Honor Code attestation.

Part 1: Fundamentals of Statistical Theory, Approach to data analysis

Tues 8/22 – Class 1
- Lectures: PUBH741 1a-1c
- Topics: Class overview (format, schedule, syllabus, books, data sets, problems, exam, introduction to Stata)
- Definitions (descriptive/inferential statistics, types of data, graphs)
- Sort yourselves into groups (we will help!)

Thurs 8/24 – Demo 1: Stata Training Session
[https://www.youtube.com/watch?v=KkCKEK7lwuo&list=PLRYSxJ3XjgQM342QrBkzek8clHa5ue45d](https://www.youtube.com/watch?v=KkCKEK7lwuo&list=PLRYSxJ3XjgQM342QrBkzek8clHa5ue45d)

Tues 8/29 - Class 2
- Lectures: PUBH741 1d-1g
- Topics: Central tendency and dispersion, Probability, Binomial distribution
- Read Chapter 1: Biostatistics- an Applied Introduction for the Public Health Practitioner

Enhancements:
- Measures of central tendency - [https://www.youtube.com/watch?v=W0NwkiSjn8w](https://www.youtube.com/watch?v=W0NwkiSjn8w)
 Measures of dispersion - https://www.youtube.com/watch?v=8WQKz8ly3G4

**In class lecture: Probability distributions**

**Begin Exercise 1 (1): Approach to the dataset, Descriptive statistics**

**Thurs 8/31—Class 3 ** **Quiz 1**

**Lectures: 1h-1j**

**Topics:** Gaussian distributions

Central Limit Theorem

Statistical inference (population vs. sample, standard error)

**Enhancements:**

Introduction to z-scores, parts 1 and 2 (for understanding the theoretical underpinnings of z-scores. We will use software rather than tables to make the calculations.

https://www.youtube.com/watch?v=Obd6ZUZ_oh8

https://www.youtube.com/watch?v=oHjHQ0DFYmU

https://www.youtube.com/watch?v=Kzqm8F9Le_4&t=313s

Central Limit Theorem:

https://www.youtube.com/watch?v=wjVhOBC3Hgk&list=PLRYSxJ3XjgQN_cpQOMcvoU4HTi6-yCJQH&index=11

Sampling distribution of the mean:

http://onlinestatbook.com/2/sampling_distributions/samp_dist_mean.html

Excellent videos on key concepts for those who want more of the theoretical underpinnings:

https://www.youtube.com/results?search_query=jbstatistics

**In class lecture: Central limit theorem, z scores, and t-scores**

Continue Exercise 1 (2)

**Tues 9/05 — Class 4**

**Lectures: 1k-1n**

**Topics:** Hypothesis testing, Type I and Type II errors, confidence intervals, sample size

**Enhancements:**

Estimation:


Hypothesis Testing:

http://onlinestatbook.com/2/logic_of_hypothesis_testing/logic_hypothesis.html

https://www.youtube.com/watch?v=9UfAYdVxN7Q

Continue Exercise 1 (3)

**DEMO_2: Descriptive data analysis, Missing data, Using do files in Stata**

Skim Article 1 and Article 2 for Demo—we will look at Methods and Table 1 in class
Part 2: Bivariable Analysis

Thurs 9/07 – Class 5 **Quiz 2
Lectures: 2a, 2b, 2c, 2d,

Read Chapter 4: Biostatistics- an Applied Introduction for the Public Health Practitioner
Topics: 2 x 2 tables (Pearson’s chi-square, Fisher’s exact test); Measures of association (risk difference, risk ratios/odds ratios)

In class lecture: Hypothesis testing and confidence intervals

Complete Exercise 1 and Report (4)

Tues 9/12 – Class 6 **Exercise 1 due before class
Lectures: 2e, 2f, 2g, 2h, 2i

Topics: Stratified analysis, Sample size calculation (difference of two proportions, risk ratios, odds ratios), R x C tables, Kappa


Enhancements: Difference of two proportions:
https://www.youtube.com/watch?v=rMbOD1A4R1g&list=PLRYSxJ3XjgQN_cP0Mcv0U4H

DEMO_3: Review Exercise 1; Approach to analysis of categorical data in Stata

Begin Exercise 2: Bivariable analysis (1)

Thurs 9/14 – Class 7

Read Chapter 2: Biostatistics- an Applied Introduction for the Public Health Practitioner
Lectures: 2j, 2k, 2l

Topics: Comparing means of one group to standard; Comparing two groups (2 sample t-tests); Sample size calculation (difference of two means)

Continue Exercise 2 (2)

Tues 9/19 – Class 8

Enhancement: Effect size calculation (see
http://onlinestatbook.com/2/effect_size/two_means.html
http://blog.stata.com/tag/hedge-g/)

DEMO_4: Approach to bivariable analysis of continuous data in Stata

Continue Exercise 2 (3)

Thurs 9/21 – Class 9 ***Quiz 3***
Lectures: 2m, 2n, 2o

Topics: Comparing more than two groups (one-way analysis of variance – ANOVA)
Multiple testing, Two repeated observations (paired t-test)
Sample size calculation (paired t-test), 2-sample t-test of differences

Continue Exercise 2 and analysis report (4)
Part 3: Linear Regression Analysis

Tues 9/26 – Class 10
Lectures: 3a-3b
Topics: Simple linear regression, correlation
Read Chapter 3: Biostatistics- an Applied Introduction for the Public Health Practitioner
DEMO_5A: Approach to linear regression and correlation analyses
Complete Exercise 2 and analysis report (5)

Thurs 9/28– Class 11 ***Exercise 2 due ***
Lectures 3c, 5a, 5b
Topics: Multiple linear regression (model, interpretation of betas, partial F, predictive values)
DEMO: Dissecting a research paper and using immediate commands in Stata
Begin Exercise 3: Regression analysis with a continuous outcome (1)

Tues 10/3 – Class 12 ***Quiz 4***
Lectures: 3d
Topics: Multiple regression (indicator variables, polynomials, interaction terms)
DEMO_5B: Multiple linear regression and model diagnostics
Continue Exercise 3 (2) Pick out articles for review—Groups 1-5

Thurs 10/5 --Class 13
Lectures: 3f-3h
Analysis of covariance; Regression modeling strategies (ANCOVA)
DEMO_5C: Approach to an etiologic model: Analysis of Covariance.
Continue Exercise 3 (3)

Tues 10/10 – Class 14 ***Quiz 5***
REVIEW of Mean Difference procedures.
Continue Exercise 3 (4)

Thurs 10/12 – Class 15 (OPTIONAL—UNIVERSITY DAY) Complete Exercise 3 (5)

Tues 10/17– Class 16 **Exercise 3 due before class***
Article reviews—Groups 1-5 (You will sign in to a survey to evaluate)

Thurs 10/19-FALL BREAK

August 9, 2017
Part 4: Logistic Regression / Survival Analysis

Tues 10/24 – Class 17
Lectures: 4a, 4b
Topics: Logistic regression rationale, “predictive” models
Read Chapter 5: Biostatistics- an Applied Introduction for the Public Health Practitioner
Begin Exercise 4 (plus Analysis Report): Logistic Regression (1)
DEMO_6A: Approach to Logistic Regression

Thurs 10/26 – Class 18 ***Quiz 6 Logistic regression***
Lectures: 4c, 4d
Topics: Logistic regression – continued (modeling disease-exposure relationship, confounding, adjusted proportions)
DEMO_6B: Diagnostics for Logistic models
Continue Exercise 4 (2)

Tues 10/31 – Class 19
DEMO_6C: Review of modeling strategies
Continue Exercise 4 (Analysis Report) (3)

Thurs 11/2 – Class 20 ***Quiz 8 modeling strategies***
Lectures: 4e
Topics: Survival analysis (Censored data, survivor function, hazard functions)
Read Chapter 7: Biostatistics- an Applied Introduction for the Public Health Practitioner
DEMO_7A: Approach to survival analysis
Continue Exercise 4 (Analysis Report) (4)

Tues 11/07 – Class 21 (Kim away) **Exercise 4 due before class****
Review Exercise 4 and expectations for Analysis Reports
Begin Exercise 5 (Analysis Report) (1)

Thurs 11/9 – Class 22
Lectures: 4f, 4g
Topics: Survival analysis – continued (Kaplan-Meier plot, log-rank test)
Continue Exercise 5 (Analysis Report) (2)

Tues 11/14 – Class 23 ***Quiz 7 survival analysis***
Lectures: 4hi (Cox model examples, proportional hazards assumption, “predictive” Cox model) DEMO_7B: Evaluating assumptions of Cox models
Continue Exercise 5 (Analysis Report) (3)

Part 5: Miscellaneous Topics

August 9, 2017
Thurs 11/16– Class 24
   In class lecture: additional analyses you might see in the literature
   Pick out articles for review—Groups 6-10
Complete Exercise 5 (plus Analysis Report) (4)

Tues 11/21– Class 25 ***Exercise 5 Due***
   DEMO_9: >>>>Using DAGGITY to assess confounding<<<<

Thurs 11/24 – THANKSGIVING

Tues 11/28 – Class 26
   Article reviews—Groups 6-10

Thurs 11/30 –*** FINAL EXAM posted***
   DEMO_10: Course review (TAs)

Tues 12/05 –Instructors available to answer questions about final exam (in classroom)

TBA—****FINAL EXAM DUE**** To enable blinded grading, please print your honor code attestation with your PID. Print your completed exam Sakai. You will put the last 5 digits of your PID number on each page of your exam rather than your name.

Extras:
Two-way ANOVA:
https://www.youtube.com/watch?v=lZFmFuZGQTk
https://www.youtube.com/watch?v=cNII9bOnY
https://www.youtube.com/watch?v=ajLdnsLPErE

Poisson regression:
Lectures 4x-y
Chapter 6: Biostatistics- an Applied Introduction for the Public Health Practitioner